Matrix Theory

Linear algebra for 21-242 at Carnegie Mellon

21-242

A proof-based treatment of linear algebra written alongside 21-242 Matrix Theory. It follows the course closely and covers a good deal more; anything outside the course is marked.

Contents

1 Vector spaces 6
1.1 Fields 6
1.2 Vector spaces 21
1.3 Lists, linear combinations, and span 39
1.4 Subspaces 51
1.5 Linear independence 64
1.6 Bases and coordinates 80
1.7 Proof patterns 95
Problems for Chapter 1 101

2 Linear maps and matrices 106
2.1 Linear maps 106
2.2 Linear maps are determined by a basis 121
2.3 Matrices of linear maps 135
2.4 Matrix multiplication 151
2.5 Invertibility 174
2.6 Kernel and image 191
2.7 The Fundamental Theorems 207
2.8 Proof patterns 222
Problems for Chapter 2 230

3 Linear systems and row reduction 235
3.1 Linear systems 235
3.2 Elementary matrices 250
3.3 Gauss–Jordan elimination 266
3.4 Solving linear systems 291
3.5 Rank and the Fundamental Theorems 311
3.6 Proof patterns 331
Problems for Chapter 3 338

4 Dimension 344
4.1 Dimension 344
4.2 Rank–nullity 365
4.3 Rank inequalities and the transpose 386
4.4 Sums and direct sums 405
4.5 Isomorphisms and change of basis 425
4.6 Products and quotients 446
4.7 Affine geometry 466
4.8 Proof patterns 491

Problems for Chapter 4 498

5 Determinants 503
5.1 What a determinant should be 503
5.2 Permutations 516
5.3 Existence and uniqueness 535
5.4 Determinants and invertibility 552
5.5 Cofactor expansion 573
5.6 Proof patterns 596
Problems for Chapter 5 604

6 Inner product spaces 610
6.1 Inner products and norms 610
6.2 Cauchy–Schwarz and the triangle inequality 631
6.3 Orthogonality and orthonormal bases 649
6.4 Orthogonal projection and Gram–Schmidt 673
6.5 Least squares 699
6.6 Adjoints 722
6.7 Proof patterns 743
Problems for Chapter 6 750

7 Eigenvalues and diagonalization 756
7.1 Polynomials 756
7.2 Eigenvalues, eigenvectors, and invariant subspaces 782
7.3 The characteristic polynomial 801
7.4 Diagonalization 822
7.5 Upper-triangular form and Schur's theorem 840
7.6 Minimal polynomials, Cayley–Hamilton, and commuting operators 858
7.7 Proof patterns 881
Problems for Chapter 7 889

8 The spectral theorems 893
8.1 Self-adjoint and normal operators 893
8.2 The spectral theorems 908
8.3 Positive operators 928
8.4 Isometries and unitary operators 947
8.5 Proof patterns 970

9 The singular value decomposition 971
9.1 Singular values and the SVD 971
9.2 Consequences of the SVD 971
9.3 Matrix factorizations 971
9.4 Proof patterns 971

10 Generalized eigenspaces and Jordan form 972
10.1 Generalized eigenvectors and nilpotent operators 972
10.2 The generalized eigenspace decomposition 972
10.3 Jordan canonical form 972
10.4 Proof patterns 972

11 Markov chains and positive matrices 973
11.1 Stochastic matrices and Markov chains 973
11.2 Perron–Frobenius 973
11.3 Reversibility 973

11.4 Proof patterns 973
12 Bilinear and quadratic forms 974
12.1 Bilinear forms 974
12.2 Quadratic forms 974
12.3 Positive definiteness and min–max 974
12.4 Inner products beyond $\mathbb{R}$ and $\mathbb{C}$ 974
12.5 Proof patterns 974
13 Duality and multilinear algebra 975
13.1 Dual spaces 975
13.2 Duality in inner product spaces 975
13.3 Multilinear maps and tensors 975
13.4 Alternating forms and the determinant revisited 975
13.5 Proof patterns 975
14 Linear algebra and abstract algebra 976
14.1 Groups through matrices 976
14.2 Fields and field extensions 976
14.3 Finite fields 976
14.4 Modules over F[x] and canonical forms 976
14.5 Proof patterns 976
15 Infinite-dimensional vector spaces 977
15.1 Bases in infinite dimensions 977
15.2 Linear algebra inside infinite-dimensional spaces 977
15.3 Inner product spaces in infinite dimensions 977
15.4 Proof patterns 977
16 Applications to data and machine learning 978
16.1 Regression, revisited 978
16.2 Principal component analysis 978
16.3 Low-rank models 978
16.4 Optimization of quadratic objectives 978
16.5 Graphs and spectral methods 978
16.6 Linear layers 978
A Sets, functions, and proofs 979
A.1 Sets and functions 979
A.2 Induction and well-ordering 987
A.3 Equivalence relations 991
A.4 Zorn's lemma 996
B Complex numbers 1001
B.1 Complex numbers 1001
C The Fundamental Theorem of Algebra 1020
C.1 Proofs of the Fundamental Theorem of Algebra 1020
Index of techniques 1021
Solutions 1024

Reading this book

Proofs are preceded, where it helps, by what they use, the techniques they rely on, hints, and a sketch. The web version reveals these one at a time.

Proofs and exercises carry a difficulty:

ROUTINE	Follows directly from the definitions; a few lines.
STANDARD	Typical homework level; needs one real idea.
HARD	Upper end of course level; several ideas or a delicate argument.
CHALLENGE	Harder than the course expects. Skip it if you are targeting course-level problems.

Material marked BEYOND 21-242 goes past what the course covers and can be skipped.

Chapter 1

Vector spaces

Fields, the vector space axioms, span, subspaces, linear independence, bases, and coordinates.

1.1 Fields

In $\mathbb{R}$ we can add, subtract, multiply, and divide by any nonzero number. In $\mathbb{R}^2$ we can add two vectors and scale a vector by a real number, but there is no useful way to multiply or divide two vectors. Linear algebra keeps these roles apart. Vectors only need to be added and scaled, while the numbers that do the scaling, the scalars, need the full arithmetic of $\mathbb{R}$. Before we define vector spaces in Section 1.2, this section makes "full arithmetic" precise. The resulting notion, a field, is all we will ever ask of our scalars, so the theory applies verbatim over $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and number systems as unfamiliar as $\mathbb{F}_2$, where $1 + 1 = 0$.

1.1.1 The field axioms

Intuition (What we need from scalars). The basic computational problem of linear algebra is solving systems of linear equations. Take the smallest case: one equation $ax + c = d$ in one unknown, with $a \neq 0$. Over $\mathbb{R}$ we subtract c and divide by a without a second thought. Spelled out, subtracting c means adding $-c$ to both sides and simplifying the left side with

$$(ax + c) + (-c) = ax + (c + (-c)) = ax + 0 = ax,$$

and dividing by a means multiplying by a^{-1} and simplifying with

$$a^{-1}(ax) = \big(a^{-1}a\big)x = 1x = x.$$

Gauss–Jordan elimination (Section 3.3) solves systems of any size with the same moves: it adds a multiple of one equation to another and divides by nonzero pivots. So we want two operations that can be undone, that can be regrouped and reordered freely, that have neutral elements 0 and 1, and that are linked by distributivity, so that multiplying a sum means multiplying each term. The field axioms are exactly this list.

Definition 1.1.1 (Field). A **field** is a set F together with two operations, **addition** $+ : F \times F \to F$ and **multiplication** $\cdot : F \times F \to F$, that satisfy the following **field axioms**.

- *Commutativity.* For all $a, b \in F$, $a + b = b + a$ and $ab = ba$.
- *Associativity.* For all $a, b, c \in F$, $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$.
- *Distributivity.* For all $a, b, c \in F$, $a(b + c) = ab + ac$.
- *Additive identity.* There is an element $0 \in F$ such that $a + 0 = a$ for all $a \in F$.
- *Multiplicative identity.* There is an element $1 \in F$ with $1 \neq 0$ such that $a \cdot 1 = a$ for all $a \in F$.
- *Additive inverses.* For every $a \in F$ there is some $b \in F$ with $a + b = 0$.
- *Multiplicative inverses.* For every $a \in F$ with $a \neq 0$ there is some $b \in F$ with $ab = 1$.

The elements of F are called **scalars**. We write ab for $a \cdot b$.

Remark 1.1.2. Since $+$ and $\cdot$ are functions $F \times F \to F$, sums and products of elements of F lie in F. When we check that a candidate set is a field, this closure has to be checked along with the axioms.

We stated the identity, inverse, and distributive axioms on one side only. Commutativity supplies the other side: $0 + a = a$, $1a = a$, $b + a = 0$ and $ba = 1$ for the inverses b above, and

$$(a + b)c = c(a + b) = ca + cb = ac + bc.$$

We use these mirrored forms freely and cite them by the axiom they come from.

Finally, read the quantifiers carefully. The identity axioms say that *some* element 0 satisfies $a + 0 = a$ for every a, and similarly for 1. The axioms after them then refer to "0" and "1" as if these were determined. If a field could contain two different additive identities, the condition $a \neq 0$ would be ambiguous. Theorem 1.1.9 shows that this cannot happen, and the proof of its part (a) uses only commutativity and the identity axioms, so the argument is not circular.

Example 1.1.3 ($\mathbb{Q}$ and $\mathbb{R}$). $\mathbb{R}$, with its usual addition and multiplication, is a field; we take the arithmetic of real numbers as known. So is $\mathbb{Q}$: sums, products, and negatives of rational numbers are rational, $1/q$ is rational for every nonzero rational q, and identities such as $a(b + c) = ab + ac$ hold in $\mathbb{Q}$ because they hold in $\mathbb{R}$. (Proposition 1.1.33 makes this last argument general.)

Example 1.1.4 (The field with two elements). Let $\mathbb{F}_2 = \{0, 1\}$, with

$$0 + 0 = 1 + 1 = 0, \quad 0 + 1 = 1 + 0 = 1, \quad 1 \cdot 1 = 1, \quad 0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0.$$

These are the parity rules for integers if we read 0 as "even" and 1 as "odd": odd plus odd is even, odd times odd is odd, and so on. Both operations are commutative by inspection, 0 and 1 are identities with $0 \neq 1$, the additive inverse of 0 is 0 and that of 1 is 1 (as $1 + 1 = 0$), and 1 is its own multiplicative inverse. Associativity and distributivity can be checked case by case, or deduced from Lemma 1.1.21 below, since $\mathbb{F}_2$ is $\mathbb{Z}/2\mathbb{Z}$.

So a field's arithmetic can look quite unlike that of $\mathbb{R}$. An argument about general fields that silently assumes $1 + 1 \neq 0$ is not valid in $\mathbb{F}_2$.

Non-example 1.1.5 ($\mathbb{Z}$). The integers satisfy every field axiom except the existence of multiplicative inverses. That one fails for $a = 2$: for every $b \in \mathbb{Z}$ the product $2b$ is even, so $2b \neq 1$. Correspondingly, the equation $2x = 1$ has no solution in $\mathbb{Z}$.

Non-example 1.1.6 (Componentwise multiplication on $\mathbb{R}^2$). Give $\mathbb{R}^2$ the componentwise operations

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b)(c, d) = (ac, bd).$$

Every axiom except multiplicative inverses holds, because each holds in each coordinate separately; the identities are $(0, 0)$ and $(1, 1)$. But $(1, 0) \neq (0, 0)$ has no inverse, since $(1, 0)(c, d) = (c, 0) \neq (1, 1)$ for every (c, d). So the obvious way to multiply vectors in $\mathbb{R}^2$ does not give a field. Section 1.1.4 gives $\mathbb{R}^2$ a less obvious multiplication that does.

Intuition (What each axiom is for). Here is what goes wrong when one axiom is dropped.

- Without multiplicative inverses we cannot divide by pivots. In $\mathbb{Z}$, the equation $2x = 1$ has no solution.
- Without additive inverses we cannot subtract one equation from another. In $\{0, 1, 2, \ldots\}$, the equation $x + 1 = 0$ has no solution.
- Without associativity, multiplying $ax = b$ by a^{-1} gives $a^{-1}(ax) = a^{-1}b$, and we could not regroup the left side as $\left(a^{-1}a\right)x = x$.
- Without commutativity of multiplication, $ax = b$ and $xa = b$ are different equations. Square matrices multiply like this (Section 2.4).
- Without distributivity, nothing links $+$ to $\cdot$, since it is the only axiom that mentions both. Even $0a = 0$ (Proposition 1.1.15) can fail: Exercise 1.1.55 builds a structure satisfying every other axiom in which $0 \cdot 0 \neq 0$.
- Without $1 \neq 0$, the one-element set $\{0\}$ with $0 + 0 = 0 \cdot 0 = 0$ would be a field. It would also be the only new example, since $1 = 0$ forces $F = \{0\}$ (Exercise 1.1.39).

1.1.2 Consequences of the axioms

From now on, F denotes a field. All we know about F is the axioms, so facts that are second nature in $\mathbb{R}$, such as $0a = 0$ and $(-a)(-b) = ab$, need proofs. In exchange, each such proof establishes the fact in every field at once.

Technique 1.1.7 (Proving from axioms). When objects are defined by axioms, as fields are here and vector spaces will be in Section 1.2, the only facts available are the axioms and what has already been proved from them. A fact that is obvious for real numbers is not available until it is proved, because a general field need not look like $\mathbb{R}$.

The move is to write the argument as a chain of equalities that changes one thing at a time, and to justify each link by the axiom or earlier result it uses. Watch for steps that move a term past another (commutativity) or regroup (associativity): in $\mathbb{R}$ we make these without noticing.

Reach for it whenever the hypotheses are an axiom system and the claim is an identity or an elementary property. If you are stuck, ask which axioms can connect the objects in the claim. For example, distributivity is the only axiom that involves both $+$ and $\cdot$, so any fact relating 0 (an additive object) to multiplication must go through it. The canonical example is $0a = 0$ (Proposition 1.1.15).

Used in. Theorem 1.1.9, Proposition 1.1.13, Proposition 1.1.15, Proposition 1.1.16, Proposition 1.1.17, Proposition 1.1.30, Proposition 1.1.33, Exercise 1.1.38, Exercise 1.1.39, Exercise 1.1.58, Proposition 1.2.5, Proposition 1.2.10, Proposition 1.2.22, Proposition 1.2.27, Proposition 1.2.28, Proposition 1.2.30, Proposition 1.2.32, Proposition 1.2.33, Proposition 1.2.34, Proposition 1.2.35, Exercise 1.2.54, Exercise 1.2.55, Exercise 1.2.56, Exercise 1.7.14, Problem 1.2, Lemma 2.1.22, Proposition 4.6.2, Theorem 4.6.22, Proposition 6.1.4.

Uniqueness

The identity and inverse axioms assert that certain elements exist, not that they are unique. Uniqueness is what entitles us to speak of *the* zero of F and *the* inverse of a.

Technique 1.1.8 (Uniqueness via two candidates). To show that at most one object has a property P, suppose x and x' both have P and prove $x = x'$. Usually one evaluates a single well-chosen expression in two ways, once using that x has P and once using that x' has P.

Reach for it whenever a claim says "the", "unique", or "exactly one". Existence and uniqueness are separate claims: prove existence by producing an object, and uniqueness by this technique. The canonical example is the uniqueness of 0 in Theorem 1.1.9, where the expression is $0 + 0'$.

Used in. Theorem 1.1.9, Exercise 1.1.11, Corollary 1.1.14, Proposition 1.1.16, Exercise 1.1.37, Exercise 1.1.46, Exercise 1.1.47, Proposition 1.2.27, Proposition 1.2.28, Proposition 1.2.30, Proposition 1.2.34, Exercise 1.2.43, Exercise 1.2.52, Proposition 1.6.16, Exercise 1.7.6, Exercise 1.7.7, Exercise 1.7.11, Problem 1.3, Problem 1.17, Proposition 2.1.4, Theorem 2.2.1, Proposition 2.3.10, Proposition 2.3.17, Proposition 2.4.3, Theorem 2.5.10, Exercise 2.5.37, Exercise 2.5.49, Exercise 2.6.42, Exercise 2.6.43, Theorem 3.3.17, Corollary 3.3.18, Proposition 3.4.27, Problem 3.2, Problem 3.17, Exercise 4.3.44, Proposition 4.4.31, Proposition 4.5.16, Proposition 4.6.7, Theorem 4.6.30, Theorem 4.7.22, Proposition 4.7.32, Proposition 5.2.5, Proposition 5.3.17, Theorem 6.4.6, Theorem 6.4.24, Theorem 6.4.38, Proposition 6.5.27, Theorem 6.6.3, Proposition 6.6.9, Theorem 7.1.14, Corollary 7.1.17, Proposition 7.1.29, Theorem 7.1.45, Exercise 7.1.51, Theorem 7.6.2, Notation 8.3.10, Theorem 8.3.23, Theorem A.1.14, Proposition A.2.4, Theorem A.3.10, Proposition B.1.2, Exercise 2.6.49, Exercise 4.6.39, Exercise 4.6.54, Exercise 6.6.37, Exercise 7.1.58, Exercise A.2.9.

Theorem 1.1.9 (Uniqueness of identities and inverses). Let F be a field.

(a) F has exactly one additive identity and exactly one multiplicative identity.

(b) Every $a \in F$ has exactly one additive inverse, and every nonzero $a \in F$ has exactly one multiplicative inverse.

Proof. ROUTINE

Uses. Definition 1.1.1 (Field)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)
Hint 1. Suppose 0 and $0'$ are both additive identities. Find one expression that involves both.
Hint 2. Evaluate $0 + 0'$ twice: once using that $0'$ is an identity, once using that 0 is.
Hint 3. For inverses, suppose $a + b = 0$ and $a + b' = 0$. Start from $b = b + 0$, replace 0 by $a + b'$, and regroup.
Sketch. Each identity absorbs the other: $0 = 0 + 0' = 0' + 0 = 0'$. If b and b' are additive inverses of a, insert $0 = a + b'$ into $b = b + 0$ and regroup to get $(b + a) + b' = 0 + b' = b'$. The multiplicative statements are the same arguments with $\cdot$ and 1 in place of $+$ and 0.

(a) Suppose 0 and $0'$ both satisfy the additive identity axiom. Then

$$\begin{aligned} 0 &= 0 + 0' && (\text{as } \ 0' \ \text{ is an additive identity}) \\ &= 0' + 0 && (\text{commutativity}) \\ &= 0' && (\text{as } \ 0 \ \text{ is an additive identity}). \end{aligned}$$

The same argument, with 1 and $1'$ in place of 0 and $0'$ and multiplication in place of addition, shows that the multiplicative identity is unique. Realize that we used only commutativity and the identity axioms. So from here on "0" and "1" are unambiguous, and the remaining axioms say what they were meant to say.

(b) Let $a \in F$, and suppose $a + b = 0$ and $a + b' = 0$. Then

$$\begin{aligned} b &= b + 0 && (\text{additive identity}) \\ &= b + (a + b') && (\text{as } \ a + b' = 0\ ) \\ &= (b + a) + b' && (\text{associativity}) \\ &= 0 + b' && (\text{as } \ b + a = a + b = 0\ ) \\ &= b' && (\text{additive identity}), \end{aligned}$$

as desired. If $a \neq 0$ and $ab = ab' = 1$, the same chain with multiplication in place of addition and 1 in place of 0 gives $b = b'$. □

Notation 1.1.10. By Theorem 1.1.9 we may write $-a$ for the additive inverse of a, and a^{-1} for the multiplicative inverse of $a \neq 0$. **Subtraction** and **division** are defined by

$$a - b = a + (-b), \quad \frac{a}{b} = ab^{-1} \quad (b \neq 0).$$

For a positive integer n, a^n is the product of n copies of a, and $a^0 = 1$.

Exercise 1.1.11. ROUTINE Show that $-0 = 0$, and that $-(a+b) = (-a)+(-b)$ for all $a, b \in F$.

Solution on page 1024

Remark 1.1.12. Regrouping arguments like the solution of Exercise 1.1.11 get long, so we record a convenience. Associativity implies that all ways of parenthesizing a sum $a_1 + a_2 + \cdots + a_n$ give the same element: by strong induction on n, each of them equals $(\cdots((a_1 + a_2) + a_3) + \cdots) + a_n$. Commutativity then implies that the order of the terms does not matter either, since any reordering is a sequence of swaps of adjacent terms. The same holds for products. From now on we drop such parentheses and reorder sums and products freely.

Arithmetic in a field

We now return to the problem that motivated the definition.

Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$). Let $a, b \in F$.

(a) The equation $a + x = b$ has exactly one solution $x \in F$, namely $x = b - a$.

(b) If $a \neq 0$, the equation $ax = b$ has exactly one solution $x \in F$, namely $x = b/a$.

Proof. ROUTINE

Uses. Definition 1.1.1 (Field), Notation 1.1.10
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Existence and uniqueness are separate claims. For existence, substitute the proposed solution.
Hint 2. For uniqueness, take any solution x of $a + x = b$, add $-a$ to both sides on the left, and regroup to isolate x.
Sketch. Substituting $x = b - a$ and regrouping gives $a + (b - a) = (a + (-a)) + b = b$. Conversely, if $a + x = b$, then $x = ((-a) + a) + x = (-a) + (a + x) = (-a) + b = b - a$. Part (b) is the same argument with a^{-1} and 1 in place of $-a$ and 0.

(a) First, $x = b - a$ is a solution:

$$\begin{aligned} a + (b - a) &= a + ((-a) + b) && \text{(definition of subtraction, commutativity)} \\ &= (a + (-a)) + b && \text{(associativity)} \\ &= 0 + b = b && \text{(additive inverse, additive identity).} \end{aligned}$$

Conversely, suppose $a + x = b$. Then

$$\begin{aligned} x &= 0 + x && \text{(additive identity)} \\ &= ((-a) + a) + x && \text{(additive inverse)} \\ &= (-a) + (a + x) && \text{(associativity)} \\ &= (-a) + b = b - a && \text{(as } a + x = b \text{, commutativity),} \end{aligned}$$

so $b - a$ is the only solution.

(b) Suppose $a \neq 0$, so that a^{-1} exists. Then $x = b/a = ba^{-1}$ is a solution, since by commutativity, associativity, and the inverse and identity axioms,

$$a\big(ba^{-1}\big) = a\big(a^{-1}b\big) = \big(aa^{-1}\big)b = 1b = b.$$

Conversely, if $ax = b$, the same axioms give

$$x = 1x = \big(a^{-1}a\big)x = a^{-1}(ax) = a^{-1}b = b/a,$$

so b/a is the only solution. □

Part (b) is what the section set out to guarantee: in any field, a linear equation in one unknown with a nonzero coefficient has exactly one solution. The uniqueness half is often used in the following form.

Corollary 1.1.14 (Cancellation). Let $a, b, c \in F$. If $a + b = a + c$, then $b = c$. If $ab = ac$ and $a \neq 0$, then $b = c$.

Proof. ROUTINE

Uses. Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Both b and c solve one equation of the kind in Proposition 1.1.13. Which one?
Hint 2. They both solve $a + x = d$, where $d = a + b$. Apply the uniqueness in part (a) of Proposition 1.1.13.
Sketch. b and c are two solutions of $a + x = a + b$, which has only one solution. Likewise b and c both solve $ax = ab$ when $a \neq 0$.

Suppose $a + b = a + c$. Then b and c are both solutions of the equation $a + x = a + b$, which has exactly one solution by part (a) of Proposition 1.1.13. Hence $b = c$. If $ab = ac$ with $a \neq 0$, then b and c both solve $ax = ab$, and part (b) of Proposition 1.1.13 gives $b = c$. □

Intuition. No axiom says anything about a product with 0. The only fact we have about 0 is additive, namely $0 + 0 = 0$, and distributivity is the only axiom that converts additive facts into multiplicative ones. So multiply $0 + 0 = 0$ by a and see what distributivity says.

Proposition 1.1.15 (Multiplying by zero). For every $a \in F$, $0a = a0 = 0$.

Proof. ROUTINE

Uses. Definition 1.1.1 (Field), Corollary 1.1.14 (Cancellation)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Which axiom could say anything about a product involving 0? Only one axiom mentions both $+$ and $\cdot$.
Hint 2. Multiply both sides of $0 + 0 = 0$ by a and expand with distributivity.
Hint 3. You get $a0 + a0 = a0$. Compare it with $a0 + 0 = a0$ and cancel.
Sketch. Distributivity turns $0 + 0 = 0$ into $a0 + a0 = a0 = a0 + 0$. Cancelling $a0$ (Corollary 1.1.14) gives $a0 = 0$, and commutativity gives $0a = 0$.

Let $a \in F$. Then

$$\begin{aligned} a0 + a0 &= a(0 + 0) && \text{(distributivity)} \\ &= a0 && \text{(additive identity)} \\ &= a0 + 0 && \text{(additive identity).} \end{aligned}$$

Cancelling $a0$ from both sides (Corollary 1.1.14) gives $a0 = 0$. By commutativity, $0a = a0 = 0$. □

The proof converted the additive fact $0 + 0 = 0$ into the multiplicative fact $a0 + a0 = a0$. The same trick shows that $0\mathbf{v} = \mathbf{0}$ for vectors in Section 1.2.

Intuition. Every sign rule below comes from one observation: to show that an element equals $-x$, it is enough to show that it is *an* additive inverse of x, since there is only one. In particular $(-1)a$ is an additive inverse of a, and once $-a = (-1)a$ is known, signs are just factors of -1 that can be moved around.

Proposition 1.1.16 (Sign rules). For all $a, b \in F$:

(a) $-(-a) = a$;

(b) $(-1)a = -a$;

(c) $(-a)b = -(ab) = a(-b)$;

(d) $(-a)(-b) = ab$.

Proof. ROUTINE

Uses. Theorem 1.1.9 (Uniqueness of identities and inverses), Proposition 1.1.15 (Multiplying by zero), Definition 1.1.1 (Field)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)

Hint 1. To show that an element equals $-x$, show that it is an additive inverse of x and quote Theorem 1.1.9.
Hint 2. For (b), compute $a + (-1)a$ by writing $a = 1a$ and using distributivity and Proposition 1.1.15.
Hint 3. For (c) and (d), rewrite each negative as (-1) times something using (b), and move the factors -1 around with associativity and commutativity.
Sketch. a is an additive inverse of $-a$, which gives (a). Distributivity and $0a = 0$ give $a + (-1)a = (1 + (-1))a = 0$, which gives (b). Then (c) is (b) plus associativity and commutativity, and (d) is (c) twice followed by (a).

(a) Since $(-a) + a = a + (-a) = 0$, the element a is an additive inverse of $-a$. By Theorem 1.1.9, $a = -(-a)$.

(b) We compute

$$\begin{aligned} a + (-1)a &= 1a + (-1)a && \text{(multiplicative identity)} \\ &= (1 + (-1))a && \text{(distributivity)} \\ &= 0a && \text{(additive inverse)} \\ &= 0, \end{aligned}$$

where the last step is Proposition 1.1.15. So $(-1)a$ is an additive inverse of a, and $(-1)a = -a$ by Theorem 1.1.9.

(c) By (b), associativity, and (b) again, $(-a)b = ((-1)a)b = (-1)(ab) = -(ab)$. Using commutativity and what we just proved, $a(-b) = (-b)a = -(ba) = -(ab)$.

(d) By (c) twice and then (a), $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$. □

In particular $(-1)(-1) = 1$, and $(a - b)c = ac - bc$ by distributivity and (c). One more consequence is used constantly, for instance whenever we factor an expression and set each factor equal to zero.

Proposition 1.1.17 (No zero divisors). If $a, b \in F$ and $ab = 0$, then $a = 0$ or $b = 0$. Equivalently, a product of nonzero elements of F is nonzero.

Proof. ROUTINE
Uses. Definition 1.1.1 (Field), Proposition 1.1.15 (Multiplying by zero)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. To prove "$a = 0$ or $b = 0$", assume one of them is nonzero and prove that the other is zero.
Hint 2. If $a \neq 0$, then a^{-1} exists. Multiply $ab = 0$ by it.
Sketch. If $a \neq 0$, then $b = 1b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0$.

Suppose $ab = 0$. If $a = 0$ we are done, so suppose $a \neq 0$. Then a^{-1} exists, and

$$\begin{aligned} b &= 1b && \text{(multiplicative identity)} \\ &= (a^{-1}a)b && \text{(multiplicative inverse)} \\ &= a^{-1}(ab) && \text{(associativity)} \\ &= a^{-1}0 && (\text{as } ab = 0) \\ &= 0, \end{aligned}$$

where the last step is Proposition 1.1.15. So $a = 0$ or $b = 0$. □

The integers also have no zero divisors, so this property alone does not make a field. In Section 1.1.3 we meet number systems that have zero divisors, and there the uniqueness in part (b) of Proposition 1.1.13 fails.

Exercise 1.1.18. ROUTINE Find all $x \in F$ with $x^2 = x$.

Solution on page 1024

1.1.3 The fields $\mathbb{F}_p$

$\mathbb{F}_2$ is the first of an infinite family of finite fields, built from arithmetic with remainders.

Definition 1.1.19 (Integers modulo n). Let $n \geq 2$ be an integer. For $x \in \mathbb{Z}$, write $x \bmod n$ for the remainder of x on division by n, that is, the unique $r \in \{0, 1, \ldots, n-1\}$ such that $n \mid x - r$. The **integers modulo** n are the set $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ with the operations

$$a +_n b = (a + b) \bmod n, \quad a \cdot_n b = (ab) \bmod n.$$

This is clock arithmetic: in $\mathbb{Z}/12\mathbb{Z}$, $9 +_{12} 5 = 2$ and $5 \cdot_{12} 5 = 1$. Outside of proofs we drop the subscripts and write $a + b$ and ab, once it is clear which n we work with.

All the field axioms except multiplicative inverses hold in $\mathbb{Z}/n\mathbb{Z}$ for every n. Structures like this are common enough to have a name.

Definition 1.1.20 (Commutative ring). A **commutative ring** is a set with an addition and a multiplication that satisfy all the field axioms, except possibly the existence of multiplicative inverses and the condition $1 \neq 0$.

So every field is a commutative ring. The integers and componentwise $\mathbb{R}^2$ (Non-example 1.1.6) are commutative rings that are not fields, and so are polynomials with coefficients in a field (Section 7.1).

Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$). Let $n \geq 2$.

(a) For all $x, y \in \mathbb{Z}$, $(x \bmod n) +_n (y \bmod n) = (x + y) \bmod n$ and $(x \bmod n) \cdot_n (y \bmod n) = (xy) \bmod n$.

(b) $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring with identities 0 and 1. The additive inverse of $a \neq 0$ is $n - a$.

Proof. STANDARD

Uses. Definition 1.1.19 (Integers modulo n), Definition 1.1.20 (Commutative ring), the division algorithm: $x \bmod n = y \bmod n$ if and only if $n \mid x - y$

Hint 1. Every axiom except inverses is either an identity, like associativity, or asks for an identity element. The identities hold in $\mathbb{Z}$; you need a way to carry them over to $\mathbb{Z}/n\mathbb{Z}$.

Hint 2. For (a), x and $x \bmod n$ differ by a multiple of n. What does that say about $(x + y)$ versus $(x \bmod n) + (y \bmod n)$?

Hint 3. For (b), use (a) to write both sides of each identity in $\mathbb{Z}/n\mathbb{Z}$ as the remainder of the corresponding side of the same identity in $\mathbb{Z}$.

Sketch. Since x and $x \bmod n$ differ by a multiple of n, so do $x + y$ and $(x \bmod n) + (y \bmod n)$, and likewise for products; this is (a). By (a), each side of an identity such as associativity in $\mathbb{Z}/n\mathbb{Z}$ is the remainder of the corresponding side in $\mathbb{Z}$, where the identity holds. The identities are 0 and 1, and $a +_n (n - a) = 0$.

Throughout we use the fact, from the division algorithm, that $x \bmod n = y \bmod n$ if and only if $n \mid x - y$. In particular $n \mid x - (x \bmod n)$ for every $x \in \mathbb{Z}$.

(a) Let $x, y \in \mathbb{Z}$, and write $x' = x \bmod n$ and $y' = y \bmod n$. Then

$$(x + y) - (x' + y') = (x - x') + (y - y')$$

is a sum of multiples of n, so $(x' + y') \bmod n = (x + y) \bmod n$. That is, $x' +_n y' = (x + y) \bmod n$. Similarly,

$$xy - x'y' = (x - x')y + x'(y - y')$$

is a multiple of n, so $x' \cdot_n y' = (xy) \bmod n$.

(b) Every element of $\mathbb{Z}/n\mathbb{Z}$ is its own remainder, $a \bmod n = a$, so (a) applies to elements of $\mathbb{Z}/n\mathbb{Z}$ directly. Let $a, b, c \in \mathbb{Z}/n\mathbb{Z}$. For associativity of addition, (a) gives

$$(a +_n b) +_n c = ((a + b) \bmod n) +_n (c \bmod n) = ((a + b) + c) \bmod n,$$

and likewise $a +_n (b +_n c) = (a + (b + c)) \bmod n$. These agree because $(a + b) + c = a + (b + c)$ in $\mathbb{Z}$. Associativity of $\cdot_n$, and commutativity of both operations, follow in the same way from the corresponding laws in $\mathbb{Z}$. For distributivity, by (a) twice,

$$a \cdot_n (b +_n c) = (a(b + c)) \bmod n = (ab + ac) \bmod n = (a \cdot_n b) +_n (a \cdot_n c).$$

Next, $a +_n 0 = a \bmod n = a$ and $a \cdot_n 1 = a \bmod n = a$, and $0 \neq 1$ because $n \geq 2$. Finally, $0 +_n 0 = 0$, and for $1 \leq a \leq n-1$ the integer $n - a$ lies in $\mathbb{Z}/n\mathbb{Z}$ and $a +_n (n-a) = n \bmod n = 0$. So every axiom of a commutative ring holds. □

Remark 1.1.22. Look back at which results of Section 1.1.2 used multiplicative inverses. Uniqueness of the identities and of additive inverses (Theorem 1.1.9), part (a) of Proposition 1.1.13, additive cancellation, Proposition 1.1.15, and Proposition 1.1.16 did not, so they hold in every commutative ring with the same proofs. We will use this in Section 7.3, where determinants have polynomial entries. Proposition 1.1.17 and multiplicative cancellation did use inverses, and Non-example 1.1.23 shows that they can fail.

Compare the multiplication tables of $\mathbb{Z}/5\mathbb{Z}$ and $\mathbb{Z}/6\mathbb{Z}$.

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Table 1.1.1. Multiplication in $\mathbb{Z}/5\mathbb{Z}$. Every nonzero row contains each element exactly once, in particular a 1 (highlighted), so every nonzero element has an inverse.

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Table 1.1.2. Multiplication in $\mathbb{Z}/6\mathbb{Z}$. The rows of 2, 3, and 4 contain 0 in a nonzero column (highlighted) and never contain 1.

Non-example 1.1.23 ($\mathbb{Z}/6\mathbb{Z}$). In $\mathbb{Z}/6\mathbb{Z}$ we have $2 \cdot 3 = 6 \bmod 6 = 0$ with $2 \neq 0$ and $3 \neq 0$. Fields have no zero divisors (Proposition 1.1.17), so $\mathbb{Z}/6\mathbb{Z}$ is not a field. The failure is visible in Table 1.1.2: the row of 2 is $0, 2, 4, 0, 2, 4$, which never contains 1, so 2 has no inverse. Linear equations misbehave accordingly: $2x = 1$ has no solution, and $2x = 4$ has two, $x = 2$ and $x = 5$.

Intuition. By Lemma 1.1.21, the only question is which nonzero a have inverses, that is, which rows of the multiplication table contain a 1. If $n = rs$ with $1 < r, s < n$, then $rs = 0$ in $\mathbb{Z}/n\mathbb{Z}$, a zero divisor, which no field has. If n is prime, we give two proofs that every nonzero a is invertible. The first constructs a^{-1} from Bézout's identity. The second never finds a^{-1}: it shows that the row of a has no repeated entries, and then counting forces every element, 1 included, to appear in it.

Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime). Let $n \geq 2$ be an integer. Then $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.

Proof (via Bézout's identity). STANDARD

Uses. Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.17 (No zero divisors), Bézout's identity: if $\gcd(a, n) = 1$, then $ax + ny = 1$ for some $x, y \in \mathbb{Z}$
Hint 1. By Lemma 1.1.21, only multiplicative inverses are in question. If $n = rs$ with $1 < r, s < n$, look at $r \cdot_n s$.
Hint 2. For $n = p$ prime, an inverse of a is an integer b with $p \mid ab - 1$, that is, $ab + py = 1$ for some integer y. Which theorem produces such integers?
Hint 3. Bézout's identity applies because $\gcd(a, p) = 1$. Reduce the resulting x modulo p.
Sketch. If $n = rs$ with $1 < r, s < n$, then $r \cdot_n s = 0$ is a product of nonzero elements, impossible in a field. If $n = p$ is prime and $0 < a < p$, then $\gcd(a, p) = 1$, so $ax + py = 1$ for some integers x, y, and $x \bmod p$ is an inverse of a by part (a) of Lemma 1.1.21.

By Lemma 1.1.21, $\mathbb{Z}/n\mathbb{Z}$ satisfies every field axiom except possibly the existence of multiplicative inverses. So it is a field if and only if every nonzero element has a multiplicative inverse.

(⇒) We prove the contrapositive. Suppose n is not prime. Since $n \geq 2$, we can write $n = rs$ with integers $1 < r, s < n$. Then r and s are nonzero elements of $\mathbb{Z}/n\mathbb{Z}$ with $r \cdot_n s = n \bmod n = 0$. A field has no zero divisors (Proposition 1.1.17), so $\mathbb{Z}/n\mathbb{Z}$ is not a field.

(⇐) Suppose $n = p$ is prime, and let $a \in \{1, \ldots, p - 1\}$. The only positive divisors of p are 1 and p, and $p \nmid a$ because $0 < a < p$, so $\gcd(a, p) = 1$. By Bézout's identity there are integers x, y with $ax + py = 1$. Let $b = x \bmod p$. By part (a) of Lemma 1.1.21,

$$a \cdot_p b = (a \bmod p) \cdot_p (x \bmod p) = (ax) \bmod p = (1 - py) \bmod p = 1,$$

where the last step holds because $1 - py$ and 1 differ by a multiple of p. So b is a multiplicative inverse of a, as desired. □

Proof (by counting). STANDARD

Uses. Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.17 (No zero divisors), Euclid's lemma: if p is prime and $p \mid ac$, then $p \mid a$ or $p \mid c$, an injective function from a finite set to itself is surjective
Hint 1. Fix $a \neq 0$ and look at its row in the multiplication table, that is, the function $x \mapsto a \cdot_p x$. You want 1 to appear in it.
Hint 2. Show that the row has no repeated entries: if $a \cdot_p x = a \cdot_p y$, then $p \mid a(x - y)$. Use Euclid's lemma.
Hint 3. A row with p entries taken from a set of p elements, with no repeats, contains every element.
Sketch. The direction (⇒) is as in the first proof. For p prime and $a \neq 0$, the map $x \mapsto a \cdot_p x$ on $\mathbb{Z}/p\mathbb{Z}$ is injective, since $p \mid a(x - y)$ and $p \nmid a$ force $p \mid x - y$, hence $x = y$. An injective map from a finite set to itself is surjective, so $a \cdot_p b = 1$ for some b.

The direction (⇒) is proved as in the first proof. For (⇐), let p be prime and $a \in \{1, \ldots, p - 1\}$, and define $m : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ by $m(x) = a \cdot_p x$. We will show that m is injective. Suppose $m(x) = m(y)$, that is, $(ax) \bmod p = (ay) \bmod p$. Then $p \mid ax - ay = a(x - y)$. Since p is prime and $p \nmid a$, Euclid's lemma gives $p \mid x - y$. But $x, y \in \{0, \ldots, p - 1\}$, so $|x - y| < p$, and the only multiple of p strictly between $-p$ and p is 0. Hence $x = y$.

So m is an injective function from the finite set $\mathbb{Z}/p\mathbb{Z}$ to itself. Its image then has p elements, so it is all of $\mathbb{Z}/p\mathbb{Z}$, and in particular $1 = m(b) = a \cdot_p b$ for some b. That is, b is a multiplicative inverse of a. □

The two proofs differ in what they produce. The first computes a^{-1}, and running the Euclidean algorithm to find Bézout's x is how one inverts elements of $\mathbb{F}_p$ in practice. The second only proves that a^{-1} exists, but it uses just two facts about $\mathbb{Z}/p\mathbb{Z}$: it is finite, and $ax = ay$ with $a \neq 0$ forces $x = y$. On a finite set, a function to itself is injective if and only if it is surjective, so a multiplication that can be cancelled can also be undone. The same principle returns in Section 4.2 with dimension in place of the number of elements: a linear map from a finite-dimensional space to itself is injective if and only if it is surjective.

Notation 1.1.25 ($\mathbb{F}_p$). For a prime p, the field $\mathbb{Z}/p\mathbb{Z}$ is denoted $\mathbb{F}_p$. Later chapters use a bare $\mathbb{F}$ to mean $\mathbb{R}$ or $\mathbb{C}$; $\mathbb{F}_p$ always carries its subscript, so the two uses never clash.

Example 1.1.26 (Arithmetic in $\mathbb{F}_5$). From Table 1.1.1, the inverses in $\mathbb{F}_5$ are $1^{-1} = 1$, $2^{-1} = 3$, $3^{-1} = 2$, and $4^{-1} = 4$. To solve $3x + 4 = 1$ in $\mathbb{F}_5$, follow Proposition 1.1.13: first $3x = 1 - 4 = -3 = 2$, and then $x = 3^{-1} \cdot 2 = 2 \cdot 2 = 4$. Check: $3 \cdot 4 + 4 = 16 \bmod 5 = 1$.

Exercise 1.1.27. ROUTINE Find the multiplicative inverse of every nonzero element of $\mathbb{F}_7$. Then solve $3x + 5 = 2$ in $\mathbb{F}_7$.

Solution on page 1024

ASIDE: *Characteristic*

In $\mathbb{F}_p$, adding 1 to itself p times gives 0, while in $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ no number of 1's adds up to 0. For a positive integer m and $a \in F$, write $m \cdot a$ for $a + a + \cdots + a$ (m terms). This is repeated addition, not multiplication by an element of F.

Definition 1.1.28 (Characteristic). The **characteristic** of a field F, written $\operatorname{char} F$, is the least positive integer m such that $m \cdot 1 = 0$, or 0 if there is no such m.

So $\operatorname{char} \mathbb{F}_p = p$ and $\operatorname{char} \mathbb{Q} = \operatorname{char} \mathbb{R} = \operatorname{char} \mathbb{C} = 0$. A nonzero characteristic is always prime (Exercise 1.1.48). If $\operatorname{char} F = p > 0$, then $p \cdot a = (p \cdot 1)a = 0$ for every $a \in F$, by distributivity.

The case that matters most later is characteristic 2, where $1 + 1 = 0$. There $-a = a$ for every a, and we cannot divide by $1 + 1$. Some later arguments do divide by $1 + 1$: the converse of "alternating implies antisymmetric" for determinants (Section 5.1), and the passage between quadratic forms and symmetric matrices (Section 12.2). Those results assume $1 + 1 \neq 0$, and $\mathbb{F}_2$ is where to look for counterexamples when the assumption is dropped.

1.1.4 The complex numbers

Componentwise multiplication does not make $\mathbb{R}^2$ a field (Non-example 1.1.6), but a different multiplication does. We give only what the next few chapters need: the definition and the field axioms. Conjugation, the modulus, polar form, and roots of unity are developed in Section B.1.

Definition 1.1.29 (Complex numbers). The **complex numbers** are the set $\mathbb{C} = \mathbb{R}^2$, whose element (a, b) we write as $a + bi$, with the operations

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

The **real part** of $a + bi$ is a and its **imaginary part** is b. We identify the real number a with $a + 0i$, and write bi for $0 + bi$ and i for $0 + 1i$.

Intuition. The multiplication rule is forced. Suppose we want a field that contains $\mathbb{R}$ and an element i with $i^2 = -1$. Then the field axioms leave no choice:

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

So the definition is the only candidate, and what remains is to check that it satisfies the axioms. In particular $i^2 = (0 \cdot 0 - 1 \cdot 1) + (0 \cdot 1 + 1 \cdot 0)i = -1$.

The identification of $a \in \mathbb{R}$ with $a + 0i$ respects the operations: $(a + 0i) + (c + 0i) = (a + c) + 0i$ and $(a + 0i)(c + 0i) = ac + 0i$. Also $a(c + di) = ac + adi$, so multiplying by a real number scales both coordinates.

Proposition 1.1.30 ($\mathbb{C}$ is a field). $\mathbb{C}$ is a field, with identities $0 = 0 + 0i$ and $1 = 1 + 0i$. The additive inverse of $a + bi$ is $-a - bi$, and if $a + bi \neq 0$, then

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i.$$

Proof, part 1 of 2: the algebraic laws. ROUTINE

> *Uses.* Definition 1.1.29 (Complex numbers), the field axioms of $\mathbb{R}$
> *Techniques.* Proving from axioms (Technique 1.1.7)
> *Hint 1.* Write $z = a + bi$, $w = c + di$, and $u = e + fi$ with real $a, \ldots, f$, and compute both sides of each law from Definition 1.1.29.
> *Hint 2.* Addition is coordinatewise, so its laws hold coordinate by coordinate. For multiplication, expand both sides and compare real parts and imaginary parts separately.
> *Sketch.* Addition is coordinatewise addition in $\mathbb{R}^2$, so its laws come from those of $\mathbb{R}$. For each multiplicative law, expand both sides with the definition; the real parts agree, and so do the imaginary parts, by the field axioms of $\mathbb{R}$.

Let $z = a + bi$, $w = c + di$, and $u = e + fi$, where $a, b, c, d, e, f \in \mathbb{R}$. Addition in $\mathbb{C}$ is coordinatewise addition in $\mathbb{R}^2$, so it is commutative and associative because addition in $\mathbb{R}$ is.

For commutativity of multiplication, $wz = (ca - db) + (cb + da)i$, which equals $zw = (ac - bd) + (ad + bc)i$ by commutativity in $\mathbb{R}$. For associativity, expanding with the definition gives

$$\begin{aligned}(zw)u &= ((ac - bd)e - (ad + bc)f) + ((ac - bd)f + (ad + bc)e)i \\ &= (ace - adf - bcf - bde) + (acf + ade + bce - bdf)i,\end{aligned}$$

and

$$\begin{aligned}z(wu) &= (a(ce - df) - b(cf + de)) + (a(cf + de) + b(ce - df))i \\ &= (ace - adf - bcf - bde) + (acf + ade + bce - bdf)i,\end{aligned}$$

so $(zw)u = z(wu)$. For distributivity,

$$\begin{aligned}z(w + u) &= (a(c + e) - b(d + f)) + (a(d + f) + b(c + e))i \\ &= ((ac - bd) + (ae - bf)) + ((ad + bc) + (af + be))i \\ &= zw + zu.\end{aligned}$$

> *Intuition.* The remaining axioms ask for identities and inverses, and only the multiplicative inverse needs an idea. For $z = a + bi$, multiplying by $a - bi$ produces a real number,
>
> $$(a + bi)(a - bi) = \left(a^2 + b^2\right) + (-ab + ba)i = a^2 + b^2,$$
>
> and a nonzero real number can be divided by. Dividing $a - bi$ by $a^2 + b^2$ gives the inverse in the statement.

Proof, part 2 of 2: identities and inverses. ROUTINE

> *Uses.* Definition 1.1.29 (Complex numbers), $x^2 + y^2 > 0$ for real $(x, y) \neq (0, 0)$
> *Hint 1.* Check directly that $0 + 0i$ and $1 + 0i$ are identities and that $-a - bi$ is an additive inverse of $a + bi$.
> *Hint 2.* For the multiplicative inverse, first explain why $a^2 + b^2 \neq 0$ when $a + bi \neq 0$. Then multiply $a + bi$ by the proposed inverse.
> *Sketch.* The identities and additive inverses are immediate from the definition. If $a + bi \neq 0$, then $a^2 + b^2 > 0$, and multiplying out shows that $(a - bi)/\left(a^2 + b^2\right)$ is an inverse.

For $z = a + bi$ we have $z + (0 + 0i) = z$ and $z(1 + 0i) = (a \cdot 1 - b \cdot 0) + (a \cdot 0 + b \cdot 1)i = z$, and $0 + 0i \neq 1 + 0i$. Also $(a + bi) + (-a - bi) = 0 + 0i$.

Now suppose $z \neq 0$, that is, $(a, b) \neq (0, 0)$. Then $D = a^2 + b^2 > 0$, since squares of real numbers are nonnegative and at least one of a^2 and b^2 is positive. Let $w = a/D - (b/D)i$. Then

$$zw = \left(a \cdot \frac{a}{D} - b \cdot \frac{-b}{D}\right) + \left(a \cdot \frac{-b}{D} + b \cdot \frac{a}{D}\right)i = \frac{a^2 + b^2}{D} + 0i = 1,$$

so w is a multiplicative inverse of z. □

From now on we compute in $\mathbb{C}$ with the field axioms and $i^2 = -1$, as in the intuition above; the product formula never needs to be memorized.

Example 1.1.31 (Complex arithmetic). $(2+3i)(1-i) = 2-2i+3i-3i^2 = 5+i$. By Proposition 1.1.30, $(1+2i)^{-1} = (1-2i)/5$, so the unique solution of $(1+2i)z = 3-i$ (Proposition 1.1.13) is

$$z = \frac{(3-i)(1-2i)}{5} = \frac{3-6i-i+2i^2}{5} = \frac{1-7i}{5}.$$

The number $a-bi$ in the proof is the complex conjugate of $a+bi$, and $\sqrt{a^2+b^2}$ is the modulus of $a+bi$. Both reappear in Chapter 6, where inner products over $\mathbb{C}$ need them.

1.1.5 Subfields

The operations of $\mathbb{C}$, restricted to the real numbers $a+0i$, are the operations of $\mathbb{R}$, and the operations of $\mathbb{R}$ restricted to $\mathbb{Q}$ are those of $\mathbb{Q}$. Fields sitting inside other fields like this come up whenever we compare the same vectors over two fields, for instance over $\mathbb{R}$ and over $\mathbb{C}$ (Section 1.5).

Definition 1.1.32 (Subfield). Let F be a field. A subset $K \subseteq F$ is a **subfield** of F if $0, 1 \in K$ and K is closed under the operations of F: for all $a, b \in K$, we have $a+b \in K$, $ab \in K$, $-a \in K$, and $a^{-1} \in K$ if $a \neq 0$.

Proposition 1.1.33 (A subfield is a field). A subfield K of a field F is a field under the operations of F restricted to K.

Proof. ROUTINE

Uses. Definition 1.1.32 (Subfield), Definition 1.1.1 (Field)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Sort the axioms into those of the form "for all elements, an identity holds" and those of the form "there exists an element".
Hint 2. The first kind hold in K because they hold in F. For the second kind, the element whose existence is claimed must lie in K, and that is what the closure conditions provide.
Sketch. Closure under $+$ and $\cdot$ makes them operations on K. Identities that hold for all elements of F hold for all elements of K. The identities $0, 1$ and the inverses required by the existence axioms lie in K by assumption.

Since K is closed under $+$ and $\cdot$, these restrict to operations $K \times K \to K$. Commutativity, associativity, and distributivity hold for all elements of F, so in particular for all elements of K. We have $0, 1 \in K$, the equations $a+0 = a$ and $a \cdot 1 = a$ hold for $a \in K$ because they hold in F, and $0 \neq 1$. Finally, for $a \in K$ we have $-a \in K$ with $a + (-a) = 0$, and if $a \neq 0$ then $a^{-1} \in K$ with $aa^{-1} = 1$. So every field axiom holds in K. □

Example 1.1.34 (Subfields and non-subfields). $\mathbb{Q}$ is a subfield of $\mathbb{R}$, and $\mathbb{R}$, viewed as the set of $a+0i$, is a subfield of $\mathbb{C}$; the identification above shows the closure conditions, with $(a+0i)^{-1} = a^{-1} + 0i$. On the other hand, $\{0, 1\}$ is not a subfield of $\mathbb{R}$, since $1+1 = 2 \notin \{0, 1\}$, even though $\{0, 1\}$ with the operations of $\mathbb{F}_2$ is a field: a subfield must use the operations of the bigger field. Similarly $\mathbb{Z}$ is not a subfield of $\mathbb{Q}$, since $2^{-1} = 1/2 \notin \mathbb{Z}$.

Exercises

Exercise 1.1.35. ROUTINE Solve the system

$$x + 2y = 3, \quad 3x + 4y = 1$$

(a) over $\mathbb{F}_7$, and (b) over $\mathbb{F}_2$, reading each integer as its remainder modulo 2.

Solution on page 1025

Exercise 1.1.36. ROUTINE Write each of the following in the form $a+bi$ with $a, b \in \mathbb{R}$: (a) $(3-2i)(1+4i)$; (b) $(2+i)^{-1}$; (c) $(1+i)/(1-i)$; (d) $(1+i)^8$.

Solution on page 1025

Exercise 1.1.37. ROUTINE Let $a, b \in F$ be nonzero. Show that $a^{-1} \neq 0$ and $\left(a^{-1}\right)^{-1} = a$, and that $ab \neq 0$ and $(ab)^{-1} = a^{-1}b^{-1}$.

Solution on page 1025

Exercise 1.1.38. ROUTINE Let $a, b, c, d \in F$ with $b \neq 0$ and $d \neq 0$. Show that

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd},$$

and that $a/b = c/d$ if and only if $ad = bc$.

Solution on page 1026

Exercise 1.1.39. ROUTINE Let F be a set with operations $+$ and $\cdot$ satisfying all the field axioms except that $1 \neq 0$ is not required, and suppose that in fact $1 = 0$. Show that F has exactly one element.

Solution on page 1026

Exercise 1.1.40. ROUTINE Prove or disprove: in every field, if $a^2 = b^2$, then $a = b$ or $a = -b$.

Solution on page 1026

Exercise 1.1.41. ROUTINE Prove or disprove: for every prime p, the equation $x^2 + 1 = 0$ has a solution in $\mathbb{F}_p$.

Solution on page 1026

Exercise 1.1.42. ROUTINE Find the error in the following argument.

Claim. In every field, the equation $x^2 = 1$ has exactly two solutions.

Proof. We have $x^2 - 1 = (x - 1)(x + 1)$, so by Proposition 1.1.17, $x^2 = 1$ holds if and only if $x = 1$ or $x = -1$. These are the two solutions.

Solution on page 1027

Exercise 1.1.43. ROUTINE Let $a, b \in F$, and define $f : F \to F$ by $f(x) = ax + b$. Show that f is a bijection if and only if $a \neq 0$, and find f^{-1} in that case.

Solution on page 1027

Exercise 1.1.44. STANDARD

(a) Show that the only complex numbers z with $z^2 = -1$ are i and $-i$.

(b) Find all $z \in \mathbb{C}$ with $z^2 = i$.

Solution on page 1027

Exercise 1.1.45. STANDARD Find the error in the following argument.

Claim. $\mathbb{Z}/n\mathbb{Z}$ is a field for every $n \geq 2$.

Proof. By Lemma 1.1.21, only multiplicative inverses need checking. Let $a \neq 0$ in $\mathbb{Z}/n\mathbb{Z}$, and define $m(x) = ax$. If $m(x) = m(y)$, then $ax = ay$, so $x = y$ by cancellation (Corollary 1.1.14). So m is an injective function from a finite set to itself, hence surjective, and $ax = 1$ for some x.

Solution on page 1027

Exercise 1.1.46. STANDARD

(a) Let F be a field with exactly three elements, $F = \{0, 1, c\}$. Show that

$$1 + 1 = c, \quad 1 + c = 0, \quad c + c = 1, \quad c \cdot c = 1.$$

Conclude that the addition and multiplication tables of F are those of $\mathbb{F}_3$ with c in place of 2.

(b) Find all fields F in which $x^3 = x$ for every $x \in F$.

Solution on page 1028

Exercise 1.1.47. STANDARD Show that the following are equivalent for a field F:

(i) $1 + 1 = 0$;

(ii) $a = -a$ for every $a \in F$;

(iii) the map $F \to F$, $x \mapsto x^2$, is injective.

Solution on page 1028

Exercise 1.1.48. STANDARD Show that the characteristic of a field is either 0 or a prime number.

Solution on page 1029

Exercise 1.1.49. STANDARD Let $\mathbb{Q}(\sqrt{2}) = \left\{a + b\sqrt{2} \;\middle|\; a, b \in \mathbb{Q}\right\}$. Show that $\mathbb{Q}(\sqrt{2})$ is a subfield of $\mathbb{R}$.

Solution on page 1029

Exercise 1.1.50. STANDARD Let F be a field with $1 + 1 \neq 0$. Write $2 = 1 + 1$ and $4 = 2 \cdot 2$, and let $a, b, c \in F$ with $a \neq 0$.

(a) Show that $ax^2 + bx + c = 0$ has a solution $x \in F$ if and only if $b^2 - 4ac = d^2$ for some $d \in F$, and that in that case the solutions are $x = (-b \pm d)/(2a)$.

(b) Show that (a) fails without the hypothesis $1 + 1 \neq 0$: find $a, b, c \in \mathbb{F}_2$ with $a \neq 0$ such that $b^2 - 4ac$ is a square but $ax^2 + bx + c$ has no root in $\mathbb{F}_2$.

Solution on page 1029

Exercise 1.1.51. STANDARD Let p be prime and let $a \in \mathbb{F}_p$ be nonzero.

(a) Show that the products $a \cdot 1, a \cdot 2, \ldots, a \cdot (p-1)$ are the elements $1, 2, \ldots, p-1$ of $\mathbb{F}_p$ in some order.

(b) Deduce that $a^{p-1} = 1$ in $\mathbb{F}_p$. (This is Fermat's little theorem.)

Solution on page 1030

Exercise 1.1.52. STANDARD Let p be an odd prime. How many elements of $\mathbb{F}_p$ are squares, that is, of the form x^2 for some $x \in \mathbb{F}_p$?

Solution on page 1030

Exercise 1.1.53. HARD Prove or disprove: the set $K = \left\{a + b\sqrt[3]{2} \;\middle|\; a, b \in \mathbb{Q}\right\}$ is a subfield of $\mathbb{R}$.

Solution on page 1030

Exercise 1.1.54. HARD Fix $c \in \mathbb{R}$ and give $\mathbb{R}^2$ the componentwise addition and the multiplication

$$(a, b)(a', b') = (aa' + cbb', ab' + ba').$$

(For $c = -1$ this is $\mathbb{C}$.) For which c is $\mathbb{R}^2$ with these operations a field?

Solution on page 1031

Exercise 1.1.55. HARD Construct a set with operations $+$ and $\cdot$ that satisfy every field axiom except distributivity, in which $0 \cdot 0 \neq 0$. (So distributivity cannot be removed from the proof of Proposition 1.1.15.)

Solution on page 1031

Exercise 1.1.56. HARD BEYOND 21-242 $\mathbb{Z}/4\mathbb{Z}$ is not a field, but there is a field with four elements. Let $\mathbb{F}_4 = \{0, 1, \alpha, \beta\}$, four distinct symbols, with the following operations. Addition: 0 is an identity, $x + x = 0$ for every x, and the sum of two different elements of $\{1, \alpha, \beta\}$ is the third. Multiplication: $0x = x0 = 0$, $1x = x1 = x$, $\alpha^2 = \beta$, $\beta^2 = \alpha$, and $\alpha\beta = \beta\alpha = 1$.

(a) Show that $\mathbb{F}_4$ is a field.

(b) Show that $\operatorname{char} \mathbb{F}_4 = 2$, and explain why $\mathbb{F}_4$ is not $\mathbb{Z}/4\mathbb{Z}$ with its elements renamed.

Solution on page 1032

Exercise 1.1.57. HARD Let F be a finite field. Show that the sum of all the elements of F is 0, unless F has exactly two elements.

Solution on page 1033

Exercise 1.1.58. CHALLENGE Show that commutativity of addition follows from the other field axioms. That is, suppose F has operations $+$ and $\cdot$ satisfying every field axiom except possibly $a + b = b + a$, and show that $a + b = b + a$ for all $a, b \in F$.

Solution on page 1033

1.2 Vector spaces

Section 1.1 made precise what we ask of scalars. This section does the same for vectors. In $\mathbb{R}^2$ we can add two vectors and scale a vector by a real number, and these operations obey a short list of rules. A vector space is a set with an addition and a scalar multiplication that obey the same rules, with scalars taken from any field. Tuples, matrices, functions, and polynomials all form vector spaces, and a fact proved from the rules alone holds in each of them.

1.2.1 The axioms

The definition lists the rules that addition and scaling obey in $\mathbb{R}^2$, and nothing more. We number them (A1)–(A8) and cite them by number.

Definition 1.2.1 (Vector space). Let F be a field. A **vector space over** F is a set V together with two operations, **vector addition** $+ : V \times V \to V$ and **scalar multiplication** $\cdot : F \times V \to V$, that satisfy the following **vector space axioms**. We write $a\mathbf{v}$ for $a \cdot \mathbf{v}$.

- (A1) *Commutativity.* For all $\mathbf{v}, \mathbf{w} \in V$, $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.
- (A2) *Associativity.* For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
- (A3) *Additive identity.* There is an element $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.
- (A4) *Additive inverses.* For every $\mathbf{v} \in V$ there is some $\mathbf{w} \in V$ with $\mathbf{v} + \mathbf{w} = \mathbf{0}$.
- (A5) *Multiplicative identity.* For all $\mathbf{v} \in V$, $1\mathbf{v} = \mathbf{v}$, where 1 is the multiplicative identity of F.
- (A6) *Associativity of scaling.* For all $a, b \in F$ and $\mathbf{v} \in V$, $(ab)\mathbf{v} = a(b\mathbf{v})$.
- (A7) *Distributivity over vector addition.* For all $a \in F$ and $\mathbf{v}, \mathbf{w} \in V$, $a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$.
- (A8) *Distributivity over scalar addition.* For all $a, b \in F$ and $\mathbf{v} \in V$, $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$.

The elements of V are called **vectors**, and $\mathbf{0}$ is called the **zero vector**. We also say that V is an F-**vector space**. A vector space over $\mathbb{R}$ is a **real vector space**, and a vector space over $\mathbb{C}$ is a **complex vector space**.

Remark 1.2.2 (Reading the axioms). Closure is part of the definition. Since $+$ and $\cdot$ are functions into V, sums and scalar multiples of vectors lie in V. When we test whether a candidate is a vector space, closure comes first: if a sum leaves the set, there is no addition on the set whose axioms could be checked (Non-example 1.2.21).

The symbols are overloaded. In (A8), the $+$ on the left adds scalars and the $+$ on the right adds vectors. In (A6), ab is a product in F, while the other two products are scalar multiplications. Bold type separates the scalar $0 \in F$ from the vector $\mathbf{0} \in V$. No axiom multiplies two vectors or adds a scalar to a vector, because neither operation exists.

Quantifier order matters. (A3) has the form "there is $\mathbf{0}$ such that for all $\mathbf{v}$": one vector serves every $\mathbf{v}$ at once. The statement with the quantifiers swapped, that each $\mathbf{v}$ has some $\mathbf{z}$ with $\mathbf{v} + \mathbf{z} = \mathbf{v}$, is weaker; part (c) of Example 1.2.25 satisfies it and fails (A3). In (A4) the order is "for every $\mathbf{v}$ there is $\mathbf{w}$", so the inverse $\mathbf{w}$ may depend on $\mathbf{v}$. Also, (A4) mentions $\mathbf{0}$, the vector promised by (A3). If two different vectors satisfied (A3), then (A4) would be ambiguous. Proposition 1.2.27 rules this out, and its proof uses only (A1) and (A3), so there is no circularity.

The field is part of the data. The same set with the same addition can be a vector space over one field and not over another (Example 1.2.20). When the field matters, we say explicitly that V is a vector space over F.

1.2.2 Examples

To show that a candidate is a vector space, we specify the two operations, check closure, and check the eight axioms. In the standard examples below, the operations act one coordinate (or one input) at a time, and each axiom reduces to the field axiom of the same name.

Example 1.2.3 (The zero space). Let $V = \{\mathbf{0}\}$ be a set with one element, with $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for all $a \in F$. Each axiom is either an equation between elements of V, which holds because V has only one element, or asks for an element of V, which $\mathbf{0}$ supplies. So $\{\mathbf{0}\}$ is a vector space over every field, called the **zero space**. By (A3), every vector space has at least one element, so the zero space is as small as a vector space can be.

Tuples and matrices

Definition 1.2.4 (The space F^n). Let $n \geq 1$. The set F^n consists of all columns

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

with $x_1, \ldots, x_n \in F$, called the **coordinates** of $\mathbf{x}$. Two columns are equal when all their coordinates are equal. The operations are coordinatewise:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad a\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}.$$

In running text we write $(x_1, \ldots, x_n)$ for the column with coordinates $x_1, \ldots, x_n$.

Proposition 1.2.5 (F^n is a vector space). F^n is a vector space over F. The column $(0, \ldots, 0)$ is a zero vector, and $(-x_1, \ldots, -x_n)$ is an additive inverse of $(x_1, \ldots, x_n)$.

Proof. ROUTINE

> *Uses.* Definition 1.2.4 (The space F^n), Definition 1.1.1 (Field)
> *Techniques.* Proving from axioms (Technique 1.1.7)
> *Hint 1.* Two columns are equal exactly when their i-th coordinates are equal for every i. So compute the i-th coordinate of each side of each axiom.
> *Hint 2.* Coordinatewise, each axiom becomes a field axiom. For example, the i-th coordinate of (A7) reads $a(x_i + y_i) = ax_i + ay_i$.
> *Sketch.* Sums and multiples of columns with entries in F have entries in F. Comparing i-th coordinates, each of (A1)–(A8) becomes the field axiom of the same name applied to x_i, y_i, z_i, with $(0, \dots, 0)$ and $(-x_1, \dots, -x_n)$ supplying (A3) and (A4).

The coordinates of $\mathbf{x} + \mathbf{y}$ and $a\mathbf{x}$ are sums and products in F, so they lie in F, and the operations map into F^n. Let $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, and $\mathbf{z} = (z_1, \dots, z_n)$ be in F^n, let $a, b \in F$, and fix $i \in \{1, \dots, n\}$. Two columns are equal when their i-th coordinates agree for every i. By the definition of the operations, the i-th coordinates of the two sides of each axiom are as follows, and they agree by the field axiom named on the right:

(A1)	$x_i + y_i = y_i + x_i$	commutativity,
(A2)	$(x_i + y_i) + z_i = x_i + (y_i + z_i)$	associativity,
(A3)	$x_i + 0 = x_i$	additive identity,
(A4)	$x_i + (-x_i) = 0$	additive inverses,
(A5)	$1x_i = x_i$	multiplicative identity,
(A6)	$(ab)x_i = a(bx_i)$	associativity,
(A7)	$a(x_i + y_i) = ax_i + ay_i$	distributivity,
(A8)	$(a + b)x_i = ax_i + bx_i$	distributivity.

In (A3) we compare $\mathbf{x} + (0, \dots, 0)$ with $\mathbf{x}$, and in (A4) we compare $\mathbf{x} + (-x_1, \dots, -x_n)$ with $(0, \dots, 0)$. So all eight axioms hold. □

Example 1.2.6 (Small fields). $\mathbb{F}_2^3$ has $2^3 = 8$ elements, and in it $(1, 1, 0) + (0, 1, 1) = (1, 0, 1)$, because $1 + 1 = 0$ in $\mathbb{F}_2$. For the same reason, $\mathbf{v} + \mathbf{v} = \mathbf{0}$ for every $\mathbf{v} \in \mathbb{F}_2^n$. In $\mathbb{C}^2$, $i(1, i) = (i, -1)$ and $(1 + i)(1, -i) = (1 + i, 1 - i)$.

For $n = 1$, F^1 is F with its own addition and multiplication. So every field is a vector space over itself.

Definition 1.2.7 (Matrices). Let $m, n \geq 1$. An $m \times n$ **matrix** over F is a rectangular array

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

of elements $a_{ij} \in F$, the **entries** of A. The entry a_{ij} sits in row i and column j, and we also write $A = (a_{ij})$. Two matrices are equal when all their entries are equal. The set of $m \times n$ matrices over F is denoted $F^{m \times n}$, and its operations are entrywise:

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \quad c(a_{ij}) = (ca_{ij}).$$

A matrix is mn entries arranged in a rectangle, and the operations ignore the arrangement. So the proof of Proposition 1.2.5, run entry by entry, shows that $F^{m \times n}$ is a vector space over F (Remark 1.2.11 below makes this precise). Its zero vector is the **zero matrix**, all of whose entries are 0. The rectangular shape starts to matter when matrices are multiplied (Section 2.4). An $n \times 1$ matrix is a column, so $F^{n \times 1}$ is F^n. For example, in $\mathbb{R}^{2 \times 3}$,

$$2\begin{pmatrix}1 & 0 & -1\\ 3 & 2 & 0\end{pmatrix}+\begin{pmatrix}0 & 4 & 1\\ -1 & 1 & 1\end{pmatrix}=\begin{pmatrix}2 & 4 & -1\\ 5 & 5 & 1\end{pmatrix}.$$

Exercise 1.2.8. ROUTINE

(a) In $\mathbb{F}_3^3$, find $\mathbf{x}$ with $2\mathbf{x}+(1,2,0)=(2,0,1)$.

(b) In $\mathbb{F}_5^{2\times 2}$, find X with $3X+A=B$, where $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$ and $B=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$.

(c) In $\mathbb{C}^2$, find $\mathbf{v}$ with $(1+i)\mathbf{v}+(1,0)=(0,2i)$.

Solution on page 1033

Functions

Matrices, functions, and polynomials keep their usual names A, f, p even when we treat them as vectors, and so do numbers in the few examples whose vectors are numbers. Bold letters are reserved for vectors of a general vector space and of F^n. The zero vector is written $\mathbf{0}$, whether it is the zero matrix, the zero function, or the zero polynomial; only when the vectors are numbers do we write it as a number.

Definition 1.2.9 (Function spaces). Let X be a set. The set F^X consists of all functions $f : X \to F$. For $f, g \in F^X$ and $a \in F$, the functions $f+g$ and af are defined pointwise:

$$(f+g)(t)=f(t)+g(t),\quad (af)(t)=af(t)\quad (t\in X).$$

Proposition 1.2.10 (F^X is a vector space). F^X is a vector space over F. The zero function $\mathbf{0}$, given by $\mathbf{0}(t)=0$ for all $t\in X$, is a zero vector, and the function $t\mapsto -f(t)$ is an additive inverse of f.

Proof. ROUTINE

Uses. Definition 1.2.9 (Function spaces), Definition 1.1.1 (Field), Proposition 1.2.5 (F^n is a vector space)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Two functions $X\to F$ are equal exactly when they take the same value at every $t\in X$.
Hint 2. Evaluate both sides of each axiom at t. This is the proof of Proposition 1.2.5, with the value $f(t)$ in place of the coordinate x_i.
Sketch. Pointwise operations give functions $X\to F$ again. Evaluating both sides of each axiom at $t\in X$ turns it into the field axiom of the same name for the values $f(t), g(t), h(t)$, and the zero function and $t\mapsto -f(t)$ supply (A3) and (A4).

For $f, g\in F^X$ and $a\in F$, the rules above assign an element of F to each $t\in X$, so $f+g$ and af are functions $X\to F$, that is, elements of F^X. Two functions $X\to F$ are equal when they agree at every $t\in X$. Let $f, g, h\in F^X$, $a, b\in F$, and $t\in X$. Evaluating both sides of each axiom at t gives

(A1) $f(t)+g(t)=g(t)+f(t)$,
(A2) $(f(t)+g(t))+h(t)=f(t)+(g(t)+h(t))$,
(A3) $f(t)+0=f(t)$,
(A4) $f(t)+(-f(t))=0$,
(A5) $1f(t)=f(t)$,
(A6) $(ab)f(t)=a(bf(t))$,
(A7) $a(f(t)+g(t))=af(t)+ag(t)$,
(A8) $(a+b)f(t)=af(t)+bf(t)$,

where in (A3) the second summand is $\mathbf{0}(t)=0$, and in (A4) it is the value at t of $t\mapsto -f(t)$. Each line is a field axiom, as in the proof of Proposition 1.2.5. Since t was arbitrary, both sides of each axiom are the same function. □

Remark 1.2.11 (Tuples and matrices are functions). A column $(x_1, \ldots, x_n)$ carries the same information as the function $\{1, \ldots, n\} \to F$, $i \mapsto x_i$, and adding or scaling columns coordinatewise is adding or scaling these functions pointwise. So F^n is F^X for $X = \{1, \ldots, n\}$, under another name. In the same way, $F^{m\times n}$ is F^X for $X = \{1, \ldots, m\} \times \{1, \ldots, n\}$, a matrix being the function $(i, j) \mapsto a_{ij}$. In particular, Proposition 1.2.10 shows that $F^{m\times n}$ is a vector space. Other choices of X give the space $\mathbb{R}^{\mathbb{R}}$ of all functions $\mathbb{R} \to \mathbb{R}$, and the space of sequences $(x_0, x_1, x_2, \ldots)$ in F, which is F^X for $X = \{0, 1, 2, \ldots\}$.

Many vector spaces of interest sit inside some F^X and use its operations. Then only closure needs a real argument.

Example 1.2.12 (Continuous functions). Let $C[0, 1]$ be the set of continuous functions $[0, 1] \to \mathbb{R}$, with the operations of $\mathbb{R}^{[0,1]}$. Sums and constant multiples of continuous functions are continuous, so the operations map into $C[0, 1]$. The equations in (A1), (A2), and (A5)–(A8) hold for all elements of $\mathbb{R}^{[0,1]}$, so in particular for continuous ones. The zero function is continuous, which gives (A3), and if f is continuous then so is $-f$, which gives (A4). So $C[0, 1]$ is a real vector space.

The same argument applies to

$$\left\{ f \in C[0, 1] \;\middle|\; \int_0^1 f(t)\,\mathrm{d}t = 0 \right\},$$

since $\int_0^1 (f + g) = \int_0^1 f + \int_0^1 g$ and $\int_0^1 af = a\int_0^1 f$: sums and multiples of functions with integral 0 have integral 0, and so do $\mathbf{0}$ and $-f$. On the other hand, the functions with integral 1 do not form a vector space under these operations: if $\int_0^1 f = \int_0^1 g = 1$, then $\int_0^1 (f + g) = 2$, so addition is not an operation on that set.

The pattern of this example, a subset of a known vector space that contains $\mathbf{0}$ and is closed under the operations, is common enough that Section 1.4 turns it into a criterion.

Polynomials

In school, a polynomial is an expression $a_0 + a_1 x + \cdots + a_d x^d$. As a definition this is not quite right: 5 and $5 + 0x$ are different expressions, and we want them to be the same polynomial. What determines a polynomial is its list of coefficients, all but finitely many of which are 0. So we define a polynomial to be that list, and recover the expressions afterwards.

Definition 1.2.13 (Polynomials). A **polynomial** over F is a sequence $p = (a_0, a_1, a_2, \ldots)$ of elements of F such that only finitely many a_k are nonzero. The a_k are the **coefficients** of p, and two polynomials are equal when all their coefficients are equal. The set of polynomials over F is denoted $F[x]$. Its operations are coefficientwise:

$$(a_0, a_1, \ldots) + (b_0, b_1, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots), \quad c(a_0, a_1, \ldots) = (ca_0, ca_1, \ldots).$$

Proposition 1.2.14 ($F[x]$ is a vector space). $F[x]$ is a vector space over F. Its zero vector is the **zero polynomial** $\mathbf{0} = (0, 0, \ldots)$.

Proof. ROUTINE

Uses. Definition 1.2.13 (Polynomials), Proposition 1.2.10 (F^X is a vector space), Remark 1.2.11 (Tuples and matrices are functions), Proposition 1.1.15 (Multiplying by zero)

Hint 1. A sequence is a function on $\{0, 1, 2, \ldots\}$, and the operations of $F[x]$ are those of F^X for $X = \{0, 1, 2, \ldots\}$. Argue as for $C[0, 1]$ in Example 1.2.12.

Hint 2. The only real work is closure. If $a_k = 0$ for $k > m$ and $b_k = 0$ for $k > n$, which coefficients of the sum must vanish?

Sketch. $F[x]$ is the set of finitely supported elements of F^X, $X = \{0, 1, 2, \ldots\}$, with the same operations. Sums and multiples of finitely supported sequences are finitely supported, as are the zero sequence and $(-a_k)$; the equational axioms are inherited from F^X.

A sequence $(a_0, a_1, \ldots)$ is the function $k \mapsto a_k$ on $X = \{0, 1, 2, \ldots\}$, and the operations of $F[x]$ are those of F^X. So $F[x] \subseteq F^X$, and we argue as in Example 1.2.12.

Closure. Let $p = (a_k)$ and $q = (b_k)$ be polynomials, and choose m and n with $a_k = 0$ for $k > m$ and $b_k = 0$ for $k > n$. For $k > \max(m, n)$ we have $a_k + b_k = 0 + 0 = 0$ and $ca_k = c0 = 0$ (Proposition 1.1.15). So $p + q$ and cp have only finitely many nonzero coefficients.

The axioms. The equations in (A1), (A2), and (A5)–(A8) hold for all elements of F^X by Proposition 1.2.10, so in particular for polynomials. The zero sequence has no nonzero coefficients, so it lies in $F[x]$ and gives (A3). For $p = (a_k)$, the sequence $q = (-a_0, -a_1, \ldots)$ has its nonzero coefficients in the same positions as p (as $-a_k = 0$ only when $a_k = 0$), so $q \in F[x]$, and $p + q = \mathbf{0}$ gives (A4). □

Notation 1.2.15. For $k \geq 0$, x^k denotes the polynomial whose k-th coefficient is 1 and whose other coefficients are 0. We also write 1 for x^0 and x for x^1. If $a_k = 0$ for all $k > d$, then

$$(a_0, a_1, a_2, \ldots) = a_0 + a_1 x + \cdots + a_d x^d = \sum_{k=0}^{d} a_k x^k,$$

where the right side is computed with the operations of $F[x]$: its j-th coefficient is a_j for $j \leq d$ and 0 for $j > d$.

So the familiar expressions are honest sums of vectors in $F[x]$, and the problem with 5 and $5 + 0x$ disappears: both are the sequence $(5, 0, 0, \ldots)$. A polynomial can be written with any d that is at least as large as the index of its last nonzero coefficient. To decide whether two expressions are the same polynomial, compare coefficients. We multiply polynomials in Section 7.1; here only the vector space operations matter.

Definition 1.2.16 (Degree). Let $p = \sum_{k=0}^{d} a_k x^k \in F[x]$. If $p \neq \mathbf{0}$, the **degree** $\deg p$ is the largest k with $a_k \neq 0$, and $a_{\deg p}$ is the **leading coefficient** of p. The zero polynomial has degree $\deg \mathbf{0} = -\infty$, with the convention that $-\infty < n$ for every integer n. For $n \geq 0$,

$$F[x]_{\leq n} = \{p \in F[x] \mid \deg p \leq n\}$$

is the set of polynomials $a_0 + a_1 x + \cdots + a_n x^n$ with $a_0, \ldots, a_n \in F$ (any of which may be 0).

Definition 1.2.17 (Evaluation). Let $p = \sum_{k=0}^{d} a_k x^k \in F[x]$ and $c \in F$. The **value** of p at c is

$$p(c) = \sum_{k=0}^{d} a_k c^k \in F.$$

The function $F \to F$, $c \mapsto p(c)$, is the **polynomial function** of p.

The value does not depend on which d we use to write p, since the extra terms are $0 \cdot c^k = 0$. For example, over $\mathbb{R}$, $\deg(3 - x^2) = 2$ and $(3 - x^2)(2) = -1$.

Proposition 1.2.18 (Degree and evaluation versus the operations). Let $p, q \in F[x]$ and $a, c \in F$.

(a) $\deg(p + q) \leq \max(\deg p, \deg q)$, with equality if $\deg p \neq \deg q$.

(b) If $a \neq 0$, then $\deg(ap) = \deg p$.

(c) $(p + q)(c) = p(c) + q(c)$ and $(ap)(c) = ap(c)$.

Proof. ROUTINE

Uses. Definition 1.2.16 (Degree), Definition 1.2.17 (Evaluation), Proposition 1.1.17 (No zero divisors), Definition 1.1.1 (Field)

Hint 1. Write $p = (a_k)$ and $q = (b_k)$. For (a), which coefficients of $p + q$ are forced to vanish? If $\deg p > \deg q$, look at the coefficient of $p + q$ in position $\deg p$.

Hint 2. For (b), when is aa_k zero? For (c), write p and q with the same d and expand.
Sketch. (a) Beyond $\max(\deg p, \deg q)$ both coefficient sequences vanish, so the sum's does; if $\deg p > \deg q$, the coefficient of $p+q$ in position $\deg p$ is $a_{\deg p} + 0 \neq 0$. (b) $aa_k = 0$ if and only if $a_k = 0$. (c) Distributivity and reordering of finite sums.

Write $p = (a_k)$ and $q = (b_k)$.

(a) Let $n = \max(\deg p, \deg q)$. For every integer $k > n$ we have $a_k = b_k = 0$, so the k-th coefficient $a_k + b_k$ of $p+q$ is 0. Hence $\deg(p+q) \leq n$; this includes the case $p+q = \mathbf{0}$, and the case $n = -\infty$, where $p = q = \mathbf{0}$. Now suppose $\deg p \neq \deg q$; by symmetry we may assume $\deg p > \deg q$. Then $p \neq \mathbf{0}$, and with $m = \deg p$ we have $b_m = 0$ because $m > \deg q$. So the m-th coefficient of $p+q$ is $a_m + 0 = a_m \neq 0$, and $\deg(p+q) \geq m$. With the first part, $\deg(p+q) = m = \max(\deg p, \deg q)$.

(b) The k-th coefficient of ap is aa_k. Since $a \neq 0$, Proposition 1.1.17 gives $aa_k = 0$ if and only if $a_k = 0$. So ap and p have their nonzero coefficients in the same positions, and in particular the same degree.

(c) Choose d with $a_k = b_k = 0$ for all $k > d$. Then, by distributivity and by reordering the finite sum,

$$(p+q)(c) = \sum_{k=0}^{d} (a_k + b_k)c^k = \sum_{k=0}^{d} a_k c^k + \sum_{k=0}^{d} b_k c^k = p(c) + q(c),$$

and $(ap)(c) = \sum_{k=0}^{d} (aa_k)c^k = a\sum_{k=0}^{d} a_k c^k = ap(c)$. □

Parts (a) and (b) say that sums and nonzero multiples of polynomials of degree at most n again have degree at most n (and $0p = \mathbf{0}$ has degree $-\infty$). So $F[x]_{\leq n}$ is closed under the operations of $F[x]$, and it is a vector space by the argument of Example 1.2.12. Part (c) says that taking polynomial functions is compatible with the operations of $F[x]$ and of F^F.

ASIDE: *Polynomials versus polynomial functions*

Different polynomials can have the same polynomial function. Over $\mathbb{F}_2$, the polynomial $f = x^2 + x$ has coefficients $(0, 1, 1, 0, \ldots)$, so $f \neq \mathbf{0}$ and $\deg f = 2$. But $f(0) = 0$ and $f(1) = 1 + 1 = 0$, so the polynomial function of f is the zero function $\mathbb{F}_2 \to \mathbb{F}_2$, the same as that of $\mathbf{0}$. Similarly, for every prime p, the polynomial $x^p - x \in \mathbb{F}_p[x]$ is nonzero but vanishes at every element of $\mathbb{F}_p$ (Exercise 1.2.51).

This is one reason to define polynomials by their coefficients. If a polynomial were its function, then $x^2 + x$ and $\mathbf{0}$ would be the same polynomial over $\mathbb{F}_2$, and degree would not be well defined. Later we will also substitute matrices and linear maps into polynomials (Section 7.2), which a function $F \to F$ cannot accommodate.

Over an infinite field such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, nothing is lost by thinking of functions. A nonzero polynomial of degree d has at most d roots (Section 7.1). So if p and q have the same polynomial function, then $(p-q)(c) = p(c) - q(c) = 0$ for every $c \in F$ (Proposition 1.2.18(c)); a nonzero $p - q$ would have infinitely many roots, so $p - q = \mathbf{0}$. This is why some books, LADR among them, can define polynomials over $\mathbb{R}$ and $\mathbb{C}$ as functions.

Changing the field

The field is part of the data of a vector space, and shrinking the field is always allowed.

Proposition 1.2.19 (Restriction of scalars). Let K be a subfield of F, and let V be a vector space over F. Then V, with the same addition and with scalar multiplication restricted to $K \times V$, is a vector space over K.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), Definition 1.1.32 (Subfield), Proposition 1.1.33 (A subfield is a field), Theorem 1.1.9 (Uniqueness of identities and inverses)
Hint 1. Sort the axioms into those about addition alone, those that hold "for all scalars", and (A5).
Hint 2. For (A5), you need to know that the multiplicative identity of the field K is the 1 of F.

Sketch. (A1)–(A4) only involve addition, which is unchanged. (A6)–(A8) hold for all scalars in F, hence for those in K, whose sums and products are computed as in F. The identity of K is the 1 of F, by uniqueness of identities in the field K, so (A5) holds.

The restricted scalar multiplication is a function $K \times V \to V$. Axioms (A1)–(A4) involve only the addition of V, which has not changed. Axioms (A6)–(A8) hold for all $a, b \in F$, so in particular for all $a, b \in K$; here $a + b$ and ab mean the same elements whether computed in K or in F, because K carries the operations of F. For (A5), K is a field (Proposition 1.1.33), and $1 \in K$ by Definition 1.1.32. Since $a \cdot 1 = a$ for every $a \in K$, the element 1 is a multiplicative identity of K, so it is the multiplicative identity of K by Theorem 1.1.9. As $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$, (A5) holds over K. □

Example 1.2.20 (Real and rational scalars). $\mathbb{C}$ is a vector space over itself (Example 1.2.6), hence over its subfield $\mathbb{R}$. For $a \in \mathbb{R}$, $a(x + yi) = ax + ayi$, so as a real vector space $\mathbb{C}$ is $\mathbb{R}^2$ with the pair (x, y) written as $x + yi$. Likewise $\mathbb{C}^n$ is a real vector space.

$\mathbb{R}$ is a vector space over $\mathbb{Q}$, with scalar multiplication the usual product. The change of field changes the geometry. Over $\mathbb{R}$, every real number is a scalar multiple of 1. Over $\mathbb{Q}$, $\sqrt{2}$ is not a rational multiple of 1, so 1 and $\sqrt{2}$ point in different directions. Section 1.5 makes this precise.

In the other direction, $\mathbb{Q}$ with the usual product is not a vector space over $\mathbb{R}$: $\sqrt{2} \cdot 1 \notin \mathbb{Q}$, so the product is not a function $\mathbb{R} \times \mathbb{Q} \to \mathbb{Q}$. For the same reason, $\mathbb{R}$ with the usual product is not a vector space over $\mathbb{C}$. Whether some other scalar multiplication works is a different question (Exercise 1.2.47).

1.2.3 Non-examples and repairs

To show that a candidate is not a vector space, one failure suffices: a sum or multiple that leaves the set, or a specific choice of vectors and scalars for which an axiom fails. The candidates below fail in different ways, and the first can be repaired by changing its operations.

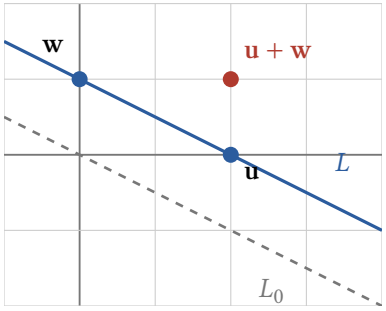


Figure 1.2.1. The line L: $x + 2y = 2$ (solid) and the line L_0: $x + 2y = 0$ through the origin (dashed). The vectors $\mathbf{u} = (2, 0)$ and $\mathbf{w} = (0, 1)$ lie on L, but their sum $(2, 1)$ does not.

Non-example 1.2.21 (A line not through the origin). Let $L = \{(x, y) \in \mathbb{R}^2 \mid x + 2y = 2\}$, with the operations of $\mathbb{R}^2$. This is not a real vector space, and the first problem is that the operations of $\mathbb{R}^2$ are not operations on L. The vectors $\mathbf{u} = (2, 0)$ and $\mathbf{w} = (0, 1)$ lie on L, but $\mathbf{u} + \mathbf{w} = (2, 1)$ does not, since $2 + 2 \cdot 1 = 4$ (Figure 1.2.1). Similarly $0\mathbf{u} = (0, 0) \notin L$.

Even setting closure aside, (A3) fails. Notice what showing this requires. Observing that $(0, 0) \notin L$ is not enough on its own, because the zero vector of a structure on L need not be $(0, 0)$ (see Example 1.2.23). We must show that no $\mathbf{z} \in L$ satisfies $\mathbf{v} + \mathbf{z} = \mathbf{v}$ for all $\mathbf{v} \in L$. Here that is quick: if $\mathbf{v} + \mathbf{z} = \mathbf{v}$ in $\mathbb{R}^2$ for even one $\mathbf{v}$, then subtracting $\mathbf{v}$ coordinatewise gives $\mathbf{z} = (0, 0)$, which is not in L.

Intuition. L is the line L_0 through the origin, shifted by $\mathbf{p} = (2, 0)$: every point of L is $\mathbf{p} + \mathbf{v}$ for exactly one $\mathbf{v} \in L_0$ (Figure 1.2.1). The line L_0 is closed under the operations of $\mathbb{R}^2$, so the only defect of L is where

its origin is. To repair L, we can add and scale as if $\mathbf{p}$ were the origin. The cleanest way to carry this out is to copy the operations of a known vector space across a bijection.

Proposition 1.2.22 (Transport of structure). Let V be a vector space over F, let W be a set, and let $\varphi : V \to W$ be a bijection. Define operations on W by

$$\mathbf{w}_1 \oplus \mathbf{w}_2 = \varphi\big(\varphi^{-1}(\mathbf{w}_1) + \varphi^{-1}(\mathbf{w}_2)\big), \quad a \odot \mathbf{w} = \varphi\big(a\varphi^{-1}(\mathbf{w})\big).$$

Then W with $\oplus$ and $\odot$ is a vector space over F in which $\varphi(\mathbf{0})$ is a zero vector, and for all $\mathbf{v}, \mathbf{v}' \in V$ and $a \in F$,

$$\varphi(\mathbf{v} + \mathbf{v}') = \varphi(\mathbf{v}) \oplus \varphi(\mathbf{v}'), \quad \varphi(a\mathbf{v}) = a \odot \varphi(\mathbf{v}). \tag{1.2.1}$$

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), a bijection φ has an inverse φ^{-1} with $\varphi^{-1}(\varphi(\mathbf{v})) = \mathbf{v}$ and $\varphi\big(\varphi^{-1}(\mathbf{w})\big) = \mathbf{w}$
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. First prove (1.2.1) directly from the definitions. Then every element of W has the form $\varphi(\mathbf{v})$.
Hint 2. With (1.2.1), each side of an axiom in W, written for elements $\varphi(\mathbf{v})$, becomes φ applied to the corresponding side of the same axiom in V.
Sketch. The definitions give $\varphi(\mathbf{v}) \oplus \varphi(\mathbf{v}') = \varphi(\mathbf{v} + \mathbf{v}')$ and $a \odot \varphi(\mathbf{v}) = \varphi(a\mathbf{v})$. Since φ is surjective, it suffices to check each axiom on elements $\varphi(\mathbf{v})$; pushing $\oplus$ and $\odot$ inside φ turns it into φ applied to the same axiom in V. The zero vector is $\varphi(\mathbf{0})$, and $\varphi(\mathbf{v}')$ is an inverse of $\varphi(\mathbf{v})$ whenever $\mathbf{v} + \mathbf{v}' = \mathbf{0}$.

The operations take values in W because φ does. Substituting $\mathbf{w}_1 = \varphi(\mathbf{v})$, $\mathbf{w}_2 = \varphi(\mathbf{v}')$, and $\mathbf{w} = \varphi(\mathbf{v})$ into the definitions and using $\varphi^{-1}(\varphi(\mathbf{v})) = \mathbf{v}$ gives (1.2.1). Since φ is surjective, every element of W is $\varphi(\mathbf{v})$ for some $\mathbf{v} \in V$, so it suffices to verify each axiom for elements of this form. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in F$. Each of the following chains uses (1.2.1) to move the operations inside φ, applies the axiom of V named on the right, and moves the operations back out:

$$\begin{aligned}
\varphi(\mathbf{v}) \oplus \varphi(\mathbf{w}) &= \varphi(\mathbf{v} + \mathbf{w}) = \varphi(\mathbf{w} + \mathbf{v}) \\
&= \varphi(\mathbf{w}) \oplus \varphi(\mathbf{v}) && \text{(A1)}, \\
(\varphi(\mathbf{u}) \oplus \varphi(\mathbf{v})) \oplus \varphi(\mathbf{w}) &= \varphi((\mathbf{u} + \mathbf{v}) + \mathbf{w}) = \varphi(\mathbf{u} + (\mathbf{v} + \mathbf{w})) \\
&= \varphi(\mathbf{u}) \oplus (\varphi(\mathbf{v}) \oplus \varphi(\mathbf{w})) && \text{(A2)}, \\
1 \odot \varphi(\mathbf{v}) &= \varphi(1\mathbf{v}) = \varphi(\mathbf{v}) && \text{(A5)}, \\
(ab) \odot \varphi(\mathbf{v}) &= \varphi((ab)\mathbf{v}) = \varphi(a(b\mathbf{v})) = a \odot (b \odot \varphi(\mathbf{v})) && \text{(A6)}, \\
a \odot (\varphi(\mathbf{v}) \oplus \varphi(\mathbf{w})) &= \varphi(a(\mathbf{v} + \mathbf{w})) = \varphi(a\mathbf{v} + a\mathbf{w}) \\
&= (a \odot \varphi(\mathbf{v})) \oplus (a \odot \varphi(\mathbf{w})) && \text{(A7)}, \\
(a + b) \odot \varphi(\mathbf{v}) &= \varphi((a + b)\mathbf{v}) = \varphi(a\mathbf{v} + b\mathbf{v}) \\
&= (a \odot \varphi(\mathbf{v})) \oplus (b \odot \varphi(\mathbf{v})) && \text{(A8)}.
\end{aligned}$$

For (A3), $\varphi(\mathbf{v}) \oplus \varphi(\mathbf{0}) = \varphi(\mathbf{v} + \mathbf{0}) = \varphi(\mathbf{v})$ for every $\mathbf{v}$, so $\varphi(\mathbf{0})$ is a zero vector of W. For (A4), given $\mathbf{v}$, choose $\mathbf{v}'$ with $\mathbf{v} + \mathbf{v}' = \mathbf{0}$; then $\varphi(\mathbf{v}) \oplus \varphi(\mathbf{v}') = \varphi(\mathbf{0})$. □

The map φ is a dictionary: it renames each vector of V, and W inherits the arithmetic of V under the new names. In the language of Section 4.5, (1.2.1) says that φ is an isomorphism from V to W with its new operations.

Example 1.2.23 (Repairing the line). Parametrize L by $\varphi : \mathbb{R} \to L$, $\varphi(t) = (2 - 2t, t)$. This is a bijection: each $(x, y) \in L$ has $x = 2 - 2y$, so $(x, y) = \varphi(y)$, and $\varphi^{-1}(x, y) = y$. Transport the structure of $\mathbb{R}$, a real vector space by Example 1.2.6, along φ. By Proposition 1.2.22, L becomes a real vector space with

$$(x, y) \oplus (x', y') = \varphi(y + y') = (2 - 2(y + y'), y + y'),$$

$$a \odot (x, y) = \varphi(ay) = (2 - 2ay, ay).$$

Its zero vector is $\varphi(0) = (2, 0) = \mathbf{p}$, and an additive inverse of (x, y) is $\varphi(-y) = (2 + 2y, -y)$.

These formulas can be written without φ. For $\mathbf{u} = (x, y)$ and $\mathbf{u}' = (x', y')$ in L we have $x = 2 - 2y$ and $x' = 2 - 2y'$, and substituting shows

$$\mathbf{u} \oplus \mathbf{u}' = \mathbf{u} + \mathbf{u}' - \mathbf{p}, \quad a \odot \mathbf{u} = \mathbf{p} + a(\mathbf{u} - \mathbf{p}),$$

where the right sides use the operations of $\mathbb{R}^2$. For instance, the first coordinate of $\mathbf{u} + \mathbf{u}' - \mathbf{p}$ is $x + x' - 2 = 2 - 2(y + y')$. So $\oplus$ and $\odot$ are the operations of $\mathbb{R}^2$ recentered at $\mathbf{p}$: subtract $\mathbf{p}$, compute as usual, and add $\mathbf{p}$ back. Any other point of L could serve as the origin instead and gives another valid structure; the line itself has no preferred origin. Such sets are studied as affine spaces in Section 4.7.

Exercise 1.2.24. ROUTINE Find the error in the following argument, and show that its conclusion is false.

Claim. No operations make the interval $(1, 2) = \{t \in \mathbb{R} \mid 1 < t < 2\}$ a real vector space.

Proof. Every vector space contains a zero vector, and the zero vector of a set of real numbers is 0. Since $0 \notin (1, 2)$, the interval cannot be a vector space.

Solution on page 1034

The next example changes one operation of $\mathbb{R}^2$ at a time. Each variant fails the axioms listed and satisfies the others, which shows what those axioms are for.

Example 1.2.25 (Changing one operation on $\mathbb{R}^2$).

(a) *Usual addition,* $a \odot (x, y) = (ax, 0)$. Axioms (A1)–(A4) are those of $\mathbb{R}^2$. Axioms (A6)–(A8) hold, because both sides agree in the first coordinate as in $\mathbb{R}^2$ and both have second coordinate 0. But (A5) fails: $1 \odot (0, 1) = (0, 0) \neq (0, 1)$. So (A5) does not follow from the other axioms; it is what stops scalar multiplication from discarding part of a vector. (The extreme case $a \odot \mathbf{v} = \mathbf{0}$ for all a and $\mathbf{v}$ also satisfies every axiom except (A5).)

(b) *Usual addition,* $a \odot (x, y) = (ax, y)$. Now (A5), (A6), and (A7) hold; for instance $(ab) \odot (x, y) = (abx, y) = a \odot (bx, y)$. But (A8) fails:

$$(1 + 1) \odot (0, 1) = (0, 1), \quad \text{while} \quad 1 \odot (0, 1) + 1 \odot (0, 1) = (0, 2).$$

Together with (A5), (A8) forces $2 \odot \mathbf{v} = \mathbf{v} + \mathbf{v}$, so scalar multiplication cannot leave a coordinate untouched.

(c) *Midpoint addition,* $(x, y) \oplus (x', y') = ((x + x')/2, (y + y')/2)$, *usual scalar multiplication.* (A1) holds, but (A2) fails: with $\mathbf{u} = \mathbf{v} = (0, 0)$ and $\mathbf{w} = (4, 0)$,

$$(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = (2, 0), \quad \text{while} \quad \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}) = (1, 0).$$

(A3) fails too, and this failure shows the quantifier order at work (Remark 1.2.2). We have $\mathbf{v} \oplus \mathbf{z} = \mathbf{v}$ exactly when $\mathbf{z} = \mathbf{v}$. So every $\mathbf{v}$ has a vector that acts like zero for it, namely $\mathbf{v}$ itself, but no single $\mathbf{z}$ works for every $\mathbf{v}$. Without a zero vector, (A4) cannot even be stated. Finally, (A8) fails: $(1 + 1)(1, 0) = (2, 0)$, while $(1, 0) \oplus (1, 0) = (1, 0)$. Axioms (A5)–(A7) hold.

The last non-example has operations that stay inside the set and satisfy most of the axioms.

Non-example 1.2.26 (Addition modulo 1). For $t \in \mathbb{R}$, write $t \bmod 1 = t - \lfloor t \rfloor \in [0, 1)$ for the fractional part of t. Then $s \bmod 1 = t \bmod 1$ exactly when $s - t \in \mathbb{Z}$. Let $V = [0, 1)$ with

$$s \oplus t = (s + t) \bmod 1, \quad a \odot t = (at) \bmod 1 \quad (a \in \mathbb{R}).$$

For instance $0.7 \oplus 0.5 = 0.2$. Think of V as a circle of circumference 1, on which adding means walking around the circle. Both operations land in V.

Replacing a real number by its fractional part changes it by an integer, and numbers that differ by an integer have the same fractional part. So, as in Lemma 1.1.21, both sides of (A2) equal $(s+t+u) \bmod 1$, and (A1) holds because $s+t = t+s$. The number 0 is a zero vector, an inverse of $t \neq 0$ is $1-t$, and 0 is its own inverse. (A5) holds since $1 \odot t = t \bmod 1 = t$ for $t \in [0,1)$. (A8) holds since $(a+b)t = at+bt$ differs from $(a \odot t) + (b \odot t)$ by an integer, so the fractional part of the latter, which is $(a \odot t) \oplus (b \odot t)$, equals that of the former, which is $(a+b) \odot t$. But (A7) fails: with $a = 1/2$ and $s = t = 1/2$,

$$\frac{1}{2} \odot \left(\frac{1}{2} \oplus \frac{1}{2}\right) = \frac{1}{2} \odot 0 = 0, \quad \text{while} \quad \left(\frac{1}{2} \odot \frac{1}{2}\right) \oplus \left(\frac{1}{2} \odot \frac{1}{2}\right) = \frac{1}{4} \oplus \frac{1}{4} = \frac{1}{2}.$$

(A6) fails as well: with $a = 1/2$, $b = 2$, and $t = 1/2$, we get $(ab) \odot t = 1 \odot t = 1/2$, while $a \odot (b \odot t) = a \odot 0 = 0$.

One might hope that a cleverer scalar multiplication repairs this. It cannot, and the reason (Example 1.2.40) is a consequence of the axioms that involves only the addition.

1.2.4 Consequences of the axioms

From now on, V is a vector space over a field F. The axioms are all we know about V, so as with fields, we prove each elementary fact from them (Technique 1.1.7). Each fact then holds in every example of the previous subsections.

Uniqueness

(A3) and (A4) assert existence only. As for fields, uniqueness is proved by comparing two candidates (Technique 1.1.8).

Proposition 1.2.27 (Uniqueness of the zero vector). A vector space has exactly one zero vector. That is, if $\mathbf{0}, \mathbf{0}' \in V$ both satisfy (A3), then $\mathbf{0} = \mathbf{0}'$.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)
Hint 1. Suppose $\mathbf{0}$ and $\mathbf{0}'$ both satisfy (A3). Find one expression that involves both.
Hint 2. Evaluate $\mathbf{0} + \mathbf{0}'$ twice: once using that $\mathbf{0}'$ satisfies (A3), and once, after (A1), using that $\mathbf{0}$ does.
Sketch. Each absorbs the other: $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}'$.

Suppose $\mathbf{0}$ and $\mathbf{0}'$ both satisfy (A3). Then

$$\begin{aligned} \mathbf{0} &= \mathbf{0} + \mathbf{0}' && (\text{as } \mathbf{0}' \text{ satisfies (A3)}) \\ &= \mathbf{0}' + \mathbf{0} && (\text{A1}) \\ &= \mathbf{0}' && (\text{as } \mathbf{0} \text{ satisfies (A3)}), \end{aligned}$$

as desired. □

The proof used only (A1) and (A3), so (A4) refers to a definite vector. When several vector spaces are in play, we write $\mathbf{0}_V$ for the zero vector of V. The proof is word for word that of Theorem 1.1.9(a), because (A1)–(A4) are the additive axioms of a field with vectors in place of scalars.

"Exactly one additive inverse" is two claims. Existence is (A4). Uniqueness becomes precise once we name two candidates.

Proposition 1.2.28 (Uniqueness of additive inverses). Every vector has exactly one additive inverse. That is, if $\mathbf{v}, \mathbf{w}, \mathbf{w}' \in V$ satisfy $\mathbf{v} + \mathbf{w} = \mathbf{0}$ and $\mathbf{v} + \mathbf{w}' = \mathbf{0}$, then $\mathbf{w} = \mathbf{w}'$.

Proof. ROUTINE

> *Uses.* Definition 1.2.1 (Vector space), Proposition 1.2.27 (Uniqueness of the zero vector)
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)
> *Hint 1.* Start from $\mathbf{w} = \mathbf{w} + \mathbf{0}$ and replace $\mathbf{0}$ by something involving $\mathbf{w}'$.
> *Hint 2.* Replace $\mathbf{0}$ by $\mathbf{v} + \mathbf{w}'$, then regroup so that $\mathbf{w}$ meets $\mathbf{v}$.
> *Sketch.* Insert $\mathbf{0} = \mathbf{v} + \mathbf{w}'$ into $\mathbf{w} = \mathbf{w} + \mathbf{0}$ and regroup: $\mathbf{w} = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{0} + \mathbf{w}' = \mathbf{w}'$.

Suppose $\mathbf{v} + \mathbf{w} = \mathbf{0}$ and $\mathbf{v} + \mathbf{w}' = \mathbf{0}$. Then

$$\begin{aligned} \mathbf{w} &= \mathbf{w} + \mathbf{0} && \text{(A3)} \\ &= \mathbf{w} + (\mathbf{v} + \mathbf{w}') && (\text{as } \mathbf{v} + \mathbf{w}' = \mathbf{0}) \\ &= (\mathbf{w} + \mathbf{v}) + \mathbf{w}' && \text{(A2)} \\ &= \mathbf{0} + \mathbf{w}' && (\text{as } \mathbf{w} + \mathbf{v} = \mathbf{v} + \mathbf{w} = \mathbf{0} \text{ by (A1)}) \\ &= \mathbf{w}' && \text{(A1 and A3),} \end{aligned}$$

as desired. □

Notation 1.2.29. The additive inverse of $\mathbf{v}$ is written $-\mathbf{v}$, and **subtraction** is defined by $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$.

The additive facts about fields carry over with the same proofs, and we record the ones we use.

Proposition 1.2.30 (Additive arithmetic). Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(a) The equation $\mathbf{u} + \mathbf{x} = \mathbf{w}$ has exactly one solution $\mathbf{x} \in V$, namely $\mathbf{x} = \mathbf{w} - \mathbf{u}$.

(b) If $\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{w}$, then $\mathbf{v} = \mathbf{w}$.

(c) $-(-\mathbf{v}) = \mathbf{v}$, and $-\mathbf{0} = \mathbf{0}$.

Proof. ROUTINE

> *Uses.* Definition 1.2.1 (Vector space), Proposition 1.2.28 (Uniqueness of additive inverses), Notation 1.2.29, Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$)
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)
> *Hint 1.* For (a), check that $\mathbf{w} - \mathbf{u}$ is a solution. For uniqueness, add $-\mathbf{u}$ on the left of $\mathbf{u} + \mathbf{x}$ and regroup.
> *Hint 2.* (b) is (a) with $\mathbf{w}$ replaced by $\mathbf{u} + \mathbf{v}$. For (c), to show a vector equals $-\mathbf{y}$, show it is an additive inverse of $\mathbf{y}$.
> *Sketch.* This is Proposition 1.1.13(a) and its consequences, whose proofs used only the additive axioms. $\mathbf{u} + (\mathbf{w} - \mathbf{u}) = (\mathbf{u} + (-\mathbf{u})) + \mathbf{w} = \mathbf{w}$, and any solution satisfies $\mathbf{x} = (-\mathbf{u} + \mathbf{u}) + \mathbf{x} = -\mathbf{u} + \mathbf{w}$. For (c), $\mathbf{v}$ is an inverse of $-\mathbf{v}$, and $\mathbf{0} + \mathbf{0} = \mathbf{0}$.

(a) First, $\mathbf{w} - \mathbf{u}$ is a solution:

$$\begin{aligned} \mathbf{u} + (\mathbf{w} + (-\mathbf{u})) &= \mathbf{u} + ((-\mathbf{u}) + \mathbf{w}) && \text{(A1)} \\ &= (\mathbf{u} + (-\mathbf{u})) + \mathbf{w} && \text{(A2)} \\ &= \mathbf{0} + \mathbf{w} = \mathbf{w} && \text{(A4, then A1 and A3).} \end{aligned}$$

Conversely, suppose $\mathbf{u} + \mathbf{x} = \mathbf{w}$. Then

$$\begin{aligned} \mathbf{x} &= \mathbf{0} + \mathbf{x} && \text{(A1 and A3)} \\ &= ((-\mathbf{u}) + \mathbf{u}) + \mathbf{x} && \text{(A4 and A1)} \\ &= (-\mathbf{u}) + (\mathbf{u} + \mathbf{x}) && \text{(A2)} \\ &= (-\mathbf{u}) + \mathbf{w} = \mathbf{w} - \mathbf{u} && \text{(hypothesis, then A1),} \end{aligned}$$

so $\mathbf{w} - \mathbf{u}$ is the only solution.

(b) Both $\mathbf{v}$ and $\mathbf{w}$ solve $\mathbf{u} + \mathbf{x} = \mathbf{u} + \mathbf{v}$, which has only one solution by (a). So $\mathbf{v} = \mathbf{w}$.

(c) By (A1) and (A4), $(-\mathbf{v}) + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$, so $\mathbf{v}$ is an additive inverse of $-\mathbf{v}$, and $\mathbf{v} = -(-\mathbf{v})$ by Proposition 1.2.28. Since $\mathbf{0} + \mathbf{0} = \mathbf{0}$ by (A3), $\mathbf{0}$ is an additive inverse of $\mathbf{0}$, so $-\mathbf{0} = \mathbf{0}$. □

Remark 1.2.31. A set with an operation satisfying (A1)–(A4) is called an **abelian group** (Section 14.1). Everything in this subsection is a fact about abelian groups: scalars have not appeared yet, and the same statements hold in $(F, +)$.

Scalars and vectors together

Intuition. Every fact so far used only (A1)–(A4). A fact relating the additive objects 0 or $\mathbf{0}$ to scalar multiplication must use (A7) or (A8), the only axioms that involve addition and scalar multiplication together. For the scalar 0, the additive fact to start from is $0 + 0 = 0$ in F, and (A8) carries it over to vectors, just as distributivity did for $0a = 0$ in a field (Proposition 1.1.15).

Proposition 1.2.32 (The scalar zero). For every $\mathbf{v} \in V$, $0\mathbf{v} = \mathbf{0}$.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.30 (Additive arithmetic)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Which axiom can turn the scalar equation $0 + 0 = 0$ into a statement about vectors?
Hint 2. Apply (A8) to $(0 + 0)\mathbf{v}$. You get $0\mathbf{v} + 0\mathbf{v} = 0\mathbf{v}$. Compare with $0\mathbf{v} + \mathbf{0} = 0\mathbf{v}$ and cancel.
Sketch. (A8) turns $0 + 0 = 0$ into $0\mathbf{v} + 0\mathbf{v} = 0\mathbf{v} = 0\mathbf{v} + \mathbf{0}$, and cancelling $0\mathbf{v}$ (Proposition 1.2.30(b)) gives $0\mathbf{v} = \mathbf{0}$.

Let $\mathbf{v} \in V$. Then

$$\begin{aligned} 0\mathbf{v} + 0\mathbf{v} &= (0 + 0)\mathbf{v} && \text{(A8)} \\ &= 0\mathbf{v} && (\text{as } \ 0 + 0 = 0 \ \text{ in } \ F\,) \\ &= 0\mathbf{v} + \mathbf{0} && \text{(A3).} \end{aligned}$$

Cancelling $0\mathbf{v}$ from both sides (Proposition 1.2.30(b)) gives $0\mathbf{v} = \mathbf{0}$. □

The companion statement is about the zero *vector* multiplied by any scalar. It is a different statement, and its first proof below uses the other distributive law. The second proof reduces it to Proposition 1.2.32.

Proposition 1.2.33 (The zero vector). For every $a \in F$, $a\mathbf{0} = \mathbf{0}$.

Proof (via (A7)). ROUTINE

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.30 (Additive arithmetic)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Imitate the proof of Proposition 1.2.32, starting from the vector equation $\mathbf{0} + \mathbf{0} = \mathbf{0}$.
Hint 2. Apply (A7) to $a(\mathbf{0} + \mathbf{0})$ and cancel $a\mathbf{0}$.
Sketch. (A7) turns $\mathbf{0} + \mathbf{0} = \mathbf{0}$ into $a\mathbf{0} + a\mathbf{0} = a\mathbf{0} = a\mathbf{0} + \mathbf{0}$; cancel $a\mathbf{0}$.

Let $a \in F$. Then

$$\begin{aligned} a\mathbf{0} + a\mathbf{0} &= a(\mathbf{0} + \mathbf{0}) && \text{(A7)} \\ &= a\mathbf{0} && \text{(A3)} \\ &= a\mathbf{0} + \mathbf{0} && \text{(A3),} \end{aligned}$$

and cancelling $a\mathbf{0}$ (Proposition 1.2.30(b)) gives $a\mathbf{0} = \mathbf{0}$. □

Proof (from $0\mathbf{v} = \mathbf{0}$). ROUTINE

Uses. Proposition 1.2.32 (The scalar zero), Proposition 1.1.15 (Multiplying by zero), Definition 1.2.1 (Vector space)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Proposition 1.2.32 says every vector times the scalar 0 is $\mathbf{0}$. Apply it to the vector $\mathbf{0}$ to rewrite $\mathbf{0}$.
Hint 2. Write $a\mathbf{0} = a(0\mathbf{0})$ and use (A6) to move the scalar 0 next to a.
Sketch. $a\mathbf{0} = a(0\mathbf{0}) = (a0)\mathbf{0} = 0\mathbf{0} = \mathbf{0}$, by Proposition 1.2.32, (A6), $a0 = 0$ in F, and Proposition 1.2.32 again.

Let $a \in F$. By Proposition 1.2.32 applied to the vector $\mathbf{0}$, we have $\mathbf{0} = 0\mathbf{0}$. Using this, then (A6), then $a0 = 0$ in F (Proposition 1.1.15), and finally Proposition 1.2.32 once more,

$$a\mathbf{0} = a(0\mathbf{0}) = (a0)\mathbf{0} = 0\mathbf{0} = \mathbf{0}.$$ □

The first proof is the template of Proposition 1.2.32 with the other distributive law. The second uses no distributive law directly: it writes $\mathbf{0}$ as $0\mathbf{0}$ and moves the scalar 0 through (A6), so the work is done by the field ($a0 = 0$) and by the earlier result, through which (A8) enters. Reducing a new claim to one already proved is often shorter than starting again from the axioms.

Intuition. To show that a vector equals $-\mathbf{v}$, it suffices to show that it is *an* additive inverse of $\mathbf{v}$, by Proposition 1.2.28. For $(-1)\mathbf{v}$, write $\mathbf{v} = 1\mathbf{v}$ and combine the two terms with (A8).

Proposition 1.2.34 (Negatives and the scalar -1). For all $a \in F$ and $\mathbf{v} \in V$:

(a) $(-1)\mathbf{v} = -\mathbf{v}$;

(b) $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$.

Proof. ROUTINE

Uses. Proposition 1.2.28 (Uniqueness of additive inverses), Proposition 1.2.32 (The scalar zero), Proposition 1.1.16 (Sign rules), Definition 1.2.1 (Vector space)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Proving from axioms (Technique 1.1.7)
Hint 1. For (a), compute $\mathbf{v} + (-1)\mathbf{v}$. Write $\mathbf{v}$ as $1\mathbf{v}$ first.
Hint 2. (A8) gives $(1 + (-1))\mathbf{v} = 0\mathbf{v}$, and Proposition 1.2.32 finishes. Conclude with Proposition 1.2.28.
Hint 3. For (b), write $-a = (-1)a$ (Proposition 1.1.16) and move the factor -1 with (A6).
Sketch. $\mathbf{v} + (-1)\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$, so $(-1)\mathbf{v} = -\mathbf{v}$ by uniqueness of inverses. Then $(-a)\mathbf{v} = ((-1)a)\mathbf{v} = (-1)(a\mathbf{v}) = -(a\mathbf{v})$, and $a(-\mathbf{v}) = a((-1)\mathbf{v}) = (a(-1))\mathbf{v} = (-a)\mathbf{v}$.

(a) We compute

$$\begin{aligned} \mathbf{v} + (-1)\mathbf{v} &= 1\mathbf{v} + (-1)\mathbf{v} && \text{(A5)} \\ &= (1 + (-1))\mathbf{v} && \text{(A8)} \\ &= 0\mathbf{v} = \mathbf{0}, \end{aligned}$$

where the last step is Proposition 1.2.32. So $(-1)\mathbf{v}$ is an additive inverse of $\mathbf{v}$, and $(-1)\mathbf{v} = -\mathbf{v}$ by Proposition 1.2.28.

(b) By Proposition 1.1.16, $-a = (-1)a = a(-1)$. Using (A6) and (a),

$$(-a)\mathbf{v} = ((-1)a)\mathbf{v} = (-1)(a\mathbf{v}) = -(a\mathbf{v}),$$

and

$$a(-\mathbf{v}) = a((-1)\mathbf{v}) = (a(-1))\mathbf{v} = (-a)\mathbf{v}.$$ □

In particular $a(\mathbf{v} - \mathbf{w}) = a\mathbf{v} + a(-\mathbf{w}) = a\mathbf{v} - a\mathbf{w}$ by (A7) and (b), and similarly $(a - b)\mathbf{v} = a\mathbf{v} - b\mathbf{v}$ by (A8) and (b).

Intuition. In a field, $ab = 0$ forces $a = 0$ or $b = 0$ (Proposition 1.1.17), and the proof multiplied by a^{-1}. The vector version has the same proof. It is the first place where the multiplicative inverses of F matter for vectors.

Proposition 1.2.35 (Zero products). Let $a \in F$ and $\mathbf{v} \in V$. If $a\mathbf{v} = \mathbf{0}$, then $a = 0$ or $\mathbf{v} = \mathbf{0}$. Consequently, $a\mathbf{v} = \mathbf{0}$ if and only if $a = 0$ or $\mathbf{v} = \mathbf{0}$.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.33 (The zero vector), Proposition 1.2.32 (The scalar zero), Definition 1.1.1 (Field)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. To prove "$a = 0$ or $\mathbf{v} = \mathbf{0}$", assume $a \neq 0$ and prove $\mathbf{v} = \mathbf{0}$.

Hint 2. If $a \neq 0$, then a^{-1} exists. Write $\mathbf{v} = 1\mathbf{v} = (a^{-1}a)\mathbf{v}$ and use (A6).
Sketch. If $a \neq 0$, then $\mathbf{v} = 1\mathbf{v} = (a^{-1}a)\mathbf{v} = a^{-1}(a\mathbf{v}) = a^{-1}\mathbf{0} = \mathbf{0}$. The converse is Proposition 1.2.32 and Proposition 1.2.33.

Suppose $a\mathbf{v} = \mathbf{0}$. If $a = 0$ we are done, so suppose $a \neq 0$. Then a^{-1} exists, and

$$\begin{aligned} \mathbf{v} &= 1\mathbf{v} && \text{(A5)} \\ &= (a^{-1}a)\mathbf{v} && \text{(multiplicative inverse in } F\text{)} \\ &= a^{-1}(a\mathbf{v}) && \text{(A6)} \\ &= a^{-1}\mathbf{0} && \text{(as } a\mathbf{v} = \mathbf{0}\text{)} \\ &= \mathbf{0}, \end{aligned}$$

where the last step is Proposition 1.2.33. Conversely, if $a = 0$ then $a\mathbf{v} = \mathbf{0}$ by Proposition 1.2.32, and if $\mathbf{v} = \mathbf{0}$ then $a\mathbf{v} = \mathbf{0}$ by Proposition 1.2.33. □

Corollary 1.2.36 (Cancelling scalars and vectors). Let $a, b \in F$ and $\mathbf{v}, \mathbf{w} \in V$.

(a) If $a\mathbf{v} = a\mathbf{w}$ and $a \neq 0$, then $\mathbf{v} = \mathbf{w}$.

(b) If $a\mathbf{v} = b\mathbf{v}$ and $\mathbf{v} \neq \mathbf{0}$, then $a = b$.

Proof. ROUTINE

Uses. Proposition 1.2.35 (Zero products), Proposition 1.2.34 (Negatives and the scalar -1), Proposition 1.2.30 (Additive arithmetic)
Hint 1. Move everything to one side: $a\mathbf{v} = a\mathbf{w}$ is equivalent to $a\mathbf{v} - a\mathbf{w} = \mathbf{0}$. Factor.
Hint 2. $a\mathbf{v} - a\mathbf{w} = a(\mathbf{v} - \mathbf{w})$ and $a\mathbf{v} - b\mathbf{v} = (a - b)\mathbf{v}$. Now use Proposition 1.2.35.
Sketch. $a(\mathbf{v} - \mathbf{w}) = a\mathbf{v} - a\mathbf{w} = \mathbf{0}$ with $a \neq 0$ forces $\mathbf{v} - \mathbf{w} = \mathbf{0}$; and $(a - b)\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$ forces $a - b = 0$.

(a) By (A7) and Proposition 1.2.34(b), $a(\mathbf{v} - \mathbf{w}) = a\mathbf{v} + a(-\mathbf{w}) = a\mathbf{v} - a\mathbf{w}$, which is $\mathbf{0}$ because $a\mathbf{v} = a\mathbf{w}$. Since $a \neq 0$, Proposition 1.2.35 gives $\mathbf{v} - \mathbf{w} = \mathbf{0}$. Hence, by (A3), (A4) with (A1), (A2), and (A1) with (A3),

$$\mathbf{v} = \mathbf{v} + ((-\mathbf{w}) + \mathbf{w}) = (\mathbf{v} - \mathbf{w}) + \mathbf{w} = \mathbf{0} + \mathbf{w} = \mathbf{w}.$$

(b) By (A8) and Proposition 1.2.34(b), $(a - b)\mathbf{v} = a\mathbf{v} + (-b)\mathbf{v} = a\mathbf{v} - b\mathbf{v} = \mathbf{0}$. Since $\mathbf{v} \neq \mathbf{0}$, Proposition 1.2.35 gives $a - b = 0$, so $a = b$. □

Part (b) says that if $\mathbf{v} \neq \mathbf{0}$, the multiples $a\mathbf{v}$ for different scalars a are different vectors. So a vector space other than $\{\mathbf{0}\}$ has at least as many elements as F (Exercise 1.2.42).

Exercise 1.2.37. ROUTINE Let $a \in F$ and $\mathbf{u}, \mathbf{w} \in V$.

(a) Suppose $a \neq 0$. Show that the equation $a\mathbf{x} + \mathbf{u} = \mathbf{w}$ has exactly one solution $\mathbf{x} \in V$, and find it.

(b) Suppose $a = 0$. For which $\mathbf{u}$ and $\mathbf{w}$ does the equation have a solution, and how many?

Solution on page 1034

Testing claims on small cases

Facts that are second nature in $\mathbb{R}^2$ can fail in other vector spaces, and the quickest way to find out is to compute in a small one.

Technique 1.2.38 (Try small cases and small fields). Before proving a general claim, test it on the smallest examples available: the zero space, F^1 and F^2, and small fields, especially $\mathbb{F}_2$, where $1 + 1 = 0$. Small cases are where counterexamples live. When the claim survives them, the computation often shows why it is true and which hypothesis a proof must use.

Reach for it when a claim feels obvious from experience with $\mathbb{R}^2$, and whenever a hypothesis such as $1 + 1 \neq 0$ or $V \neq \{\mathbf{0}\}$ might be needed. The canonical example is Example 1.2.39: $\mathbf{v} = -\mathbf{v}$ forces $\mathbf{v} = \mathbf{0}$ over $\mathbb{R}$, but not over $\mathbb{F}_2$.

Used in. Exercise 1.2.49, Exercise 1.3.30, Exercise 1.3.43, Exercise 1.4.42, Exercise 1.7.6, Exercise 1.7.13, Exercise 1.7.15, Exercise 2.8.5, Problem 2.8, Problem 2.14, Problem 2.20, Problem 3.6, Problem 4.17, Exercise 5.6.10, Exercise 5.6.15, Exercise 1.3.46, Exercise 1.4.32, Exercise 1.5.36, Exercise 7.1.53, Exercise 7.5.31, Exercise 7.6.41.

Example 1.2.39 (When does $\mathbf{v} = -\mathbf{v}$ force $\mathbf{v} = \mathbf{0}$?). Over $\mathbb{R}$, a vector equal to its own negative is $\mathbf{0}$. Over $\mathbb{F}_2$ this fails: in $\mathbb{F}_2^2$, $(1,0) + (1,0) = (1+1,0) = (0,0)$, so $(1,0) = -(1,0)$, although $(1,0) \neq \mathbf{0}$. So the answer depends on the field, and the computation suggests why. By Proposition 1.2.28, $\mathbf{v} = -\mathbf{v}$ holds exactly when $\mathbf{v} + \mathbf{v} = \mathbf{0}$, and by (A5) and (A8),

$$\mathbf{v} + \mathbf{v} = 1\mathbf{v} + 1\mathbf{v} = (1+1)\mathbf{v}.$$

If $1 + 1 \neq 0$ in F, then $(1+1)\mathbf{v} = \mathbf{0}$ forces $\mathbf{v} = \mathbf{0}$ by Proposition 1.2.35. If $1 + 1 = 0$ in F, then $\mathbf{v} + \mathbf{v} = 0\mathbf{v} = \mathbf{0}$ for every $\mathbf{v}$, so every vector is its own negative. The test over $\mathbb{F}_2$ produced both the counterexample and the hypothesis $1 + 1 \neq 0$ under which the claim holds (compare Aside (Characteristic)).

Example 1.2.40 (Addition modulo 1, again). We show that no scalar multiplication makes $V = [0,1)$ a real vector space with the addition $\oplus$ of Non-example 1.2.26. In every real vector space, $\mathbf{v} + \mathbf{v} = \mathbf{0}$ forces $\mathbf{v} = \mathbf{0}$, by Example 1.2.39, because $1 + 1 = 2 \neq 0$ in $\mathbb{R}$. But in V,

$$\frac{1}{2} \oplus \frac{1}{2} = 1 \bmod 1 = 0,$$

and $1/2$ is not the zero vector 0 (the zero vector is unique, Proposition 1.2.27, and 0 satisfies (A3) for $\oplus$). The argument uses only consequences of the axioms that every candidate scalar multiplication must satisfy, so it rules them all out at once.

This is a different way to show that something is not a vector space. Instead of finding an axiom that fails for given operations, find a consequence of the axioms that involves only the part we cannot change, here the addition, and check that it fails.

Exercises

Exercise 1.2.41. ROUTINE Let $p, q \in F[x]$ be nonzero. Prove or disprove each statement.

(a) $\deg(p+q) = \max(\deg p, \deg q)$.

(b) If $\deg(p+q) < \deg p$, then $\deg p = \deg q$, and the leading coefficients of p and q add to 0.

(c) The set of polynomials in $\mathbb{R}[x]$ of degree exactly 2, with the operations of $\mathbb{R}[x]$, is a real vector space.

Solution on page 1035

Exercise 1.2.42. ROUTINE Suppose $V \neq \{\mathbf{0}\}$. Show that V has at least as many elements as F. In particular, a vector space over an infinite field is either $\{\mathbf{0}\}$ or infinite, and if V is finite, then so is F.

Solution on page 1035

Exercise 1.2.43. STANDARD

(a) Find the error in the following argument.

Claim. If $\mathbf{u}, \mathbf{v} \in V$ and $\mathbf{u} + \mathbf{v} = \mathbf{u}$, then $\mathbf{v} = \mathbf{0}$.

Proof. By Proposition 1.2.27, $\mathbf{0}$ is the only vector $\mathbf{z}$ with $\mathbf{u} + \mathbf{z} = \mathbf{u}$. Since $\mathbf{u} + \mathbf{v} = \mathbf{u}$, we get $\mathbf{v} = \mathbf{0}$.

(b) Prove the claim correctly.

(c) Give an example of a set with an addition satisfying (A1)–(A3), in which $\mathbf{u} + \mathbf{v} = \mathbf{u}$ for some $\mathbf{u}$ and some $\mathbf{v}$ that is not the identity element of (A3).

Solution on page 1035

Exercise 1.2.44. STANDARD

(a) For which fields F is there a vector space over F with exactly two elements?

(b) Show that no vector space over $\mathbb{F}_2$ has exactly three elements, and no vector space over $\mathbb{F}_3$ has exactly four.

Solution on page 1035

Exercise 1.2.45. STANDARD For a positive integer n and $\mathbf{v} \in V$, write $n * \mathbf{v} = \mathbf{v} + \cdots + \mathbf{v}$ (n terms), and similarly $n * 1 = 1 + \cdots + 1 \in F$ (written $n \cdot 1$ in Definition 1.1.28; we use $*$ here because juxtaposition already means scalar multiplication). Suppose $V \neq \{\mathbf{0}\}$. Show that the following are equivalent:

(i) $n * 1 = 0$ in F;

(ii) $n * \mathbf{v} = \mathbf{0}$ for every $\mathbf{v} \in V$;

(iii) $n * \mathbf{v} = \mathbf{0}$ for some $\mathbf{v} \neq \mathbf{0}$ in V.

Deduce that $\operatorname{char} F$ (Definition 1.1.28) can be read off from the addition of V: it is the least $n \geq 1$ with $n * \mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in V$, or 0 if there is none.

Solution on page 1036

Exercise 1.2.46. STANDARD

(a) Is there a scalar multiplication that makes $\mathbb{Z}$, with its usual addition, a vector space over $\mathbb{F}_p$?

(b) Is there a field F and a scalar multiplication that make $\mathbb{Z}/4\mathbb{Z}$, with its usual addition, a vector space over F?

Solution on page 1036

Exercise 1.2.47. STANDARD Is there a scalar multiplication $\odot : \mathbb{C} \times \mathbb{R} \to \mathbb{R}$ that makes $\mathbb{R}$, with its usual addition, a complex vector space, and that agrees with the usual product for real scalars, that is, $a \odot x = ax$ for $a \in \mathbb{R}$?

Solution on page 1037

Exercise 1.2.48. STANDARD Keep the usual addition on $\mathbb{C}^2$. For $a = x + yi \in \mathbb{C}$ with $x, y \in \mathbb{R}$, write $\bar{a} = x - yi$, $|a| = \sqrt{x^2 + y^2}$, and $\operatorname{Re} a = x$. Which of the following scalar multiplications make $\mathbb{C}^2$ a complex vector space?

(a) $a \odot \mathbf{v} = \bar{a}\mathbf{v}$.

(b) $a \odot \mathbf{v} = |a|\mathbf{v}$.

(c) $a \odot \mathbf{v} = (\operatorname{Re} a)\mathbf{v}$.

Solution on page 1037

Exercise 1.2.49. STANDARD Let F be a field. On F^2, keep the usual addition and define $a \odot (x, y) = (ax, a^2 y)$. For which fields F is this a vector space over F?

Solution on page 1037

Exercise 1.2.50. STANDARD Let $P = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$ be the parabola. Find operations that make P a real vector space in which $(1, 1) \oplus (2, 4) = (3, 9)$. In your structure, find the zero vector, the additive inverse of $(2, 4)$, and $3 \odot (-1, 1)$.

Solution on page 1038

Exercise 1.2.51. STANDARD

(a) Find all $f \in \mathbb{F}_2[x]_{\leq 3}$ whose polynomial function $\mathbb{F}_2 \to \mathbb{F}_2$ is the zero function. How many polynomials does $\mathbb{F}_2[x]_{\leq 3}$ contain, and how many functions $\mathbb{F}_2 \to \mathbb{F}_2$ are there? When do two polynomials in $\mathbb{F}_2[x]_{\leq 3}$ have the same polynomial function?

(b) Let p be a prime. Show that $x^p - x \in \mathbb{F}_p[x]$ is a nonzero polynomial whose polynomial function is zero.

Solution on page 1038

Exercise 1.2.52. STANDARD Let V be a vector space over $\mathbb{Q}$.

(a) Show that for every integer $n \geq 1$ and $\mathbf{w} \in V$, the scalar multiple $n\mathbf{w}$ equals $\mathbf{w} + \cdots + \mathbf{w}$ (n terms).

(b) Let $q = m/n$ with $m, n \in \mathbb{Z}$ and $n \geq 1$. Show that $q\mathbf{v}$ is the only vector $\mathbf{w} \in V$ with $n\mathbf{w} = m\mathbf{v}$.

(c) Deduce that the scalar multiplication of a $\mathbb{Q}$-vector space is determined by its addition: if $\odot$ and $\odot'$ both make $(V, +)$ a vector space over $\mathbb{Q}$, then $q \odot \mathbf{v} = q \odot' \mathbf{v}$ for all q and $\mathbf{v}$.

Solution on page 1038

Exercise 1.2.53. HARD Let $V = (-1, 1)$, and define $u \oplus v = (u + v)/(1 + uv)$. (In units where the speed of light is 1, this is how velocities along a line add in special relativity.)

(a) Show that $u \oplus v \in V$ for $u, v \in V$.

(b) Find a scalar multiplication $\odot$ that makes $(V, \oplus, \odot)$ a real vector space. Compute $2 \odot v$ and $(1/2) \odot v$ explicitly, and identify the zero vector and the additive inverse of v.

Solution on page 1039

Exercise 1.2.54. HARD

(a) Find the error in the following argument.

Claim. (A4) follows from the other axioms.

Proof. Let $\mathbf{v} \in V$. By (A5), (A8), and Proposition 1.2.32,

$$\mathbf{v} + (-1)\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}.$$

So $(-1)\mathbf{v}$ is an additive inverse of $\mathbf{v}$.

(b) Show that the claim is false: construct a set with operations that satisfy every axiom except (A4).

Solution on page 1040

Exercise 1.2.55. HARD Suppose V, with operations $+$ and $\cdot$, satisfies every vector space axiom except possibly (A5). Let

$$W = \{\mathbf{v} \in V \mid 1\mathbf{v} = \mathbf{v}\}, \quad N = \{\mathbf{v} \in V \mid a\mathbf{v} = \mathbf{0} \text{ for all } a \in F\}.$$

(a) Show that every $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{n}$ with $\mathbf{w} \in W$ and $\mathbf{n} \in N$, in exactly one way.

(b) Show that W, with the operations of V, is a vector space over F.

(c) Show that (A5) holds if and only if $N = \{\mathbf{0}\}$.

Solution on page 1040

Exercise 1.2.56. HARD Suppose V, with operations $+$ and $\cdot$, satisfies (A2)–(A8), but not necessarily (A1). (Note that (A3) and (A4) only say $\mathbf{v}+\mathbf{0}=\mathbf{v}$ and $\mathbf{v}+\mathbf{w}=\mathbf{0}$, with $\mathbf{0}$ and $\mathbf{w}$ on the right. Without (A1) we cannot prove that the zero vector is unique, so fix one element $\mathbf{0}$ satisfying (A3) for which (A4) holds, and let $\mathbf{0}$ always mean this element.) Show that $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ for all $\mathbf{u},\mathbf{v}\in V$.

Solution on page 1041

Exercise 1.2.57. CHALLENGE Let $f:\mathbb{R}\to\mathbb{R}$ be a function, and on $V=\mathbb{R}$ with its usual addition define $a\odot x=f(a)x$.

(a) Show that $(V,+,\odot)$ is a real vector space if and only if $f(1)=1$, $f(a+b)=f(a)+f(b)$, and $f(ab)=f(a)f(b)$ for all $a,b\in\mathbb{R}$.

(b) Show that the only such f is the identity. So the usual product is the only scalar multiplication of this form.

Solution on page 1042

1.3 Lists, linear combinations, and span

A vector space has two operations, and each of them combines at most two vectors. Applying them over and over to finitely many vectors $\mathbf{v}_1,\ldots,\mathbf{v}_k$ produces sums $a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k$, and every expression built from the $\mathbf{v}_i$ with the two operations can be brought into this form (Lemma 1.3.9). Two questions about these sums run through the rest of the book: which vectors they produce, and in how many ways. This section answers the first question with the span; the second leads to linear independence in Section 1.5. We begin by making "finitely many vectors $\mathbf{v}_1,\ldots,\mathbf{v}_k$" precise, in a form where the order is fixed and repetitions are allowed.

1.3.1 Lists

Definition 1.3.1 (List). Let V be a set and let $k\geq 0$ be an integer. A **list of length** k in V is a function $\{1,\ldots,k\}\to V$. Writing $\mathbf{v}_i$ for its value at i, we write the list as

$$\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k,$$

without parentheses, and call $\mathbf{v}_i$ its i**-th entry**. Two lists are equal when they have the same length and the same i-th entry for every i. Lists are named by calligraphic letters, as in $\mathcal{L}=\mathbf{v}_1,\ldots,\mathbf{v}_k$, and $|\mathcal{L}|=k$ is the **length** of $\mathcal{L}$. For $k=0$ the domain is empty, and there is exactly one function from the empty set to V. It is the **empty list**, written $()$.

So a list is a finite sequence. Defining it as a function settles what equality of lists means, and it makes the empty list an ordinary list rather than an exception. A list of length one is written $\mathbf{v}$, and a list of length two $\mathbf{v},\mathbf{w}$. Columns in F^n keep their parentheses, so $(1,0),(0,1)$ is a list of length two in $\mathbb{R}^2$, and $(1,0)$ alone is a list of length one.

Example 1.3.2 (Lists versus sets). In $\mathbb{R}^2$, the lists $(1,0),(0,1)$ and $(0,1),(1,0)$ are different, since their first entries differ, while the sets $\{(1,0),(0,1)\}$ and $\{(0,1),(1,0)\}$ are equal. In any vector space, $\mathbf{0},\mathbf{0},\mathbf{0}$ is a list of length 3. It is different from the list $\mathbf{0}$ of length 1 and from the empty list $()$, although $\{\mathbf{0},\mathbf{0},\mathbf{0}\}=\{\mathbf{0}\}$ as sets.

We work with lists because a linear combination gives each entry its own coefficient. In the list $\mathbf{v},\mathbf{v}$, the vector $\mathbf{v}$ receives two coefficients, and $1\mathbf{v}+(-1)\mathbf{v}$ produces $\mathbf{0}$ in a way that the set $\{\mathbf{v}\}$ cannot record. This kind of redundancy is what linear independence detects (Section 1.5), and the order of a list will fix the order of coordinates (Section 1.6).

Notation 1.3.3. Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V.

- We write $\mathbf{v} \in \mathcal{L}$ if $\mathbf{v} = \mathbf{v}_i$ for some i. For a subset $S \subseteq V$, we write $\mathcal{L} \subseteq S$ if every entry of $\mathcal{L}$ lies in S.
- For a list $\mathcal{M} = \mathbf{w}_1, ..., \mathbf{w}_m$, the **concatenation** $\mathcal{L}, \mathcal{M}$ is the list $\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{w}_1, ..., \mathbf{w}_m$ of length $k + m$. In particular, $\mathcal{L}, \mathbf{w}$ is $\mathcal{L}$ with $\mathbf{w}$ appended.
- A **sublist** of $\mathcal{L}$ is a list $\mathbf{v}_{i_1}, ..., \mathbf{v}_{i_m}$ with $1 \le i_1 < \cdots < i_m \le k$. It is obtained by deleting some entries and keeping the rest in order. A **reordering** of $\mathcal{L}$ is a list $\mathbf{v}_{\sigma(1)}, ..., \mathbf{v}_{\sigma(k)}$, where σ is a bijection from $\{1, ..., k\}$ to itself.

1.3.2 Linear combinations

From now on, V is a vector space over a field F.

Definition 1.3.4 (Linear combination). Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V. A **linear combination** of $\mathcal{L}$ is an expression

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k, \quad \text{also written} \quad \sum_{i=1}^{k} a_i\mathbf{v}_i,$$

where $a_1, ..., a_k \in F$ are its **coefficients**. Two linear combinations of $\mathcal{L}$ are the same exactly when their coefficients are the same. Carrying out the operations of V in the expression produces a vector of V, and we say that the linear combination **represents** that vector. The empty list has exactly one linear combination, the empty sum, and it represents $\mathbf{0}$.

To carry out the operations, compute each term $a_i\mathbf{v}_i$ and add the terms from left to right:

$$\sum_{i=1}^{0} a_i\mathbf{v}_i = \mathbf{0}, \quad \sum_{i=1}^{k} a_i\mathbf{v}_i = \left(\sum_{i=1}^{k-1} a_i\mathbf{v}_i\right) + a_k\mathbf{v}_k \quad (k \ge 1). \tag{1.3.1}$$

By (A1) and (A2), neither the order nor the grouping of the terms affects the result; the argument is the one for sums in a field (Section 1.1.2). The empty sum must be $\mathbf{0}$ if the recursion is to give $\sum_{i=1}^{1} a_i\mathbf{v}_i = a_1\mathbf{v}_1$.

The definition separates two things that everyday usage runs together: the expression, which is determined by its coefficients once $\mathcal{L}$ is fixed, and the vector it represents. Formally, a linear combination of $\mathcal{L}$ is just its list of coefficients $a_1, ..., a_k$. When we write $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, the equation says that the linear combination on the right represents $\mathbf{w}$. We also say "$\mathbf{w}$ is a linear combination of $\mathcal{L}$" to mean that some linear combination of $\mathcal{L}$ represents $\mathbf{w}$. This phrase is standard and harmless, as long as we keep track of which object is meant. The next example shows why the distinction matters.

Example 1.3.5 (One vector, many representations). In $\mathbb{R}^2$, let $\mathcal{L}_1 = (0,0), (1,-1), (-\pi, \pi)$. The linear combinations

$$0(0,0) + 1(1,-1) + 0(-\pi,\pi) \quad \text{and} \quad \sqrt{87}(0,0) + 0(1,-1) + \left(-\frac{1}{\pi}\right)(-\pi,\pi)$$

are different, because their coefficients differ, but both represent $(1,-1)$. So "the linear combination equal to $(1,-1)$" does not name a single linear combination of $\mathcal{L}_1$. By contrast, in F^2 the vector (x,y) is represented by exactly one linear combination of $(1,0), (0,1)$: the combination $a(1,0) + b(0,1)$ represents (a,b), which is (x,y) only if $a = x$ and $b = y$. Lists with this uniqueness property are the linearly independent ones (Section 1.5).

Exercise 1.3.6. ROUTINE For the list $\mathcal{L}_1$ of Example 1.3.5, find all linear combinations of $\mathcal{L}_1$ that represent $(1,-1)$, and all that represent $\mathbf{0}$. How are the two families related?

Solution on page 1042

Example 1.3.7 (Standard lists). For $1 \le i \le n$, let $\mathbf{e}_i \in F^n$ be the column with 1 in position i and 0 elsewhere. For every $\mathbf{x} = (x_1, \dots, x_n) \in F^n$,

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

because by the coordinatewise operations, the i-th coordinate of the right side is $x_1 \cdot 0 + \cdots + x_i \cdot 1 + \cdots + x_n \cdot 0 = x_i$. So every vector of F^n is a linear combination of $\mathbf{e}_1, \dots, \mathbf{e}_n$, the standard basis of F^n (Definition 1.6.8).

In $F[x]$, Notation 1.2.15 says that the polynomial with coefficients $(a_0, \dots, a_d, 0, 0, \dots)$ is represented by the linear combination $a_0 \cdot 1 + a_1 x + \cdots + a_d x^d$ of the list $1, x, \dots, x^d$. The coefficients of the polynomial and those of the linear combination coincide.

Example 1.3.8 (Is a vector a linear combination?). Is $(1, 2, 3) \in \mathbb{R}^3$ a linear combination of $(1, 1, 0), (0, 1, 1)$? An arbitrary linear combination represents

$$a(1, 1, 0) + b(0, 1, 1) = (a, a + b, b).$$

For this to be $(1, 2, 3)$, we need $a = 1$, $b = 3$, and $a + b = 2$. The first two equations force $a + b = 4$, so no linear combination works, and the answer is no. In contrast, $(1, 4, 3) = 1(1, 1, 0) + 3(0, 1, 1)$.

The answer depends on the field. Read in $\mathbb{F}_2^3$, the same coordinates give the vector $(1, 0, 1)$, and the equations become $a = 1$, $b = 1$, and $a + b = 0$. These hold, since $1 + 1 = 0$ in $\mathbb{F}_2$. Indeed, $(1, 1, 0) + (0, 1, 1) = (1, 2, 1) = (1, 0, 1)$ in $\mathbb{F}_2^3$.

In general, deciding whether $\mathbf{w}$ is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_k$ means deciding whether the equation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{w}$ can be solved for the coefficients. In F^n this is a system of linear equations, and Section 3.1 develops the general method (Technique 3.1.9).

Computing with linear combinations relies on the following rules.

Lemma 1.3.9 (Arithmetic of linear combinations). Let $\mathcal{L} = \mathbf{v}_1, \dots, \mathbf{v}_k$ be a list in V, and let $a_1, \dots, a_k, b_1, \dots, b_k, c \in F$. Then

(a) $\sum_{i=1}^k a_i\mathbf{v}_i + \sum_{i=1}^k b_i\mathbf{v}_i = \sum_{i=1}^k (a_i + b_i)\mathbf{v}_i$;

(b) $c\sum_{i=1}^k a_i\mathbf{v}_i = \sum_{i=1}^k (ca_i)\mathbf{v}_i$;

(c) $\sum_{i=1}^k 0\mathbf{v}_i = \mathbf{0}$.

Here each sum denotes the vector that the linear combination represents.

Proof. ROUTINE

Uses. (1.3.1), Definition 1.2.1 (Vector space), Proposition 1.2.32 (The scalar zero), Proposition 1.2.33 (The zero vector), induction on k

Hint 1. Use induction on k, splitting off the last term of each sum as in (1.3.1).

Hint 2. In the inductive step of (a), regroup $(\mathbf{x} + a\mathbf{v}) + (\mathbf{y} + b\mathbf{v})$ as $(\mathbf{x} + \mathbf{y}) + (a\mathbf{v} + b\mathbf{v})$ and apply (A8) to the second part. For (b), apply (A7) and then (A6).

Sketch. Induction on k. For $k = 0$ every sum is $\mathbf{0}$, and $\mathbf{0} + \mathbf{0} = \mathbf{0}$, $c\mathbf{0} = \mathbf{0}$. In the inductive step, write each sum as the sum of its first k terms plus its last term. Then (a) follows by regrouping with (A1) and (A2) and applying (A8) to the last terms, (b) follows from (A7) and (A6), and (c) from $0\mathbf{v}_{k+1} = \mathbf{0}$.

For a list of scalars $a = a_1, a_2, \dots$, write $S_k(a) = \sum_{i=1}^k a_i\mathbf{v}_i$, and write $a + b$ and ca for the lists with entries $a_i + b_i$ and ca_i. By (1.3.1), $S_0(a) = \mathbf{0}$ and $S_{k+1}(a) = S_k(a) + a_{k+1}\mathbf{v}_{k+1}$. We use induction on k.

For $k = 0$ all the sums are empty. Then (a) reads $\mathbf{0} + \mathbf{0} = \mathbf{0}$, which is (A3); (b) reads $c\mathbf{0} = \mathbf{0}$, which is Proposition 1.2.33; and (c) reads $\mathbf{0} = \mathbf{0}$.

Now suppose that (a)–(c) hold for lists of length k, and consider a list of length $k + 1$. For (a),

$$\begin{aligned} S_{k+1}(a) + S_{k+1}(b) &= (S_k(a) + a_{k+1}\mathbf{v}_{k+1}) + (S_k(b) + b_{k+1}\mathbf{v}_{k+1}) \\ &= (S_k(a) + S_k(b)) + (a_{k+1}\mathbf{v}_{k+1} + b_{k+1}\mathbf{v}_{k+1}) \quad \text{(A1, A2)} \end{aligned}$$

$$= S_k(a+b) + (a_{k+1} + b_{k+1})\mathbf{v}_{k+1} \qquad \text{(induction, A8)}$$
$$= S_{k+1}(a+b).$$

For (b),

$$cS_{k+1}(a) = cS_k(a) + c(a_{k+1}\mathbf{v}_{k+1}) \qquad \text{(A7)}$$
$$= S_k(ca) + (ca_{k+1})\mathbf{v}_{k+1} = S_{k+1}(ca) \qquad \text{(induction, A6).}$$

For (c), $S_{k+1}(0) = S_k(0) + 0\mathbf{v}_{k+1} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, by the inductive hypothesis, Proposition 1.2.32, and (A3). □

Together with $-\mathbf{v} = (-1)\mathbf{v}$ (Proposition 1.2.34), the lemma rewrites any expression built from $\mathbf{v}_1, ..., \mathbf{v}_k$ with addition, subtraction, and scalar multiplication as a single linear combination of $\mathbf{v}_1, ..., \mathbf{v}_k$. For example, reading $\mathbf{v}_1 + 3\mathbf{v}_2$ and $\mathbf{v}_1$ as the linear combinations of $\mathbf{v}_1, \mathbf{v}_2$ with coefficients $1, 3$ and $1, 0$,

$$2(\mathbf{v}_1 + 3\mathbf{v}_2) - \mathbf{v}_1 = (2\mathbf{v}_1 + 6\mathbf{v}_2) + ((-1)\mathbf{v}_1 + 0\mathbf{v}_2) = \mathbf{v}_1 + 6\mathbf{v}_2.$$

1.3.3 Span

Definition 1.3.10 (Span). The **span** of a list $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ in V is the set of vectors represented by linear combinations of $\mathcal{L}$:

$$\operatorname{span}(\mathcal{L}) = \{a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \mid a_1, ..., a_k \in F\} \subseteq V.$$

We also write $\operatorname{span}(\mathbf{v}_1, ..., \mathbf{v}_k)$ for $\operatorname{span}(\mathcal{L})$. In particular, $\operatorname{span}(()) = \{\mathbf{0}\}$, since the only linear combination of the empty list represents $\mathbf{0}$.

So $\mathbf{w} \in \operatorname{span}(\mathcal{L})$ means that there are scalars $a_1, ..., a_k$ with $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. To show that $\mathbf{w}$ is in the span, exhibit the scalars. To show that it is not, show that no choice of scalars works. Usually this means finding a property that every vector of the span has and $\mathbf{w}$ lacks. In Example 1.3.8, every vector $(a, a+b, b)$ of the span has middle coordinate equal to the sum of the outer two, and $(1, 2, 3)$ does not.

Example 1.3.11 (Small spans). *One vector.* $\operatorname{span}(\mathbf{v}) = \{a\mathbf{v} \mid a \in F\}$ is the set of multiples of $\mathbf{v}$. In $\mathbb{R}^2$ with $\mathbf{v} \neq \mathbf{0}$, it is the line through $\mathbf{0}$ and $\mathbf{v}$. In $\mathbb{F}_2^n$ it is $\{\mathbf{0}, \mathbf{v}\}$, since the only scalars are 0 and 1.

Zero. $\operatorname{span}(\mathbf{0}) = \{a\mathbf{0} \mid a \in F\} = \{\mathbf{0}\}$ by Proposition 1.2.33. So the lists $()$, $\mathbf{0}$, and $\mathbf{0}, \mathbf{0}$ are different, but they have the same span.

Standard lists. By Example 1.3.7, $\operatorname{span}(\mathbf{e}_1, ..., \mathbf{e}_n) = F^n$. Similarly $\operatorname{span}\left(1, x, ..., x^d\right) = F[x]_{\leq d}$: a linear combination $a_0 + a_1x + \cdots + a_dx^d$ has no nonzero coefficients beyond position d, so it has degree at most d, and conversely Definition 1.2.16 describes $F[x]_{\leq d}$ as the set of such combinations.

Functions. In $\mathbb{R}^{\mathbb{R}}$, the identity $\sin(t+c) = \cos c \sin t + \sin c \cos t$ shows that every shift $t \mapsto \sin(t+c)$ of the sine function lies in $\operatorname{span}(\sin, \cos)$. The constant function 1 does not: if $1 = a\sin + b\cos$ as functions, then evaluating at $t = 0$ and at $t = \pi$ gives $1 = b$ and $1 = -b$, a contradiction.

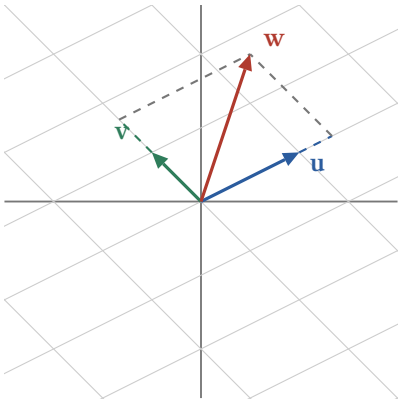


Figure 1.3.1. The faint lines are the combinations $a\mathbf{u} + b\mathbf{v}$ of $\mathbf{u} = (2, 1)$ and $\mathbf{v} = (-1, 1)$ in which a or b is an integer. Every vector of $\mathbb{R}^2$ appears to be a combination of $\mathbf{u}$ and $\mathbf{v}$; for instance, $\mathbf{w} = (1, 3) = \frac{4}{3}\mathbf{u} + \frac{5}{3}\mathbf{v}$.

Figure 1.3.1 suggests that two vectors pointing in different directions span the whole plane; Exercise 1.3.37 makes this precise.

Definition 1.3.12 (Spanning list). A list $\mathcal{L}$ in V **spans** V, and is a **spanning list** of V, if $\operatorname{span}(\mathcal{L}) = V$.

By Example 1.3.11, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ spans F^n, the list $1, x, \ldots, x^d$ spans $F[x]_{\leq d}$, and the empty list spans the zero space $\{\mathbf{0}\}$. A spanning list may contain entries that are not needed: $(1, 0), (0, 1), (1, 1)$ also spans $\mathbb{R}^2$. Spanning lists without such entries are the bases of Section 1.6.

Double containment

Describing a span explicitly means proving that two sets are equal: the span, whose elements are the vectors that some linear combination represents, and a set described in another way, typically by equations. Such proofs follow a fixed pattern, built from two moves.

Technique 1.3.13 (Fix an arbitrary element and chase it). To prove $A \subseteq B$, let x be an arbitrary element of A. Unpack the definition of A until x is described concretely, manipulate that description until it matches the definition of B, and conclude that $x \in B$. Since x was arbitrary, $A \subseteq B$.

Reach for it for every containment between sets defined by properties. Once both definitions are written out, the proof has a fixed shape: the first line is "let $x \in A$", the next uses the definition of A, and the last uses the definition of B. For spans, unpacking produces scalars ("let $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, say $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$"), and repacking means exhibiting scalars. The canonical example is each half of Example 1.3.15.
Used in. Example 1.3.15, Exercise 1.3.16, Proposition 1.3.18, Proposition 1.4.14, Proposition 1.4.16, Proposition 1.4.22, Proposition 1.4.24, Exercise 1.7.5, Proposition 2.5.4, Proposition 2.5.5, Exercise 2.5.33, Proposition 2.6.8, Theorem 2.6.11, Proposition 2.6.15, Proposition 2.7.7, Problem 2.1, Proposition 4.2.3, Proposition 4.5.5, Theorem 4.5.17, Theorem 4.6.5, Proposition 4.7.3, Proposition 4.7.10, Proposition 6.3.9, Proposition 7.2.30, Proposition A.1.5, Theorem A.1.12, Exercise A.1.26, Exercise 2.6.36, Exercise 2.6.39, Exercise 2.6.51, Exercise 2.6.52.

Technique 1.3.14 (Double containment). To prove that two sets A and B are equal, prove $A \subseteq B$ and $B \subseteq A$ separately, each by Technique 1.3.13.

Reach for it whenever a claim says that two sets are equal. The two containments usually need different ideas, and one is often much easier than the other. For a span, "$\operatorname{span}(\mathcal{L}) \subseteq B$" is a computation with an arbitrary linear combination, while "$B \subseteq \operatorname{span}(\mathcal{L})$" requires constructing coefficients for each element of B. A common mistake is to prove one containment carefully and forget the other (Exercise 1.3.30). The canonical example is Example 1.3.15.

Used in. Example 1.3.15, Exercise 1.3.16, Corollary 1.3.19, Exercise 1.3.26, Exercise 1.3.30, Exercise 1.3.32, Exercise 1.3.33, Exercise 1.3.36, Corollary 1.4.17, Exercise 1.4.37, Exercise 1.7.5, Exercise 1.7.15, Proposition 2.1.16, Corollary 2.6.31, Exercise 2.6.51, Proposition 3.1.19, Exercise 3.1.41, Exercise 3.1.45, Proposition 3.2.30, Exercise 3.2.46, Problem 3.1, Proposition 4.4.2, Proposition 4.6.11, Lemma 4.6.14, Proposition 4.7.3, Exercise 6.2.42, Proposition 6.3.9, Exercise 6.3.11, Theorem 6.4.15, Corollary 6.4.25, Theorem 6.4.26, Exercise 6.4.47, Proposition 6.5.9, Proposition 6.6.23, Proposition A.1.5, Lemma A.3.5, Exercise 1.4.35, Exercise 1.4.36, Exercise 1.4.44, Exercise 2.6.39, Exercise 4.6.40, Exercise A.2.10, Exercise A.3.20.

Example 1.3.15 (A span that is a line). In $\mathbb{R}^2$, let $\mathcal{L}_1 = (0,0), (1,-1), (-\pi,\pi)$ as in Example 1.3.5, and let

$$W = \{(c_1, c_2) \in \mathbb{R}^2 \mid c_2 = -c_1\}.$$

Then $\operatorname{span}(\mathcal{L}_1) = W$. In particular, $(2,0) \notin \operatorname{span}(\mathcal{L}_1)$, since $0 \neq -2$.

Proof. ROUTINE

Uses. Definition 1.3.10 (Span), Definition 1.3.4 (Linear combination), Definition 1.2.4 (The space F^n)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Write out both definitions. What does an element of $\operatorname{span}(\mathcal{L}_1)$ look like, and what condition defines W?
Hint 2. For $\operatorname{span}(\mathcal{L}_1) \subseteq W$, compute an arbitrary linear combination coordinatewise. For $W \subseteq \operatorname{span}(\mathcal{L}_1)$, you need coefficients for $(c_1, -c_1)$, and one entry of $\mathcal{L}_1$ suffices.
Sketch. Every linear combination of $\mathcal{L}_1$ represents $(a_2 - \pi a_3, -(a_2 - \pi a_3))$, which lies in W. Conversely, $(c_1, -c_1) = 0(0,0) + c_1(1,-1) + 0(-\pi,\pi)$.

We proceed by double containment.

$\operatorname{span}(\mathcal{L}_1) \subseteq W$. Let $\mathbf{w} \in \operatorname{span}(\mathcal{L}_1)$. Then some linear combination of $\mathcal{L}_1$ represents $\mathbf{w}$, so for some $a_1, a_2, a_3 \in \mathbb{R}$,

$$\mathbf{w} = a_1(0,0) + a_2(1,-1) + a_3(-\pi,\pi) = (a_2 - \pi a_3, -a_2 + \pi a_3).$$

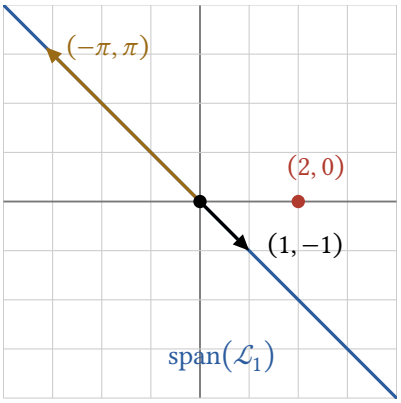
The second coordinate of $\mathbf{w}$ is $-a_2 + \pi a_3 = -(a_2 - \pi a_3)$, the negative of the first, so $\mathbf{w} \in W$.

$W \subseteq \operatorname{span}(\mathcal{L}_1)$. Let $(c_1, c_2) \in W$. By the definition of W, $c_2 = -c_1$, so

$$(c_1, c_2) = (c_1, -c_1) = 0(0,0) + c_1(1,-1) + 0(-\pi,\pi).$$

This linear combination of $\mathcal{L}_1$ represents (c_1, c_2), so $(c_1, c_2) \in \operatorname{span}(\mathcal{L}_1)$, as desired. □

The two containments needed different kinds of argument. The first was a computation that applies to every linear combination at once. The second needed one well-chosen linear combination for each element of W, and it used only the entry $(1,-1)$. The other two entries of $\mathcal{L}_1$ contribute nothing to the span, and Example 1.3.21 explains why without computing anything.



The span of $\mathcal{L}_1$ is the line $c_2 = -c_1$, and $(2, 0)$ is not on it. In the interactive version, $(-\pi, \pi)$ is replaced by $(-3, 3)$, which has the same span. Drag an entry off the line: the span becomes all of $\mathbb{R}^2$, and the explorer shows coefficients that represent $\mathbf{w} = (2, 0)$. (Interactive in the web version.)

Exercise 1.3.16. ROUTINE Show that

$$\operatorname{span}((1, 0, -1), (0, 1, -1)) = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}.$$

Solution on page 1043

1.3.4 Comparing spans

A span is usually an infinite set, so comparing two spans element by element means checking infinitely many vectors. The results of this subsection reduce such a comparison to finitely many checks.

Proposition 1.3.17 (A list lies in its span). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V. Then $\mathbf{0} \in \operatorname{span}(\mathcal{L})$, and $\mathbf{v}_j \in \operatorname{span}(\mathcal{L})$ for each j. That is, $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L})$.

Proof. ROUTINE

Uses. Definition 1.3.10 (Span), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.32 (The scalar zero)
Hint 1. Membership in a span is shown by exhibiting coefficients.
Hint 2. Use the coefficients $0, \ldots, 0$ for $\mathbf{0}$. For $\mathbf{v}_j$, use the coefficient 1 on $\mathbf{v}_j$ and 0 on every other entry.
Sketch. $\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k$, and $\mathbf{v}_j$ is represented by the combination with coefficient 1 in position j and 0 elsewhere, since $0\mathbf{v}_i = \mathbf{0}$ and $1\mathbf{v}_j = \mathbf{v}_j$.

By Lemma 1.3.9(c), the linear combination with all coefficients 0 represents $\mathbf{0}$, so $\mathbf{0} \in \operatorname{span}(\mathcal{L})$. Now fix j, and consider the linear combination with $a_j = 1$ and $a_i = 0$ for $i \neq j$. Its terms are $1\mathbf{v}_j = \mathbf{v}_j$, by (A5), and $0\mathbf{v}_i = \mathbf{0}$ for $i \neq j$, by Proposition 1.2.32. Adding copies of $\mathbf{0}$ changes nothing, by (A3), so this linear combination represents $\mathbf{v}_j$, and $\mathbf{v}_j \in \operatorname{span}(\mathcal{L})$. □

Intuition (Recipes of recipes). Think of a linear combination of $\mathcal{L}$ as a recipe that says how much of each $\mathbf{v}_i$ to use. Suppose that every entry of $\mathcal{M} = \mathbf{w}_1, \ldots, \mathbf{w}_m$ has a recipe in terms of $\mathcal{L}$. A linear combination of $\mathcal{M}$ is then a recipe whose ingredients are themselves recipes, and expanding it gives a recipe in terms of $\mathcal{L}$. For instance, if $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{w}_2 = \mathbf{v}_1 - \mathbf{v}_2$, then

$$3\mathbf{w}_1 + 2\mathbf{w}_2 = (3\mathbf{v}_1 + 3\mathbf{v}_2) + (2\mathbf{v}_1 - 2\mathbf{v}_2) = 5\mathbf{v}_1 + \mathbf{v}_2.$$

So as soon as the finitely many entries of $\mathcal{M}$ lie in $\operatorname{span}(\mathcal{L})$, every vector of $\operatorname{span}(\mathcal{M})$ does.

Proposition 1.3.18 (Comparing spans). Let $\mathcal{L}$ and $\mathcal{M}$ be lists in V. Then $\operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$ if and only if $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$.

Proof. STANDARD

Uses. Proposition 1.3.17 (A list lies in its span), Lemma 1.3.9 (Arithmetic of linear combinations), Definition 1.3.10 (Span), induction on m
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. One direction follows at once from Proposition 1.3.17. For the other, let $\mathbf{u} \in \operatorname{span}(\mathcal{M})$ and write $\mathbf{u} = \sum_j c_j \mathbf{w}_j$.
Hint 2. Replace each $\mathbf{w}_j$ by a linear combination of $\mathcal{L}$ that represents it, say $\mathbf{w}_j = \sum_i b_{ij}\mathbf{v}_i$. Then use Lemma 1.3.9 to collect the coefficient of each $\mathbf{v}_i$.
Sketch. ($\Rightarrow$) $\mathcal{M} \subseteq \operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$. ($\Leftarrow$) If $\mathbf{w}_j = \sum_i b_{ij}\mathbf{v}_i$, then by Lemma 1.3.9, $\sum_j c_j \mathbf{w}_j = \sum_j \sum_i (c_j b_{ij})\mathbf{v}_i = \sum_i \bigl(\sum_j c_j b_{ij}\bigr)\mathbf{v}_i$, which is a linear combination of $\mathcal{L}$.

Write $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathcal{M} = \mathbf{w}_1, \ldots, \mathbf{w}_m$.

($\Rightarrow$) By Proposition 1.3.17, $\mathcal{M} \subseteq \operatorname{span}(\mathcal{M})$, and $\operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$ by assumption.

($\Leftarrow$) Suppose that $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$. For each j, choose scalars $b_{1j}, \ldots, b_{kj}$ with

$$\mathbf{w}_j = \sum_{i=1}^{k} b_{ij}\mathbf{v}_i.$$

Let $\mathbf{u} \in \operatorname{span}(\mathcal{M})$, say $\mathbf{u} = \sum_{j=1}^{m} c_j \mathbf{w}_j$ with $c_1, \ldots, c_m \in F$. By Lemma 1.3.9(b), $c_j \mathbf{w}_j = \sum_{i=1}^{k} (c_j b_{ij})\mathbf{v}_i$ for each j. Adding these m vectors one at a time with Lemma 1.3.9(a), that is, by induction on m, gives

$$\mathbf{u} = \sum_{j=1}^{m} \sum_{i=1}^{k} (c_j b_{ij})\mathbf{v}_i = \sum_{i=1}^{k} \left(\sum_{j=1}^{m} c_j b_{ij} \right) \mathbf{v}_i.$$

(For $m = 0$ both sides are $\mathbf{0}$, by Lemma 1.3.9(c).) So $\mathbf{u}$ is represented by the linear combination of $\mathcal{L}$ with coefficients $d_i = \sum_j c_j b_{ij}$, and $\mathbf{u} \in \operatorname{span}(\mathcal{L})$. Since $\mathbf{u}$ was arbitrary, $\operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$, as desired. □

The forward direction is Proposition 1.3.17. The backward direction is the useful one: to show that one span contains another, it suffices to check the finitely many entries of one list. Section 1.4 gives a second proof. There we show that $\operatorname{span}(\mathcal{L})$ is the smallest subspace of V containing the entries of $\mathcal{L}$ (Proposition 1.4.14), and the backward direction follows at once.

Corollary 1.3.19 (Equal spans). Let $\mathcal{L}$ and $\mathcal{M}$ be lists in V. Then $\operatorname{span}(\mathcal{L}) = \operatorname{span}(\mathcal{M})$ if and only if $\mathcal{L} \subseteq \operatorname{span}(\mathcal{M})$ and $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$.

Proof. ROUTINE

Uses. Proposition 1.3.18 (Comparing spans)
Techniques. Double containment (Technique 1.3.14)
Hint 1. Split the equality of sets into two containments.
Hint 2. Apply Proposition 1.3.18 to each containment.
Sketch. $\operatorname{span}(\mathcal{L}) \subseteq \operatorname{span}(\mathcal{M})$ if and only if $\mathcal{L} \subseteq \operatorname{span}(\mathcal{M})$, and $\operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$ if and only if $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$.

The sets $\operatorname{span}(\mathcal{L})$ and $\operatorname{span}(\mathcal{M})$ are equal exactly when each contains the other. By Proposition 1.3.18, $\operatorname{span}(\mathcal{L}) \subseteq \operatorname{span}(\mathcal{M})$ holds exactly when $\mathcal{L} \subseteq \operatorname{span}(\mathcal{M})$, and $\operatorname{span}(\mathcal{M}) \subseteq \operatorname{span}(\mathcal{L})$ holds exactly when $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$. □

Corollary 1.3.20 (Operations that preserve the span). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V.

(a) Every reordering of $\mathcal{L}$ has the same span as $\mathcal{L}$.

(b) If $\mathcal{L}'$ is a sublist of $\mathcal{L}$, then $\operatorname{span}(\mathcal{L}') \subseteq \operatorname{span}(\mathcal{L})$.

(c) For $\mathbf{w} \in V$, $\operatorname{span}(\mathcal{L}, \mathbf{w}) = \operatorname{span}(\mathcal{L})$ if and only if $\mathbf{w} \in \operatorname{span}(\mathcal{L})$. In particular, appending $\mathbf{0}$ or a second copy of an entry of $\mathcal{L}$ does not change the span.

(d) Let $i \neq j$, and let $a, c \in F$ with $c \neq 0$. Replacing $\mathbf{v}_j$ by $c\mathbf{v}_j$, or by $\mathbf{v}_j + a\mathbf{v}_i$, does not change the span.

Proof. ROUTINE

Uses. Corollary 1.3.19 (Equal spans), Proposition 1.3.18 (Comparing spans), Proposition 1.3.17 (A list lies in its span), Notation 1.3.3
Hint 1. Each part compares the spans of two lists. By Corollary 1.3.19, it suffices to check that the entries of each list lie in the span of the other, and Proposition 1.3.17 handles the entries the two lists share.
Hint 2. For (d), write $\mathbf{v}_j$ in terms of the new entries: $\mathbf{v}_j = c^{-1}(c\mathbf{v}_j)$, and $\mathbf{v}_j = (\mathbf{v}_j + a\mathbf{v}_i) + (-a)\mathbf{v}_i$.
Sketch. (a) The two lists have the same entries. (b) $\mathcal{L}' \subseteq \mathcal{L} \subseteq \operatorname{span}(\mathcal{L})$. (c) If $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, then every entry of $\mathcal{L}, \mathbf{w}$ lies in $\operatorname{span}(\mathcal{L})$, and $\mathcal{L}$ is a sublist of $\mathcal{L}, \mathbf{w}$; conversely $\mathbf{w} \in \operatorname{span}(\mathcal{L}, \mathbf{w})$. (d) The new entry is a linear combination of $\mathcal{L}$, and $\mathbf{v}_j$ is a linear combination of the new list because $c \neq 0$ and $i \neq j$.

(a) A reordering $\mathcal{L}'$ of $\mathcal{L}$ has the same entries as $\mathcal{L}$. By Proposition 1.3.17, $\mathcal{L}' \subseteq \operatorname{span}(\mathcal{L})$ and $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L}')$, so the spans are equal by Corollary 1.3.19.

(b) Every entry of $\mathcal{L}'$ is an entry of $\mathcal{L}$, so $\mathcal{L}' \subseteq \operatorname{span}(\mathcal{L})$ by Proposition 1.3.17, and $\operatorname{span}(\mathcal{L}') \subseteq \operatorname{span}(\mathcal{L})$ by Proposition 1.3.18.

(c) Suppose that $\mathbf{w} \in \operatorname{span}(\mathcal{L})$. The entries of $\mathcal{L}, \mathbf{w}$ are the $\mathbf{v}_i$, which lie in $\operatorname{span}(\mathcal{L})$ by Proposition 1.3.17, and $\mathbf{w}$, which lies there by assumption. So $\operatorname{span}(\mathcal{L}, \mathbf{w}) \subseteq \operatorname{span}(\mathcal{L})$ by Proposition 1.3.18. Conversely, $\mathcal{L}$ is a sublist of $\mathcal{L}, \mathbf{w}$, so $\operatorname{span}(\mathcal{L}) \subseteq \operatorname{span}(\mathcal{L}, \mathbf{w})$ by (b). Now suppose instead that $\operatorname{span}(\mathcal{L}, \mathbf{w}) = \operatorname{span}(\mathcal{L})$. Then $\mathbf{w} \in \operatorname{span}(\mathcal{L}, \mathbf{w}) = \operatorname{span}(\mathcal{L})$ by Proposition 1.3.17. The last statement of (c) follows, since $\mathbf{0}$ and the entries of $\mathcal{L}$ lie in $\operatorname{span}(\mathcal{L})$ by Proposition 1.3.17.

(d) Let $\mathcal{L}'$ be $\mathcal{L}$ with $\mathbf{v}_j$ replaced by $\mathbf{v}'_j$, where $\mathbf{v}'_j = c\mathbf{v}_j$ or $\mathbf{v}'_j = \mathbf{v}_j + a\mathbf{v}_i$. The entries of $\mathcal{L}'$ other than $\mathbf{v}'_j$ are entries of $\mathcal{L}$, and $\mathbf{v}'_j$ is a linear combination of $\mathcal{L}$: the one with coefficient c on $\mathbf{v}_j$ and 0 elsewhere, or the one with coefficient 1 on $\mathbf{v}_j$, a on $\mathbf{v}_i$, and 0 elsewhere (this uses $i \neq j$). So $\mathcal{L}' \subseteq \operatorname{span}(\mathcal{L})$. In the other direction, the entries of $\mathcal{L}$ other than $\mathbf{v}_j$ are entries of $\mathcal{L}'$, and

$$\mathbf{v}_j = c^{-1}(c\mathbf{v}_j) \quad \text{or} \quad \mathbf{v}_j = (\mathbf{v}_j + a\mathbf{v}_i) + (-a)\mathbf{v}_i,$$

by (A6) and (A5), or by (A2), (A8), Proposition 1.2.32, and (A3). Since $c \neq 0$, and since $\mathbf{v}_i$ is still an entry of $\mathcal{L}'$ (as $i \neq j$), these are linear combinations of $\mathcal{L}'$. So $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L}')$, and Corollary 1.3.19 finishes the proof. □

Combining (a) and (c): the entry $\mathbf{v}_j$ can be deleted from $\mathcal{L}$ without changing the span exactly when $\mathbf{v}_j$ lies in the span of the remaining entries. (Reorder so that $\mathbf{v}_j$ comes last, and apply (c).) Parts (a) and (d) are the three moves of row reduction, applied to the entries of a list: swap two entries, scale an entry by a nonzero scalar, and add a multiple of one entry to another. Section 3.3 performs them on the rows of a matrix. Both hypotheses in (d) are needed: scaling by $c = 0$, or replacing $\mathbf{v}_j$ by $\mathbf{v}_j + (-1)\mathbf{v}_j$, turns $\mathbf{v}_j$ into $\mathbf{0}$, and in $\mathbb{R}^2$ this shrinks $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) = \mathbb{R}^2$ to the line $\operatorname{span}(\mathbf{e}_1)$.

Example 1.3.21 (A span that is a line, again). In $\mathcal{L}_1 = (0,0), (1,-1), (-\pi, \pi)$, the third entry is $-\pi(1,-1)$ and the first is $\mathbf{0}$. By Corollary 1.3.20(c), together with (a) to move entries to the end, we may delete $(-\pi, \pi)$, since it lies in $\operatorname{span}((0,0),(1,-1))$, and then delete $(0,0)$. Hence

$$\operatorname{span}(\mathcal{L}_1) = \operatorname{span}((1,-1)) = \{c(1,-1) \mid c \in \mathbb{R}\} = \{(c,-c) \mid c \in \mathbb{R}\},$$

which is the set W of Example 1.3.15. Two observations about entries have replaced the double containment.

Example 1.3.22 (Equal spans, and a small field). Let $\mathcal{L} = (1,1,0), (0,1,1)$ and $\mathcal{M} = (1,2,1), (1,0,-1)$ in F^3. We transform $\mathcal{L}$ into $\mathcal{M}$ with the moves of Corollary 1.3.20(d):

$$(1,1,0),(0,1,1) \quad \rightsquigarrow \quad (1,2,1),(0,1,1) \quad \rightsquigarrow \quad (1,2,1),(0,-2,-2) \quad \rightsquigarrow \quad (1,2,1),(1,0,-1).$$

The first move adds the second entry to the first, the second scales the second entry by -2, and the third adds the first entry to the second. Each move preserves the span, provided that $-2 \neq 0$ in F, which holds exactly when $1 + 1 \neq 0$. So over $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{F}_3$, $\operatorname{span}(\mathcal{L}) = \operatorname{span}(\mathcal{M})$.

Over $\mathbb{F}_2$ the conclusion is false. There $(1,2,1) = (1,0,-1) = (1,0,1)$, so $\mathcal{M}$ is the list $(1,0,1),(1,0,1)$, and $\operatorname{span}(\mathcal{M}) = \operatorname{span}((1,0,1)) = \{\mathbf{0},(1,0,1)\}$ by Corollary 1.3.20(c) and Example 1.3.11. This set does not contain the vector $(1,1,0)$ of $\operatorname{span}(\mathcal{L})$. The move that fails is the scaling by $-2 = 0$. Testing the claim over $\mathbb{F}_2$ (Technique 1.2.38) is how one notices that the hypothesis $c \neq 0$ in (d) is doing real work.

Remark 1.3.23 (What the span forgets). By Corollary 1.3.20, the span of a list depends neither on the order of the entries, nor on repetitions, nor on entries that lie in the span of the others. In particular, it depends only on the set of entries (Exercise 1.3.24), and lists of different lengths can have the same span. The span also does not record how many linear combinations represent each of its vectors. Section 1.5 shows that some entry of a list lies in the span of the other entries exactly when some vector of its span has more than one representation.

Exercise 1.3.24. ROUTINE Let $\mathcal{L}$ and $\mathcal{M}$ be lists in V with the same set of entries; that is, $\mathbf{v} \in \mathcal{L}$ if and only if $\mathbf{v} \in \mathcal{M}$. Show that $\operatorname{span}(\mathcal{L}) = \operatorname{span}(\mathcal{M})$.

Solution on page 1043

Exercise 1.3.25. ROUTINE Let $\mathbf{v}_1,\ldots,\mathbf{v}_k \in V$, and let $\mathbf{w}_j = \mathbf{v}_1 + \cdots + \mathbf{v}_j$ for $1 \leq j \leq k$. Show that $\operatorname{span}(\mathbf{w}_1,\ldots,\mathbf{w}_k) = \operatorname{span}(\mathbf{v}_1,\ldots,\mathbf{v}_k)$.

Solution on page 1043

ASIDE: *Spans of infinite sets* BEYOND 21-242

Every linear combination has finitely many terms. A vector space only adds two vectors at a time, and an infinite sum would need a notion of convergence, which a general vector space does not have. For an arbitrary subset $S \subseteq V$, possibly infinite, one defines $\operatorname{span}(S)$ as the set of vectors represented by linear combinations of finite lists in S; the empty list is allowed, so $\operatorname{span}(\varnothing) = \{\mathbf{0}\}$. For a finite set S, this is the span of any list whose set of entries is S (Exercise 1.3.24). For example, $\operatorname{span}(\{1, x, x^2, \ldots\}) = F[x]$, since each polynomial is a linear combination of $1, x, \ldots, x^d$ for some d, while no finite list spans $F[x]$ (Section 1.6). Spanning sets of this kind are the subject of Section 15.1.

Exercises

Exercise 1.3.26. ROUTINE Let $\mathcal{L} = (1,2,0),(0,1,1)$ in $\mathbb{R}^3$.

(a) Show that $\operatorname{span}(\mathcal{L}) = \{(x,y,z) \in \mathbb{R}^3 \mid y = 2x + z\}$.

(b) Decide whether $(1,1,1)$ and $(2,5,1)$ lie in $\operatorname{span}(\mathcal{L})$. For each one that does, find a linear combination of $\mathcal{L}$ that represents it.

Solution on page 1043

Exercise 1.3.27. ROUTINE Let $\mathcal{L} = (1,1,0),(0,1,2)$ in $\mathbb{F}_3^3$.

(a) List the elements of $\operatorname{span}(\mathcal{L})$, and decide whether $(1,0,1)$ and $(1,1,1)$ belong to it.

(b) Let $\mathcal{M} = (1,1,0),(0,1,2),(1,2,2)$. Show that $\operatorname{span}(\mathcal{M}) = \operatorname{span}(\mathcal{L})$. How many linear combinations of $\mathcal{M}$ are there, and how many of them represent $(1,0,1)$?

Solution on page 1044

Exercise 1.3.28. ROUTINE

(a) Find all lists $\mathcal{L}$ in V with $\operatorname{span}(\mathcal{L}) = \{\mathbf{0}\}$. Which lists span the zero space?

(b) Let $\mathbf{v},\mathbf{w} \in V$. Show that $\operatorname{span}(\mathbf{v}) = \operatorname{span}(\mathbf{w})$ if and only if $\mathbf{w} = c\mathbf{v}$ for some nonzero $c \in F$.

Solution on page 1044

Exercise 1.3.29. ROUTINE

(a) Find the error in the following argument.

Claim. Let $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $a_1, a_2, b_1, b_2 \in F$. If $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = b_1\mathbf{v}_1 + b_2\mathbf{v}_2$, then $a_1 = b_1$ and $a_2 = b_2$.

Proof. Two linear combinations of $\mathbf{v}_1, \mathbf{v}_2$ are the same exactly when they have the same coefficients (Definition 1.3.4). The two sides of the hypothesis are equal, so their coefficients are equal.

(b) Show that the claim is false in general, and that it is true for $\mathbf{v}_1 = (1, 0)$ and $\mathbf{v}_2 = (0, 1)$ in F^2.

Solution on page 1045

Exercise 1.3.30. STANDARD

(a) Find the gap in the following argument.

Claim. $\operatorname{span}((1, 1), (1, -1)) = \mathbb{R}^2$.

Proof. Let $\mathbf{w} \in \operatorname{span}((1, 1), (1, -1))$, say $\mathbf{w} = a(1, 1) + b(1, -1) = (a + b, a - b)$. Both coordinates are real numbers, so $\mathbf{w} \in \mathbb{R}^2$. Hence $\operatorname{span}((1, 1), (1, -1)) = \mathbb{R}^2$.

(b) Prove the claim.

(c) Is $\operatorname{span}((1, 1), (1, -1)) = \mathbb{F}_2^2$ in $\mathbb{F}_2^2$? Is $\operatorname{span}((1, 1), (1, -1)) = \mathbb{F}_3^2$ in $\mathbb{F}_3^2$?

Solution on page 1045

Exercise 1.3.31. STANDARD Work in the real vector space $\mathbb{R}^{\mathbb{R}}$ of functions $\mathbb{R} \to \mathbb{R}$.

(a) Show that the function $t \mapsto \cos^2 t$ lies in $\operatorname{span}(1, c_2)$, where $c_2(t) = \cos 2t$.

(b) Show that $\sin \notin \operatorname{span}(1, \cos)$.

(c) Does the function $t \mapsto \sin^2 t$ lie in $\operatorname{span}(1, \cos)$?

Solution on page 1046

Exercise 1.3.32. STANDARD Let $W = \{p \in \mathbb{R}[x]_{\leq 3} \mid p(1) = 0\}$. Find a list of length 3 that spans W, and prove that it does.

Solution on page 1046

Exercise 1.3.33. STANDARD

(a) Let $p_0, p_1, \ldots, p_n \in F[x]$ with $\deg p_k = k$ for each k. Show that $\operatorname{span}(p_0, \ldots, p_n) = F[x]_{\leq n}$.

(b) Write x^2 as a linear combination of $1, x - 1, x^2 - 2x + 1$ in $\mathbb{R}[x]$.

Solution on page 1046

Exercise 1.3.34. STANDARD Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V, and let $\mathbf{w} \in \operatorname{span}(\mathcal{L})$ with $\mathbf{w} \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$. Show that

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{w}) = \operatorname{span}(\mathcal{L}).$$

In other words, $\mathbf{w}$ can be exchanged for $\mathbf{v}_k$ without changing the span.

Solution on page 1047

Exercise 1.3.35. STANDARD Let $\mathcal{L}$ and $\mathcal{M}$ be lists in V, and let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$. Prove or disprove each statement.

(a) $\operatorname{span}(\mathcal{L}) \cup \operatorname{span}(\mathcal{M}) = \operatorname{span}(\mathcal{L}, \mathcal{M})$.

(b) $\operatorname{span}(\mathcal{L}) \cap \operatorname{span}(\mathcal{M}) = \operatorname{span}(\mathcal{N})$, where $\mathcal{N}$ is a list of the vectors that are entries of both $\mathcal{L}$ and $\mathcal{M}$.

(c) If $\operatorname{span}(\mathcal{L}) = \operatorname{span}(\mathcal{M})$, then $|\mathcal{L}| = |\mathcal{M}|$.

(d) If $\mathbf{u} \in \operatorname{span}(\mathbf{v}, \mathbf{w})$ and $\mathbf{v} \in \operatorname{span}(\mathbf{u}, \mathbf{w})$, then $\mathbf{w} \in \operatorname{span}(\mathbf{u}, \mathbf{v})$.

(e) If $\mathbf{u} \in \operatorname{span}(\mathbf{v}, \mathbf{w})$ and $\mathbf{u} \notin \operatorname{span}(\mathbf{w})$, then $\mathbf{v} \in \operatorname{span}(\mathbf{u}, \mathbf{w})$.

Solution on page 1047

Exercise 1.3.36. STANDARD Let $\mathcal{L}$ and $\mathcal{M}$ be lists in V. Show that

$$\operatorname{span}(\mathcal{L}, \mathcal{M}) = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in \operatorname{span}(\mathcal{L}), \mathbf{w} \in \operatorname{span}(\mathcal{M})\}.$$

Solution on page 1048

Exercise 1.3.37. STANDARD Let $\mathbf{u} = (a, b)$ and $\mathbf{v} = (c, d)$ be vectors in F^2.

(a) Show that $\mathbf{u}, \mathbf{v}$ spans F^2 if and only if $ad - bc \neq 0$.

(b) Suppose $\mathbf{u} \neq \mathbf{0}$. Show that $\mathbf{u}, \mathbf{v}$ spans F^2 if and only if $\mathbf{v} \notin \operatorname{span}(\mathbf{u})$.

Solution on page 1048

Exercise 1.3.38. STANDARD Let F be a finite field with q elements.

(a) Show that no list of length 1 spans F^2.

(b) Show that exactly $\big(q^2 - 1\big)\big(q^2 - q\big)$ lists of length 2 span F^2.

(c) For $q = 2$, how many lists of length 3 span $\mathbb{F}_2^2$?

Solution on page 1049

Exercise 1.3.39. STANDARD Let F be a finite field with q elements, and let $\mathcal{L}$ be a list of length k in V.

(a) Show that $\operatorname{span}(\mathcal{L})$ has at most q^k elements.

(b) Deduce that no list of length less than n spans F^n.

Solution on page 1049

Exercise 1.3.40. HARD Let F be a finite field with q elements, and let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V. For $\mathbf{w} \in V$, let $N(\mathbf{w})$ be the number of linear combinations of $\mathcal{L}$ that represent $\mathbf{w}$.

(a) Show that $N(\mathbf{w}) = N(\mathbf{0})$ for every $\mathbf{w} \in \operatorname{span}(\mathcal{L})$.

(b) Deduce that $q^k = |\operatorname{span}(\mathcal{L})| \cdot N(\mathbf{0})$. In particular, $|\operatorname{span}(\mathcal{L})|$ divides q^k.

(c) If $F = \mathbb{F}_p$ with p prime, show that $\operatorname{span}(\mathcal{L})$ has p^r elements for some integer $0 \leq r \leq k$.

Solution on page 1049

Exercise 1.3.41. STANDARD Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and let $D = \operatorname{span}(\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w})$.

(a) Show that $D = \operatorname{span}(\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u})$ and $D \subseteq \operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

(b) Prove or disprove: $D = \operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

(c) Show that $D = \operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ if and only if $\mathbf{u} \in D$.

Solution on page 1050

Exercise 1.3.42. HARD Let $n \geq 2$, let $\mathbf{s} = (1, 1, \ldots, 1) \in F^n$, and let $\mathbf{u}_i = \mathbf{s} - \mathbf{e}_i$ be the vector with 0 in position i and 1 elsewhere. For a positive integer m, write $m * 1 = 1 + \cdots + 1$ (m terms) in F, as in Exercise 1.2.45. Show that $\mathbf{u}_1, \ldots, \mathbf{u}_n$ spans F^n if and only if $(n-1) * 1 \neq 0$ in F.

Solution on page 1050

Exercise 1.3.43. HARD Let $n \geq 3$. Let $\mathcal{W}_2$ be a list of all vectors in F^n with exactly two coordinates equal to 1 and all others 0, and define $\mathcal{W}_3$ similarly with three coordinates equal to 1.

(a) Show that if $1 + 1 \neq 0$ in F, then $\mathcal{W}_2$ spans F^n.

(b) Over $\mathbb{F}_2$, show that $\operatorname{span}(\mathcal{W}_2) = \{\mathbf{x} \in \mathbb{F}_2^n \mid x_1 + \cdots + x_n = 0\}$.

(c) Over $\mathbb{F}_2$, show that $\mathcal{W}_3$ spans $\mathbb{F}_2^n$ when $n \geq 4$. What is $\operatorname{span}(\mathcal{W}_3)$ when $n = 3$?

Solution on page 1051

Exercise 1.3.44. STANDARD Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$ (Example 1.2.20). Show that $\sqrt{3} \notin \operatorname{span}\left(1, \sqrt{2}\right)$. You may use that $\sqrt{2}$, $\sqrt{3}$, and $\sqrt{6}$ are irrational.

Solution on page 1052

Exercise 1.3.45. HARD BEYOND 21-242 Show that no finite list spans $\mathbb{R}$ as a vector space over $\mathbb{Q}$. (Use countability.)

Solution on page 1052

Exercise 1.3.46. CHALLENGE A row of $n \geq 1$ lamps, each on or off, is recorded as a vector $\mathbf{x} \in \mathbb{F}_2^n$, with $x_i = 1$ when lamp i is on. Pressing lamp i switches lamp i and its neighbors $i-1$ and $i+1$ (those that exist). That is, it adds $\mathbf{t}_i = \mathbf{e}_{i-1} + \mathbf{e}_i + \mathbf{e}_{i+1}$ to $\mathbf{x}$, where $\mathbf{e}_0 = \mathbf{e}_{n+1} = \mathbf{0}$.

(a) Show that every configuration can be switched entirely off by pressing lamps if and only if $\mathbf{t}_1, \ldots, \mathbf{t}_n$ spans $\mathbb{F}_2^n$.

(b) Show that this happens if and only if n is not of the form $3m + 2$.

Solution on page 1052

1.4 Subspaces

Several of the vector spaces in Section 1.2 live inside a larger one and use its operations: $C[0,1]$ inside $\mathbb{R}^{[0,1]}$, and $F[x]_{\leq n}$ inside $F[x]$. So does every span. The span of a list in V is a subset of V, and by Lemma 1.3.9, sums and scalar multiples of linear combinations of the list are again linear combinations of it. This section names such subsets and gives a quick test for recognizing them. The lecture defines subspaces later (L11–12), alongside dimension. We meet them now because the span is their prototype, and because the kernel and the image of a linear map (Section 2.6) are subspaces the moment they are defined.

1.4.1 The subspace criterion

Throughout, V is a vector space over a field F.

Definition 1.4.1 (Subspace). A subset $W \subseteq V$ is a **subspace** of V if W is a vector space over F under the operations of V, restricted to W. That is, for $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$, the sum $\mathbf{w} + \mathbf{w}'$ and the scalar multiple $a\mathbf{w}$ are computed in V. A subspace $W \neq V$ is a **proper** subspace.

So a subspace is a vector space in its own right, and its operations are not new: we add and scale elements of W exactly as we do in V. In particular, sums and scalar multiples of elements of W must lie in W, because closure is part of being a vector space (Remark 1.2.2).

Non-example 1.4.2 (A line that misses the origin). In $\mathbb{R}^2$, the line $L = \{(c_1, c_2) \mid c_1 + c_2 = 1\}$ is not a subspace. The operations of $\mathbb{R}^2$ are not even operations on L: the vectors $(1, 0)$ and $(0, 1)$ lie on L, but their sum $(1, 1)$ does not, since $1 + 1 \neq 1$. Non-example 1.2.21 made the same observation for the line $x + 2y = 2$. By contrast, the line $\{(r, r) \mid r \in \mathbb{R}\}$ through the origin is a subspace (Example 1.4.8).

Checking all eight axioms for every candidate would be tedious, and most of the work is redundant.

> *Intuition (What has to be checked).* Suppose $W \subseteq V$ is closed under the operations of V. Six of the axioms, (A1), (A2), and (A5)–(A8), are identities that hold *for all* vectors and scalars. An identity that holds for all vectors of V holds in particular for all vectors of W, so these six come for free. The other two axioms assert that something *exists*: (A3) a zero vector, and (A4) an additive inverse of each vector. These are where a subset can fail, because the vector that exists in V might not lie in W. Scalar multiplication produces both of them: $0\mathbf{w} = \mathbf{0}$ and $(-1)\mathbf{w} = -\mathbf{w}$ (Proposition 1.2.32, Proposition 1.2.34). For the first we need some $\mathbf{w} \in W$ to start from, so W must be nonempty. We used the same reasoning for subfields (Proposition 1.1.33) and for $C[0, 1]$ (Example 1.2.12).

The lecture packages this as a criterion with two equivalent tests. The first speaks about linear combinations, which suits spans; the second is the one we check in practice.

> **Theorem 1.4.3** (Subspace criterion). Let $W \subseteq V$. The following are equivalent.
>
> (I) W is a subspace of V.
>
> (II) $W \neq \emptyset$, and W is **stable under linear combinations**: for every $k \geq 1$, all $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$, and all $a_1, \ldots, a_k \in F$, the vector $a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$ lies in W.
>
> (III) $\mathbf{0} \in W$, and W is **closed under addition and scalar multiplication**: $\mathbf{w} + \mathbf{w}' \in W$ and $a\mathbf{w} \in W$ for all $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$.
>
> Here $\mathbf{0}$ is the zero vector of V.

Proof. ROUTINE

> *Uses.* Definition 1.4.1 (Subspace), Definition 1.2.1 (Vector space), Proposition 1.2.32 (The scalar zero), Proposition 1.2.34 (Negatives and the scalar -1), (1.3.1), induction on k
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* Sort the axioms (A1)–(A8) into identities that hold "for all" vectors and scalars, and statements that something exists. Which kind can fail for a subset that is closed under the operations?
> *Hint 2.* For (I) ⇒ (III), the subspace W has a zero vector of its own, but (III) asks for the zero vector of V. Multiply the zero vector of W by the scalar 0.
> *Hint 3.* For (III) ⇒ (II), induct on k and split off the last term. For (II) ⇒ (I), produce the zero vector and negatives inside W as $0\mathbf{w}_0$ and $(-1)\mathbf{w}$.
> *Sketch.* (I) ⇒ (III): the operations of W land in W, and $\mathbf{0} = 0 \cdot \mathbf{0}_W \in W$. (III) ⇒ (II): induction on k, writing a combination of length $k + 1$ as one of length k plus a single term. (II) ⇒ (I): sums and multiples are combinations of length 2 and 1, so the operations restrict to W; the identities (A1), (A2), (A5)–(A8) are inherited from V; and $\mathbf{0} = 0\mathbf{w}_0$ and $-\mathbf{w} = (-1)\mathbf{w}$ lie in W and supply (A3) and (A4).

We prove (I) ⇒ (III) ⇒ (II) ⇒ (I).

(I) ⇒ (III). Suppose W is a subspace. Since W is a vector space under the operations of V, these operations map $W \times W$ and $F \times W$ into W; that is, W is closed under addition and scalar multiplication. By (A3) for the vector space W, there is a zero vector $\mathbf{0}_W \in W$. By closure under scalar multiplication, $0 \cdot \mathbf{0}_W \in W$, and $0 \cdot \mathbf{0}_W = \mathbf{0}$ by Proposition 1.2.32, applied in V. Hence $\mathbf{0} \in W$.

(III) ⇒ (II). Since $\mathbf{0} \in W$, $W \neq \emptyset$. We show by induction on $k \geq 1$ that $a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k \in W$ for all $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ and $a_1, \ldots, a_k \in F$. For $k = 1$ this is closure under scalar multiplication. Suppose it holds for some $k \geq 1$, and let $\mathbf{w}_1, \ldots, \mathbf{w}_{k+1} \in W$ and $a_1, \ldots, a_{k+1} \in F$. By (1.3.1),

$$\sum_{i=1}^{k+1} a_i\mathbf{w}_i = \left(\sum_{i=1}^{k} a_i\mathbf{w}_i\right) + a_{k+1}\mathbf{w}_{k+1}.$$

The first summand lies in W by the inductive hypothesis, and the second by closure under scalar multiplication. So their sum lies in W, by closure under addition.

(II) ⇒ (I). Let $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$. Then $\mathbf{w} + \mathbf{w}' = 1\mathbf{w} + 1\mathbf{w}'$ by (A5), and $a\mathbf{w}$ are linear combinations of elements of W, of lengths 2 and 1, so they lie in W by (II). Hence the operations of V restrict to functions $W \times W \to W$ and $F \times W \to W$. Axioms (A1), (A2), and (A5)–(A8) are identities that hold for all vectors of V and all scalars, so in particular for all vectors of W. It remains to check (A3) and (A4).

(A3). Since $W \neq \emptyset$, we may choose some $\mathbf{w}_0 \in W$. By (II), $0\mathbf{w}_0 \in W$, and $0\mathbf{w}_0 = \mathbf{0}$ by Proposition 1.2.32. So $\mathbf{0} \in W$, and $\mathbf{w} + \mathbf{0} = \mathbf{w}$ for every $\mathbf{w} \in W$, because this holds in V.

(A4). Let $\mathbf{w} \in W$. By (II), $(-1)\mathbf{w} \in W$, and $\mathbf{w} + (-1)\mathbf{w} = \mathbf{w} + (-\mathbf{w}) = \mathbf{0}$ by Proposition 1.2.34, where $\mathbf{0}$ is the zero vector of W just found.

So W satisfies all eight axioms, and W is a subspace of V. □

Remark 1.4.4 (The zero vector of a subspace). Let W be a subspace of V. The proof of (I) ⇒ (III) showed that $\mathbf{0} \in W$. Since $\mathbf{w} + \mathbf{0} = \mathbf{w}$ for every $\mathbf{w} \in W$, the vector $\mathbf{0}$ is a zero vector of the vector space W, so it is *the* zero vector of W (Proposition 1.2.27, applied in W). That is, $\mathbf{0}_W = \mathbf{0}_V$. Similarly, the additive inverse of $\mathbf{w}$ in W is the vector $-\mathbf{w}$ of V, by Proposition 1.2.28. So the vectors whose existence (A3) and (A4) promise are also inherited from V, and from now on we write $\mathbf{0}$ for the zero vector of every subspace of V.

Remark 1.4.5 (Why nonemptiness matters). The empty set is closed under addition and scalar multiplication, because there are no elements to test: "for all $\mathbf{w}, \mathbf{w}' \in \emptyset$" is vacuously true. Likewise, it is stable under linear combinations of lists of length $k \geq 1$. But $\emptyset$ is not a vector space, since (A3) requires an element. So the conditions $W \neq \emptyset$ in (II) and $\mathbf{0} \in W$ in (III) cannot be dropped, and the proof shows where they enter: nonemptiness was used exactly once, to produce the zero vector of W as $0\mathbf{w}_0$.

In practice, check $\mathbf{0} \in W$. It proves nonemptiness, and it is also the fastest test in the other direction: a set that does not contain $\mathbf{0}$ is not a subspace, whatever else it does. If (II) also allowed $k = 0$, the empty linear combination, which represents $\mathbf{0}$ (Definition 1.3.4), would force $\mathbf{0} \in W$, and the condition $W \neq \emptyset$ would be automatic. Exercise 1.4.15 states the criterion in this form.

> **Exercise 1.4.6.** ROUTINE Show that $W \subseteq V$ is a subspace of V if and only if $\mathbf{0} \in W$ and $a\mathbf{w} + \mathbf{w}' \in W$ for all $\mathbf{w}, \mathbf{w}' \in W$ and all $a \in F$.

Solution on page 1053

This one-line form is convenient when W is defined by a condition that is visibly compatible with linear combinations, as for kernels in Section 2.6.

1.4.2 Examples

To show that a subset is a subspace, we check (III). To show that it is not, one failure suffices: $\mathbf{0} \notin W$, or particular vectors and scalars whose sum or scalar multiple leaves W.

Example 1.4.7 (Trivial subspaces). $\{\mathbf{0}\}$ and V are subspaces of V. For V itself there is nothing to check. The set $\{\mathbf{0}\}$ contains $\mathbf{0}$, and $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ by (A3) and Proposition 1.2.33. With the operations of V, $\{\mathbf{0}\}$ is the zero space (Example 1.2.3). By (III), every subspace of V contains it.

Example 1.4.8 (One equation in F^n). Let $a_1, \dots, a_n \in F$, and let

$$H = \{\mathbf{x} \in F^n \mid a_1x_1 + \cdots + a_nx_n = 0\}.$$

Then H is a subspace of F^n. It contains $\mathbf{0}$, since $a_1 0 + \cdots + a_n 0 = 0$. If $\mathbf{x}, \mathbf{y} \in H$ and $c \in F$, then by the coordinatewise operations, distributivity, and regrouping,

$$\sum_{i=1}^{n} a_i(x_i + y_i) = \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} a_i y_i = 0 + 0 = 0,$$
$$\sum_{i=1}^{n} a_i(cx_i) = c \sum_{i=1}^{n} a_i x_i = c \cdot 0 = 0,$$

so $\mathbf{x} + \mathbf{y} \in H$ and $c\mathbf{x} \in H$. For $n = 2$ and the equation $x_1 - x_2 = 0$ over $\mathbb{R}$, H is the line $\{(r, r) \mid r \in \mathbb{R}\}$. In $\mathbb{R}^3$, the equation $x_1 + x_2 + x_3 = 0$ gives a plane through $\mathbf{0}$.

With a nonzero right-hand side, the solution set of $a_1x_1 + \cdots + a_nx_n = b$ is never a subspace, since it does not contain $\mathbf{0}$: the left side at $\mathbf{0}$ is $0 \neq b$. The solution set of a homogeneous system of m equations is the set of vectors that lie in m subspaces of the form H at once, one for each equation; it is a subspace by Proposition 1.4.16 below. Corollary 3.1.16 reaches the same conclusion by recognizing it as a kernel.

Example 1.4.9 (Polynomials). For $n \geq 0$, $F[x]_{\leq n}$ is a subspace of $F[x]$. It contains $\mathbf{0}$, whose degree is $-\infty$, and by Proposition 1.2.18(a) and (b), sums and nonzero multiples of polynomials of degree at most n have degree at most n; the multiple $0p = \mathbf{0}$ does too.

Fix $c \in F$. The polynomials that vanish at c form a subspace $\{p \in F[x] \mid p(c) = 0\}$ of $F[x]$. Indeed $\mathbf{0}(c) = 0$, and if $p(c) = q(c) = 0$ and $a \in F$, then $(p + q)(c) = p(c) + q(c) = 0$ and $(ap)(c) = ap(c) = 0$ by Proposition 1.2.18(c).

The polynomials of degree exactly n do not form a subspace, since $\mathbf{0}$ has degree $-\infty$. Neither do the polynomials with $p(c) = 1$, since $\mathbf{0}(c) = 0$.

Example 1.4.10 (Functions). Let X be a set and $t_0 \in X$. The functions vanishing at t_0 form a subspace of F^X: the zero function vanishes at t_0, and if $f(t_0) = g(t_0) = 0$, then $(f + g)(t_0) = 0 + 0 = 0$ and $(af)(t_0) = a \cdot 0 = 0$ by the pointwise operations. The functions with $f(t_0) = 1$ do not, since the zero function is not among them.

In $\mathbb{R}^{\mathbb{R}}$, the **even** functions $E = \{f \mid f(-t) = f(t) \text{ for all } t\}$ and the **odd** functions $O = \{f \mid f(-t) = -f(t) \text{ for all } t\}$ form subspaces. For O: the zero function is odd, as $0 = -0$. If $f, g \in O$ and $a \in \mathbb{R}$, then for every t,

$$(f + g)(-t) = f(-t) + g(-t) = -f(t) - g(t) = -(f + g)(t), \quad (af)(-t) = af(-t) = -(af)(t).$$

The argument for E is the same without the signs. Finally, $C[0, 1]$ is a subspace of $\mathbb{R}^{[0,1]}$ (Example 1.2.12), and so are the continuous functions with integral 0.

Example 1.4.11 (Matrices). In $F^{n\times n}$, the **symmetric** matrices, those with $a_{ij} = a_{ji}$ for all i, j, form a subspace S, and so do the **trace-zero** matrices, those with $a_{11} + \cdots + a_{nn} = 0$. For $n = 2$,

$$S = \left\{ \begin{pmatrix} a & b \\ b & d \end{pmatrix} \,\middle|\, a, b, d \in F \right\}, \quad T = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \,\middle|\, a, b, c \in F \right\}.$$

Both are cut out by equations among the entries that the zero matrix satisfies, and that survive entrywise sums and scalar multiples: if $a_{ij} = a_{ji}$ and $b_{ij} = b_{ji}$, then $a_{ij} + b_{ij} = a_{ji} + b_{ji}$ and $ca_{ij} = ca_{ji}$, and the trace-zero condition is the equation of Example 1.4.8 in the n^2 entries. In the language of Section 2.4, S consists of the A with $A^{\mathsf{T}} = A$, and T of the A with $\operatorname{tr} A = 0$.

Example 1.4.12 (The subspaces of $\mathbb{F}_2^2$). Over $\mathbb{F}_2$, the only scalars are 0 and 1, and $0\mathbf{w} = \mathbf{0}$, $1\mathbf{w} = \mathbf{w}$. So a subset of an $\mathbb{F}_2$-vector space that contains $\mathbf{0}$ is automatically closed under scalar multiplication, and the subspaces are exactly the subsets that contain $\mathbf{0}$ and are closed under addition.

The space $\mathbb{F}_2^2$ has four elements, $\mathbf{0}$, $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$, and $\mathbf{e}_1 + \mathbf{e}_2 = (1,1)$. Each vector satisfies $\mathbf{v} + \mathbf{v} = \mathbf{0}$, and the sum of any two different nonzero vectors is the third one. So a subset containing $\mathbf{0}$ and at most one nonzero vector $\mathbf{v}$ is closed under addition, since $\mathbf{v} + \mathbf{v} = \mathbf{0}$ and $\mathbf{v} + \mathbf{0} = \mathbf{v}$. A subset containing two different nonzero vectors contains their sum, hence all three nonzero vectors, and it is all of $\mathbb{F}_2^2$. Hence $\mathbb{F}_2^2$ has exactly five subspaces:

$$\{\mathbf{0}\}, \quad \{\mathbf{0}, (1,0)\}, \quad \{\mathbf{0}, (0,1)\}, \quad \{\mathbf{0}, (1,1)\}, \quad \mathbb{F}_2^2.$$

The middle three are the spans of the nonzero vectors (Example 1.3.11), the analogues of lines through the origin. Section 1.4.6 shows that every F^2 looks like this.

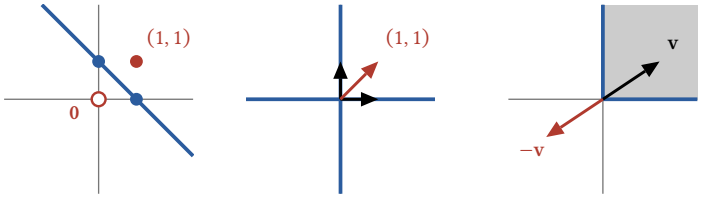


Figure 1.4.1. Three subsets of $\mathbb{R}^2$ that fail to be subspaces, each in a different way. Left: the line $c_1 + c_2 = 1$ misses $\mathbf{0}$; it also contains $(1,0)$ and $(0,1)$ but not their sum. Middle: the union of the axes contains $\mathbf{0}$ and every multiple of its vectors, but not $(1,0) + (0,1)$. Right: the closed first quadrant contains $\mathbf{0}$ and all sums of its vectors, but not $(-1)\mathbf{v}$.

Non-example 1.4.13 (Three ways to fail). Each of the following subsets of $\mathbb{R}^2$ violates a different condition of (III) (Figure 1.4.1).

(a) The line $L = \{(c_1, c_2) \mid c_1 + c_2 = 1\}$ does not contain $\mathbf{0}$. (It fails the other two conditions as well.)

(b) The union of the axes, $A = \{(x, y) \mid xy = 0\}$, contains $\mathbf{0}$ and is closed under scalar multiplication: if $xy = 0$, then $(ax)(ay) = a^2xy = 0$. But $(1,0)$ and $(0,1)$ lie in A, while their sum $(1,1)$ does not.

(c) The first quadrant $Q = \{(x, y) \mid x \geq 0, y \geq 0\}$ contains $\mathbf{0}$ and is closed under addition, since sums of nonnegative numbers are nonnegative. But $(1,0) \in Q$, while $(-1)(1,0) = (-1,0) \notin Q$.

The empty set satisfies both closure conditions and fails only $\mathbf{0} \in W$ (Remark 1.4.5). Together with (b) and (c), this shows that none of the three conditions in (III) follows from the other two.

The set A in (b) is the union of two subspaces, the two axes. We return to it in Section 1.4.4.

1.4.3 The span is the smallest subspace

Intuition. Let U be a subspace of V that contains the entries of a list $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$. By (II) of Theorem 1.4.3, U contains every linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$, so $\operatorname{span}(\mathcal{L}) \subseteq U$. If $\operatorname{span}(\mathcal{L})$ is itself a subspace, it is therefore the smallest subspace containing the entries of $\mathcal{L}$: it lies inside every other one. And it is a subspace, because sums and scalar multiples of linear combinations of $\mathcal{L}$ are linear combinations of $\mathcal{L}$. So the span can be described from the inside, as the set of all linear combinations, or from the outside, as the least subspace that the entries of $\mathcal{L}$ force on us.

Proposition 1.4.14 (The span is the smallest subspace containing a list). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V.

(a) $\operatorname{span}(\mathcal{L})$ is a subspace of V, and $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L})$.

(b) If U is a subspace of V and $\mathcal{L} \subseteq U$, then $\operatorname{span}(\mathcal{L}) \subseteq U$.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion), Proposition 1.3.17 (A list lies in its span), Lemma 1.3.9 (Arithmetic of linear combinations), Definition 1.3.10 (Span)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For (a), check (III) of Theorem 1.4.3. What is the sum of two linear combinations of $\mathcal{L}$?
Hint 2. For (b), an element of $\operatorname{span}(\mathcal{L})$ is a linear combination of vectors that lie in U. Which condition of Theorem 1.4.3 speaks about linear combinations? Treat the empty list separately.
Sketch. (a) $\mathbf{0}$ and the entries of $\mathcal{L}$ lie in $\operatorname{span}(\mathcal{L})$ by Proposition 1.3.17, and by Lemma 1.3.9, $\sum a_i\mathbf{v}_i + \sum b_i\mathbf{v}_i = \sum(a_i + b_i)\mathbf{v}_i$ and $c\sum a_i\mathbf{v}_i = \sum(ca_i)\mathbf{v}_i$. (b) The subspace U satisfies (II) of Theorem 1.4.3 and contains $\mathbf{v}_1, \ldots, \mathbf{v}_k$, so it contains every linear combination of them; if $k = 0$, $\operatorname{span}(\mathcal{L}) = \{\mathbf{0}\} \subseteq U$.

(a) We check (III) of Theorem 1.4.3. By Proposition 1.3.17, $\mathbf{0} \in \operatorname{span}(\mathcal{L})$, and also $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L})$. Let $\mathbf{u}, \mathbf{w} \in \operatorname{span}(\mathcal{L})$ and $c \in F$, say $\mathbf{u} = \sum_{i=1}^{k} a_i\mathbf{v}_i$ and $\mathbf{w} = \sum_{i=1}^{k} b_i\mathbf{v}_i$. By Lemma 1.3.9,

$$\mathbf{u} + \mathbf{w} = \sum_{i=1}^{k}(a_i + b_i)\mathbf{v}_i \quad \text{and} \quad c\mathbf{u} = \sum_{i=1}^{k}(ca_i)\mathbf{v}_i,$$

which are linear combinations of $\mathcal{L}$. So $\mathbf{u} + \mathbf{w}$ and $c\mathbf{u}$ lie in $\operatorname{span}(\mathcal{L})$, and $\operatorname{span}(\mathcal{L})$ is a subspace.

(b) Let U be a subspace with $\mathcal{L} \subseteq U$, and let $\mathbf{w} \in \operatorname{span}(\mathcal{L})$. If $k = 0$, then $\mathbf{w} = \mathbf{0}$, which lies in U by (III) of Theorem 1.4.3. If $k \geq 1$, then $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ for some scalars a_i, where each $\mathbf{v}_i$ lies in U. By (II) of Theorem 1.4.3, $\mathbf{w} \in U$. Hence $\operatorname{span}(\mathcal{L}) \subseteq U$. □

In particular, $\operatorname{span}(\mathcal{L})$ is a vector space in its own right, and $\mathcal{L}$ spans it. Part (b) gives the second proof of Proposition 1.3.18 promised in Section 1.3.4. If $\mathcal{M} \subseteq \operatorname{span}(\mathcal{L})$, then $\operatorname{span}(\mathcal{L})$ is a subspace, by (a), that contains the entries of $\mathcal{M}$, so it contains $\operatorname{span}(\mathcal{M})$, by (b). The first proof expanded every linear combination of $\mathcal{M}$ into one of $\mathcal{L}$ and collected coefficients. The second needs only that $\operatorname{span}(\mathcal{L})$ is closed under sums and scalar multiples; Theorem 1.4.3 upgrades this to closure under all linear combinations, by an induction done once and for all.

Part (b) also changes how we prove that a span lies in a given subspace: it suffices to check the finitely many entries of the list. For instance, the plane $P = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$ is a subspace (Example 1.4.8) that contains $(1, 0, -1)$ and $(0, 1, -1)$, so it contains their span. In Exercise 1.3.16 the same containment took a computation with an arbitrary linear combination.

Exercise 1.4.15. ROUTINE Show that a subset $W \subseteq V$ is a subspace of V if and only if $\operatorname{span}(\mathcal{L}) \subseteq W$ for every list $\mathcal{L}$ in W, the empty list included.

Solution on page 1053

1.4.4 Intersections and unions

Given two subspaces, we would like to build new ones from them. Intersections always work; unions almost never do.

Proposition 1.4.16 (Intersections of subspaces). Let $(U_i)_{i \in I}$ be a family of subspaces of V, indexed by any set I. Then

$$\bigcap_{i \in I} U_i = \{\mathbf{v} \in V \mid \mathbf{v} \in U_i \text{ for every } i \in I\}$$

is a subspace of V.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Check the three conditions of (III) in Theorem 1.4.3 for the intersection.
Hint 2. A vector lies in the intersection exactly when it lies in every U_i, and each U_i satisfies (III) on its own.
Sketch. $\mathbf{0}$ lies in every U_i. If $\mathbf{u}, \mathbf{w}$ lie in every U_i, then so do $\mathbf{u}+\mathbf{w}$ and $a\mathbf{u}$, since each U_i is closed under the operations.

Write $X = \bigcap_{i \in I} U_i$. We check (III) of Theorem 1.4.3. Each U_i is a subspace, so $\mathbf{0} \in U_i$ for every i, and $\mathbf{0} \in X$. Let $\mathbf{u}, \mathbf{w} \in X$ and $a \in F$, and fix $i \in I$. Then $\mathbf{u}, \mathbf{w} \in U_i$, and since U_i is closed under addition and scalar multiplication, $\mathbf{u}+\mathbf{w} \in U_i$ and $a\mathbf{u} \in U_i$. As i was arbitrary, $\mathbf{u}+\mathbf{w} \in X$ and $a\mathbf{u} \in X$. So X is a subspace. □

The family may be infinite, or even empty; for $I = \emptyset$ the condition "for every $i \in I$" is vacuous, the intersection is V, and the proof still applies. The polynomials of degree at most n that vanish at c form a subspace, as the intersection of two subspaces from Example 1.4.9, and the solution set of a homogeneous system is the intersection of the subspaces H of Example 1.4.8, one for each equation.

Intersections give a second description of the span, from the outside.

Corollary 1.4.17 (The span as an intersection). Let $\mathcal{L}$ be a list in V. Then $\operatorname{span}(\mathcal{L})$ is the intersection of all subspaces of V that contain the entries of $\mathcal{L}$.

Proof. ROUTINE

Uses. Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 1.4.16 (Intersections of subspaces)
Techniques. Double containment (Technique 1.3.14)
Hint 1. Let X be the intersection. Is $\operatorname{span}(\mathcal{L})$ one of the subspaces being intersected?
Hint 2. Each subspace in the family contains $\operatorname{span}(\mathcal{L})$, by Proposition 1.4.14(b).
Sketch. Since $\operatorname{span}(\mathcal{L})$ belongs to the family, $X \subseteq \operatorname{span}(\mathcal{L})$. Since every member contains $\operatorname{span}(\mathcal{L})$, $\operatorname{span}(\mathcal{L}) \subseteq X$.

Let X be the intersection of all subspaces U of V with $\mathcal{L} \subseteq U$. We proceed by double containment. By Proposition 1.4.14(a), $\operatorname{span}(\mathcal{L})$ is a member of the family, so $X \subseteq \operatorname{span}(\mathcal{L})$. By Proposition 1.4.14(b), every member U satisfies $\operatorname{span}(\mathcal{L}) \subseteq U$, so $\operatorname{span}(\mathcal{L}) \subseteq X$. □

The description in Definition 1.3.10 builds the span by listing its elements, and it is the one we compute with. The description in Corollary 1.4.17 uses no linear combinations at all, and so it makes sense for any subset $S \subseteq V$, finite or not: by Proposition 1.4.16, the intersection of all subspaces containing S is a subspace, and it is the smallest one containing S.

Unions

Now take two subspaces U and W. The union $U \cup W$ contains $\mathbf{0}$, and it is closed under scalar multiplication, since a multiple of a vector of U lies in U and a multiple of a vector of W lies in W. Addition is the problem. The union of the two axes of $\mathbb{R}^2$ (Figure 1.4.1, middle) is not closed under addition: a vector on one axis plus a vector on the other lies on neither. Reading off exactly why this example fails is a move worth naming.

Technique 1.4.18 (Finding counterexamples and minimal examples). To disprove a claim about all objects of some kind, a single object for which it fails is enough. Look for one by shrinking everything: the smallest space, the fewest vectors, the simplest field (Technique 1.2.38). Then read the counterexample. Ask which features made it work; often they point to the exact hypothesis under which the claim does hold, and the counterexample, run on general objects, turns into a proof.

Reach for it whenever a claim seems plausible but no proof suggests itself, and in every "prove or disprove". The canonical example is the claim that the union of two subspaces is a subspace. It fails for the two axes of $\mathbb{R}^2$, since $(1,0)+(0,1)$ lies on neither (Non-example 1.4.13). The failure used only one vector of the first subspace outside the second and one vector of the second outside the first. So nested subspaces escape the counterexample, and Proposition 1.4.20 shows that only they do.

Used in. Proposition 1.4.20, Exercise 1.4.37, Exercise 1.5.34, Exercise 1.6.35, Exercise 1.7.12, Exercise 1.7.15, Problem 1.10, Exercise 2.7.22, Exercise 2.8.14, Exercise 2.8.15, Problem 2.10, Problem 2.14, Problem 2.20, Problem 3.7, Problem 3.8, Exercise 4.1.41, Exercise 4.1.48, Exercise 4.2.38, Exercise 4.3.29, Exercise 4.3.34, Exercise 4.3.36, Exercise 4.4.34, Exercise 4.5.39, Exercise 4.5.46, Exercise 4.6.6, Exercise 4.6.51, Exercise 4.7.51, Exercise 4.8.5, Exercise 4.8.6, Exercise 4.8.11, Problem 5.4, Exercise 6.5.44, Exercise 6.7.15, Problem 6.6, Problem 6.7, Exercise 7.2.43, Exercise 7.2.53, Exercise 7.2.54, Exercise 7.3.38, Exercise 7.4.38, Exercise 7.5.33, Exercise 7.6.43, Exercise 7.7.4, Exercise 7.7.5, Exercise 7.7.6, Exercise 7.7.8, Problem 7.5, Exercise 8.3.29, Exercise 8.3.38, Exercise B.1.38, Exercise 1.4.26, Exercise 1.4.27, Exercise 1.4.40, Exercise 1.5.16, Exercise 2.4.44, Exercise 2.5.34, Exercise 2.5.35, Exercise 2.6.36, Exercise 2.6.37, Exercise 2.6.38, Exercise 2.6.40, Exercise 2.6.50, Exercise 2.6.52, Exercise 2.7.21, Exercise 2.7.27, Exercise 4.2.45, Exercise 4.3.32, Exercise 4.3.35, Exercise 5.5.38, Exercise 5.5.49, Exercise 6.2.34, Exercise 6.3.46, Exercise 6.3.47, Exercise 6.3.55, Exercise 6.4.48, Exercise 6.4.50, Exercise 6.4.63, Exercise 6.5.35, Exercise 6.6.36, Exercise 7.1.53, Exercise 7.5.37, Exercise 8.2.34, Exercise A.1.15, Exercise A.1.25, Exercise A.2.13, Exercise A.3.14, Exercise A.3.17, Exercise A.3.18.

The key step in the counterexample is that a vector on one axis plus a vector off that axis lands off that axis. This holds for every subspace.

Lemma 1.4.19 (Adding a vector outside a subspace). Let U be a subspace of V, let $\mathbf{u} \in U$, and let $\mathbf{v} \in V$ with $\mathbf{v} \notin U$. Then $\mathbf{u} + \mathbf{v} \notin U$.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion), Proposition 1.2.34 (Negatives and the scalar -1)
Hint 1. Suppose $\mathbf{u} + \mathbf{v} \in U$. Can you recover $\mathbf{v}$ from two vectors of U?
Hint 2. $\mathbf{v} = (\mathbf{u} + \mathbf{v}) + (-1)\mathbf{u}$, a linear combination of two vectors of U.
Sketch. If $\mathbf{u} + \mathbf{v} \in U$, then $\mathbf{v} = (\mathbf{u} + \mathbf{v}) + (-1)\mathbf{u}$ lies in U by closure, a contradiction.

Suppose, to the contrary, that $\mathbf{u} + \mathbf{v} \in U$. By Proposition 1.2.34 and the axioms,

$$(\mathbf{u} + \mathbf{v}) + (-1)\mathbf{u} = (\mathbf{v} + \mathbf{u}) + (-\mathbf{u}) = \mathbf{v} + (\mathbf{u} + (-\mathbf{u})) = \mathbf{v}.$$

The left side is a linear combination of $\mathbf{u} + \mathbf{v}$ and $\mathbf{u}$, which lie in U, so $\mathbf{v} \in U$ by (II) of Theorem 1.4.3. This contradicts $\mathbf{v} \notin U$. □

Proposition 1.4.20 (Unions of two subspaces). Let U and W be subspaces of V. Then $U \cup W$ is a subspace of V if and only if $U \subseteq W$ or $W \subseteq U$.

Intuition. The backward direction is immediate, since $U \cup W$ is then W or U. The forward direction concludes "$U \subseteq W$ or $W \subseteq U$", and there are two natural ways to prove a disjunction. One is to assume that the first alternative fails and deduce the second: a vector $\mathbf{u} \in U$ outside W is a lever, and adding it to each vector of W forces that vector into U. The other is the contrapositive: assume both alternatives fail, and build the counterexample of the two axes from one witness on each side.

Proof (assume one alternative fails). STANDARD

Uses. Theorem 1.4.3 (Subspace criterion), Lemma 1.4.19 (Adding a vector outside a subspace)
Hint 1. The backward direction is immediate: if $U \subseteq W$, then $U \cup W = W$.
Hint 2. For the forward direction, assume $U \not\subseteq W$ and prove $W \subseteq U$. Fix $\mathbf{u} \in U$ with $\mathbf{u} \notin W$.
Hint 3. For $\mathbf{w} \in W$, the sum $\mathbf{u} + \mathbf{w}$ lies in $U \cup W$. Which of the two can it lie in? Use Lemma 1.4.19.
Sketch. ($\Leftarrow$) $U \cup W$ is W or U. ($\Rightarrow$) Suppose $U \cup W$ is a subspace and $U \not\subseteq W$; fix $\mathbf{u} \in U \setminus W$. For $\mathbf{w} \in W$, $\mathbf{u} + \mathbf{w} \in U \cup W$ by closure, and $\mathbf{u} + \mathbf{w} \notin W$ by Lemma 1.4.19 (applied to W, with $\mathbf{w} \in W$ and $\mathbf{u} \notin W$). So $\mathbf{u} + \mathbf{w} \in U$, and $\mathbf{w} = (\mathbf{u} + \mathbf{w}) - \mathbf{u} \in U$. Hence $W \subseteq U$.

($\Leftarrow$) If $U \subseteq W$, then $U \cup W = W$, and if $W \subseteq U$, then $U \cup W = U$. Either way, $U \cup W$ is a subspace.

($\Rightarrow$) Suppose $U \cup W$ is a subspace. If $U \subseteq W$, we are done. Otherwise, fix a vector $\mathbf{u} \in U$ with $\mathbf{u} \notin W$; we show $W \subseteq U$. Let $\mathbf{w} \in W$. Both $\mathbf{u}$ and $\mathbf{w}$ lie in the subspace $U \cup W$, so $\mathbf{u} + \mathbf{w} \in U \cup W$. By Lemma 1.4.19, applied to the subspace W, the vector $\mathbf{w} \in W$, and the vector $\mathbf{u} \notin W$, we have $\mathbf{w} + \mathbf{u} \notin W$. Hence $\mathbf{u} + \mathbf{w} \in U$. Since $\mathbf{u} \in U$ as well, $\mathbf{w} = (\mathbf{u} + \mathbf{w}) + (-1)\mathbf{u} \in U$ by (II) of Theorem 1.4.3. As $\mathbf{w}$ was arbitrary, $W \subseteq U$. □

Proof (generalize the counterexample). STANDARD

> *Uses.* Theorem 1.4.3 (Subspace criterion), Lemma 1.4.19 (Adding a vector outside a subspace)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* The backward direction is immediate. For the forward direction, prove the contrapositive: if neither subspace contains the other, then $U \cup W$ is not closed under addition.
> *Hint 2.* Imitate the two axes: take $\mathbf{u} \in U$ outside W and $\mathbf{w} \in W$ outside U, and locate $\mathbf{u} + \mathbf{w}$.
> *Sketch.* If $U \not\subseteq W$ and $W \not\subseteq U$, pick $\mathbf{u} \in U \setminus W$ and $\mathbf{w} \in W \setminus U$. By Lemma 1.4.19, applied once to U and once to W, $\mathbf{u} + \mathbf{w}$ lies in neither U nor W. So $U \cup W$ is not closed under addition.

The backward direction is as in the first proof. For the forward direction, we prove the contrapositive. Suppose $U \not\subseteq W$ and $W \not\subseteq U$, and pick $\mathbf{u} \in U$ with $\mathbf{u} \notin W$ and $\mathbf{w} \in W$ with $\mathbf{w} \notin U$. By Lemma 1.4.19 applied to U, with $\mathbf{u} \in U$ and $\mathbf{w} \notin U$, we get $\mathbf{u} + \mathbf{w} \notin U$. By Lemma 1.4.19 applied to W, with $\mathbf{w} \in W$ and $\mathbf{u} \notin W$, we get $\mathbf{w} + \mathbf{u} \notin W$. So $\mathbf{u}$ and $\mathbf{w}$ lie in $U \cup W$, but $\mathbf{u} + \mathbf{w}$ does not. Hence $U \cup W$ is not closed under addition, and it is not a subspace by Theorem 1.4.3. □

The second proof is the counterexample of the two axes, with U and W in place of the axes and the two witnesses in place of $(1, 0)$ and $(0, 1)$. The first proof shows a different habit: to prove "P or Q", assume P fails and use the failure. Here the failure of $U \subseteq W$ supplied a single vector $\mathbf{u}$, and that one vector moved every vector of W into U.

Exercise 1.4.21. ROUTINE Let U and W be subspaces of V with $U \cup W = V$. Show that $U = V$ or $W = V$.

Solution on page 1054

Unions of chains

A union of subspaces can still be a subspace when the subspaces are nested. Closure under addition involves only two vectors at a time, so only two members of the family at a time, and for nested subspaces Proposition 1.4.20 applies to each pair.

Proposition 1.4.22 (Unions of chains). Let $(U_i)_{i \in I}$ be a family of subspaces of V, indexed by a nonempty set I, such that for all $i, j \in I$, $U_i \subseteq U_j$ or $U_j \subseteq U_i$. Then $\bigcup_{i \in I} U_i$ is a subspace of V. In particular, if $U_1 \subseteq U_2 \subseteq U_3 \subseteq \cdots$ are subspaces of V, then $\bigcup_{n \geq 1} U_n$ is a subspace of V.

Proof. ROUTINE

> *Uses.* Theorem 1.4.3 (Subspace criterion)
> *Techniques.* Fix an arbitrary element and chase it (Technique 1.3.13)
> *Hint 1.* Check (III) of Theorem 1.4.3. Two vectors of the union lie in two members of the family.
> *Hint 2.* If $\mathbf{u} \in U_i$ and $\mathbf{w} \in U_j$, one of U_i, U_j contains both vectors.
> *Sketch.* $\mathbf{0}$ lies in any member. If $\mathbf{u} \in U_i$ and $\mathbf{w} \in U_j$, say with $U_i \subseteq U_j$, then $\mathbf{u} + \mathbf{w} \in U_j$. Multiples of $\mathbf{u}$ stay in U_i. For an increasing sequence, $U_m \subseteq U_n$ whenever $m \leq n$, so any two members are nested.

Write $X = \bigcup_{i \in I} U_i$. We check (III) of Theorem 1.4.3. Since I is nonempty, there is some $i_0 \in I$, and $\mathbf{0} \in U_{i_0} \subseteq X$. Let $\mathbf{u}, \mathbf{w} \in X$ and $a \in F$, say $\mathbf{u} \in U_i$ and $\mathbf{w} \in U_j$. Then $a\mathbf{u} \in U_i \subseteq X$. By hypothesis, $U_i \subseteq U_j$ or $U_j \subseteq U_i$; exchanging the names of $\mathbf{u}$ and $\mathbf{w}$ if necessary, we may assume $U_i \subseteq U_j$. Then $\mathbf{u}$ and $\mathbf{w}$ both lie in the subspace U_j, so $\mathbf{u} + \mathbf{w} \in U_j \subseteq X$. So X is a subspace.

For an increasing sequence $U_1 \subseteq U_2 \subseteq \cdots$, we have $U_m \subseteq U_n$ whenever $m \leq n$, by induction on $n - m$, so any two members are nested and the first part applies. □

Example 1.4.23 ($F[x]$ as a union of a chain). The subspaces $F[x]_{\leq 0} \subseteq F[x]_{\leq 1} \subseteq F[x]_{\leq 2} \subseteq \cdots$ of Example 1.4.9 form an increasing sequence, and their union is all of $F[x]$, since every polynomial has a degree, which is $-\infty$ or some $n \geq 0$. Each $F[x]_{\leq n}$ is a proper subspace, as $x^{n+1} \notin F[x]_{\leq n}$. So $F[x]$ is the union of infinitely many proper subspaces. By Exercise 1.4.21, two proper subspaces never suffice, and over an infinite field no finite number does (Exercise 1.4.43).

1.4.5 Sums of subspaces

Since $U \cup W$ is usually not a subspace, we ask instead for the smallest subspace that contains both U and W. Any subspace containing U and W contains every sum $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$, and these sums already form a subspace.

Proposition 1.4.24 (The sum of two subspaces). Let U and W be subspaces of V, and let

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W\}.$$

(a) $U + W$ is a subspace of V that contains U and W.

(b) If X is a subspace of V with $U \subseteq X$ and $W \subseteq X$, then $U + W \subseteq X$.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion), Definition 1.2.1 (Vector space)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For (a), check (III) of Theorem 1.4.3. Regroup $(\mathbf{u} + \mathbf{w}) + (\mathbf{u}' + \mathbf{w}')$ so that the vectors of U are together.
Hint 2. To see $U \subseteq U + W$, write $\mathbf{u} = \mathbf{u} + \mathbf{0}$. For (b), use that X is closed under addition.
Sketch. (a) $\mathbf{0} = \mathbf{0} + \mathbf{0}$; $(\mathbf{u} + \mathbf{w}) + (\mathbf{u}' + \mathbf{w}') = (\mathbf{u} + \mathbf{u}') + (\mathbf{w} + \mathbf{w}')$ and $a(\mathbf{u} + \mathbf{w}) = a\mathbf{u} + a\mathbf{w}$; and $\mathbf{u} = \mathbf{u} + \mathbf{0}$, $\mathbf{w} = \mathbf{0} + \mathbf{w}$. (b) If $\mathbf{u}, \mathbf{w} \in X$, then $\mathbf{u} + \mathbf{w} \in X$.

(a) We check (III) of Theorem 1.4.3. Since $\mathbf{0} \in U$ and $\mathbf{0} \in W$, $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$. Let $\mathbf{u} + \mathbf{w}$ and $\mathbf{u}' + \mathbf{w}'$ be elements of $U + W$, with $\mathbf{u}, \mathbf{u}' \in U$ and $\mathbf{w}, \mathbf{w}' \in W$, and let $a \in F$. By (A1) and (A2),

$$(\mathbf{u} + \mathbf{w}) + (\mathbf{u}' + \mathbf{w}') = (\mathbf{u} + \mathbf{u}') + (\mathbf{w} + \mathbf{w}'),$$

with $\mathbf{u} + \mathbf{u}' \in U$ and $\mathbf{w} + \mathbf{w}' \in W$, and by (A7), $a(\mathbf{u} + \mathbf{w}) = a\mathbf{u} + a\mathbf{w}$, with $a\mathbf{u} \in U$ and $a\mathbf{w} \in W$. So both lie in $U + W$, and $U + W$ is a subspace. Finally, for $\mathbf{u} \in U$ we have $\mathbf{u} = \mathbf{u} + \mathbf{0} \in U + W$, and for $\mathbf{w} \in W$ we have $\mathbf{w} = \mathbf{0} + \mathbf{w} \in U + W$, by (A3) and (A1).

(b) Let $\mathbf{u} + \mathbf{w} \in U + W$, with $\mathbf{u} \in U \subseteq X$ and $\mathbf{w} \in W \subseteq X$. Since X is closed under addition, $\mathbf{u} + \mathbf{w} \in X$. □

So $U + W$ is the smallest subspace containing $U \cup W$, and $U \cup W \subseteq U + W$ with equality exactly when one of U, W contains the other (Proposition 1.4.20). Sums of spans are spans: $\operatorname{span}(\mathcal{L}) + \operatorname{span}(\mathcal{M}) = \operatorname{span}(\mathcal{L}, \mathcal{M})$ by Exercise 1.3.36. For example, in F^3 the sum of the x_1-axis $\operatorname{span}(\mathbf{e}_1)$ and the x_2-axis $\operatorname{span}(\mathbf{e}_2)$ is $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$, the plane $x_3 = 0$, while their union is just the two axes. Sums of more than two subspaces, the case where every vector of $U + W$ has only one decomposition $\mathbf{u} + \mathbf{w}$, and the dimension of $U + W$ are the subject of Section 4.4.

1.4.6 Subspaces of the plane

In $\mathbb{R}^2$ we have met three kinds of subspace: $\{\mathbf{0}\}$, lines through the origin, and $\mathbb{R}^2$ itself. There are no others, over any field.

Intuition. A subspace that contains one nonzero vector contains the whole line through it, and a subspace that contains two vectors pointing in different directions contains the whole plane. There is no room in between. The proof makes this precise with Proposition 1.4.14(b), which turns "contains these vectors" into "contains their span", and with Exercise 1.3.37, which says when two vectors span F^2.

Theorem 1.4.25 (The subspaces of F^2). The subspaces of F^2 are $\{\mathbf{0}\}$, the lines $\operatorname{span}(\mathbf{v})$ with $\mathbf{v} \neq \mathbf{0}$, and F^2. Two such lines $\operatorname{span}(\mathbf{v})$ and $\operatorname{span}(\mathbf{v}')$ are equal if and only if $\mathbf{v}' = c\mathbf{v}$ for some $c \neq 0$.

Proof. STANDARD

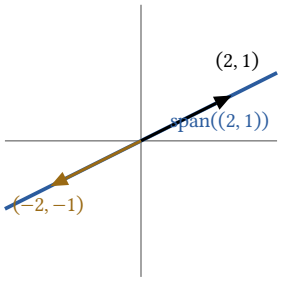
Uses. Proposition 1.4.14 (The span is the smallest subspace containing a list), Exercise 1.3.37, Exercise 1.3.28, Example 1.4.7 (Trivial subspaces)
Hint 1. Each of the listed sets is a subspace. For the converse, let U be a subspace and split into cases according to how much of F^2 it contains.

Hint 2. If $U \neq \{\mathbf{0}\}$, pick $\mathbf{v} \in U$ with $\mathbf{v} \neq \mathbf{0}$. Then $\operatorname{span}(\mathbf{v}) \subseteq U$. If this is not all of U, pick $\mathbf{w} \in U$ outside $\operatorname{span}(\mathbf{v})$.
Hint 3. By Exercise 1.3.37(b), $\mathbf{v}, \mathbf{w}$ spans F^2. Now apply Proposition 1.4.14(b) again.
Sketch. The listed sets are subspaces by Example 1.4.7 and Proposition 1.4.14(a). Let U be a subspace other than $\{\mathbf{0}\}$, and pick $\mathbf{v} \neq \mathbf{0}$ in U, so $\operatorname{span}(\mathbf{v}) \subseteq U$. Either $U = \operatorname{span}(\mathbf{v})$, or there is $\mathbf{w} \in U \setminus \operatorname{span}(\mathbf{v})$, and then $F^2 = \operatorname{span}(\mathbf{v}, \mathbf{w}) \subseteq U$ by Exercise 1.3.37(b). The statement about equal lines is Exercise 1.3.28(b).

The sets $\{\mathbf{0}\}$ and F^2 are subspaces (Example 1.4.7), and so is every span (Proposition 1.4.14(a)). Conversely, let U be a subspace of F^2. If $U = \{\mathbf{0}\}$, we are done. Otherwise, choose $\mathbf{v} \in U$ with $\mathbf{v} \neq \mathbf{0}$. By Proposition 1.4.14(b), $\operatorname{span}(\mathbf{v}) \subseteq U$. If $U = \operatorname{span}(\mathbf{v})$, then U is a line. Otherwise, choose $\mathbf{w} \in U$ with $\mathbf{w} \notin \operatorname{span}(\mathbf{v})$. Since $\mathbf{v} \neq \mathbf{0}$ and $\mathbf{w} \notin \operatorname{span}(\mathbf{v})$, the list $\mathbf{v}, \mathbf{w}$ spans F^2 by Exercise 1.3.37(b). Both entries lie in U, so by Proposition 1.4.14(b),

$$F^2 = \operatorname{span}(\mathbf{v}, \mathbf{w}) \subseteq U,$$

and $U = F^2$. The criterion for two lines to be equal is Exercise 1.3.28(b). □



Both vectors lie on the line $\operatorname{span}((2, 1))$, so their span is that line. In the interactive version, drag the second vector off the line: the span jumps from the line to all of $\mathbb{R}^2$, with nothing in between. (Interactive in the web version.)

Over $\mathbb{R}$, the lines $\operatorname{span}(\mathbf{v})$ are the lines through the origin. Over $\mathbb{F}_2$ each line has two elements, and the theorem recovers the five subspaces of Example 1.4.12. Over a field with q elements, F^2 has $q + 1$ lines (Exercise 1.4.33).

The proof was a count in disguise: U contains no direction, one direction, or two. Section 4.1 makes this count precise. Every subspace U of F^n has a dimension between 0 and n, with $U = F^n$ only for dimension n (Proposition 4.1.21). In particular, the subspaces of $\mathbb{R}^3$ are $\{\mathbf{0}\}$, the lines and the planes through $\mathbf{0}$, and $\mathbb{R}^3$.

Exercises

Exercise 1.4.26. ROUTINE Decide which of the following are subspaces, and justify each answer.

(a) $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1 = 2x_3\}$.

(b) $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1\}$.

(c) $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1 x_2 = 0\}$.

(d) $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1 \leq x_2\}$.

(e) $\{\mathbf{x} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 0\}$, and the subset of $\mathbb{C}^3$ defined by the same equation.

Solution on page 1054

Exercise 1.4.27. ROUTINE Decide which of the following are subspaces, and justify each answer. Write $p = \sum_k a_k x^k$ for polynomials.

(a) $\left\{f \in \mathbb{R}^{\mathbb{R}} \;\middle|\; f(0) = f(1)\right\}$.

(b) $\left\{f \in \mathbb{R}^{\mathbb{R}} \;\middle|\; f(1) = f(0)^2\right\}$.

(c) The bounded functions in $\mathbb{R}^{\mathbb{R}}$, those f for which there is M with $|f(t)| \leq M$ for all t.

(d) $\{p \in \mathbb{R}[x] \mid p(1)p(2) = 0\}$.

(e) $\{p \in \mathbb{R}[x] \mid a_k = 0 \text{ for every odd } k\}$.

(f) $\left\{A \in F^{2\times 2} \;\middle|\; a_{11}a_{22} = a_{12}a_{21}\right\}$.

Solution on page 1054

Exercise 1.4.28. STANDARD By Proposition 1.2.19, $\mathbb{C}^2$ is a vector space over $\mathbb{C}$ and also over $\mathbb{R}$, with the same addition. For $z = a + bi$ with $a, b \in \mathbb{R}$, write $\overline{z} = a - bi$.

(a) Show that every subspace of $\mathbb{C}^2$ over $\mathbb{C}$ is also a subspace of $\mathbb{C}^2$ over $\mathbb{R}$.

(b) For each of the sets $\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}$ and $D = \{(z, \overline{z}) \mid z \in \mathbb{C}\}$, decide whether it is a subspace of $\mathbb{C}^2$ over $\mathbb{R}$, and whether it is a subspace over $\mathbb{C}$.

Solution on page 1055

Exercise 1.4.29. ROUTINE Find the error in the following argument.

Claim. The set $P = \{t \in \mathbb{R} \mid t > 0\}$ is a subspace of the real vector space $\mathbb{R}$.

Proof. For $s, t \in P$ and $c \in \mathbb{R}$, define $s \oplus t = st$ and $c \odot t = t^c$. With these operations, P is a real vector space, with zero vector 1. Hence P is a subspace of $\mathbb{R}$.

(You may take for granted that P is a real vector space under $\oplus$ and $\odot$.)

Solution on page 1055

Exercise 1.4.30. ROUTINE Find the error in the following argument.

Claim. If U and W are subspaces of V, then $U \cup W$ is a subspace of V.

Proof. Since $\mathbf{0} \in U$, we have $\mathbf{0} \in U \cup W$. Let $\mathbf{v} \in U \cup W$ and $a \in F$. Then $\mathbf{v}$ lies in U or in W, and in either case $a\mathbf{v}$ lies in the same subspace, hence in $U \cup W$. Finally, let $\mathbf{v}, \mathbf{w} \in U \cup W$. Each of them lies in U or in W, and U and W are closed under addition, so $\mathbf{v} + \mathbf{w} \in U$ or $\mathbf{v} + \mathbf{w} \in W$. Hence $\mathbf{v} + \mathbf{w} \in U \cup W$.

Solution on page 1055

Exercise 1.4.31. STANDARD

(a) Let p be a prime, and let V be a vector space over $\mathbb{F}_p$. Show that a nonempty subset $W \subseteq V$ that is closed under addition is a subspace.

(b) Show that (a) fails over $\mathbb{Q}$: find a nonempty subset of $\mathbb{Q}^2$ that is closed under addition and under $\mathbf{v} \mapsto -\mathbf{v}$, but is not a subspace of the $\mathbb{Q}$-vector space $\mathbb{Q}^2$.

Solution on page 1056

Exercise 1.4.32. STANDARD For which primes p is $S_p = \left\{(x, y) \in \mathbb{F}_p^2 \;\middle|\; x^2 = y^2\right\}$ a subspace of $\mathbb{F}_p^2$?

Solution on page 1056

Exercise 1.4.33. STANDARD

(a) List all subspaces of $\mathbb{F}_3^2$, writing each line as a set of three vectors.

(b) Let F be a finite field with q elements. Show that F^2 has exactly $q+1$ lines $\operatorname{span}(\mathbf{v})$, $\mathbf{v} \neq \mathbf{0}$, and hence exactly $q+3$ subspaces.

Solution on page 1056

Exercise 1.4.34. STANDARD Let $\sigma : F^2 \to F^2$ swap the coordinates, $\sigma(x, y) = (y, x)$. Find all subspaces U of F^2 such that $\sigma(\mathbf{u}) \in U$ for every $\mathbf{u} \in U$. The answer depends on whether $1 + 1 = 0$ in F.

Solution on page 1057

Exercise 1.4.35. STANDARD Let E and O be the subspaces of even and odd functions in $\mathbb{R}^{\mathbb{R}}$ (Example 1.4.10).

(a) Show that $E \cap O = \{\mathbf{0}\}$ and $E + O = \mathbb{R}^{\mathbb{R}}$.

(b) Over $\mathbb{F}_2$, call $f : \mathbb{F}_2 \to \mathbb{F}_2$ even if $f(-t) = f(t)$ for all t, and odd if $f(-t) = -f(t)$ for all t. Show that every such f is both even and odd. Which steps of (a) fail over $\mathbb{F}_2$?

Solution on page 1057

Exercise 1.4.36. STANDARD Let $n \geq 2$. In $F^{n\times n}$, let S be the symmetric matrices, N the strictly upper triangular matrices ($a_{ij} = 0$ whenever $i \geq j$), and K the skew-symmetric matrices ($a_{ij} = -a_{ji}$ for all i, j).

(a) Show that N and K are subspaces, that $S \cap N = \{\mathbf{0}\}$, and that $S + N = F^{n\times n}$. This holds over every field.

(b) Show that $S + K = F^{n\times n}$ if and only if $1 + 1 \neq 0$ in F.

Solution on page 1058

Exercise 1.4.37. STANDARD Let U, W, and X be subspaces of V. Prove or disprove each statement.

(a) If $U + W = U + X$, then $W = X$.

(b) $U \cap (W + X) = (U \cap W) + (U \cap X)$.

(c) If $W \subseteq U$, then $U \cap (W + X) = W + (U \cap X)$.

Solution on page 1058

Exercise 1.4.38. STANDARD Let U be a proper subspace of V. Show that the only subspace of V that contains every vector outside U is V itself.

Solution on page 1059

Exercise 1.4.39. STANDARD

(a) Let U be a subspace of $F[x]$ that contains a polynomial of degree k for every $k \geq 0$. Show that $U = F[x]$.

(b) Find a proper subspace of $F[x]$ that contains polynomials of infinitely many different degrees.

Solution on page 1059

Exercise 1.4.40. HARD A function $f : \mathbb{R} \to \mathbb{R}$ has **period** $p > 0$ if $f(t + p) = f(t)$ for all $t \in \mathbb{R}$, and f is **periodic** if it has some period.

(a) Show that the functions with period 1 form a subspace of $\mathbb{R}^{\mathbb{R}}$.

(b) Is the set of all periodic functions a subspace of $\mathbb{R}^{\mathbb{R}}$?

Solution on page 1059

Exercise 1.4.41. HARD Let $n \geq 2$. Call a subspace U of F^n *symmetric* if permuting the coordinates of any vector of U gives a vector of U. Let $\mathbf{s} = (1, 1, ..., 1)$ and $Z = \{\mathbf{x} \in F^n \mid x_1 + \cdots + x_n = 0\}$.

(a) Show that the symmetric subspaces of F^n are exactly $\{\mathbf{0}\}$, $\text{span}(\mathbf{s})$, Z, and F^n.

(b) When is $\text{span}(\mathbf{s}) \subseteq Z$?

Solution on page 1060

Exercise 1.4.42. HARD

(a) Show that $\mathbb{F}_2^2$ is the union of three subspaces, none of which contains the other two.

(b) Suppose F has at least three elements, and let U_1, U_2, U_3 be subspaces of V whose union is a subspace. Show that one of the U_i contains the other two.

Solution on page 1061

Exercise 1.4.43. CHALLENGE BEYOND 21-242 Let V be a vector space over F, and let $U_1, ..., U_n$ be proper subspaces of V, where $n \geq 1$.

(a) Show that if F has at least n elements, then $U_1 \cup \cdots \cup U_n \neq V$.

(b) Deduce that over an infinite field, no vector space is a union of finitely many proper subspaces.

(c) Show that (a) is sharp: if F has q elements, then F^2 is a union of $q + 1$ proper subspaces.

Solution on page 1061

Exercise 1.4.44. ROUTINE

(a) Find the smallest subspace of $\mathbb{R}^3$ containing $(1, 1, 0)$ and $(1, -1, 0)$, and describe it by one equation.

(b) Find the smallest subspace of $\mathbb{R}[x]$ containing $1 + x^2$ and $1 - x^2$.

Solution on page 1062

1.5 Linear independence

A linear combination of a list represents one vector, but one vector can be represented by many linear combinations. Section 1.3 asked which vectors the linear combinations of a list represent, and the span answered it. This section asks in how many ways. In Example 1.3.5, the vector $(1, -1)$ has two different representations by the list $(0, 0), (1, -1), (-\pi, \pi)$, while every vector of F^2 has exactly one representation by $(1, 0), (0, 1)$. Lists of the second kind are called linearly independent. A linearly independent list that also spans V names every vector of V by a unique list of coefficients, and these coefficients are the coordinates of Section 1.6.

1.5.1 Unique representations

Throughout, V is a vector space over a field F.

Definition 1.5.1 (Linear independence). A list $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ in V is **linearly independent** if every vector of $\text{span}(\mathcal{L})$ is represented by exactly one linear combination of $\mathcal{L}$. Otherwise $\mathcal{L}$ is **linearly**

dependent. We also say that the vectors $\mathbf{v}_1, ..., \mathbf{v}_k$ are linearly independent (or dependent), and we often drop the word "linearly".

Every vector of $\operatorname{span}(\mathcal{L})$ is represented by at least one linear combination of $\mathcal{L}$, by Definition 1.3.10, so independence asks for at most one. In terms of coefficients, $\mathcal{L}$ is linearly independent when, for all scalars $a_1, ..., a_k$ and $b_1, ..., b_k$,

$$\sum_{i=1}^{k} a_i \mathbf{v}_i = \sum_{i=1}^{k} b_i \mathbf{v}_i \quad \text{implies} \quad a_1 = b_1, \ ..., \ a_k = b_k.$$

This is the claim of Exercise 1.3.29, which holds for some lists and fails for others. Order and repetitions are part of a list, and independence is a property of the list, not of its set of entries: we will see that a list with a repeated entry is never independent (Proposition 1.5.15).

Example 1.5.2 (The lists $\mathcal{L}_0$ and $\mathcal{L}_1$). In $\mathbb{R}^2$, let $\mathcal{L}_0 = (1,0), (0,1)$ and $\mathcal{L}_1 = (0,0), (1,-1), (-\pi, \pi)$.

$\mathcal{L}_0$ is linearly independent. The linear combination with coefficients a_1, a_2 represents $a_1(1,0) + a_2(0,1) = (a_1, a_2)$. So if the linear combinations with coefficients a_1, a_2 and b_1, b_2 both represent (c_1, c_2), then $(a_1, a_2) = (c_1, c_2) = (b_1, b_2)$, that is, $a_1 = c_1 = b_1$ and $a_2 = c_2 = b_2$. The two linear combinations were the same all along. The same computation shows that $\mathbf{e}_1, ..., \mathbf{e}_n$ is linearly independent in F^n, since the linear combination with coefficients $a_1, ..., a_n$ represents $(a_1, ..., a_n)$ (Example 1.3.7).

$\mathcal{L}_1$ is linearly dependent. By Example 1.3.15, $\operatorname{span}(\mathcal{L}_1)$ consists of the vectors $(c, -c)$ with $c \in \mathbb{R}$. Each of them has the two representations

$$0(0,0) + c(1,-1) + 0(-\pi, \pi) \quad \text{and} \quad \sqrt{87}(0,0) + 0(1,-1) + \left(-\frac{c}{\pi}\right)(-\pi, \pi),$$

which are different linear combinations, as their first coefficients differ. So no vector of $\operatorname{span}(\mathcal{L}_1)$ has a unique representation.

Intuition (Independence as the absence of redundancy). The two representations of $(c, -c)$ by $\mathcal{L}_1$ show where the redundancy lies. Any coefficient can be placed on $(0,0)$ at no cost, and since $(-\pi, \pi) = -\pi(1,-1)$, an amount of $(1,-1)$ can be traded for a multiple of $(-\pi, \pi)$. An independent list allows no such trade. From the vector that a linear combination represents we can recover every coefficient, so no entry can be made up for by the others. Theorem 1.5.5 and Lemma 1.5.10 turn these two readings, "the coefficients can be recovered" and "no entry is made up for by the others", into criteria we can check.

1.5.2 Testing at zero

Checking Definition 1.5.1 directly means checking every vector of the span, usually infinitely many. One vector suffices.

Intuition. Two representations of the same vector $\mathbf{w}$ differ by a representation of $\mathbf{0}$: subtracting coefficients gives a linear combination that represents $\mathbf{w} - \mathbf{w} = \mathbf{0}$. In the other direction, adding the coefficients of a representation of $\mathbf{0}$ to those of a representation of $\mathbf{w}$ gives another representation of $\mathbf{w}$ (Exercise 1.3.6). So every vector of the span has as many representations as $\mathbf{0}$ has, and $\mathbf{0}$ is the vector to test. It always has the representation with all coefficients 0; the question is whether it has any other.

Definition 1.5.3 (Linear relation). A **linear relation** of a list $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ is a linear combination of $\mathcal{L}$ that represents $\mathbf{0}$, that is, a list of scalars $a_1, ..., a_k$ with

$$a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k = \mathbf{0}.$$

The **trivial relation** is the one with $a_1 = \cdots = a_k = 0$; it represents $\mathbf{0}$ by Lemma 1.3.9(c). A linear relation with at least one nonzero coefficient is **nontrivial**, and is also called a **dependence relation** of $\mathcal{L}$.

The empty list has one linear combination, the empty one, which represents $\mathbf{0}$ and has no nonzero coefficient. So its only linear relation is the trivial one.

The lecture proves the equivalence below by proving one direction in a contrapositive form. The move deserves a name.

Technique 1.5.4 (Contrapositive). An implication "P implies Q" is logically equivalent to its contrapositive, "not Q implies not P". Prove the contrapositive when "not Q" gives you more to work with than P does. The typical case is a conclusion Q asserting that something holds *for every* object, or that some object is *unique*. Then "not Q" hands you a specific object, a counterexample or a second candidate, and you can compute with it.

Reach for it when the hypothesis is hard to use directly and the conclusion is a "for all" or a uniqueness statement. The canonical example is (II) ⇒ (I) in Theorem 1.5.5. Statement (I) asserts uniqueness for every vector of the span. Its negation provides one vector with two different representations, and subtracting them shows that (II) fails.

Used in. Theorem 1.5.5, Corollary 1.5.13, Corollary 1.5.17, Problem 1.1, Problem 1.3, Corollary 2.5.11, Proposition 2.6.14, Proposition 3.4.15, Proposition 3.4.16, Corollary 4.1.13, Theorem 4.7.22, Theorem 5.4.7, Corollary 7.1.23, Exercise 1.5.41, Exercise A.1.24.

Theorem 1.5.5 (Independence is decided at zero). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V. The following are equivalent.

(I) $\mathcal{L}$ is linearly independent.

(II) $\mathcal{L}$ has no nontrivial linear relation: if $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, then $a_1 = \cdots = a_k = 0$.

Proof. ROUTINE

Uses. Definition 1.5.1 (Linear independence), Definition 1.5.3 (Linear relation), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.34 (Negatives and the scalar -1)

Techniques. Contrapositive (Technique 1.5.4)

Hint 1. (I) ⇒ (II) is a special case of the definition. Which vector of the span is in question, and which representation does it always have?

Hint 2. For (II) ⇒ (I), prove the contrapositive. If $\mathcal{L}$ is dependent, some $\mathbf{w} \in \operatorname{span}(\mathcal{L})$ has two different representations.

Hint 3. Subtract the two representations: compute $\mathbf{w} + (-1)\mathbf{w}$ with Lemma 1.3.9. Why is one of the resulting coefficients nonzero?

Sketch. (I) ⇒ (II): $\mathbf{0} \in \operatorname{span}(\mathcal{L})$, and the trivial relation represents it, so by uniqueness it is the only linear combination representing $\mathbf{0}$. (II) ⇒ (I), by contrapositive: if $\sum a_i\mathbf{v}_i$ and $\sum b_i\mathbf{v}_i$ are different representations of $\mathbf{w}$, then $\sum(a_i - b_i)\mathbf{v}_i = \mathbf{w} - \mathbf{w} = \mathbf{0}$, and $a_j \neq b_j$ for some j, so this relation is nontrivial.

(I) ⇒ (II). Suppose that $\mathcal{L}$ is linearly independent. The trivial relation represents $\mathbf{0}$, so $\mathbf{0} \in \operatorname{span}(\mathcal{L})$, and by (I) no other linear combination of $\mathcal{L}$ represents $\mathbf{0}$. That is, every linear relation of $\mathcal{L}$ is trivial.

(II) ⇒ (I). We prove the contrapositive: if $\mathcal{L}$ is linearly dependent, then $\mathcal{L}$ has a nontrivial linear relation. Suppose that $\mathcal{L}$ is linearly dependent. Then some $\mathbf{w} \in \operatorname{span}(\mathcal{L})$ is represented by two different linear combinations of $\mathcal{L}$, say with coefficients $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$. By (A4) and Proposition 1.2.34, then Lemma 1.3.9(b) with $c = -1$, and Lemma 1.3.9(a),

$$\begin{aligned}\mathbf{0} = \mathbf{w} + (-1)\mathbf{w} &= \sum_{i=1}^{k} a_i\mathbf{v}_i + (-1)\sum_{i=1}^{k} b_i\mathbf{v}_i \\ &= \sum_{i=1}^{k} a_i\mathbf{v}_i + \sum_{i=1}^{k}(-b_i)\mathbf{v}_i = \sum_{i=1}^{k}(a_i - b_i)\mathbf{v}_i.\end{aligned}$$

The two linear combinations are different, so $a_j \neq b_j$ for some j, and then $a_j - b_j \neq 0$. So the scalars $a_1 - b_1, ..., a_k - b_k$ form a nontrivial linear relation of $\mathcal{L}$, and (II) fails. □

In practice, to show that $\mathcal{L}$ is independent, suppose that $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ and deduce that every a_i is 0. To show that $\mathcal{L}$ is dependent, exhibit one nontrivial relation. A relation is a certificate: anyone can check it by computing a single linear combination, without knowing how it was found.

A single relation also produces many representations. If $c_1, ..., c_k$ is a dependence relation and $b_1, ..., b_k$ represents $\mathbf{w}$, then for every $t \in F$ the coefficients $b_1 + tc_1, ..., b_k + tc_k$ represent $\mathbf{w} + t\mathbf{0} = \mathbf{w}$, by Lemma 1.3.9. Different values of t give different linear combinations, because $c_j \neq 0$ for some j. So over an infinite field, a dependent list represents each vector of its span in infinitely many ways, and over a finite field with q elements in at least q ways.

Example 1.5.6 (A list whose independence depends on the field). Consider $\mathcal{L} = (1, 2), (2, 1)$ in F^2, where $2 = 1 + 1$ in F. Suppose that $a(1, 2) + b(2, 1) = \mathbf{0}$, that is,

$$a + 2b = 0 \quad \text{and} \quad 2a + b = 0.$$

Subtracting twice the first equation from the second gives $b - 4b = 0$, that is, $-3b = 0$, where $3 = 1 + 1 + 1$.

If $3 \neq 0$ in F, as in $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_2$, and $\mathbb{F}_5$, then $-3 \neq 0$ and $b = 0$ (Proposition 1.1.17), and then $a = -2b = 0$. So the only linear relation is trivial, and $\mathcal{L}$ is linearly independent by Theorem 1.5.5. (Over $\mathbb{F}_2$ the list is just $(1, 0), (0, 1)$.)

If $3 = 0$, as in $\mathbb{F}_3$, both equations hold for $a = b = 1$, since $1 + 2 = 3 = 0$. So $(1, 2) + (2, 1) = \mathbf{0}$ in $\mathbb{F}_3^2$ is a dependence relation, and $\mathcal{L}$ is linearly dependent. In fact $(2, 1) = 2(1, 2)$ in $\mathbb{F}_3^2$, since $2 \cdot 2 = 4 = 1$.

Example 1.5.7 (Monomials). In $F[x]$, the list $1, x, x^2, ..., x^n$ is linearly independent. By Notation 1.2.15, the linear combination $a_0 + a_1x + \cdots + a_nx^n$ represents the polynomial with coefficients $(a_0, a_1, ..., a_n, 0, 0, ...)$, and by Definition 1.2.13 this is the zero polynomial only if $a_0 = a_1 = \cdots = a_n = 0$. So the only linear relation is trivial. This holds over every field, $\mathbb{F}_2$ included, because polynomials are compared coefficient by coefficient (Aside (Polynomials versus polynomial functions)).

> **Exercise 1.5.8.** ROUTINE Let $\mathbf{v}, \mathbf{w} \in V$ be linearly independent. Show that $\mathbf{v} + \mathbf{w}, \mathbf{v} - \mathbf{w}$ is linearly independent if and only if $1 + 1 \neq 0$ in F.

Solution on page 1062

1.5.3 The linear dependence lemma

Theorem 1.5.5 detects dependence through a relation. The next result reads off from a relation which entry is redundant.

> *Intuition.* A dependence relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ can be solved for any entry whose coefficient is nonzero, and this writes that entry as a linear combination of the others. To control which entries appear on the other side, solve for the *last* entry with a nonzero coefficient. All later entries have coefficient 0, so only earlier entries remain.
>
> Here is a picture. Build $\text{span}(\mathcal{L})$ by adding the entries one at a time: first $\text{span}(\mathbf{v}_1)$, then $\text{span}(\mathbf{v}_1, \mathbf{v}_2)$, and so on. At each step the span either grows or stays the same, and it stays the same exactly when the new entry already lies in the span of the earlier ones (Corollary 1.3.20(c)). The lemma says that for a dependent list, some step does not make the span grow.

The proof chooses an index that is as large as possible. This is the first of many proofs built on an extreme choice.

Technique 1.5.9 (Extremal choice and minimal counterexamples). When a proof needs an object with special properties, choose one that is extreme in some respect: the largest index with a nonzero coefficient, a polynomial of largest degree, a dependence relation with the fewest nonzero coefficients, a smallest counterexample. Extremality is a hypothesis you get for free. "The largest j with $a_j \neq 0$" also says that every later coefficient is 0. "A relation with the fewest nonzero coefficients" also says that every relation with fewer nonzero coefficients is trivial.

It combines with contradiction. To show that no object has some property, suppose that some do, take an extreme one, and use it to build another object with the property that is more extreme still. This contradicts the choice.

Reach for it when a relation or a sum has many terms and the argument would be easy if most of them vanished. The canonical example is the linear dependence lemma (Lemma 1.5.10), which solves for the largest index with a nonzero coefficient. The same choice, of a largest degree, proves Proposition 1.5.24, and the contradiction form appears in Exercise 1.5.44.

Used in. Lemma 1.5.10, Proposition 1.5.24, Example 1.5.27, Exercise 1.5.44, Exercise 1.5.45, Theorem 1.6.10, Exercise 1.7.8, Exercise 1.7.9, Problem 1.4, Problem 1.15, Exercise 3.6.13, Problem 3.18, Lemma 4.1.2, Theorem 4.1.10, Proposition 4.1.21, Exercise 4.1.56, Theorem 4.4.7, Exercise 4.6.55, Theorem 4.7.34, Problem 6.13, Theorem 7.1.32, Theorem 7.1.45, Theorem 7.2.21, Proposition 7.2.35, Proposition 7.2.37, Exercise 7.2.51, Theorem 7.6.2, Theorem 7.6.7, Theorem 7.6.31, Problem 7.1, Theorem 8.2.5, Exercise 8.2.44, Exercise 8.2.46, Notation 8.3.10, Exercise A.1.30, Theorem A.2.1, Proposition A.2.4, Exercise A.2.10, Theorem A.4.4, Theorem A.4.7, Exercise A.4.8, Exercise 4.2.51, Exercise 5.5.39, Exercise 6.4.61, Exercise 7.4.53, Exercise 7.6.49, Exercise 7.6.56.

Lemma 1.5.10 (Linear dependence lemma). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a linearly dependent list in V. Then there is an index $j \in \{1, \ldots, k\}$ with

$$\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}).$$

Moreover, for every such j, deleting $\mathbf{v}_j$ from $\mathcal{L}$ does not change the span:

$$\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k) = \operatorname{span}(\mathcal{L}).$$

For $j = 1$ the condition reads $\mathbf{v}_1 \in \operatorname{span}(()) = \{\mathbf{0}\}$, that is, $\mathbf{v}_1 = \mathbf{0}$.

Proof. STANDARD

Uses. Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.32 (The scalar zero), Proposition 1.2.28 (Uniqueness of additive inverses), Proposition 1.2.34 (Negatives and the scalar -1), Corollary 1.3.20 (Operations that preserve the span)

Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)

Hint 1. By Theorem 1.5.5, $\mathcal{L}$ has a nontrivial relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. Solve it for one entry. Which choice of entry leaves only earlier entries on the other side?

Hint 2. Let j be the largest index with $a_j \neq 0$. The terms after j vanish; move the others to the other side and divide by a_j.

Hint 3. For the second statement, $\mathbf{v}_j$ lies in the span of the list with $\mathbf{v}_j$ deleted. Use Corollary 1.3.20(a), (b), and (c).

Sketch. Take a nontrivial relation and the largest j with $a_j \neq 0$. Then $a_1\mathbf{v}_1 + \cdots + a_j\mathbf{v}_j = \mathbf{0}$, so $\mathbf{v}_j = \sum_{i<j}(-a_i/a_j)\mathbf{v}_i$. If $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ and $\mathcal{L}'$ is $\mathcal{L}$ without $\mathbf{v}_j$, then $\mathbf{v}_j \in \operatorname{span}(\mathcal{L}')$, so appending $\mathbf{v}_j$ to $\mathcal{L}'$ does not change the span (Corollary 1.3.20(c)); and $\mathcal{L}', \mathbf{v}_j$ is a reordering of $\mathcal{L}$.

Since $\mathcal{L}$ is dependent, (II) of Theorem 1.5.5 fails: there are scalars $a_1, \ldots, a_k$, not all 0, with $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. Let j be the largest index with $a_j \neq 0$. For $i > j$ we have $a_i = 0$, so $a_i\mathbf{v}_i = \mathbf{0}$ by Proposition 1.2.32, and dropping these terms gives

$$(a_1\mathbf{v}_1 + \cdots + a_{j-1}\mathbf{v}_{j-1}) + a_j\mathbf{v}_j = \mathbf{0}.$$

By Proposition 1.2.28, $a_j\mathbf{v}_j$ is the additive inverse of the sum in parentheses. By Proposition 1.2.34 and Lemma 1.3.9(b), that inverse is (-1) times the sum, so

$$a_j\mathbf{v}_j = (-a_1)\mathbf{v}_1 + \cdots + \big(-a_{j-1}\big)\mathbf{v}_{j-1}.$$

Since $a_j \neq 0$, we may multiply both sides by a_j^{-1}. By (A6) and (A5) on the left, and Lemma 1.3.9(b) on the right,

$$\mathbf{v}_j = \sum_{i=1}^{j-1}\left(-\frac{a_i}{a_j}\right)\mathbf{v}_i \in \operatorname{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\big).$$

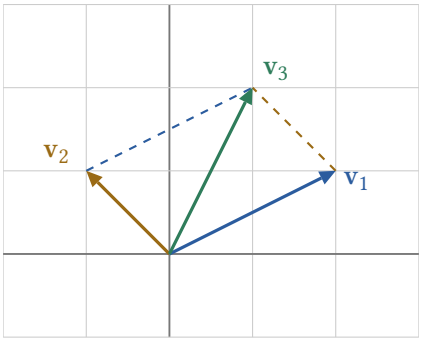
(If $j = 1$, the sum is empty and the equation says $\mathbf{v}_1 = \mathbf{0}$.)

Now let j be any index with $\mathbf{v}_j \in \operatorname{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\big)$, and let $\mathcal{L}'$ be the list obtained from $\mathcal{L}$ by deleting $\mathbf{v}_j$. The list $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ is a sublist of $\mathcal{L}'$, so $\mathbf{v}_j \in \operatorname{span}(\mathcal{L}')$ by Corollary 1.3.20(b). By Corollary 1.3.20(c), $\operatorname{span}\big(\mathcal{L}', \mathbf{v}_j\big) = \operatorname{span}(\mathcal{L}')$. The list $\mathcal{L}', \mathbf{v}_j$ is a reordering of $\mathcal{L}$, so its span is $\operatorname{span}(\mathcal{L})$ by Corollary 1.3.20(a). Hence $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$. □

The proof did not search for j. It read j off a dependence relation: any relation gives an index, the largest one with a nonzero coefficient. Different relations can give different indices, and the lemma guarantees only that *some* j works, not that a particular one does.

Example 1.5.11 (Which entries are redundant). Take $\mathcal{L}_1 = (0,0), (1,-1), (-\pi, \pi)$ in $\mathbb{R}^2$ again, and write $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ for its entries. The relation $1\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3 = \mathbf{0}$ has largest nonzero index 1, and indeed $\mathbf{v}_1 = \mathbf{0} \in \operatorname{span}(())$. The relation $0\mathbf{v}_1 + \pi\mathbf{v}_2 + 1\mathbf{v}_3 = \mathbf{0}$ gives $j = 3$, and indeed $\mathbf{v}_3 = -\pi\mathbf{v}_2$. No relation gives $j = 2$, because $(1,-1) \notin \operatorname{span}((0,0)) = \{\mathbf{0}\}$. Deleting $\mathbf{v}_1$ and then, in the remaining list $(1,-1), (-\pi, \pi)$, deleting the last entry, the lemma recovers $\operatorname{span}\big(\mathcal{L}_1\big) = \operatorname{span}((1,-1))$, the computation of Example 1.3.21.

The last entry of a dependent list need not be redundant. In $\mathbb{R}^2$, the list $(1,0), (2,0), (0,1)$ is dependent, since $2(1,0) + (-1)(2,0) + 0(0,1) = \mathbf{0}$, but $(0,1)$ does not lie in $\operatorname{span}((1,0), (2,0))$, which is the first axis. Here $j = 2$ is the only index that works.



The list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 = (2,1), (-1,1), (1,2)$ is dependent: $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. So $\mathbf{v}_3 \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$, and by Lemma 1.5.10 deleting $\mathbf{v}_3$ leaves the span unchanged. In the interactive version, the explorer reports a dependence relation and the representations of $\mathbf{w} = (0,3)$. Delete one vector and drag the other two onto a common line to see when two vectors are dependent (Proposition 1.5.15). With three vectors in $\mathbb{R}^2$, no position makes the list independent; Section 4.1 explains why. (Interactive in the web version.)

Exercise 1.5.12. ROUTINE In $\mathbb{R}^3$, let $\mathcal{L}$ be the list

$$\mathbf{v}_1, \ldots, \mathbf{v}_5 = (1,1,0), (2,2,0), (0,1,1), (1,2,1), (0,0,1).$$

Find every j with $\mathbf{v}_j \in \operatorname{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\big)$, and for each such j, a dependence relation whose largest index with a nonzero coefficient is j.

Solution on page 1063

The lemma has a converse. If some entry lies in the span of the earlier ones, solving for it backwards produces a relation in which that entry has coefficient -1. So the condition of the lemma characterizes dependence.

Corollary 1.5.13 (Independence, one entry at a time). A list $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ in V is linearly independent if and only if

$$\mathbf{v}_j \notin \operatorname{span}(\mathbf{v}_1, ..., \mathbf{v}_{j-1}) \quad \text{for every } j \in \{1, ..., k\}.$$

Proof. ROUTINE

Uses. Lemma 1.5.10 (Linear dependence lemma), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.34 (Negatives and the scalar -1), Proposition 1.2.32 (The scalar zero)
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. One direction is the contrapositive of Lemma 1.5.10.
Hint 2. For the other, suppose $\mathbf{v}_j = b_1\mathbf{v}_1 + \cdots + b_{j-1}\mathbf{v}_{j-1}$. Write down a linear relation of $\mathcal{L}$ in which $\mathbf{v}_j$ has coefficient -1.
Sketch. ($\Leftarrow$) If $\mathcal{L}$ were dependent, Lemma 1.5.10 would give a j violating the condition. ($\Rightarrow$) If $\mathbf{v}_j = \sum_{i<j} b_i\mathbf{v}_i$, then the coefficients $b_1, ..., b_{j-1}, -1, 0, ..., 0$ form a nontrivial relation, so $\mathcal{L}$ is dependent.

($\Leftarrow$) This is the contrapositive of Lemma 1.5.10: if $\mathcal{L}$ is dependent, then $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, ..., \mathbf{v}_{j-1})$ for some j.

($\Rightarrow$) We prove the contrapositive. Suppose that $\mathbf{v}_j = b_1\mathbf{v}_1 + \cdots + b_{j-1}\mathbf{v}_{j-1}$ for some j and some scalars b_i. Consider the linear combination of $\mathcal{L}$ with coefficients $b_1, ..., b_{j-1}$ on the first $j-1$ entries, -1 on $\mathbf{v}_j$, and 0 on the entries after $\mathbf{v}_j$. By Proposition 1.2.34, Proposition 1.2.32, and (A4), it represents

$$\big(b_1\mathbf{v}_1 + \cdots + b_{j-1}\mathbf{v}_{j-1}\big) + (-1)\mathbf{v}_j = \mathbf{v}_j + \big(-\mathbf{v}_j\big) = \mathbf{0}.$$

Its coefficient on $\mathbf{v}_j$ is $-1 \neq 0$, so it is a nontrivial relation, and $\mathcal{L}$ is dependent by Theorem 1.5.5. □

Condition (II) of Theorem 1.5.5 does not refer to the order of the entries, so a reordering of an independent list is independent (Proposition 1.5.18(a)). Applying Corollary 1.5.13 to a reordering in which a chosen entry comes last gives the symmetric form: $\mathcal{L}$ is linearly dependent if and only if some entry lies in the span of the other entries. In that form, the entry is the one that the intuition at the start of the section called "made up for by the others".

The corollary also says what happens when an independent list is lengthened by one vector.

Proposition 1.5.14 (Appending a vector). Let $\mathcal{L}$ be a linearly independent list in V, and let $\mathbf{w} \in V$. Then $\mathcal{L}, \mathbf{w}$ is linearly independent if and only if $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$.

Proof. ROUTINE

Uses. Corollary 1.5.13 (Independence, one entry at a time), Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations)
Hint 1. For ($\Rightarrow$), the last entry of $\mathcal{L}, \mathbf{w}$ is $\mathbf{w}$. Apply Corollary 1.5.13.
Hint 2. For ($\Leftarrow$), take a relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + b\mathbf{w} = \mathbf{0}$, and split into the cases $b \neq 0$ and $b = 0$.
Sketch. If $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, the last entry of $\mathcal{L}, \mathbf{w}$ lies in the span of the earlier ones. Conversely, in a relation $\sum a_i\mathbf{v}_i + b\mathbf{w} = \mathbf{0}$, a nonzero b would give $\mathbf{w} = \sum(-a_i/b)\mathbf{v}_i \in \operatorname{span}(\mathcal{L})$; so $b = 0$, and then all $a_i = 0$ by independence of $\mathcal{L}$.

Write $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$.

($\Rightarrow$) If $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, then the last entry of $\mathcal{L}, \mathbf{w}$ lies in the span of the entries before it, so $\mathcal{L}, \mathbf{w}$ is dependent by Corollary 1.5.13.

($\Leftarrow$) Suppose that $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and let $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + b\mathbf{w} = \mathbf{0}$. If $b \neq 0$, then, as in the proof of Lemma 1.5.10,

$$\mathbf{w} = \sum_{i=1}^{k} \Big(-\frac{a_i}{b}\Big)\mathbf{v}_i \in \operatorname{span}(\mathcal{L}),$$

a contradiction. So $b = 0$, and $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. Since $\mathcal{L}$ is independent, all $a_i = 0$ by Theorem 1.5.5. So the relation is trivial, and $\mathcal{L}, \mathbf{w}$ is independent. □

So an independent list that does not span V can always be lengthened, keeping it independent: append any vector outside its span. Repeating this is how independent lists are extended to bases in Section 4.1 (Theorem 4.1.16). Whether the process stops is a real question. In $F[x]$ it never does: $x^{n+1} \notin \operatorname{span}(1, x, \ldots, x^n)$, since every polynomial in that span has degree at most n (Example 1.3.11), so appending the next power of x never stops.

1.5.4 First consequences

Proposition 1.5.15 (Short lists, zeros, and repeats). Let V be a vector space over F.

(a) The empty list () is linearly independent.

(b) A list $\mathbf{v}$ of length one is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$.

(c) A list containing $\mathbf{0}$ is linearly dependent. A list with a repeated entry, $\mathbf{v}_i = \mathbf{v}_j$ for some $i \neq j$, is linearly dependent.

(d) A list $\mathbf{v}, \mathbf{w}$ of length two is linearly dependent if and only if one of $\mathbf{v}$, $\mathbf{w}$ is a scalar multiple of the other.

Proof. ROUTINE

Uses. Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.35 (Zero products), Proposition 1.2.32 (The scalar zero), Proposition 1.2.33 (The zero vector), Proposition 1.2.34 (Negatives and the scalar -1)
Hint 1. Use (II) of Theorem 1.5.5 throughout. For dependence, write down a nontrivial relation; for independence, start from an arbitrary relation.
Hint 2. For (b) and the forward direction of (d), the key fact is that $a\mathbf{v} = \mathbf{0}$ forces $a = 0$ or $\mathbf{v} = \mathbf{0}$ (Proposition 1.2.35).
Hint 3. In (d), given $a\mathbf{v} + b\mathbf{w} = \mathbf{0}$ with $(a, b) \neq (0, 0)$, split into the cases $a \neq 0$ and $a = 0$.
Sketch. (a) The only relation of () is trivial. (b) $1\mathbf{0} = \mathbf{0}$ is nontrivial; if $\mathbf{v} \neq \mathbf{0}$, then $a\mathbf{v} = \mathbf{0}$ forces $a = 0$. (c) $1\mathbf{0} = \mathbf{0}$ and $\mathbf{v}_i - \mathbf{v}_j = \mathbf{0}$, padded with zero coefficients. (d) $\mathbf{v} = c\mathbf{w}$ gives the relation $\mathbf{v} - c\mathbf{w} = \mathbf{0}$; conversely, from $a\mathbf{v} + b\mathbf{w} = \mathbf{0}$ with $a \neq 0$ we get $\mathbf{v} = (-b/a)\mathbf{w}$, and with $a = 0$ we get $b\mathbf{w} = \mathbf{0}$, $b \neq 0$, so $\mathbf{w} = \mathbf{0} = 0\mathbf{v}$.

We use (II) of Theorem 1.5.5.

(a) The only linear relation of () is the empty one, which is trivial.

(b) If $\mathbf{v} = \mathbf{0}$, then $1\mathbf{v} = \mathbf{0}$ by Proposition 1.2.33, a nontrivial relation. If $\mathbf{v} \neq \mathbf{0}$ and $a\mathbf{v} = \mathbf{0}$, then $a = 0$ by Proposition 1.2.35, so the only relation is trivial.

(c) If $\mathbf{v}_j = \mathbf{0}$, the linear combination with coefficient 1 on $\mathbf{v}_j$ and 0 elsewhere represents $1\mathbf{0} = \mathbf{0}$, by Proposition 1.2.33 and Proposition 1.2.32; it is a nontrivial relation. If $\mathbf{v}_i = \mathbf{v}_j$ with $i \neq j$, the linear combination with coefficient 1 on $\mathbf{v}_i$, -1 on $\mathbf{v}_j$, and 0 elsewhere represents $\mathbf{v}_i + (-\mathbf{v}_i) = \mathbf{0}$, by Proposition 1.2.34; it is nontrivial as well.

(d) If $\mathbf{v} = c\mathbf{w}$, then $1\mathbf{v} + (-c)\mathbf{w} = c\mathbf{w} + (-c\mathbf{w}) = \mathbf{0}$ is a nontrivial relation; similarly if $\mathbf{w} = c\mathbf{v}$. Conversely, suppose that $a\mathbf{v} + b\mathbf{w} = \mathbf{0}$ with a and b not both 0. If $a \neq 0$, then $\mathbf{v} = (-b/a)\mathbf{w}$, by the argument in the proof of Lemma 1.5.10. If $a = 0$, then $b \neq 0$ and $b\mathbf{w} = \mathbf{0}$, so $\mathbf{w} = \mathbf{0} = 0\mathbf{v}$ by Proposition 1.2.35 and Proposition 1.2.32. Either way, one vector is a multiple of the other. □

In (d), "one of them is a multiple of the other" cannot be strengthened to "each is a multiple of the other". If $\mathbf{v} \neq \mathbf{0}$, the list $\mathbf{v}, \mathbf{0}$ is dependent, and $\mathbf{0} = 0\mathbf{v}$, but $\mathbf{v}$ is not a multiple of $\mathbf{0}$. In $\mathbb{R}^2$, (d) says that two vectors are dependent exactly when they lie on a common line through the origin. Part (c) shows once more that independence belongs to lists: the list $\mathbf{v}, \mathbf{v}$ is dependent, although its set of entries is $\{\mathbf{v}\}$.

Exercise 1.5.16. ROUTINE Find the error in the following argument.

Claim. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$, and suppose that none of them is a scalar multiple of another. Then $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly independent.

Proof. Suppose that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0}$ and that some $a_i \neq 0$, say $a_1 \neq 0$. Then $\mathbf{v}_1 = (-a_2/a_1)\mathbf{v}_2 + (-a_3/a_1)\mathbf{v}_3$, so $\mathbf{v}_1$ is a scalar multiple of $\mathbf{v}_2$ or of $\mathbf{v}_3$, a contradiction. So all $a_i = 0$.

Solution on page 1063

Corollary 1.5.17 (Sublists). Every sublist of a linearly independent list is linearly independent. Equivalently, if some sublist of a list $\mathcal{L}$ is linearly dependent, then $\mathcal{L}$ is linearly dependent.

Proof. ROUTINE

Uses. Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.32 (The scalar zero), Notation 1.3.3
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. Use (II) of Theorem 1.5.5. A relation of the sublist uses only some of the entries of $\mathcal{L}$.
Hint 2. Extend the coefficients of a relation of the sublist by 0 on the entries of $\mathcal{L}$ that were deleted.
Sketch. A relation $\sum_r c_r\mathbf{v}_{i_r} = \mathbf{0}$ of the sublist, padded with zero coefficients on the deleted entries, is a relation of $\mathcal{L}$, so all its coefficients are 0. The second statement is the contrapositive of the first.

Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be linearly independent, and let $\mathcal{L}' = \mathbf{v}_{i_1}, ..., \mathbf{v}_{i_m}$ be a sublist, with $i_1 < \cdots < i_m$. Suppose that $c_1\mathbf{v}_{i_1} + \cdots + c_m\mathbf{v}_{i_m} = \mathbf{0}$. Define $a_i = c_r$ if $i = i_r$, and $a_i = 0$ if i is not one of $i_1, ..., i_m$. In $\sum_{i=1}^{k} a_i\mathbf{v}_i$, the terms with $i = i_r$ are $c_r\mathbf{v}_{i_r}$, in the same order as in the relation, and the other terms are $0\mathbf{v}_i = \mathbf{0}$ by Proposition 1.2.32. So

$$\sum_{i=1}^{k} a_i\mathbf{v}_i = \sum_{r=1}^{m} c_r\mathbf{v}_{i_r} = \mathbf{0}.$$

Since $\mathcal{L}$ is independent, every $a_i = 0$ by Theorem 1.5.5; in particular $c_r = a_{i_r} = 0$ for every r. So $\mathcal{L}'$ has only the trivial relation, and it is independent.

The second statement is the contrapositive of the first. □

The corollary goes in one direction only. A dependent list can have independent sublists, such as $(1, 0), (0, 1)$ inside the dependent list $(1, 0), (0, 1), (1, 1)$. The corollary also gives a second proof of Proposition 1.5.15(c): a list containing $\mathbf{0}$ has the dependent sublist $\mathbf{0}$, and a list with $\mathbf{v}_i = \mathbf{v}_j$ for $i < j$ has the dependent sublist $\mathbf{v}_i, \mathbf{v}_j$, by Proposition 1.5.15(b) and (d).

The span of a list is unchanged by the moves of Corollary 1.3.20(a) and (d). These moves preserve independence too, and they are the moves of row reduction (Section 3.3).

Proposition 1.5.18 (Operations that preserve independence). Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V, and let $\mathcal{L}'$ be obtained from $\mathcal{L}$ by one of the following operations:

(a) reordering the entries;

(b) replacing $\mathbf{v}_j$ by $c\mathbf{v}_j$, where $c \neq 0$;

(c) replacing $\mathbf{v}_j$ by $\mathbf{v}_j + a\mathbf{v}_i$, where $i \neq j$ and $a \in F$.

Then $\text{span}(\mathcal{L}') = \text{span}(\mathcal{L})$, and $\mathcal{L}'$ is linearly independent if and only if $\mathcal{L}$ is.

Proof. ROUTINE

Uses. Corollary 1.3.20 (Operations that preserve the span), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.1.17 (No zero divisors), Lemma 1.3.9 (Arithmetic of linear combinations)
Hint 1. The spans are handled by Corollary 1.3.20. For independence, it suffices to show that if $\mathcal{L}$ is independent then so is $\mathcal{L}'$: each operation can be undone by an operation of the same kind.
Hint 2. Take a relation of $\mathcal{L}'$, expand the modified entry, and collect coefficients to get a relation of $\mathcal{L}$. Then read off what independence of $\mathcal{L}$ says.

Sketch. Spans: Corollary 1.3.20(a) and (d). Each operation is undone by one of the same kind: reorder back, scale by c^{-1}, add $-a\mathbf{v}_i$. So it suffices to show that $\mathcal{L}$ independent implies $\mathcal{L}'$ independent. A relation of $\mathcal{L}'$ becomes a relation of $\mathcal{L}$ with coefficient $b_j c$ on $\mathbf{v}_j$ in (b), and with coefficients $b_i + ab_j$ on $\mathbf{v}_i$ and b_j on $\mathbf{v}_j$ in (c). Independence of $\mathcal{L}$ makes these 0, which forces every $b_l = 0$.

The spans agree by Corollary 1.3.20(a) and (d). Each operation can be undone by an operation of the same kind: a reordering by the reverse reordering, the scaling by c by the scaling by c^{-1}, and the replacement of $\mathbf{v}_j$ by $\mathbf{v}_j + a\mathbf{v}_i$ by the replacement of the new j-th entry by itself plus $(-a)\mathbf{v}_i$, which is allowed since $\mathbf{v}_i$ is still an entry of $\mathcal{L}'$ (as $i \neq j$). So it suffices to show that if $\mathcal{L}$ is independent, then so is $\mathcal{L}'$. Suppose that $\mathcal{L}$ is independent, and take a linear relation of $\mathcal{L}'$. We use (II) of Theorem 1.5.5.

(a) The coefficients of the relation, attached to the entries of $\mathcal{L}$ in their original positions, give a linear combination of $\mathcal{L}$ with the same terms in a different order. By (A1) and (A2) it also represents $\mathbf{0}$, so all its coefficients are 0.

(b) Let the relation be $\sum_{l \neq j} b_l \mathbf{v}_l + b_j(c\mathbf{v}_j) = \mathbf{0}$. By (A6), $b_j(c\mathbf{v}_j) = (b_j c)\mathbf{v}_j$, so this is the linear combination of $\mathcal{L}$ with coefficient $b_j c$ on $\mathbf{v}_j$ and b_l on $\mathbf{v}_l$ for $l \neq j$. By independence of $\mathcal{L}$, $b_l = 0$ for $l \neq j$ and $b_j c = 0$. Since $c \neq 0$, $b_j = 0$ by Proposition 1.1.17.

(c) Let the relation be $\sum_{l \neq j} b_l \mathbf{v}_l + b_j(\mathbf{v}_j + a\mathbf{v}_i) = \mathbf{0}$. By (A7), (A6), and Lemma 1.3.9, the left side is the linear combination of $\mathcal{L}$ with coefficient $b_i + b_j a$ on $\mathbf{v}_i$, b_j on $\mathbf{v}_j$, and b_l on $\mathbf{v}_l$ for $l \neq i, j$. By independence of $\mathcal{L}$, all of these are 0. So $b_j = 0$, then $b_i = b_i + b_j a = 0$, and $b_l = 0$ for the remaining l. □

Both hypotheses are needed, as they were for the span: replacing $\mathbf{v}_j$ by $0\mathbf{v}_j$, or by $\mathbf{v}_j + (-1)\mathbf{v}_j$, puts $\mathbf{0}$ into the list.

1.5.5 Changing the field

A linear relation is a statement about scalars, so whether a list is independent depends on which scalars are allowed. In Example 1.5.6 the same coordinates gave an independent list in $\mathbb{R}^2$ and a dependent one in $\mathbb{F}_3^2$, but these are different vector spaces. More striking is a single set of vectors, viewed over two fields.

Example 1.5.19 (The same vectors over two fields).

(a) $\mathbb{C}$ is a vector space over $\mathbb{C}$ and, by restricting scalars, over $\mathbb{R}$ (Example 1.2.20). Over $\mathbb{C}$, the list $1, i$ is dependent, since $i \cdot 1 + (-1) \cdot i = 0$. Over $\mathbb{R}$ it is independent: if $a \cdot 1 + b \cdot i = 0$ with $a, b \in \mathbb{R}$, then $a + bi = 0$, so $a = b = 0$ by Definition 1.1.29.

(b) Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. The list $1, \sqrt{2}$ is independent over $\mathbb{Q}$: if $a + b\sqrt{2} = 0$ with $a, b \in \mathbb{Q}$ and $b \neq 0$, then $\sqrt{2} = -a/b$ would be rational; so $b = 0$, and then $a = 0$. Over $\mathbb{R}$ the same list is dependent, since $\sqrt{2} \cdot 1 + (-1)\sqrt{2} = 0$.

(c) In $\mathbb{C}^2$, the list $(1, i), (i, -1)$ is dependent over $\mathbb{C}$, since $(i, -1) = i(1, i)$. Over $\mathbb{R}$ it is independent: for $a, b \in \mathbb{R}$,

$$a(1, i) + b(i, -1) = (a + bi, -b + ai),$$

and this is $\mathbf{0}$ only if $a + bi = 0$, that is, $a = b = 0$.

In each case the list was independent over the smaller field and dependent over the larger one. The opposite never happens.

Proposition 1.5.20 (Fewer scalars, fewer relations). Let K be a subfield of F, and regard V as a vector space over K as well (Proposition 1.2.19). If a list in V is linearly independent over F, then it is linearly independent over K.

Proof. ROUTINE

Uses. Proposition 1.2.19 (Restriction of scalars), Theorem 1.5.5 (Independence is decided at zero)

Hint 1. Compare the linear relations over K with those over F.
Hint 2. Scalar multiplication over K is that of F, restricted to K.
Sketch. A linear relation over K has coefficients in $K \subseteq F$ and is computed with the same operations, so it is a linear relation over F, hence trivial.

Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be independent over F, and let $a_1, ..., a_k \in K$ with $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. The scalar multiplication of the K-vector space V is that of F restricted to K, and the addition is the same, so $a_1, ..., a_k$ is also a linear relation of $\mathcal{L}$ over F. By Theorem 1.5.5, all $a_i = 0$. So $\mathcal{L}$ is independent over K. □

The converse fails, as Example 1.5.19 shows. In each of those examples, though, the vectors themselves involved the larger field: i, $\sqrt{2}$, and $(i, -1)$. For vectors whose coordinates lie in the smaller field, the extra scalars bring no new relations.

Intuition. Let $\mathcal{L}$ be a list of real vectors. Viewed in $\mathbb{C}^m$, it has more candidate relations, with complex coefficients. But a complex relation splits into its real and imaginary parts, and because the entries of $\mathcal{L}$ are real, each part is a real relation by itself.

Proposition 1.5.21 (Real lists in $\mathbb{C}^m$). Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in $\mathbb{R}^m$. Then $\mathcal{L}$ is linearly independent in the real vector space $\mathbb{R}^m$ if and only if it is linearly independent in the complex vector space $\mathbb{C}^m$.

Proof. STANDARD

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.1.29 (Complex numbers), Definition 1.2.4 (The space F^n)
Hint 1. One direction holds because every real relation is a complex relation.
Hint 2. For the other, write each complex coefficient as $c_j = a_j + b_j i$ with $a_j, b_j \in \mathbb{R}$, and look at one coordinate of the relation at a time.
Hint 3. In the t-th coordinate, $\sum_j c_j v_{jt} = \sum_j a_j v_{jt} + \left(\sum_j b_j v_{jt}\right)i$ with both sums real. When is a complex number 0?
Sketch. (⇐) A real relation is a complex one. (⇒) If $\sum c_j\mathbf{v}_j = \mathbf{0}$ with $c_j = a_j + b_j i$, then in each coordinate t the real and imaginary parts give $\sum a_j v_{jt} = 0$ and $\sum b_j v_{jt} = 0$, since the v_{jt} are real. So $\sum a_j\mathbf{v}_j = \sum b_j\mathbf{v}_j = \mathbf{0}$ in $\mathbb{R}^m$, and independence over $\mathbb{R}$ gives $a_j = b_j = 0$.

The operations of $\mathbb{C}^m$, applied to real vectors and real scalars, agree with those of $\mathbb{R}^m$, since $\mathbb{R}$ is a subfield of $\mathbb{C}$. Write $v_{jt} \in \mathbb{R}$ for the t-th coordinate of $\mathbf{v}_j$.

(⇐) Suppose that $\mathcal{L}$ is independent in $\mathbb{C}^m$, and let $a_1, ..., a_k \in \mathbb{R}$ with $\sum_j a_j\mathbf{v}_j = \mathbf{0}$ in $\mathbb{R}^m$. This equation also holds in $\mathbb{C}^m$, with the real numbers a_j regarded as complex scalars, so all $a_j = 0$ by Theorem 1.5.5.

(⇒) Suppose that $\mathcal{L}$ is independent in $\mathbb{R}^m$, and let $c_1, ..., c_k \in \mathbb{C}$ with $\sum_j c_j\mathbf{v}_j = \mathbf{0}$ in $\mathbb{C}^m$. Write $c_j = a_j + b_j i$ with $a_j, b_j \in \mathbb{R}$. Fix a coordinate $t \in \{1, ..., m\}$. Since each v_{jt} is real, $c_j v_{jt} = a_j v_{jt} + \left(b_j v_{jt}\right)i$, and the t-th coordinate of the relation reads

$$0 = \sum_{j=1}^{k} c_j v_{jt} = \left(\sum_{j=1}^{k} a_j v_{jt}\right) + \left(\sum_{j=1}^{k} b_j v_{jt}\right)i,$$

where both sums in parentheses are real numbers. By Definition 1.1.29, a complex number $x + yi$ with $x, y \in \mathbb{R}$ is 0 only if $x = y = 0$. So both sums vanish for every t, which says that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \quad \text{and} \quad b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k = \mathbf{0} \quad \text{in } \mathbb{R}^m.$$

Since $\mathcal{L}$ is independent in $\mathbb{R}^m$, all a_j and b_j are 0, so all $c_j = 0$. By Theorem 1.5.5, $\mathcal{L}$ is independent in $\mathbb{C}^m$. □

Exercise 1.5.22. ROUTINE Decide whether $(1, 1 + i), (1 - i, 2)$ is linearly independent in $\mathbb{C}^2$ over $\mathbb{C}$, and over $\mathbb{R}$. Does this contradict Proposition 1.5.21?

Solution on page 1063

Remark 1.5.23 (Rational and real scalars). The analogue of Proposition 1.5.21 holds for $\mathbb{Q} \subseteq \mathbb{R}$: a list in $\mathbb{Q}^m$ is independent over $\mathbb{Q}$ if and only if it is independent over $\mathbb{R}$. It holds for every subfield K of a field E

(Exercise 1.5.42). The proof above does not carry over directly. It split each complex coefficient into two real ones, which works because every complex number is a real linear combination of 1 and i. No finite list of real numbers spans $\mathbb{R}$ over $\mathbb{Q}$ (Exercise 1.3.45), so there is no such splitting of real coefficients into rational ones. The proof in Exercise 1.5.42 uses only the finitely many real numbers that occur as coefficients in one relation.

1.5.6 Polynomials and functions

In a space of functions, a linear relation is an identity that holds at *every* point. So to prove independence we may evaluate at any points we like, or compare other data that the functions determine, such as their behavior for large t. For polynomials, which are coefficient sequences rather than functions, the data to compare are the coefficients.

> *Intuition.* Among polynomials of different degrees, the one of largest degree has a leading term that the others cannot cancel. So a nontrivial combination of them has a nonzero top coefficient. Choosing the largest degree is the extremal move of Technique 1.5.9.

Proposition 1.5.24 (Distinct degrees). Let $p_1, \dots, p_k \in F[x]$ be nonzero polynomials with distinct degrees. Then $p_1, \dots, p_k$ is linearly independent.

Proof. STANDARD

> *Uses.* Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.13 (Polynomials), Definition 1.2.16 (Degree), Proposition 1.1.17 (No zero divisors)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* Suppose that a nontrivial combination $a_1p_1 + \cdots + a_kp_k$ is the zero polynomial. Among the indices with $a_i \neq 0$, choose one for which $\deg p_i$ is as large as possible.
> *Hint 2.* Compute the coefficient of the combination in position $d = \deg p_i$. Which terms can contribute to it?
> *Sketch.* Suppose $\sum a_ip_i = \mathbf{0}$ with some $a_i \neq 0$. Among those i, take the one with the largest degree d. Every other a_lp_l is $\mathbf{0}$ or has degree less than d, so the coefficient of the sum in position d is a_i times the leading coefficient of p_i, which is nonzero. Contradiction.

Suppose, to the contrary, that $a_1p_1 + \cdots + a_kp_k = \mathbf{0}$ with not all $a_i = 0$. Among the indices i with $a_i \neq 0$, choose one for which $d = \deg p_i$ is largest; since the degrees are distinct, it is unique. Let $c \neq 0$ be the leading coefficient of p_i. The operations of $F[x]$ are coefficientwise (Definition 1.2.13), so the coefficient of the sum in position d is the sum of the coefficients of the a_lp_l in position d. For $l = i$ this is a_ic. For $l \neq i$ with $a_l \neq 0$, we have $\deg p_l < d$ by the choice of i, so the coefficient of p_l in position d is 0, and so is that of a_lp_l. For l with $a_l = 0$, $a_lp_l = \mathbf{0}$. So the coefficient of $a_1p_1 + \cdots + a_kp_k$ in position d is a_ic, which is nonzero by Proposition 1.1.17. This contradicts $a_1p_1 + \cdots + a_kp_k = \mathbf{0}$. By Theorem 1.5.5, the list is independent. □

For example, $1, 1+x, 1+x+x^2$ and $1, x-1, x^2-2x+1$ are linearly independent in $F[x]$ over every field. The hypothesis matters: x^2+x, x^2-x, x has two entries of degree 2, and it is dependent, since $(x^2+x) - (x^2-x) + (-2)x = \mathbf{0}$.

Example 1.5.25 (Polynomials versus polynomial functions). Over $\mathbb{F}_2$, the list x, x^2 is linearly independent in $\mathbb{F}_2[x]$, by Proposition 1.5.24. Their polynomial functions $\mathbb{F}_2 \to \mathbb{F}_2$ are $c \mapsto c$ and $c \mapsto c^2$, and these are the same function, since $0^2 = 0$ and $1^2 = 1$. So in the vector space $\mathbb{F}_2^{\mathbb{F}_2}$ the two functions form a list with a repeated entry, which is dependent (Proposition 1.5.15(c)). The same "polynomials" are independent in one vector space and dependent in another, as Aside (Polynomials versus polynomial functions) warned.

Over an infinite field this cannot happen. If $a_1, \dots, a_k$ is a linear relation among the polynomial functions of $p_1, \dots, p_k$, then the polynomial $q = a_1p_1 + \cdots + a_kp_k$ satisfies $q(c) = 0$ for every $c \in F$ (Proposition 1.2.18(c)). A nonzero polynomial has only finitely many roots (Proposition 7.1.22), so $q = \mathbf{0}$, and $a_1, \dots, a_k$ is a linear relation among the polynomials themselves.

Example 1.5.26 (Trigonometric functions). In $\mathbb{R}^{\mathbb{R}}$, the list $\sin, \cos$ is linearly independent. Suppose that $a \sin + b \cos = \mathbf{0}$, the zero function. Then $a \sin t + b \cos t = 0$ for every $t \in \mathbb{R}$. At $t = 0$ this reads $b = 0$, and at $t = \pi/2$ it reads $a = 0$.

On the other hand, $1, \sin^2, \cos^2$ (where $\sin^2$ is the function $t \mapsto (\sin t)^2$) is dependent, since $\sin^2 t + \cos^2 t - 1 = 0$ for every t.

The method used for $\sin, \cos$ works in general. Evaluating a relation at a point t_0 gives one linear equation for the coefficients, so evaluating at as many points as there are coefficients gives a system of equations. If the only solution is 0, the functions are independent. If the system has nonzero solutions, nothing follows: a combination that vanishes at a few points need not vanish everywhere. To prove dependence, one needs an identity valid at *every* point, such as $\sin^2 + \cos^2 = 1$.

Example 1.5.27 (Exponential functions). BEYOND 21-242 In $\mathbb{R}^{\mathbb{R}}$, let $f_k(t) = e^{kt}$ for $k = 1, 2, 3$. Then f_1, f_2, f_3 is linearly independent. (The proofs use calculus.)

> *Intuition.* Evaluating at three points works, but it produces equations with entries such as e^2 and e^6 that hide what is going on. Two features of exponentials give cleaner proofs. Differentiation sends each f_k to a multiple of itself, $f_k' = k f_k$, with a different factor for each k; combining a relation with its derivative eliminates one term at a time. And as $t \to \infty$, the term with the largest exponent dominates all the others.

Proof (differentiate to eliminate terms). STANDARD

> *Uses.* Theorem 1.5.5 (Independence is decided at zero), the derivative of e^{kt} is ke^{kt}
> *Techniques.* Apply an operator to a relation to kill terms (Technique 4.1.25)
> *Hint 1.* Suppose $ae^t + be^{2t} + ce^{3t} = 0$ for all t. Differentiate this identity. Which combination of the identity and its derivative has no e^t term?
> *Hint 2.* Subtract the identity from its derivative. Repeat the idea on the result to eliminate the e^{2t} term, then evaluate at $t = 0$.
> *Sketch.* From $ae^t + be^{2t} + ce^{3t} = 0$, subtracting the identity from its derivative gives $be^{2t} + 2ce^{3t} = 0$; doing the same with factor 2 gives $2ce^{3t} = 0$. Evaluating at $t = 0$ in reverse order gives $c = 0$, $b = 0$, $a = 0$.

Suppose that $af_1 + bf_2 + cf_3 = \mathbf{0}$, that is,

$$ae^t + be^{2t} + ce^{3t} = 0 \quad \text{for all } t \in \mathbb{R}.$$

Both sides are differentiable functions of t, and the derivative of the zero function is 0, so

$$ae^t + 2be^{2t} + 3ce^{3t} = 0 \quad \text{for all } t.$$

Subtracting the first identity from the second eliminates the term with e^t:

$$be^{2t} + 2ce^{3t} = 0 \quad \text{for all } t.$$

Differentiating this gives $2be^{2t} + 6ce^{3t} = 0$, and subtracting twice the previous identity eliminates the term with e^{2t}: $2ce^{3t} = 0$ for all t. At $t = 0$ this gives $2c = 0$, so $c = 0$. Then $be^{2t} + 2ce^{3t} = 0$ at $t = 0$ gives $b = 0$, and the first identity at $t = 0$ gives $a + b + c = 0$, so $a = 0$. By Theorem 1.5.5, f_1, f_2, f_3 is independent. □

Proof (the dominant term). STANDARD

> *Uses.* Theorem 1.5.5 (Independence is decided at zero), $e^{-st} \to 0$ as $t \to \infty$, for $s > 0$
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* Suppose $ae^t + be^{2t} + ce^{3t} = 0$ for all t. Which term grows fastest as $t \to \infty$?
> *Hint 2.* Divide the identity by e^{3t} and let $t \to \infty$.
> *Sketch.* Dividing by e^{3t} gives $ae^{-2t} + be^{-t} + c = 0$ for all t; letting $t \to \infty$ gives $c = 0$. Then divide $ae^t + be^{2t} = 0$ by e^{2t} to get $b = 0$, and $ae^t = 0$ gives $a = 0$.

Suppose that $ae^t + be^{2t} + ce^{3t} = 0$ for all $t \in \mathbb{R}$. Multiplying by e^{-3t},

$$ae^{-2t} + be^{-t} + c = 0 \quad \text{for all } t.$$

As $t \to \infty$, $e^{-2t} \to 0$ and $e^{-t} \to 0$, so the left side tends to c. It is 0 for every t, so $c = 0$. Now $ae^t + be^{2t} = 0$ for all t; multiplying by e^{-2t} and letting $t \to \infty$ in the same way gives $b = 0$. Finally $ae^t = 0$ at $t = 0$ gives $a = 0$. □

The first proof uses only that differentiation multiplies each f_k by a different factor. It adapts to sequences, with a shift in place of the derivative (Exercise 1.5.44), and Section 4.1 names the move: apply an operator to a relation to eliminate terms (Technique 4.1.25). The second proof is the analytic counterpart of Proposition 1.5.24: there the largest degree dominated, here the largest exponent. Both extend to any number of distinct exponents (Exercise 1.5.45).

ASIDE: *Independent sets* BEYOND 21-242

A subset $S \subseteq V$, possibly infinite, is called linearly independent if every list of distinct elements of S is linearly independent. Distinctness is needed because a list with a repeated entry is never independent. For a finite set S, this agrees with independence of any list that runs through the elements of S once each (Proposition 1.5.18(a), Corollary 1.5.17). For example, $\{1, x, x^2, \ldots\}$ is linearly independent in $F[x]$, since a list of distinct monomials is a reordering of a sublist of some $1, x, \ldots, x^n$ (Example 1.5.7). By Exercise 1.5.45, the set of functions $t \mapsto e^{\lambda t}$, $\lambda \in \mathbb{R}$, is an uncountable independent subset of $\mathbb{R}^{\mathbb{R}}$. Independent sets and spanning sets (Aside (Spans of infinite sets)) lead to bases of infinite-dimensional spaces (Section 15.1).

Exercises

Exercise 1.5.28. ROUTINE Decide whether each list in $\mathbb{R}^3$ is linearly independent. For each dependent list, find a dependence relation and every j with $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$.

(a) $(1, 2, 3), (4, 5, 6), (7, 8, 9)$.

(b) $(1, 1, 1), (0, 1, 1), (0, 0, 1)$.

(c) $(1, -1, 0), (0, 1, -1), (-1, 0, 1)$.

Solution on page 1064

Exercise 1.5.29. ROUTINE For which primes p is $(1, 1, 1), (1, 2, 3), (1, 4, 9)$ linearly independent in $\mathbb{F}_p^3$? Is it linearly independent in $\mathbb{Q}^3$?

Solution on page 1064

Exercise 1.5.30. ROUTINE

(a) Find the error in the following argument.

Claim. If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is linearly dependent, then $\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$.

Proof. Take a nontrivial relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, and let j be the largest index with $a_j \neq 0$. Since k is the largest index, $j = k$. Solving for $\mathbf{v}_k$ as in Lemma 1.5.10 gives $\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$.

(b) Show that the claim does hold when $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ is linearly independent.

Solution on page 1064

Exercise 1.5.31. ROUTINE Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V, and let $S_j = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$ for $0 \leq j \leq k$, so that $S_0 = \{\mathbf{0}\}$. Show that $\mathcal{L}$ is linearly independent if and only if

$$S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_k,$$

that is, every entry makes the span strictly larger.

Solution on page 1065

Exercise 1.5.32. STANDARD Decide whether each list is linearly independent.

(a) In $\mathbb{R}^{\mathbb{R}}$: the functions 1, $t \mapsto \cos 2t$, and $t \mapsto \cos^2 t$.

(b) In $\mathbb{R}^{\mathbb{R}}$: the functions $t \mapsto \sin t$, $t \mapsto \sin 2t$, and $t \mapsto \sin 3t$.

(c) The functions $t \mapsto t$ and $t \mapsto |t|$, first in $\mathbb{R}^{\mathbb{R}}$ and then in $\mathbb{R}^{[0,\infty)}$.

Solution on page 1065

Exercise 1.5.33. STANDARD Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be linearly independent, let $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, and fix j. Let $\mathcal{L}'$ be $\mathcal{L}$ with $\mathbf{v}_j$ replaced by $\mathbf{w}$. Show that $\mathcal{L}'$ is linearly independent if and only if $a_j \neq 0$, and that in this case $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$.

Solution on page 1065

Exercise 1.5.34. STANDARD Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and let $\mathcal{L}, \mathcal{M}$ be lists in V. Prove or disprove each statement.

(a) If $\mathbf{u}, \mathbf{v}$ and $\mathbf{v}, \mathbf{w}$ are linearly independent, then so is $\mathbf{u}, \mathbf{w}$.

(b) If $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is linearly dependent, then $\mathbf{w} \in \operatorname{span}(\mathbf{u}, \mathbf{v})$.

(c) If $\mathcal{L}$ and $\mathcal{M}$ are linearly independent, then so is the concatenation $\mathcal{L}, \mathcal{M}$.

(d) If $\mathbf{v}_1, ..., \mathbf{v}_k$ is linearly independent, then $\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{v}_1 + \cdots + \mathbf{v}_k$ is linearly dependent.

(e) If $\mathbf{u}, \mathbf{v}$ is linearly independent and $\mathbf{w} \notin \operatorname{span}(\mathbf{u}) \cup \operatorname{span}(\mathbf{v})$, then $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is linearly independent.

Solution on page 1066

Exercise 1.5.35. STANDARD In $F[x]$, let

$$p_1 = x^2 - 3x + 2, \quad p_2 = x^2 - 2x, \quad p_3 = x^2 - x,$$

where $3 = 1 + 1 + 1$ and $2 = 1 + 1$ in F. Show that p_1, p_2, p_3 is linearly independent if and only if $1 + 1 \neq 0$ in F.

Solution on page 1066

Exercise 1.5.36. STANDARD Let $n \geq 2$, and let $\mathbf{v}_1, ..., \mathbf{v}_n$ be linearly independent in V. Let

$$\mathbf{w}_i = \mathbf{v}_i + \mathbf{v}_{i+1} \quad (1 \leq i < n), \quad \mathbf{w}_n = \mathbf{v}_n + \mathbf{v}_1.$$

Find all n and all fields F for which $\mathbf{w}_1, ..., \mathbf{w}_n$ is linearly independent.

Solution on page 1067

Exercise 1.5.37. STANDARD Let V be a vector space over $\mathbb{C}$, regarded also as a vector space over $\mathbb{R}$ (Proposition 1.2.19). Show that $\mathbf{v}_1, ..., \mathbf{v}_k$ is linearly independent over $\mathbb{C}$ if and only if

$$\mathbf{v}_1, i\mathbf{v}_1, \mathbf{v}_2, i\mathbf{v}_2, ..., \mathbf{v}_k, i\mathbf{v}_k$$

is linearly independent over $\mathbb{R}$.

Solution on page 1067

Exercise 1.5.38. STANDARD Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in $\mathbb{R}^m$. Show that $\mathcal{L}$ spans the real vector space $\mathbb{R}^m$ if and only if it spans the complex vector space $\mathbb{C}^m$.

Solution on page 1068

Exercise 1.5.39. STANDARD Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. Show that $\ln 2, \ln 3, \ln 5$ is linearly independent. (Use unique factorization of positive integers into primes.)

Solution on page 1068

Exercise 1.5.40. HARD Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. Show that $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ is linearly independent. You may use Exercise 1.3.44: $\sqrt{3} \notin \operatorname{span}\left(1, \sqrt{2}\right)$ over $\mathbb{Q}$.

Solution on page 1069

Exercise 1.5.41. HARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{Q}^m$ have integer coordinates, let p be a prime, and let $\overline{\mathbf{v}}_j \in \mathbb{F}_p^m$ be the vector whose coordinates are those of $\mathbf{v}_j$ reduced modulo p.

(a) Show that if $\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_k$ is linearly independent in $\mathbb{F}_p^m$, then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is linearly independent in $\mathbb{Q}^m$.

(b) Show that the converse of (a) is false.

Solution on page 1069

Exercise 1.5.42. HARD Let K be a subfield of a field E, for instance $\mathbb{Q} \subseteq \mathbb{R}$, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in K^m. Show that $\mathcal{L}$ is linearly independent in the K-vector space K^m if and only if it is linearly independent in the E-vector space E^m. This exercise uses Theorem 1.6.10 from Section 1.6.

Solution on page 1070

Exercise 1.5.43. HARD Let F be a finite field with q elements.

(a) Show that a list $\mathcal{L}$ of length k in V is linearly independent if and only if $\operatorname{span}(\mathcal{L})$ has exactly q^k elements.

(b) Show that the number of linearly independent lists of length k in F^n is

$$(q^n - 1)(q^n - q)\left(q^n - q^2\right)\cdots\left(q^n - q^{k-1}\right).$$

(c) How many lists of length 3 in $\mathbb{F}_2^3$ are linearly independent? How many of length 4?

Solution on page 1071

Exercise 1.5.44. HARD Let $X = \{0, 1, 2, \ldots\}$, so that F^X is the space of sequences $(x_0, x_1, x_2, \ldots)$ in F (Remark 1.2.11). For $\lambda \in F$, let $\mathbf{s}_\lambda = \left(1, \lambda, \lambda^2, \lambda^3, \ldots\right)$. Show that if $\lambda_1, \ldots, \lambda_k \in F$ are distinct, then $\mathbf{s}_{\lambda_1}, \ldots, \mathbf{s}_{\lambda_k}$ is linearly independent.

Solution on page 1071

Exercise 1.5.45. HARD BEYOND 21-242 Let $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ be distinct. Show that the functions $t \mapsto e^{\lambda_1 t}, \ldots, t \mapsto e^{\lambda_k t}$ are linearly independent in $\mathbb{R}^{\mathbb{R}}$.

Solution on page 1072

Exercise 1.5.46. HARD A town has n residents, numbered $1, \ldots, n$, and k clubs $C_1, \ldots, C_k \subseteq \{1, \ldots, n\}$. Every club has an odd number of members, and for $i \neq j$ the clubs C_i and C_j have an even number of members in common. (In particular, no club is listed twice.) For a club C, let $\mathbf{x}_C \in \mathbb{F}_2^n$ have t-th coordinate 1 if $t \in C$ and 0 otherwise. Show that $\mathbf{x}_{C_1}, \ldots, \mathbf{x}_{C_k}$ is linearly independent in $\mathbb{F}_2^n$.

Solution on page 1072

Exercise 1.5.47. HARD Let $a_1 < a_2 < \cdots < a_n$ be real numbers, and let $f_i(t) = |t - a_i|$. Show that the functions $1, t \mapsto t, f_1, \ldots, f_n$ are linearly independent in $\mathbb{R}^{\mathbb{R}}$. (No calculus is needed.)

Solution on page 1073

Exercise 1.5.48. CHALLENGE Let $\lambda \geq 1$ and $m \geq 2$ be integers, and let $A_1, ..., A_m$ be distinct subsets of $\{1, ..., n\}$ such that $|A_i \cap A_j| = \lambda$ for all $i \neq j$. For a subset A, let $\mathbf{x}_A \in \mathbb{R}^n$ have t-th coordinate 1 if $t \in A$ and 0 otherwise. Show that $\mathbf{x}_{A_1}, ..., \mathbf{x}_{A_m}$ is linearly independent in $\mathbb{R}^n$.

Solution on page 1073

1.6 Bases and coordinates

In F^n a vector is its list of entries: $\mathbf{x} = (x_1, ..., x_n)$ is the linear combination $x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, and no other linear combination of $\mathbf{e}_1, ..., \mathbf{e}_n$ represents it. A basis of a vector space V is a list that does for V what $\mathbf{e}_1, ..., \mathbf{e}_n$ does for F^n: every vector of V is represented by exactly one of its linear combinations. The coefficients of that combination are the coordinates of the vector. They name each vector of V by a column in F^n, and they turn computations in V into computations with columns. This section defines bases, shows that every spanning list contains one, and develops coordinates.

1.6.1 Bases

Throughout, V is a vector space over a field F.

Definition 1.6.1 (Basis). A list $\mathcal{B}$ in V is a **basis** of V if every vector of V is represented by exactly one linear combination of $\mathcal{B}$. The plural of basis is **bases**.

So $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ is a basis when, for every $\mathbf{v} \in V$, the equation

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{v}$$

has exactly one solution $(a_1, ..., a_n)$. "Exactly one" is two claims: at least one solution, and at most one. The first is a statement about the span of $\mathcal{B}$, and the second is the uniqueness that defines linear independence.

Proposition 1.6.2 (Bases are independent spanning lists). A list $\mathcal{B}$ in V is a basis of V if and only if $\mathcal{B}$ spans V and $\mathcal{B}$ is linearly independent.

Proof. ROUTINE

Uses. Definition 1.6.1 (Basis), Definition 1.3.12 (Spanning list), Definition 1.5.1 (Linear independence), Definition 1.3.10 (Span)

Hint 1. Split "exactly one linear combination represents $\mathbf{v}$" into "at least one" and "at most one". Which of the two properties of $\mathcal{B}$ is each of them?

Hint 2. Definition 1.5.1 asks for uniqueness only for vectors of $\operatorname{span}(\mathcal{B})$. If $\mathcal{B}$ spans V, which vectors are those?

Sketch. "At least one representation for every $\mathbf{v} \in V$" says $\operatorname{span}(\mathcal{B}) = V$. Given that, "at most one representation for every $\mathbf{v} \in V$" is the same as "exactly one for every vector of $\operatorname{span}(\mathcal{B})$", which is linear independence.

(⇒) Suppose that $\mathcal{B}$ is a basis of V. Every $\mathbf{v} \in V$ is represented by some linear combination of $\mathcal{B}$, so $V \subseteq \operatorname{span}(\mathcal{B})$. The span consists of vectors of V, so $\operatorname{span}(\mathcal{B}) = V$. Every vector of $\operatorname{span}(\mathcal{B}) = V$ is then represented by exactly one linear combination of $\mathcal{B}$, which is Definition 1.5.1.

(⇐) Suppose that $\mathcal{B}$ spans V and is linearly independent, and let $\mathbf{v} \in V$. Since $\mathbf{v} \in V = \operatorname{span}(\mathcal{B})$, some linear combination of $\mathcal{B}$ represents $\mathbf{v}$, and by independence no other one does. □

The lecture defines a basis by unique representation and records the proposition as an equivalent form. The proposition splits the work of checking a basis into two familiar tasks: show that every vector is a linear combination of $\mathcal{B}$ (spanning), and show that $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$ forces every $a_i = 0$ (independence, by Theorem 1.5.5). The definition suggests doing both at once: solve $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{v}$ for an arbitrary

v. If the computation shows that the a_i *must* be given by certain formulas, and these formulas do give a solution, then the solution exists and is unique.

Example 1.6.3 (First bases).

(a) In $\mathbb{R}^2$, consider $(1,1),(1,-1)$. For $(x,y)\in\mathbb{R}^2$, the equation $a(1,1)+b(1,-1)=(x,y)$ reads $a+b=x$ and $a-b=y$. Adding and subtracting, any solution has $a=(x+y)/2$ and $b=(x-y)/2$, and these values are a solution. So each (x,y) is represented by exactly one linear combination, and $(1,1),(1,-1)$ is a basis of $\mathbb{R}^2$. The same computation works over any field with $1+1\neq 0$. Over $\mathbb{F}_2$ the list is not a basis: there $(1,-1)=(1,1)$, and a list with a repeated entry is dependent (Proposition 1.5.15).

(b) Let $P=\{(x,y,z)\in\mathbb{R}^3 \mid x+y+z=0\}$. The list $(1,0,-1),(0,1,-1)$ spans P (Exercise 1.3.16). It is independent, because $a(1,0,-1)+b(0,1,-1)=(a,b,-a-b)$ is $\mathbf{0}$ only if $a=b=0$. So it is a basis of the plane P, a vector space in its own right (Definition 1.4.1).

(c) The zero space $\{\mathbf{0}\}$ has the empty list () as a basis: () spans $\{\mathbf{0}\}$ and is independent (Definition 1.3.10, Proposition 1.5.15). It is the only basis, since every nonempty list in $\{\mathbf{0}\}$ contains $\mathbf{0}$ and is therefore dependent.

Non-example 1.6.4 (One property without the other).

(a) In $\mathbb{R}^3$, the list $(1,0,1),(0,1,1)$ is linearly independent, because $a(1,0,1)+b(0,1,1)=(a,b,a+b)$, which is $\mathbf{0}$ only if $a=b=0$. It does not span $\mathbb{R}^3$: every vector $(a,b,a+b)$ of its span has third coordinate equal to the sum of the first two, and $(0,0,1)$ does not. So the vector $(0,0,1)$ has no representation at all.

(b) In $\mathbb{R}^2$, the list $(1,0),(0,1),(1,1)$ spans $\mathbb{R}^2$, as its sublist $(1,0),(0,1)$ already does. It is dependent, since $(1,0)+(0,1)-(1,1)=\mathbf{0}$. So every vector has many representations: $(1,1)=1(1,0)+1(0,1)+0(1,1)=0(1,0)+0(0,1)+1(1,1)$.

Whether a list is a basis depends on the field, since both spanning and independence do.

Example 1.6.5 ($\mathbb{C}$ over $\mathbb{R}$ and over $\mathbb{C}$). Regard $\mathbb{C}$ as a vector space over $\mathbb{R}$ (Example 1.2.20). The list $1,i$ spans it, since every complex number is $a+bi=a\cdot 1+b\cdot i$ with $a,b\in\mathbb{R}$, and it is independent over $\mathbb{R}$ (Example 1.5.19). So $1,i$ is a basis of $\mathbb{C}$ over $\mathbb{R}$. Over $\mathbb{C}$ the same list is dependent, as $i\cdot 1+(-1)\cdot i=0$, so it is not a basis of $\mathbb{C}$ over $\mathbb{C}$. Instead, the list 1 is: every $z\in\mathbb{C}$ is $z\cdot 1$ for exactly one complex scalar, namely z. So the same set has bases of length 2 over $\mathbb{R}$ and of length 1 over $\mathbb{C}$.

Exercise 1.6.6. ROUTINE Let $W=\{p\in\mathbb{R}[x]_{\leq 3} \mid p(1)=0\}$. Show that $x-1,x^2-1,x^3-1$ is a basis of W.

Solution on page 1074

1.6.2 Standard bases

The spaces F^n, $F^{m\times n}$, and $F[x]_{\leq n}$ each come with a basis in which the coefficients of a vector are the numbers used to write it down.

Proposition 1.6.7 (Standard bases).

(a) In F^n, the list $\mathbf{e}_1,\ldots,\mathbf{e}_n$ is a basis, where $\mathbf{e}_i$ has 1 in position i and 0 elsewhere.

(b) In $F^{m\times n}$, let U_{ij} be the matrix with 1 in position (i,j) and 0 elsewhere. The mn matrices U_{ij}, listed row by row as

$$U_{11},U_{12},\ldots,U_{1n},U_{21},\ldots,U_{2n},\ldots,U_{m1},\ldots,U_{mn},$$

form a basis.

(c) In $F[x]_{\leq n}$, the list $1,x,x^2,\ldots,x^n$ is a basis.

Proof. ROUTINE

Uses. Definition 1.6.1 (Basis), Definition 1.2.4 (The space F^n), Definition 1.2.7 (Matrices), Notation 1.2.15, Example 1.3.7 (Standard lists)
Hint 1. Use Definition 1.6.1 directly: compute the vector represented by an arbitrary linear combination, and compare with the vector you want to represent.
Hint 2. The combination $c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$ is the column $(c_1, \ldots, c_n)$. What are the analogous statements for $\sum c_{ij}U_{ij}$ and $c_0 + c_1x + \cdots + c_nx^n$?
Sketch. In each case the combination with coefficients c is the vector whose entries (or coefficients) are the c's: $\sum c_i\mathbf{e}_i = (c_1, \ldots, c_n)$, $\sum c_{ij}U_{ij}$ has (i,j) entry c_{ij}, and $\sum c_kx^k$ has k-th coefficient c_k. So the combination represents a given vector exactly when its coefficients are that vector's entries.

(a) By the coordinatewise operations, the linear combination with coefficients $c_1, \ldots, c_n$ represents $c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n = (c_1, \ldots, c_n)$ (Example 1.3.7). This equals $\mathbf{x} = (x_1, \ldots, x_n)$ if and only if $c_i = x_i$ for every i. So every $\mathbf{x} \in F^n$ is represented by exactly one linear combination.

(b) By the entrywise operations, the (k,l) entry of $\sum_{i,j} c_{ij}U_{ij}$ is $\sum_{i,j} c_{ij}\left(U_{ij}\right)_{kl} = c_{kl}$, since $\left(U_{ij}\right)_{kl}$ is 1 for $(i,j) = (k,l)$ and 0 otherwise. So the linear combination with coefficients c_{ij} represents $A = \left(a_{ij}\right)$ if and only if $c_{ij} = a_{ij}$ for all i, j.

(c) By Notation 1.2.15, $c_0 + c_1x + \cdots + c_nx^n$ is the polynomial with coefficients $(c_0, \ldots, c_n, 0, 0, \ldots)$. A polynomial $p \in F[x]_{\leq n}$ has the form $(a_0, \ldots, a_n, 0, 0, \ldots)$, and the two are equal if and only if $c_k = a_k$ for $0 \leq k \leq n$.□

Definition 1.6.8 (Standard bases). The bases of Proposition 1.6.7 are the **standard bases** of F^n, $F^{m\times n}$, and $F[x]_{\leq n}$. The vector $\mathbf{e}_i$ is the i-th **standard basis vector** of F^n, and the U_{ij} are the **matrix units** of $F^{m\times n}$.

The notation $\mathbf{e}_i$ does not record n, which is always clear from the context: $\mathbf{e}_1 = (1,0)$ in F^2 and $\mathbf{e}_1 = (1,0,0)$ in F^3. For matrix units we avoid the name E_{ij}, which Section 3.2 reserves for elementary matrices. Once matrices can be multiplied, $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Lemma 2.4.24). A column is an $n \times 1$ matrix, and the matrix units of $F^{n\times 1}$ are the $U_{i1} = \mathbf{e}_i$, so (b) contains (a).

The standard bases have lengths n, mn, and $n+1$, and these numbers will turn out to be the dimensions of the three spaces: we will see in Section 4.1 that all bases of a vector space have the same length (Theorem 4.1.4). The space F^n has many other bases, as Example 1.6.3(a) shows already for $n = 2$. The standard basis is the one in which coefficients can be read off without computation.

1.6.3 Every spanning list contains a basis

A spanning list fails to be a basis only by being dependent. The linear dependence lemma finds an entry whose deletion does not change the span, so we can delete it and still have a spanning list. Repeating this, the list gets shorter at each step, so the process stops, and it can only stop at an independent list. To turn "repeat until it stops" into a proof, we induct on the length of the list.

Technique 1.6.9 (Induction on the length of a list). Many statements are about all lists in V, or about linear combinations of every length. To prove one, induct on the length k. The base case is usually the empty list, where the claim tends to be immediate: $\operatorname{span}(()) = \{\mathbf{0}\}$, the empty list is independent, and the empty sum is $\mathbf{0}$. In the inductive step, take a list of length $k+1$, produce from it a list of length k (by deleting a well-chosen entry, or by splitting off the last term of a sum), apply the inductive hypothesis to the shorter list, and carry the conclusion back to the original list.

Two things need care. First, the inductive hypothesis must cover *every* list of length k, since the shorter list is produced by the argument and is not known in advance. Second, if the statement has hypotheses (for instance, that the list spans V), the shorter list must satisfy them too. Deleting an arbitrary entry of a spanning list can destroy spanning, which is why the deleted entry has to be chosen (Exercise 1.6.30).

Reach for it when removing one entry or one term reduces a claim to the same claim for a shorter list. The canonical example is Theorem 1.6.10, which deletes an entry supplied by Lemma 1.5.10. The arithmetic of linear combinations (Lemma 1.3.9) was proved by splitting off the last term, and so is Proposition 1.6.19(c) below. Later, the exchange lemma (Lemma 4.1.2) and Gram–Schmidt (Algorithm 6.4.14) are built the same way.

Used in. Theorem 1.6.10, Proposition 1.6.19, Exercise 1.6.30, Exercise 1.7.10, Exercise 1.7.11, Problem 1.17, Theorem 2.1.3, Lemma 4.1.2, Theorem 4.1.16, Corollary 4.1.35, Proposition 4.7.30, Exercise 4.7.61, Corollary 6.2.13, Corollary 6.3.6, Theorem 6.4.15, Proposition 7.1.22, Proposition 7.1.27, Exercise 2.6.50, Exercise 6.4.62, Exercise 6.6.19.

The lecture proves a slightly more general statement, about lists that need not span V. Any list spans its own span, so the general statement says that every list can be thinned to an independent list without losing any of its span.

Theorem 1.6.10 (Every spanning list contains a basis). Every list $\mathcal{L}$ in V has a linearly independent sublist $\mathcal{L}'$ with $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$, and such an $\mathcal{L}'$ is a basis of $\operatorname{span}(\mathcal{L})$. In particular, if $\mathcal{L}$ spans V, then some sublist of $\mathcal{L}$ is a basis of V.

Once an independent sublist $\mathcal{L}'$ with the same span is found, the rest is formal. The entries of $\mathcal{L}'$ lie in the subspace $\operatorname{span}(\mathcal{L})$ (Proposition 1.4.14), $\mathcal{L}'$ spans it, and $\mathcal{L}'$ is independent there, because linear relations are computed with the same operations in V and in the subspace. So $\mathcal{L}'$ is a basis of $\operatorname{span}(\mathcal{L})$ by Proposition 1.6.2, and if $\mathcal{L}$ spans V, then $\operatorname{span}(\mathcal{L}) = V$. We give three proofs that $\mathcal{L}'$ exists. The first is the lecture's induction; the second and third show what that induction is doing in two other forms.

Proof (induction on the length). STANDARD

Uses. Lemma 1.5.10 (Linear dependence lemma), Proposition 1.5.15 (Short lists, zeros, and repeats), Notation 1.3.3
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Induct on $k = |\mathcal{L}|$. What is the empty list's span, and is the empty list independent?
Hint 2. In the inductive step, if $\mathcal{L}$ is independent there is nothing to do. Otherwise, which entry can be deleted without changing the span?
Hint 3. Delete the entry given by Lemma 1.5.10 and apply the inductive hypothesis to the shorter list. Why is a sublist of the shorter list also a sublist of $\mathcal{L}$?
Sketch. Induct on $|\mathcal{L}|$. The empty list is independent. For a list of length $k+1$: if it is independent, take $\mathcal{L}' = \mathcal{L}$; if not, Lemma 1.5.10 gives an entry whose deletion keeps the span, and the inductive hypothesis, applied to the resulting list of length k, gives an independent sublist with the same span, which is also a sublist of $\mathcal{L}$.

We prove by induction on $k \geq 0$ that every list of length k in V has a linearly independent sublist with the same span.

Base case. If $k = 0$, then $\mathcal{L} = ()$, which is independent (Proposition 1.5.15(a)). Take $\mathcal{L}' = \mathcal{L}$.

Inductive step. Suppose that every list of length k in V has an independent sublist with the same span, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$. If $\mathcal{L}$ is independent, take $\mathcal{L}' = \mathcal{L}$. Otherwise, by Lemma 1.5.10 there is an index j such that the list

$$\mathcal{L}_j = \mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_{k+1}$$

satisfies $\operatorname{span}\left(\mathcal{L}_j\right) = \operatorname{span}(\mathcal{L})$. The list $\mathcal{L}_j$ has length k, so by the inductive hypothesis it has an independent sublist $\mathcal{L}'$ with $\operatorname{span}(\mathcal{L}') = \operatorname{span}\left(\mathcal{L}_j\right) = \operatorname{span}(\mathcal{L})$. Deleting entries from $\mathcal{L}_j$ amounts to deleting entries from $\mathcal{L}$, so $\mathcal{L}'$ is a sublist of $\mathcal{L}$ (Notation 1.3.3), as desired. □

Intuition (What the induction is doing). Unwinding the induction, the proof deletes redundant entries one at a time until none is left. The induction guarantees that this stops, because each deletion shortens the list. Another way to guarantee it is to look directly at the end result: a shortest sublist with the same span. Such a sublist exists because lengths are nonnegative integers, and it cannot be dependent,

since the dependence lemma would shorten it further. This is the extremal form of the same argument (Technique 1.5.9).

Proof (a shortest sublist). STANDARD

Uses. Lemma 1.5.10 (Linear dependence lemma), Notation 1.3.3
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Among all sublists of $\mathcal{L}$ with the same span as $\mathcal{L}$, choose one of the smallest length. Why is there at least one such sublist?
Hint 2. If the chosen sublist were dependent, what would Lemma 1.5.10 produce?
Sketch. The sublists of $\mathcal{L}$ with span $\operatorname{span}(\mathcal{L})$ include $\mathcal{L}$, so there is one, $\mathcal{L}'$, of smallest length. If $\mathcal{L}'$ were dependent, deleting the entry given by Lemma 1.5.10 would give a shorter sublist of $\mathcal{L}$ with the same span, contradicting minimality.

Let S be the set of sublists of $\mathcal{L}$ whose span is $\operatorname{span}(\mathcal{L})$. It is nonempty, since $\mathcal{L} \in S$. The lengths of the members of S are nonnegative integers, so some $\mathcal{L}' \in S$ has the smallest length. Suppose, to the contrary, that $\mathcal{L}'$ is dependent. By Lemma 1.5.10, deleting a suitable entry of $\mathcal{L}'$ gives a list $\mathcal{L}''$ with $\operatorname{span}(\mathcal{L}'') = \operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$. As a sublist of a sublist of $\mathcal{L}$, $\mathcal{L}''$ is a sublist of $\mathcal{L}$, so $\mathcal{L}'' \in S$, and it is shorter than $\mathcal{L}'$. This contradicts the choice of $\mathcal{L}'$. So $\mathcal{L}'$ is independent. □

Both proofs make a choice they cannot control: which redundant entry to delete, or which shortest sublist to take. The third proof makes no choices. It reads the list from left to right and keeps exactly the entries that are not in the span of the entries before them. This procedure, **sifting**, can be carried out by hand or by a computer, and its output is determined by $\mathcal{L}$. In Section 4.2 the entries it keeps reappear as the pivot columns of a matrix.

Proof (sifting). STANDARD

Uses. Corollary 1.5.13 (Independence, one entry at a time), Corollary 1.3.20 (Operations that preserve the span), Corollary 1.3.19 (Equal spans), Proposition 1.3.17 (A list lies in its span)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Let $\mathcal{L}'$ consist of the entries $\mathbf{v}_j$ with $\mathbf{v}_j \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$. Independence of $\mathcal{L}'$ should come from Corollary 1.5.13.
Hint 2. For each j, compare $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$ with the span of the entries of $\mathcal{L}'$ among $\mathbf{v}_1, \ldots, \mathbf{v}_j$. Prove that they are equal by induction on j.
Hint 3. In the inductive step, split into the cases where $\mathbf{v}_j$ is kept and where it is deleted. In the second case, Corollary 1.3.20(c) applies.
Sketch. Keep $\mathbf{v}_j$ exactly when $\mathbf{v}_j \notin S_{j-1} = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$. By induction on j, the kept entries among the first j span S_j: a kept entry is added to both sides, and a deleted entry lies in S_{j-1}, so $S_j = S_{j-1}$. Taking $j = k$ gives the span. Each kept entry lies outside S_{j-1}, which is the span of the earlier kept entries, so the kept list is independent by Corollary 1.5.13.

Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$, and write $S_j = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$ for $0 \le j \le k$, so $S_0 = \{\mathbf{0}\}$. Call $\mathbf{v}_j$ *kept* if $\mathbf{v}_j \notin S_{j-1}$ and *deleted* otherwise. Let $\mathcal{L}'$ be the sublist of kept entries, and for $0 \le j \le k$ let $\mathcal{K}_j$ be the sublist of kept entries among $\mathbf{v}_1, \ldots, \mathbf{v}_j$.

We show by induction on j that $\operatorname{span}(\mathcal{K}_j) = S_j$. For $j = 0$ both sides are $\{\mathbf{0}\}$. Let $j \ge 1$ and assume $\operatorname{span}(\mathcal{K}_{j-1}) = S_{j-1}$.

- If $\mathbf{v}_j$ is kept, then $\mathcal{K}_j = \mathcal{K}_{j-1}, \mathbf{v}_j$. The entries of $\mathcal{K}_j$ are entries of $\mathbf{v}_1, \ldots, \mathbf{v}_j$, so they lie in S_j (Proposition 1.3.17). Conversely, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ lie in $S_{j-1} = \operatorname{span}(\mathcal{K}_{j-1}) \subseteq \operatorname{span}(\mathcal{K}_j)$, and $\mathbf{v}_j$ is an entry of $\mathcal{K}_j$. By Corollary 1.3.19, $\operatorname{span}(\mathcal{K}_j) = S_j$.
- If $\mathbf{v}_j$ is deleted, then $\mathcal{K}_j = \mathcal{K}_{j-1}$ and $\mathbf{v}_j \in S_{j-1}$, so $S_j = S_{j-1}$ by Corollary 1.3.20(c). Hence $\operatorname{span}(\mathcal{K}_j) = \operatorname{span}(\mathcal{K}_{j-1}) = S_{j-1} = S_j$.

For $j = k$ this gives $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{K}_k) = S_k = \operatorname{span}(\mathcal{L})$.

For independence, let $\mathbf{v}_j$ be a kept entry. The entries of $\mathcal{L}'$ that come before it form $\mathcal{K}_{j-1}$, whose span is S_{j-1}, and $\mathbf{v}_j \notin S_{j-1}$ by definition. So no entry of $\mathcal{L}'$ lies in the span of the entries of $\mathcal{L}'$ before it, and $\mathcal{L}'$ is independent by Corollary 1.5.13. □

Example 1.6.11 (Sifting a list of polynomials). In $\mathbb{R}[x]_{\le 2}$, let

$$\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_5 = 1 + x, \quad 2 + 2x, \quad x - x^2, \quad 1 + x^2, \quad x^2.$$

We sift. The entry $\mathbf{v}_1 = 1 + x$ is nonzero, so it is kept. Next, $\mathbf{v}_2 = 2\mathbf{v}_1$ lies in $\operatorname{span}(\mathbf{v}_1)$ and is deleted. Every vector of $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2) = \operatorname{span}(1 + x)$ has degree at most 1, so $\mathbf{v}_3 = x - x^2$ is kept. Then $\mathbf{v}_4 = 1 + x^2 = (1 + x) - \big(x - x^2\big)$ is deleted. Finally, a linear combination

$$a(1 + x) + b\big(x - x^2\big) = a + (a + b)x - bx^2$$

equals x^2 only if $a = 0$, $b = -1$, and $a + b = 0$, which is impossible. So $\mathbf{v}_5 = x^2$ is kept, and sifting produces

$$\mathcal{L}' = 1 + x, \quad x - x^2, \quad x^2,$$

a basis of $\operatorname{span}(\mathcal{L})$. Here $\operatorname{span}(\mathcal{L})$ is all of $\mathbb{R}[x]_{\leq 2}$, because x^2, $x = \big(x - x^2\big) + x^2$, and $1 = (1 + x) - x$ lie in $\operatorname{span}(\mathcal{L}')$ (Proposition 1.3.18). So $\mathcal{L}'$ is a basis of $\mathbb{R}[x]_{\leq 2}$.

Other sublists of $\mathcal{L}$ are bases too. For instance, $2 + 2x, 1 + x^2, x^2$ is independent and spans $\mathbb{R}[x]_{\leq 2}$, as one checks in the same way. So a spanning list usually contains several bases, and sifting picks a particular one, determined by the order of the entries: it keeps an entry exactly when the entries before it do not already account for it.

Exercise 1.6.12. ROUTINE Let $\mathcal{L} = (1, 0), (1, 1), (2, 2), (0, 1)$ in $\mathbb{R}^2$. Find all sublists of $\mathcal{L}$ that are bases of $\mathbb{R}^2$.

Solution on page 1074

Minimal spanning lists and maximal independent lists

The first proof deletes entries until no entry can be deleted without shrinking the span. The next result says that this stopping condition characterizes bases, and that so does the mirror-image condition for independent lists, to which no vector can be appended.

Proposition 1.6.13 (Bases are minimal spanning lists and maximal independent lists). Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a list in V. The following are equivalent.

(i) $\mathcal{B}$ is a basis of V.
(ii) $\mathcal{B}$ spans V, and for every j, the list obtained by deleting $\mathbf{v}_j$ from $\mathcal{B}$ does not span V.
(iii) $\mathcal{B}$ is linearly independent, and for every $\mathbf{w} \in V$, the list $\mathcal{B}, \mathbf{w}$ is linearly dependent.

Proof. STANDARD

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Lemma 1.5.10 (Linear dependence lemma), Proposition 1.5.14 (Appending a vector), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Prove (i) ⇔ (ii) and (i) ⇔ (iii). In each, one half of the definition of a basis is common to both statements.
Hint 2. For (ii) ⇒ (i), suppose $\mathcal{B}$ were dependent and apply Lemma 1.5.10. For (i) ⇒ (ii), if deleting $\mathbf{v}_j$ leaves a spanning list, write $\mathbf{v}_j$ as a combination of the others.
Hint 3. For (iii), use Proposition 1.5.14: for independent $\mathcal{B}$, the list $\mathcal{B}, \mathbf{w}$ is independent exactly when $\mathbf{w} \notin \operatorname{span}(\mathcal{B})$.
Sketch. (i) ⇒ (ii): if the list without $\mathbf{v}_j$ spans V, then $\mathbf{v}_j = \sum_{i \neq j} b_i \mathbf{v}_i$, and the coefficients b_i and -1 form a dependence relation of $\mathcal{B}$. (ii) ⇒ (i): if $\mathcal{B}$ were dependent, Lemma 1.5.10 would delete an entry without changing the span V. (i) ⇔ (iii): for independent $\mathcal{B}$, by Proposition 1.5.14, $\mathcal{B}, \mathbf{w}$ is dependent for every $\mathbf{w}$ exactly when every $\mathbf{w}$ lies in $\operatorname{span}(\mathcal{B})$, that is, when $\mathcal{B}$ spans V.

Throughout we use Proposition 1.6.2: (i) says that $\mathcal{B}$ spans V and is independent. Write $\mathcal{B}_j$ for $\mathcal{B}$ with $\mathbf{v}_j$ deleted.

(i) ⇒ (ii). Suppose that $\mathcal{B}$ is a basis, and suppose that $\mathcal{B}_j$ spans V for some j. Then $\mathbf{v}_j \in \operatorname{span}\big(\mathcal{B}_j\big)$, say $\mathbf{v}_j = \sum_{i \neq j} b_i \mathbf{v}_i$. The linear combination of $\mathcal{B}$ with coefficient b_i on $\mathbf{v}_i$ for $i \neq j$ and -1 on $\mathbf{v}_j$ represents $\mathbf{v}_j - \mathbf{v}_j = \mathbf{0}$, and it is nontrivial. This contradicts the independence of $\mathcal{B}$ (Theorem 1.5.5). So no $\mathcal{B}_j$ spans V.

(ii) ⇒ (i). Suppose (ii). Then $\mathcal{B}$ spans V, and it remains to show that $\mathcal{B}$ is independent. If it were dependent, Lemma 1.5.10 would give an index j with $\operatorname{span}\big(\mathcal{B}_j\big) = \operatorname{span}(\mathcal{B}) = V$, contradicting (ii).

(i) ⇔ (iii). Both statements include the independence of $\mathcal{B}$, so assume it. By Proposition 1.5.14, for $\mathbf{w} \in V$ the list $\mathcal{B}, \mathbf{w}$ is dependent if and only if $\mathbf{w} \in \operatorname{span}(\mathcal{B})$. So $\mathcal{B}, \mathbf{w}$ is dependent for every $\mathbf{w} \in V$ if and only if $V \subseteq \operatorname{span}(\mathcal{B})$, that is, if and only if $\mathcal{B}$ spans V. □

In (ii), deleting one entry suffices: if some sublist of $\mathcal{B}$ shorter than $\mathcal{B}$ spans V, then so does any $\mathcal{B}_j$ containing it, since enlarging a list can only enlarge its span (Corollary 1.3.20(b)). So (ii) says that $\mathcal{B}$ is a spanning list from which nothing can be removed, and (iii) says that $\mathcal{B}$ is an independent list to which nothing can be added. These are statements about removing and adding entries, not about lengths. In Section 4.1 they become statements about lengths: every independent list is at most as long as every spanning list (Lemma 4.1.2).

1.6.4 Coordinates

Choosing a basis of V amounts to choosing a coordinate system for V. The lecture marks the next definition as important, and it is the reason bases are worth having: every vector of V gets a name in F^n, and the name determines the vector.

Definition 1.6.14 (Coordinates). Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V, and let $\mathbf{v} \in V$. Let $a_1, ..., a_n$ be the coefficients of the unique linear combination of $\mathcal{B}$ that represents $\mathbf{v}$, so that $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. The **coordinate vector** of $\mathbf{v}$ with respect to $\mathcal{B}$, or the **$\mathcal{B}$-coordinates** of $\mathbf{v}$, is

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n,$$

and a_i is the **i-th coordinate** of $\mathbf{v}$ with respect to $\mathcal{B}$. The function $V \to F^n$, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$, is the **coordinate map** of $\mathcal{B}$.

For $n = 0$, that is, for the basis () of $\{\mathbf{0}\}$, we use the convention that $F^0 = \{\mathbf{0}\}$ consists of the empty column, so that $[\mathbf{0}]_{()} = \mathbf{0}$. With it, everything below holds for $n = 0$ as well.

The definition makes sense because $\mathcal{B}$ is a basis: some linear combination represents $\mathbf{v}$ since $\mathcal{B}$ spans, and only one does since $\mathcal{B}$ is independent. Read in reverse, it says that for every column $(a_1, ..., a_n) \in F^n$,

$$[\mathbf{v}]_{\mathcal{B}} = (a_1, ..., a_n) \quad \text{if and only if} \quad \mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n. \tag{1.6.1}$$

The forward direction is the definition. For the backward direction, if the combination with coefficients $a_1, ..., a_n$ represents $\mathbf{v}$, then it is *the* combination representing $\mathbf{v}$, by uniqueness. This is how coordinates are computed in practice: to find $[\mathbf{v}]_{\mathcal{B}}$, find any linear combination of $\mathcal{B}$ that represents $\mathbf{v}$. For instance, $[\mathbf{v}_j]_{\mathcal{B}} = \mathbf{e}_j$, since $\mathbf{v}_j$ is the combination with coefficient 1 on $\mathbf{v}_j$ and 0 elsewhere, and $[\mathbf{0}]_{\mathcal{B}} = \mathbf{0}$.

In the standard bases, coordinates are the familiar numbers. By the proof of Proposition 1.6.7, the coordinate vector of $\mathbf{x} \in F^n$ in the standard basis is $\mathbf{x}$ itself. That of $A \in F^{m \times n}$ is

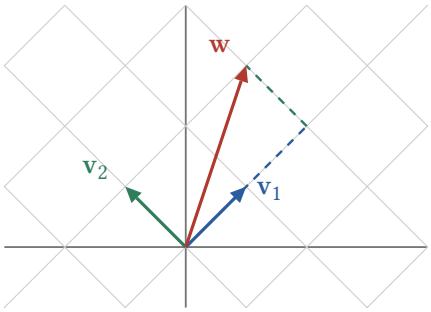
$$(a_{11}, a_{12}, ..., a_{1n}, a_{21}, ..., a_{mn}) \in F^{mn},$$

the entries of A read row by row. That of $a_0 + a_1x + \cdots + a_nx^n$ is $(a_0, a_1, ..., a_n) \in F^{n+1}$, the list of its coefficients; note the shift, as the i-th coordinate of a polynomial is its coefficient a_{i-1}. For other bases, coordinates take some work to compute, and they carry information that the entries do not show directly.

Example 1.6.15 (Coordinates in the plane). In $\mathbb{R}^2$, let $\mathcal{B} = \mathbf{v}_1, \mathbf{v}_2 = (1, 1), (-1, 1)$. Solving $a(1, 1) + b(-1, 1) = (x, y)$, that is, $a - b = x$ and $a + b = y$, gives the unique solution $a = (x + y)/2$, $b = (y - x)/2$. So $\mathcal{B}$ is a basis, and

$$[(x, y)]_{\mathcal{B}} = \begin{pmatrix} (x + y)/2 \\ (y - x)/2 \end{pmatrix}.$$

For $\mathbf{w} = (1, 3)$ this gives $[\mathbf{w}]_{\mathcal{B}} = (2, 1)$: indeed $\mathbf{w} = 2\mathbf{v}_1 + \mathbf{v}_2$. The standard coordinates $(1, 3)$ say how to reach $\mathbf{w}$ along the horizontal and vertical grid lines: one step right, three steps up. The $\mathcal{B}$-coordinates say how to reach it along the tilted grid of lines parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$: two steps along $\mathbf{v}_1$, then one along $\mathbf{v}_2$.



The basis $\mathbf{v}_1, \mathbf{v}_2 = (1, 1), (-1, 1)$ and $\mathbf{w} = (1, 3) = 2\mathbf{v}_1 + \mathbf{v}_2$, so $[\mathbf{w}]_{\mathcal{B}} = (2, 1)$. In the interactive version, the coefficients that the explorer reports for $\mathbf{w}$ are its $\mathcal{B}$-coordinates; drag $\mathbf{w}$ to see them change. Drag $\mathbf{v}_2$ onto the line through $\mathbf{v}_1$: the list becomes dependent, $\mathbf{w}$ gets many representations or none, and coordinates no longer make sense. (Interactive in the web version.)

The order of the entries of $\mathcal{B}$ matters. The list $\mathbf{v}_2, \mathbf{v}_1$ is also a basis, and in it the coordinates of $\mathbf{w}$ are $(1, 2)$. This is one reason bases are lists rather than sets.

Coordinates name every vector of V, and different vectors get different names. The lecture records this as an observation.

Proposition 1.6.16 (The coordinate map is a bijection). Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V. The coordinate map $V \to F^n$, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$, is a bijection. Its inverse is

$$F^n \to V, \quad (a_1, \ldots, a_n) \mapsto a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n.$$

Proof. ROUTINE

Uses. Definition 1.6.14 (Coordinates), (1.6.1)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Call the proposed inverse Φ. A function with a two-sided inverse is a bijection. Check $\Phi([\mathbf{v}]_{\mathcal{B}}) = \mathbf{v}$ and $[\Phi(\mathbf{a})]_{\mathcal{B}} = \mathbf{a}$.
Hint 2. The first identity is the definition of $[\mathbf{v}]_{\mathcal{B}}$. For the second, which linear combination of $\mathcal{B}$ represents $\Phi(\mathbf{a})$?
Sketch. Let $\Phi(a_1, \ldots, a_n) = \sum a_i\mathbf{v}_i$. By definition, $[\mathbf{v}]_{\mathcal{B}}$ lists coefficients representing $\mathbf{v}$, so $\Phi([\mathbf{v}]_{\mathcal{B}}) = \mathbf{v}$. The combination with coefficients a_i represents $\Phi(\mathbf{a})$, so by uniqueness $[\Phi(\mathbf{a})]_{\mathcal{B}} = \mathbf{a}$ ((1.6.1)). So the two maps are inverse to each other.

Let $\Phi : F^n \to V$ be the map $(a_1, \ldots, a_n) \mapsto a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. For $\mathbf{v} \in V$ with $[\mathbf{v}]_{\mathcal{B}} = (a_1, \ldots, a_n)$, we have $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ by Definition 1.6.14, that is, $\Phi([\mathbf{v}]_{\mathcal{B}}) = \mathbf{v}$. For $\mathbf{a} = (a_1, \ldots, a_n) \in F^n$, the combination with coefficients $a_1, \ldots, a_n$ represents $\Phi(\mathbf{a})$, so $[\Phi(\mathbf{a})]_{\mathcal{B}} = \mathbf{a}$ by (1.6.1). So Φ is a two-sided inverse of the coordinate map, and both maps are bijections. □

For a fixed basis the correspondence is perfect, but it depends heavily on the basis, as the lecture's example shows.

Example 1.6.17 (One polynomial, two bases). In $\mathbb{R}[x]_{\leq 2}$, let $\mathcal{B} = 1, x, x^2$ be the standard basis and let $\mathcal{C} = 1, 1 + x, 1 + x + x^2$. To see that $\mathcal{C}$ is a basis, solve

$$a \cdot 1 + b(1 + x) + c\big(1 + x + x^2\big) = (a + b + c) + (b + c)x + cx^2 = p_0 + p_1x + p_2x^2.$$

Comparing coefficients, $c = p_2$, then $b = p_1 - p_2$, then $a = p_0 - p_1$, and these values do solve the equation. So every polynomial of degree at most 2 is represented by exactly one linear combination of $\mathcal{C}$, and

$$[p_0 + p_1x + p_2x^2]_{\mathcal{C}} = \begin{pmatrix} p_0 - p_1 \\ p_1 - p_2 \\ p_2 \end{pmatrix}, \quad \text{while} \quad [p_0 + p_1x + p_2x^2]_{\mathcal{B}} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix}.$$

For $p = x + x^2$,

$$[x + x^2]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad [x + x^2]_{\mathcal{C}} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

and indeed $x + x^2 = -1 + 0(1 + x) + \left(1 + x + x^2\right)$. The same vector has different coordinates in different bases, and the column $(0, 1, 1)$ names $x + x^2$ in $\mathcal{B}$ but names $(1 + x) + \left(1 + x + x^2\right) = 2 + 2x + x^2$ in $\mathcal{C}$. How $[\mathbf{v}]_{\mathcal{B}}$ and $[\mathbf{v}]_{\mathcal{C}}$ are related for two bases is the subject of Section 4.5.

Over $\mathbb{R}$, the basis $1, i$ of $\mathbb{C}$ (Example 1.6.5) gives $[a + bi]_{(1,i)} = (a, b)$. This is the identification of $\mathbb{C}$ with the plane $\mathbb{R}^2$, and Proposition 1.6.16 makes it a bijection.

Exercise 1.6.18. ROUTINE Let $\mathcal{T} = 1, x - 1, x^2 - 2x + 1$ in $F[x]_{\leq 2}$.

(a) Show that $\mathcal{T}$ is a basis of $F[x]_{\leq 2}$.

(b) Show that $\left[a + bx + cx^2\right]_{\mathcal{T}} = (a + b + c, b + 2c, c)$.

(c) Explain why the first coordinate of $[p]_{\mathcal{T}}$ is $p(1)$ for every $p \in F[x]_{\leq 2}$.

Solution on page 1075

Coordinates respect linear combinations

The coordinate map is a bijection, and it is compatible with the vector space operations: adding in V and then taking coordinates gives the same column as taking coordinates and then adding in F^n. This is what makes coordinates useful for computation.

Proposition 1.6.19 (Coordinates respect linear combinations). Let $\mathcal{B}$ be a basis of V of length n.

(a) For all $\mathbf{x}, \mathbf{y} \in V$, $[\mathbf{x} + \mathbf{y}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{B}} + [\mathbf{y}]_{\mathcal{B}}$.

(b) For all $c \in F$ and $\mathbf{x} \in V$, $[c\mathbf{x}]_{\mathcal{B}} = c[\mathbf{x}]_{\mathcal{B}}$.

(c) For all $k \geq 0$, all $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$, and all $c_1, \ldots, c_k \in F$,

$$[c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k]_{\mathcal{B}} = c_1[\mathbf{x}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{x}_k]_{\mathcal{B}}.$$

On each left side the operations are those of V, and on each right side those of F^n.

Proof. ROUTINE

Uses. Definition 1.6.14 (Coordinates), (1.6.1), Lemma 1.3.9 (Arithmetic of linear combinations), Definition 1.2.4 (The space F^n), (1.3.1)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. By (1.6.1), to show $[\mathbf{w}]_{\mathcal{B}} = (d_1, \ldots, d_n)$ it suffices to exhibit $\mathbf{w}$ as $\sum d_i\mathbf{v}_i$.
Hint 2. For (a), write $\mathbf{x} = \sum a_i\mathbf{v}_i$ and $\mathbf{y} = \sum b_i\mathbf{v}_i$, and combine the two sums with Lemma 1.3.9. Do (b) the same way.
Hint 3. For (c), induct on k. Split off the last term $c_k\mathbf{x}_k$ and apply (a), (b), and the inductive hypothesis.
Sketch. If $[\mathbf{x}]_{\mathcal{B}} = (a_i)$ and $[\mathbf{y}]_{\mathcal{B}} = (b_i)$, then $\mathbf{x} + \mathbf{y} = \sum(a_i + b_i)\mathbf{v}_i$ and $c\mathbf{x} = \sum(ca_i)\mathbf{v}_i$ by Lemma 1.3.9, so by (1.6.1) the coordinates are $(a_i + b_i)$ and (ca_i), which are $[\mathbf{x}]_{\mathcal{B}} + [\mathbf{y}]_{\mathcal{B}}$ and $c[\mathbf{x}]_{\mathcal{B}}$. For (c), induct on k: the empty sum is $\mathbf{0}$ with coordinates $\mathbf{0}$, and $[\mathbf{u} + c_k\mathbf{x}_k]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} + c_k[\mathbf{x}_k]_{\mathcal{B}}$ by (a) and (b).

Write $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$.

(a) Let $[\mathbf{x}]_{\mathcal{B}} = (a_1, \ldots, a_n)$ and $[\mathbf{y}]_{\mathcal{B}} = (b_1, \ldots, b_n)$, so that $\mathbf{x} = \sum_i a_i\mathbf{v}_i$ and $\mathbf{y} = \sum_i b_i\mathbf{v}_i$. By Lemma 1.3.9(a),

$$\mathbf{x}+\mathbf{y}=\sum_{i=1}^{n}a_i\mathbf{v}_i+\sum_{i=1}^{n}b_i\mathbf{v}_i=\sum_{i=1}^{n}(a_i+b_i)\mathbf{v}_i.$$

So the combination with coefficients $a_1+b_1,\dots,a_n+b_n$ represents $\mathbf{x}+\mathbf{y}$, and by (1.6.1),

$$[\mathbf{x}+\mathbf{y}]_{\mathcal{B}}=\begin{pmatrix}a_1+b_1\\ \vdots\\ a_n+b_n\end{pmatrix}=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}+\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}=[\mathbf{x}]_{\mathcal{B}}+[\mathbf{y}]_{\mathcal{B}},$$

where the middle equality is the definition of addition in F^n (Definition 1.2.4).

(b) With $[\mathbf{x}]_{\mathcal{B}}=(a_1,\dots,a_n)$ as before, Lemma 1.3.9(b) gives $c\mathbf{x}=\sum_i(ca_i)\mathbf{v}_i$. So $[c\mathbf{x}]_{\mathcal{B}}=(ca_1,\dots,ca_n)=c[\mathbf{x}]_{\mathcal{B}}$, by (1.6.1) and the definition of scalar multiplication in F^n.

(c) We induct on k. For $k=0$ both sides are empty sums: the left side is $[\mathbf{0}]_{\mathcal{B}}=\mathbf{0}$, and the right side is $\mathbf{0}\in F^n$. Suppose that (c) holds for combinations of length k, and let $\mathbf{u}=c_1\mathbf{x}_1+\cdots+c_k\mathbf{x}_k$. By (1.3.1), then (a), then (b) and the inductive hypothesis,

$$\begin{aligned}[c_1\mathbf{x}_1+\cdots+c_{k+1}\mathbf{x}_{k+1}]_{\mathcal{B}}&=[\mathbf{u}+c_{k+1}\mathbf{x}_{k+1}]_{\mathcal{B}}=[\mathbf{u}]_{\mathcal{B}}+[c_{k+1}\mathbf{x}_{k+1}]_{\mathcal{B}}\\&=\big(c_1[\mathbf{x}_1]_{\mathcal{B}}+\cdots+c_k[\mathbf{x}_k]_{\mathcal{B}}\big)+c_{k+1}[\mathbf{x}_{k+1}]_{\mathcal{B}},\end{aligned}$$

as desired. □

The lecture calls (c) the "Mega-Prop": the coordinate map *respects linear combinations*. Functions between vector spaces with this property are the linear maps of Section 2.1 (Definition 2.1.1), and the induction in (c) shows in general that (a) and (b) are enough to check (Theorem 2.1.3). So the coordinate map of a basis is a linear map, and by Proposition 1.6.16 it is a bijection. Such a map is an isomorphism (Section 4.5): once a basis of length n is chosen, V and F^n are the same vector space with different names for the vectors.

The practical consequence is that every question about linear combinations in V can be asked in F^n instead.

Proposition 1.6.20 (Coordinates translate questions about V into F^n). Let $\mathcal{B}$ be a basis of V of length n, and let $\mathbf{x}_1,\dots,\mathbf{x}_k,\mathbf{w}\in V$. For all $c_1,\dots,c_k\in F$,

$$c_1\mathbf{x}_1+\cdots+c_k\mathbf{x}_k=\mathbf{w}\quad\text{ if and only if }\quad c_1[\mathbf{x}_1]_{\mathcal{B}}+\cdots+c_k[\mathbf{x}_k]_{\mathcal{B}}=[\mathbf{w}]_{\mathcal{B}}.$$

Consequently:

(a) $\mathbf{x}_1,\dots,\mathbf{x}_k$ is linearly independent in V if and only if $[\mathbf{x}_1]_{\mathcal{B}},\dots,[\mathbf{x}_k]_{\mathcal{B}}$ is linearly independent in F^n;

(b) $\mathbf{w}\in\operatorname{span}(\mathbf{x}_1,\dots,\mathbf{x}_k)$ if and only if $[\mathbf{w}]_{\mathcal{B}}\in\operatorname{span}\big([\mathbf{x}_1]_{\mathcal{B}},\dots,[\mathbf{x}_k]_{\mathcal{B}}\big)$;

(c) $\mathbf{x}_1,\dots,\mathbf{x}_k$ spans V if and only if $[\mathbf{x}_1]_{\mathcal{B}},\dots,[\mathbf{x}_k]_{\mathcal{B}}$ spans F^n. In particular, $\mathbf{x}_1,\dots,\mathbf{x}_k$ is a basis of V if and only if $[\mathbf{x}_1]_{\mathcal{B}},\dots,[\mathbf{x}_k]_{\mathcal{B}}$ is a basis of F^n.

Proof. ROUTINE

Uses. Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 1.6.16 (The coordinate map is a bijection), Theorem 1.5.5 (Independence is decided at zero), Definition 1.3.10 (Span), Proposition 1.6.2 (Bases are independent spanning lists)

Hint 1. For the displayed equivalence, apply the coordinate map to the left equation, and use that it is injective for the converse.

Hint 2. Then (a) is the case $\mathbf{w}=\mathbf{0}$, and (b) is a restatement. For (c), every vector of F^n is $[\mathbf{w}]_{\mathcal{B}}$ for some $\mathbf{w}\in V$.

Sketch. By Proposition 1.6.19(c), $[\sum c_i\mathbf{x}_i]_{\mathcal{B}}=\sum c_i[\mathbf{x}_i]_{\mathcal{B}}$, and the coordinate map is injective, so $\sum c_i\mathbf{x}_i=\mathbf{w}$ exactly when the coordinates agree. So the two lists have the same linear relations (take $\mathbf{w}=\mathbf{0}$), and $\mathbf{w}$ is in one span exactly when $[\mathbf{w}]_{\mathcal{B}}$ is in the other. Since the coordinate map is surjective, (b) for all $\mathbf{w}$ gives (c).

By Proposition 1.6.19(c), $[c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k]_{\mathcal{B}} = c_1[\mathbf{x}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{x}_k]_{\mathcal{B}}$. The coordinate map is injective (Proposition 1.6.16), so two vectors of V are equal if and only if their coordinate vectors are. Applying this to $\sum_i c_i\mathbf{x}_i$ and $\mathbf{w}$ proves the displayed equivalence.

(a) Since $[\mathbf{0}]_{\mathcal{B}} = \mathbf{0}$, the case $\mathbf{w} = \mathbf{0}$ says that the two lists have exactly the same linear relations. So one has only the trivial relation if and only if the other does, and (a) follows from Theorem 1.5.5.

(b) By the displayed equivalence, some $c_1, \ldots, c_k$ satisfy the left equation if and only if they satisfy the right one.

(c) Suppose that the $\mathbf{x}_i$ span V, and let $\mathbf{y} \in F^n$. The coordinate map is surjective, so $\mathbf{y} = [\mathbf{w}]_{\mathcal{B}}$ for some $\mathbf{w} \in V$. Since $\mathbf{w} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$, (b) gives $\mathbf{y} \in \operatorname{span}\big([\mathbf{x}_1]_{\mathcal{B}}, \ldots, [\mathbf{x}_k]_{\mathcal{B}}\big)$. Conversely, if the coordinate vectors span F^n, then for every $\mathbf{w} \in V$, $[\mathbf{w}]_{\mathcal{B}}$ lies in their span, so $\mathbf{w} \in \operatorname{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ by (b). The last statement follows from (a), (c), and Proposition 1.6.2. □

Example 1.6.21 (Testing polynomials through their coordinates). In $F[x]_{\leq 2}$, consider $1 + x + x^2$, $1 + 2x + 3x^2$, $1 + 4x + 9x^2$. Their coordinate vectors in the standard basis are $(1,1,1), (1,2,3), (1,4,9)$. In Exercise 1.5.29 we reduced a linear relation among these columns to $2c = 0$, using steps that are valid over every field, so the columns are independent when $1 + 1 \neq 0$. When $1 + 1 = 0$, the last two columns are both $(1,0,1)$, so they are dependent. By Proposition 1.6.20(a), the three polynomials are linearly independent if and only if $1 + 1 \neq 0$ in F. Nothing about polynomials was used beyond their coordinates. In the same way, questions about matrices in $F^{m\times n}$ become questions about columns in F^{mn}, and Section 3.1 solves such questions systematically.

> **Exercise 1.6.22.** ROUTINE Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_n$ be bases of V such that $[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{C}}$ for every $\mathbf{v} \in V$. Show that $\mathcal{B} = \mathcal{C}$.

Solution on page 1075

1.6.5 Finite and infinite dimension

Theorem 1.6.10 produces a basis from any spanning list. It applies exactly to the spaces that have a spanning list.

> **Definition 1.6.23** (Finite-dimensional). A vector space V is **finite-dimensional** if some list in V spans V. Otherwise V is **infinite-dimensional**.

The name anticipates Section 4.1, where the dimension of a finite-dimensional space is defined as the common length of its bases (Definition 4.1.5). Every space in Section 1.6.2 is finite-dimensional, as are $\{\mathbf{0}\}$ (spanned by ()), $\mathbb{C}$ over $\mathbb{R}$, and the span of any list.

> **Corollary 1.6.24** (Bases exist in finite dimensions). Every finite-dimensional vector space has a basis. More precisely, every spanning list of V contains a basis of V as a sublist.

Proof. ROUTINE

> *Uses.* Definition 1.6.23 (Finite-dimensional), Theorem 1.6.10 (Every spanning list contains a basis)
> *Hint 1.* What does the definition of finite-dimensional provide?
> *Hint 2.* Apply Theorem 1.6.10 to it.
> *Sketch.* A finite-dimensional space has a spanning list, and by Theorem 1.6.10 that list has a sublist which is a basis.

If V is finite-dimensional, some list $\mathcal{L}$ spans V. By Theorem 1.6.10, some sublist of $\mathcal{L}$ is a basis of V. The same argument applies to every spanning list of V. □

Not every vector space is finite-dimensional. The basic example is $F[x]$, and the reason is the degree.

Proposition 1.6.25 (No finite list spans $F[x]$). For every list $p_1, ..., p_k$ in $F[x]$, $\operatorname{span}(p_1, ..., p_k) \neq F[x]$. So $F[x]$ is infinite-dimensional.

Proof. ROUTINE

Uses. Example 1.4.9 (Polynomials), Proposition 1.4.14 (The span is the smallest subspace containing a list), Definition 1.2.16 (Degree)
Hint 1. Every linear combination of $p_1, ..., p_k$ has degree at most the largest of the degrees $\deg p_i$. Find a polynomial of larger degree.
Hint 2. Let d be the largest of $0, \deg p_1, ..., \deg p_k$. Then every p_i lies in the subspace $F[x]_{\leq d}$. What does Proposition 1.4.14 say about the span?
Sketch. With $d = \max(0, \deg p_1, ..., \deg p_k)$, all p_i lie in the subspace $F[x]_{\leq d}$, hence so does their span (Proposition 1.4.14). But $x^{d+1} \notin F[x]_{\leq d}$.

Let d be the largest of the integers $0, \deg p_1, ..., \deg p_k$, where an entry $\deg p_i = -\infty$ (for $p_i = \mathbf{0}$) does not affect the maximum, and $d = 0$ if $k = 0$. Then $p_1, ..., p_k \in F[x]_{\leq d}$, which is a subspace of $F[x]$ (Example 1.4.9). By Proposition 1.4.14(b), $\operatorname{span}(p_1, ..., p_k) \subseteq F[x]_{\leq d}$. The polynomial x^{d+1} has degree $d + 1 > d$, so it is not in $F[x]_{\leq d}$, and hence not in $\operatorname{span}(p_1, ..., p_k)$. So no list spans $F[x]$, and $F[x]$ is infinite-dimensional by Definition 1.6.23. □

So $F[x]$ has no basis in our sense, since a basis is a finite list that spans. Other infinite-dimensional spaces are $\mathbb{R}$ over $\mathbb{Q}$ (Exercise 1.3.45) and the space of finitely supported functions on an infinite set. Function spaces such as $\mathbb{R}^{\mathbb{R}}$ and $C[0, 1]$ are infinite-dimensional too, but proving it is easier with the tools of Section 4.1, which show that a subspace of a finite-dimensional space is finite-dimensional (Proposition 4.1.21).

Remark 1.6.26 (Thinning down and growing up). Proposition 1.6.13 suggests two ways to reach a basis. *Thinning down:* start with a spanning list and delete redundant entries, as in Theorem 1.6.10. *Growing up:* start with an independent list and, as long as it does not span V, append a vector outside its span, which keeps it independent by Proposition 1.5.14. By Proposition 1.6.13, either process can only stop at a basis.

The difference is that thinning always stops, because a list has only finitely many entries to delete, while growing need not stop. In $F[x]$ it never stops, whatever vectors are appended, because no list spans $F[x]$ (Proposition 1.6.25). For instance, starting from the empty list and appending 1, then x, then x^2, and so on, every stage $1, x, ..., x^n$ is independent (Example 1.5.7), and there is always a next power of x outside its span. In a finite-dimensional space growing does stop, because an independent list can never be longer than a spanning list (Lemma 4.1.2). That is how Section 4.1 proves that every independent list extends to a basis (Theorem 4.1.16). In infinite dimensions, making "grow forever" produce a basis needs set theory, which the next aside describes.

ASIDE: *Bases of infinite-dimensional spaces* BEYOND 21-242

Allowing bases to be infinite subsets fixes the problem. A subset $B \subseteq V$ is called a **basis** (or **Hamel basis**) of V if every vector of V is represented by exactly one linear combination of finitely many distinct elements of B, up to terms with coefficient 0. Equivalently, B is linearly independent (Aside (Independent sets)) and spans V (Aside (Spans of infinite sets)). The set $\{1, x, x^2, ...\}$ is a basis of $F[x]$ in this sense, and the coordinates of a polynomial with respect to it are its coefficient sequence, which is how Definition 1.2.13 defined $F[x]$ in the first place. Every vector space has a basis in this sense, but the proof uses Zorn's lemma and gives no way to write one down for $\mathbb{R}$ over $\mathbb{Q}$ or for $\mathbb{R}^{\mathbb{R}}$ (Section 15.1).

Exercises

Exercise 1.6.27. ROUTINE Decide whether each list is a basis of the given space, and justify your answer.

(a) $(1,2,0),(0,1,2),(2,0,1)$ in F^3. (The answer depends on F.)

(b) $1+x^2, x-x^2, 2+x+x^2$ in $\mathbb{R}[x]_{\leq 2}$.

(c) $U_{11}, U_{12}+U_{21}, U_{22}$ in $\mathbb{R}^{2\times 2}$.

Solution on page 1075

Exercise 1.6.28. ROUTINE

(a) Show that $\mathcal{B} = (1,1,1),(0,1,1),(0,0,1)$ is a basis of F^3, find $[(x,y,z)]_{\mathcal{B}}$, and find the vector $\mathbf{v} \in F^3$ with $[\mathbf{v}]_{\mathcal{B}} = (1,2,3)$.

(b) In $F^{2\times 2}$, let

$$\mathcal{C} = U_{11}, \quad U_{11}+U_{12}, \quad U_{11}+U_{12}+U_{21}, \quad U_{11}+U_{12}+U_{21}+U_{22}.$$

Show that $\mathcal{C}$ is a basis and find $[A]_{\mathcal{C}}$ for $A = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$.

Solution on page 1076

Exercise 1.6.29. ROUTINE For which $c \in \mathbb{R}$ is $1+x^2, x+x^2, 1+x+cx^2$ a basis of $\mathbb{R}[x]_{\leq 2}$? Work with coordinate vectors.

Solution on page 1076

Exercise 1.6.30. ROUTINE

(a) Find the error in the following argument.

Claim. Every spanning list of V contains a basis of V.

Proof. We induct on the length k of the spanning list $\mathcal{L}$. If $k = 0$, then $\mathcal{L} = ()$ spans V, so $V = \{\mathbf{0}\}$, and $()$ is a basis of V. Suppose the claim holds for spanning lists of length k, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ span V. If $\mathcal{L}$ is independent, it is a basis. Otherwise, delete its last entry. The list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ has length k, so by the inductive hypothesis it contains a basis of V, which is also a sublist of $\mathcal{L}$.

(b) Repair the argument.

Solution on page 1077

Exercise 1.6.31. ROUTINE The list $\mathcal{L} = (1,0),(0,1),(1,1)$ spans $\mathbb{R}^2$. A student defines $[\mathbf{v}]_{\mathcal{L}} \in \mathbb{R}^3$ to be "the coefficients of a linear combination of $\mathcal{L}$ that represents $\mathbf{v}$", and concludes that $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{L}}$ is a function $\mathbb{R}^2 \to \mathbb{R}^3$ satisfying $[\mathbf{x}+\mathbf{y}]_{\mathcal{L}} = [\mathbf{x}]_{\mathcal{L}} + [\mathbf{y}]_{\mathcal{L}}$, by the proof of Proposition 1.6.19.

(a) What is wrong with the definition, and which step of the proof of Proposition 1.6.19(a) fails for $\mathcal{L}$?

(b) Let $c : \mathbb{R}^2 \to \mathbb{R}^3$ be any function such that $c(\mathbf{v})$ is the list of coefficients of some linear combination of $\mathcal{L}$ representing $\mathbf{v}$, for every $\mathbf{v}$. Show that c is injective but not surjective.

Solution on page 1077

Exercise 1.6.32. STANDARD Let $n \geq 1$, and let $S \subseteq F^{n\times n}$ be the subspace of symmetric matrices, those with $a_{ji} = a_{ij}$ for all i, j (Example 1.4.11). Let K be the set of $A \in F^{n\times n}$ with $a_{ji} = -a_{ij}$ for all i, j.

(a) Find a basis of S of length $n(n+1)/2$.

(b) Suppose that $1+1 \neq 0$ in F. Find a basis of K of length $n(n-1)/2$.

(c) What are K and a basis of K when $F = \mathbb{F}_2$?

Solution on page 1078

Exercise 1.6.33. STANDARD Let $Z \subseteq F^{3\times 3}$ be the set of matrices in which every row and every column sums to 0. For $k, l \in \{1, 2\}$, let

$$Z_{kl} = U_{kl} - U_{k3} - U_{3l} + U_{33}.$$

Show that $\mathcal{Z} = Z_{11}, Z_{12}, Z_{21}, Z_{22}$ is a basis of Z, and that for $A \in Z$,

$$[A]_{\mathcal{Z}} = (a_{11}, a_{12}, a_{21}, a_{22}).$$

Solution on page 1078

Exercise 1.6.34. STANDARD Let $n \geq 1$, and let $p_0, ..., p_n \in F[x]_{\leq n}$.

(a) Show that if $\deg p_k = k$ for each k, then $p_0, ..., p_n$ is a basis of $F[x]_{\leq n}$.

(b) Show that if no p_k has degree n, then $p_0, ..., p_n$ is not a basis of $F[x]_{\leq n}$.

(c) Find a basis of $F[x]_{\leq n}$ all of whose entries have degree n.

Solution on page 1079

Exercise 1.6.35. STANDARD Prove or disprove each statement.

(a) If $\mathcal{B}$ is a basis of V and U is a subspace of V, then some sublist of $\mathcal{B}$ is a basis of U.

(b) If $\mathcal{B}$ and $\mathcal{C}$ are bases of V and $V \neq \{\mathbf{0}\}$, then the concatenation $\mathcal{B}, \mathcal{C}$ is not a basis of V.

(c) If $\mathbf{v}_1, ..., \mathbf{v}_n$ is a basis of V and $c_1, ..., c_n \in F$ are nonzero, then $c_1\mathbf{v}_1, ..., c_n\mathbf{v}_n$ is a basis of V.

(d) If $\mathcal{L}$ is a list in V such that every vector of V is represented by at most one linear combination of $\mathcal{L}$, then $\mathcal{L}$ is a basis of V.

(e) If $\mathcal{L}$ spans V, then exactly one sublist of $\mathcal{L}$ is a basis of V.

Solution on page 1079

Exercise 1.6.36. STANDARD Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V, let $\mathbf{w} \in V$ with $[\mathbf{w}]_{\mathcal{B}} = (a_1, ..., a_n)$, and fix j. Let $\mathcal{B}'$ be the list obtained from $\mathcal{B}$ by replacing $\mathbf{v}_j$ with $\mathbf{w}$.

(a) Show that $\mathcal{B}'$ is a basis of V if and only if $a_j \neq 0$.

(b) Deduce that if $\mathbf{w} \neq \mathbf{0}$, then $\mathbf{w}$ can replace some entry of $\mathcal{B}$ so that the result is still a basis.

Solution on page 1080

Exercise 1.6.37. STANDARD Let $f : V \to F^n$ be a bijection that respects linear combinations:

$$f(c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k) = c_1 f(\mathbf{x}_1) + \cdots + c_k f(\mathbf{x}_k)$$

for all $k \geq 0$, all $\mathbf{x}_i \in V$, and all $c_i \in F$. Show that there is exactly one basis $\mathcal{C}$ of V such that $f(\mathbf{v}) = [\mathbf{v}]_{\mathcal{C}}$ for every $\mathbf{v} \in V$. In other words, the coordinate maps of bases are exactly the bijections $V \to F^n$ that respect linear combinations.

Solution on page 1080

Exercise 1.6.38. STANDARD Let F be a finite field with q elements, and let $n \geq 1$.

(a) Show that a list in F^n is a basis of F^n if and only if it is linearly independent and has length n.

(b) How many bases does F^n have? List all bases of $\mathbb{F}_2^2$.

Solution on page 1081

Exercise 1.6.39. HARD Let $X = \{0, 1, 2, \ldots\}$, so that $\mathbb{R}^X$ is the space of real sequences $\mathbf{s} = (s_0, s_1, s_2, \ldots)$, and let

$$V = \left\{ \mathbf{s} \in \mathbb{R}^X \;\middle|\; s_{k+2} = s_{k+1} + s_k \text{ for all } k \geq 0 \right\}.$$

(a) Show that V is a subspace of $\mathbb{R}^X$, and that if $\mathbf{s} \in V$ and $s_0 = s_1 = 0$, then $\mathbf{s} = \mathbf{0}$.

(b) Let $\mathbf{g}, \mathbf{f} \in V$ be the sequences with $g_0 = 1, g_1 = 0$ and $f_0 = 0, f_1 = 1$. (So $\mathbf{f} = (0, 1, 1, 2, 3, 5, \ldots)$ is the Fibonacci sequence.) Show that $\mathbf{g}, \mathbf{f}$ is a basis of V, and that $[\mathbf{s}]_{(\mathbf{g},\mathbf{f})} = (s_0, s_1)$.

(c) Let $\varphi = \left(1 + \sqrt{5}\right)/2$ and $\psi = \left(1 - \sqrt{5}\right)/2$, the roots of $t^2 = t + 1$. Show that the geometric sequences $\mathbf{a} = \left(\varphi^k\right)_{k \geq 0}$ and $\mathbf{b} = \left(\psi^k\right)_{k \geq 0}$ form a basis of V. By finding $[\mathbf{f}]_{(\mathbf{a},\mathbf{b})}$, prove that

$$f_k = \frac{\varphi^k - \psi^k}{\sqrt{5}} \quad \text{for all } k \geq 0.$$

Solution on page 1081

Exercise 1.6.40. HARD Let $S \subseteq F^{3\times 3}$ be the set of matrices whose three row sums and three column sums are all equal; for $A \in S$, write $s(A)$ for this common value. Let $Z \subseteq S$ and $\mathcal{Z} = Z_{11}, Z_{12}, Z_{21}, Z_{22}$ be as in Exercise 1.6.33, the case $s(A) = 0$.

(a) Show that S is a subspace of $F^{3\times 3}$, that $I, Z_{11}, Z_{12}, Z_{21}, Z_{22}$ is a basis of S, where $I = U_{11} + U_{22} + U_{33}$, and that for $A \in S$ with $s = s(A)$,

$$[A]_{(I,\mathcal{Z})} = (s, \quad a_{11} - s, \quad a_{12}, \quad a_{21}, \quad a_{22} - s).$$

(b) Let J be the matrix all of whose entries are 1. Find $[J]_{(I,\mathcal{Z})}$, and determine for which fields $J \in Z$.

(c) A **permutation matrix** has exactly one entry 1 in each row and each column, and 0 elsewhere. Show that the six 3×3 permutation matrices span S, and find a dependence relation among them.

Solution on page 1082

Exercise 1.6.41. HARD Let K be a subfield of a field E, and let V be a vector space over E; regard E and V as vector spaces over K by restricting scalars (Proposition 1.2.19). Suppose that $\alpha_1, \ldots, \alpha_m$ is a basis of E over K and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of V over E.

(a) Show that the mn vectors $\alpha_i \mathbf{v}_j$, listed as

$$\alpha_1\mathbf{v}_1, \ldots, \alpha_m\mathbf{v}_1, \quad \alpha_1\mathbf{v}_2, \ldots, \alpha_m\mathbf{v}_2, \quad \ldots, \quad \alpha_1\mathbf{v}_n, \ldots, \alpha_m\mathbf{v}_n,$$

form a basis of V over K.

(b) Write down the resulting basis of $\mathbb{C}^2$ over $\mathbb{R}$, and the coordinates of $(1 + 2i, 3 - i)$ in it.

Solution on page 1083

Exercise 1.6.42. CHALLENGE Let $n \geq 1$, and let $V = \mathbb{F}_2^X$ be the space of all functions $X \to \mathbb{F}_2$, where $X = \mathbb{F}_2^n$. For a subset $S \subseteq \{1, \ldots, n\}$, let $m_S \in V$ be the function

$$m_S(y_1, \ldots, y_n) = \prod_{i \in S} y_i,$$

with $m_\varnothing = 1$ (the constant function). Show that the 2^n functions m_S, listed in any order, form a basis of V. Deduce that every function $\mathbb{F}_2^n \to \mathbb{F}_2$ can be written in exactly one way as a sum of distinct products $y_{i_1} y_{i_2} \cdots y_{i_r}$ with $i_1 < \cdots < i_r$ (allowing the empty product 1).

Solution on page 1084

1.7 Proof patterns

Most proofs in this chapter are short, and most of their steps were forced. Once a claim is written out with its definitions unpacked and its quantifiers in order, its shape usually decides the first move. This section sorts the chapter's proofs by the question they answer. For each question it compares the moves that answer it, on examples from the chapter, and says when a move is the wrong tool. The exercises at the end practice one move at a time, and the chapter's problem set (Problems for Chapter 1) mixes them.

1.7.1 Showing that two sets are equal

The default is double containment (Technique 1.3.14), with each containment proved by chasing a generic element (Technique 1.3.13). The two halves are rarely alike. In Example 1.3.15, the containment $\operatorname{span}(\mathcal{L}_1) \subseteq W$ was a computation that handled every linear combination at once, while $W \subseteq \operatorname{span}(\mathcal{L}_1)$ needed a linear combination built for each element of W. Expect one half to be a check and the other a construction, and look for the construction first, since that is where proofs get stuck.

When the sets carry structure, finitely many checks can replace a containment.

- If both sets are spans, compare entries: $\operatorname{span}(\mathcal{L}) = \operatorname{span}(\mathcal{M})$ exactly when each list lies in the span of the other (Corollary 1.3.19).
- If one set is a subspace U, then $\operatorname{span}(\mathcal{L}) \subseteq U$ as soon as the entries of $\mathcal{L}$ lie in U (Proposition 1.4.14(b)).
- If one set is defined as the smallest subspace with some property, use the property instead of the elements, as in the proof of Corollary 1.4.17.

Example 1.7.1 (One equality, two ways). Let $P = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$. Exercise 1.3.16 proved $\operatorname{span}((1,0,-1),(0,1,-1)) = P$ by chasing an element through each containment. With the structure in view, one half needs no computation: P is a subspace (Example 1.4.8) that contains $(1,0,-1)$ and $(0,1,-1)$, so it contains their span by Proposition 1.4.14(b). The other half is the construction, and with the tools of this chapter it cannot be avoided. For $(x,y,z) \in P$ we have $z = -x - y$, so

$$(x,y,z) = x(1,0,-1) + y(0,1,-1).$$

In the same way, Example 1.3.21 replaced both halves of Example 1.3.15 by two observations about entries.

Chasing elements is the wrong tool for showing that two sets *differ*. Then one element of one set outside the other is enough, and the way to find it is a property that every element of one set has. Every vector $(a, a+b, b)$ of $\operatorname{span}((1,1,0),(0,1,1))$ has middle coordinate equal to the sum of the other two, and $(1,2,3)$ does not (Example 1.3.8). For finite sets there is a third option: a containment together with equal sizes gives equality. Over a finite field this decides when a list is a basis of F^n (Exercise 1.6.38). From Section 4.1 on, dimension plays the role of size over every field (Proposition 4.1.21).

Exercise 1.7.5 practices the choice between these tools, and Exercise 1.3.30 shows what is lost when one half is forgotten.

1.7.2 Showing that something is unique

Every "the" in this chapter, such as the zero vector or the coordinate vector, rests on a uniqueness proof. There were two ways to give one.

The first is uniqueness via two candidates (Technique 1.1.8): take two candidates and show that they are equal, as for identities and inverses (Theorem 1.1.9, Proposition 1.2.27, Proposition 1.2.28). With vectors, the comparison is usually a subtraction, and the difference of two candidates solves the homogeneous version of the problem. Two representations of $\mathbf{w}$ differ by a representation of $\mathbf{0}$, which is how (II) $\Rightarrow$ (I) of Theorem 1.5.5 went, and two solutions of $a\mathbf{x} = \mathbf{w}$ differ by a solution of $a\mathbf{x} = \mathbf{0}$ (Corollary 1.2.36). So

uniqueness reduces to the statement that the homogeneous problem has only the zero solution, a question about one fixed vector. Section 3.1 builds on this reduction (Technique 3.1.18).

The second is to derive the formula: suppose x is a solution, deduce that it must equal a specific expression, and check that this expression is a solution. This proves existence and uniqueness at once, and it computes the object. Proposition 1.1.13 was proved this way.

Example 1.7.2 (A basis, two ways). Is $(1,1),(1,-1)$ a basis of $\mathbb{R}^2$? Following Proposition 1.6.2, we can check two things separately. Spanning needs coefficients for each (x,y), which still have to be found (they turn out to be $(x+y)/2$ and $(x-y)/2$), and independence needs $a+b=a-b=0$ to force $a=b=0$, which it does (add and subtract the equations). Deriving the formula instead: if $a(1,1)+b(1,-1)=(x,y)$, then $a+b=x$ and $a-b=y$, so

$$a=\frac{x+y}{2},\quad b=\frac{x-y}{2},$$

and these values are a solution. So every (x,y) has exactly one representation, and we have its coordinates as a bonus (Example 1.6.3). Both routes divide by 2, and both fail over $\mathbb{F}_2$, where the list repeats an entry.

The formula is shorter here, but it needs equations that can be solved in general. For the polynomials of Exercise 1.6.6, no formula was needed: spanning came from Exercise 1.3.32 and independence from degrees (Proposition 1.5.24).

Before proving that something is unique, check that it is. The index j of Lemma 1.5.10 is not unique (Example 1.5.11), and a vector does not determine its coefficients with respect to a spanning list that is dependent (Exercise 1.6.31). Uniqueness also says nothing about existence: a list in which every vector has at most one representation need not span (Exercise 1.6.35). Practice: Exercise 1.7.6, Exercise 1.7.7, and Exercise 1.2.43.

1.7.3 Showing that a list is independent, or dependent

Dependence is shown by a certificate: one nontrivial relation, which anyone can check by computing a single linear combination (Theorem 1.5.5). Independence is a statement about *all* relations, and the chapter proved it in several ways.

- *The test at zero.* Assume $a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k=\mathbf{0}$ and solve for the coefficients (Example 1.5.6).
- *One entry at a time.* Show that each entry lies outside the span of the entries before it (Corollary 1.5.13, Proposition 1.5.14). Usually some property is shared by the whole span of the earlier entries and fails for the new entry, such as having degree at most n.
- *An extreme term.* Look at the largest degree (Proposition 1.5.24) or the fastest-growing function (Example 1.5.27, second proof), which nothing else in a relation can cancel.
- *Evaluation and operators,* for functions: evaluate a relation at well-chosen points (Example 1.5.26), or apply an operation such as differentiation that eliminates terms (Example 1.5.27, first proof).
- *Transfer.* Pass to coordinates (Proposition 1.6.20), apply the moves of Proposition 1.5.18, recognize the list as a sublist of a list already known to be independent (Corollary 1.5.17), or change the field: independence over F gives independence over a subfield (Proposition 1.5.20), and for lists of real vectors, independence over $\mathbb{R}$ gives independence over $\mathbb{C}$ (Proposition 1.5.21).

Example 1.7.3 (Three tools on one list). Let $p_1=1+x$, $p_2=x+x^2$, and $p_3=1+x^2$ in $F[x]$.

The test at zero. The combination $ap_1+bp_2+cp_3=(a+c)+(a+b)x+(b+c)x^2$ is $\mathbf{0}$ exactly when $a+c=a+b=b+c=0$. The first two equations give $b=c=-a$, and then the third reads $-(1+1)a=0$. So if $1+1\neq 0$, the only relation is trivial and the list is independent; if $1+1=0$, then $a=b=c=1$ is a dependence relation.

One entry at a time. $p_1 \neq \mathbf{0}$. Every vector of $\operatorname{span}(p_1)$ has degree at most 1, so $p_2 \notin \operatorname{span}(p_1)$. The vectors of $\operatorname{span}(p_1, p_2)$ are $a + (a+b)x + bx^2$, and this is p_3 exactly when $a = b = 1$ and $a + b = 0$, which is possible only when $1 + 1 = 0$. Same answer, with more information: over $\mathbb{F}_2$, p_3 is the first entry in the span of the earlier ones, $p_3 = p_1 + p_2$.

An extreme term. Proposition 1.5.24 does not apply, since p_2 and p_3 both have degree 2. Its hypothesis is doing real work: the leading terms of p_2 and p_3 can cancel, and over $\mathbb{F}_2$ the relation $p_1 + p_2 + p_3 = \mathbf{0}$ cancels them. When a tool's hypothesis fails, fall back on the test at zero.

Two tempting shortcuts are invalid. For functions on an infinite set, evaluating at finitely many points can prove independence but not dependence, since a combination that vanishes at a few points need not vanish everywhere (Section 1.5.6). And a list in which no entry is a multiple of another can still be dependent (Exercise 1.5.16). Practice: Exercise 1.7.8, Exercise 1.7.9, and Exercise 1.7.12.

1.7.4 Proving a statement about every list

Claims about lists of every length, or about linear combinations of every length, are proved by induction on the length of the list (Technique 1.6.9), usually with the empty list as the base case. The arithmetic of linear combinations (Lemma 1.3.9), (III) ⇒ (II) of Theorem 1.4.3, and Proposition 1.6.19(c) split off the last term of a sum. Theorem 1.6.10 deletes a well-chosen entry.

That theorem has three proofs, and they are one idea in three forms: delete redundant entries until none is left. The induction deletes one entry and hands the shorter list to the inductive hypothesis. The extremal proof takes a shortest sublist with the same span and shows that the dependence lemma could not shorten it further. Sifting reads the list once, from left to right, and keeps exactly the entries that the earlier ones do not account for. When only existence is needed, the extremal form is the shortest. When something must be carried along from step to step, as the coefficients are in Proposition 1.6.19(c), the induction is the natural form. When the output itself matters, sifting is the one to use, because it produces a definite sublist; its kept entries return as pivot columns (Section 4.2).

Two things go wrong in inductions on length. The shorter list must satisfy the hypotheses of the statement: deleting the last entry of a spanning list need not leave a spanning list, which is why the entry to delete is supplied by Lemma 1.5.10 (Exercise 1.6.30). And the inductive step must work from the base case on, not only for long lists (Exercise 1.7.10). Finally, induction proves things about all finite lists, but it does not show that a process of adding vectors stops. Thinning a spanning list stops because its length drops at every step, while growing an independent list need not stop, and in $F[x]$ it never does (Remark 1.6.26). Practice: Exercise 1.7.10 and Exercise 1.7.11.

1.7.5 Proving a statement with "or"

Several results of the chapter conclude "P or Q", among them Proposition 1.1.17, Proposition 1.2.35, and Proposition 1.4.20, and each of these has a proof by the same move.

> **Technique 1.7.4** (Proving a disjunction). To prove "P or Q", assume that P fails and prove Q. The failure of P is a hypothesis you can use, and it usually hands you an object: an inverse a^{-1} when $a \neq 0$, or a vector of U outside W when $U \not\subseteq W$. If one alternative is easier to negate usefully, negate that one. The other route is the contrapositive: assume that both alternatives fail and derive the negation of the hypothesis (or a contradiction). It hands you two objects at once, which helps when the argument needs both.
>
> Reach for it whenever a conclusion reads "... or ...". The canonical example is Proposition 1.4.20, which has one proof of each kind.
>
> *Used in.* Exercise 1.7.13.

In Proposition 1.1.17 and Proposition 1.2.35, the conclusion is "$a = 0$ or ...", and assuming $a \neq 0$ supplies a^{-1}, which cancels a. In the first proof of Proposition 1.4.20, the failure of $U \subseteq W$ supplies one vector $\mathbf{u} \in U$ outside W, and adding $\mathbf{u}$ to each vector of W pushes that vector into U. The second proof negates both alternatives and uses the two witnesses to rebuild the counterexample of the two axes.

The invalid version of this move is a case split on two objects at once: "$\mathbf{v}$ lies in U or in W, and so does $\mathbf{w}$, so $\mathbf{v} + \mathbf{w}$ lies in U or in W". The split forgets the mixed case, $\mathbf{v} \in U$ and $\mathbf{w} \in W$, which is where the claim fails (Exercise 1.4.30). Practice: Exercise 1.7.13.

1.7.6 Disproving a claim, and proving that no object exists

To disprove "every X has property P", one X without P suffices (Technique 1.4.18), and the smallest examples are the place to look (Technique 1.2.38): the zero space, F^1 and F^2, lists of length one or two, and the field $\mathbb{F}_2$. Most of the chapter's counterexamples were of this size: $\mathbf{v} = -\mathbf{v}$ in $\mathbb{F}_2^2$ (Example 1.2.39), the scaling move of Corollary 1.3.20(d) with the scalar -2, which is 0 in $\mathbb{F}_2$ (Example 1.3.22), the two axes of $\mathbb{R}^2$ (Non-example 1.4.13), and $x^2 = 1$ in $\mathbb{F}_2$ (Exercise 1.1.42). "Small" means the smallest case in which the claim says something. The claim that a list with no entry a multiple of another is independent is true for two vectors, so its counterexample needs three (Exercise 1.5.16).

Showing that *no* object of some kind exists is a different task, and one failure is not enough. In Example 1.2.25, each candidate operation on $\mathbb{R}^2$ was refuted by one instance of one failing axiom. But to show that *no* scalar multiplication makes $[0, 1)$ with addition modulo 1 a real vector space (Example 1.2.40), testing candidates one at a time cannot finish. Instead we found a consequence of the axioms that involves only the addition, which is the part held fixed, and showed that this addition violates it. The same need to rule out every candidate appeared on a smaller scale in Non-example 1.2.21: to show that the line $x + 2y = 2$ has no zero vector under the operations of $\mathbb{R}^2$, it was not enough to note that $(0, 0)$ is not on the line; every candidate $\mathbf{z}$ had to be excluded. Practice: Exercise 1.7.15 and Exercise 1.7.14.

1.7.7 Reading the chapter's proofs

Most of the chapter's proofs were found by reading the definitions slowly. Here is what to look at first.

The definitions and the quantifiers. Unpack every definition in the claim and keep the quantifiers in order. Then the shape is visible. "For all" calls for a generic element, "there exists" for a construction, "unique" for two candidates, and "or" for negating one alternative. For instance, independence says that every vector of the span has *at most one* representation. Its negation hands us a vector with two representations, and subtracting them gives a relation. That is the proof of (II) ⇒ (I) in Theorem 1.5.5, found by negating the definition. The subspace criterion came from sorting (A1)–(A8) by their quantifiers: identities that hold for all vectors are inherited by any subset closed under the operations, and the existence statements (A3) and (A4) are where a subset can fail (the intuition before Theorem 1.4.3). The same sorting proved Proposition 1.1.33 and Proposition 1.2.19. When a fact has to connect two operations, list the axioms that mention both: distributivity for $0a = 0$ (Proposition 1.1.15), and (A7) and (A8) for vectors.

The extreme choice. A dependence relation can be solved for any entry whose coefficient is nonzero. The dependence lemma wants only earlier entries on the other side, so the proof solves for the *last* entry with a nonzero coefficient (Lemma 1.5.10). Solving for the *first* one instead puts that entry in the span of the later entries (Exercise 1.7.8). The same move, a largest index or degree, runs Proposition 1.5.24 and the second proof in Example 1.5.27, and a shortest sublist runs the second proof of Theorem 1.6.10. Extremality is a hypothesis we get for free: "the largest j with $a_j \neq 0$" also says that every later coefficient is 0.

The degenerate cases. The empty list, the zero space, and the index $j = 1$ are where the conventions get tested: $\text{span}(()) = \{\mathbf{0}\}$, the empty list is independent and is the basis of $\{\mathbf{0}\}$, and $j = 1$ in Lemma 1.5.10 means $\mathbf{v}_1 = \mathbf{0}$. The subspace criterion needs its nonemptiness condition exactly once, to produce $\mathbf{0}$ (Remark 1.4.5).

The counterexample, read closely. The natural first guess is that $U \cup W$ is a subspace whenever U and W are. The two axes of $\mathbb{R}^2$ refute it. Reading the counterexample, it used only one vector of the first axis outside the second and one vector of the second outside the first, whose sum lies on neither (Lemma 1.4.19). So the failure needs each subspace to stick out of the other. That suggests the corrected statement, Proposition 1.4.20, and its second proof is the counterexample run on general U and W. The same reading turned the counterexample over $\mathbb{F}_2$ in Example 1.2.39 into the hypothesis $1 + 1 \neq 0$: the computation $\mathbf{v} + \mathbf{v} = (1 + 1)\mathbf{v}$ located the scalar that has to be invertible.

1.7.8 Common errors

Each of these appears somewhere in the chapter, most of them as find-the-error exercises.

- *Unions.* "$U \cup W$ is closed under addition because U and W are." The case split forgets the mixed case (Exercise 1.4.30). The repair is Proposition 1.4.20, and the smallest subspace containing both is $U + W$ (Proposition 1.4.24).
- $\mathbf{0}$ *in a span versus* $\mathbf{0}$ *in a list.* $\mathbf{0}$ lies in every span, including $\operatorname{span}(()) = \{\mathbf{0}\}$, and this says nothing about independence. A list that *contains* $\mathbf{0}$ is always dependent (Proposition 1.5.15(c)), while the empty list is independent. Relatedly, the zero vector of a structure is whatever satisfies (A3) for its addition, not the number 0 or the vector $(0, 0)$ (Exercise 1.2.24, Example 1.2.23). A subspace, on the other hand, always has the zero vector of the ambient space (Remark 1.4.4), which is the point of Exercise 1.4.29.
- *Lists versus sets.* The list $\mathbf{v}, \mathbf{v}$ is dependent, although its set of entries is $\{\mathbf{v}\}$. The span forgets order and repetition (Remark 1.3.23). Independence does not forget repetition, and coordinates do not forget order: reordering a basis reorders the coordinates. A vector determines its coefficients only with respect to an independent list (Exercise 1.3.29, Exercise 1.6.31).
- *Pairwise is not enough.* A list of three vectors, none a multiple of another, can be dependent (Exercise 1.5.16). Two independent lists can have a dependent concatenation (Exercise 1.7.12).
- *The wrong redundant entry.* The last entry of a dependent list need not lie in the span of the others (Exercise 1.5.30), and an induction that deletes an arbitrary entry of a spanning list breaks (Exercise 1.6.30).
- *Half of a double containment.* Showing that a span lies in V proves nothing about whether the list spans V (Exercise 1.3.30).
- *Hidden division.* An argument that divides by $1 + 1$, assumes $1 \neq -1$, or cancels a zero divisor fails over $\mathbb{F}_2$ or $\mathbb{Z}/n\mathbb{Z}$ (Exercise 1.1.42, Exercise 1.1.45, Example 1.3.22).
- *Local versus global.* The zero vector is unique among vectors that satisfy $\mathbf{v} + \mathbf{z} = \mathbf{v}$ for *every* $\mathbf{v}$; a vector that works for one $\mathbf{v}$ needs a separate argument (Exercise 1.2.43). Deriving (A4) from $(-1)\mathbf{v}$ is circular, because the proof of $0\mathbf{v} = \mathbf{0}$ used (A4) (Exercise 1.2.54).

Exercises

Exercise 1.7.5. ROUTINE For each equality, decide which containments are finite checks and which need a construction, and prove the equality.

(a) $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \operatorname{span}(\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \mathbf{v}_3)$, for any $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$.

(b) $\{p \in F[x]_{\leq 2} \mid p(0) = p(1)\} = \operatorname{span}(1, x^2 - x)$.

Solution on page 1084

Exercise 1.7.6. ROUTINE Let $\mathbf{u}, \mathbf{w} \in V$, and consider the equation $\mathbf{m} - \mathbf{u} = \mathbf{w} - \mathbf{m}$ for an unknown $\mathbf{m} \in V$.

(a) Suppose $1 + 1 \neq 0$ in F. Show that the equation has exactly one solution, and find it.

(b) Suppose $F = \mathbb{F}_2$. Show that the equation has a solution only if $\mathbf{u} = \mathbf{w}$, and that in that case every $\mathbf{m} \in V$ is a solution.

Solution on page 1085

Exercise 1.7.7. STANDARD Let $n \geq 2$, and let $\mathbf{v}_1, ..., \mathbf{v}_n$ be a linearly dependent list in V such that every sublist of length $n-1$ is linearly independent.

(a) Show that every dependence relation $c_1, ..., c_n$ of the list has all $c_i \neq 0$.

(b) Show that any two dependence relations are proportional: if $c_1, ..., c_n$ and $d_1, ..., d_n$ are dependence relations, then $d_i = \lambda c_i$ for all i, for some $\lambda \in F$.

(c) Find such a list of length 3 in $\mathbb{R}^2$, and all of its linear relations.

Solution on page 1086

Exercise 1.7.8. ROUTINE Let $\mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V. Show that it is linearly independent if and only if

$$\mathbf{v}_j \notin \operatorname{span}(\mathbf{v}_{j+1}, ..., \mathbf{v}_k) \quad \text{for every } j \in \{1, ..., k\}.$$

For the backward direction, argue directly with an extremal choice of index, without reordering the list and quoting Corollary 1.5.13.

Solution on page 1086

Exercise 1.7.9. STANDARD Let X be a set, let $f_1, ..., f_n \in F^X$, and let $t_1, ..., t_n \in X$ be points such that

$$f_i(t_i) \neq 0 \text{ for every } i, \quad \text{and} \quad f_i(t_j) = 0 \text{ whenever } j < i.$$

(a) Show that $f_1, ..., f_n$ is linearly independent.

(b) For $i \geq 0$, let $r_i \in \mathbb{R}^{\mathbb{R}}$ be the ramp $r_i(t) = \max(0, t - i)$. Show that $1, r_0, r_1, ..., r_{n-1}$ is linearly independent in $\mathbb{R}^{\mathbb{R}}$ for every $n \geq 1$, where 1 is the constant function.

Solution on page 1087

Exercise 1.7.10. ROUTINE Find the error in the following argument, and show that its claim is false. Then show that the argument does prove the following statement: if every two of the nonzero vectors $\mathbf{v}_1, ..., \mathbf{v}_k$ lie on a common line, then all of them do.

Claim. Any nonzero vectors $\mathbf{v}_1, ..., \mathbf{v}_k \in V$, with $k \geq 1$, lie on a common line: there is a vector $\mathbf{u}$ with $\mathbf{v}_1, ..., \mathbf{v}_k \in \operatorname{span}(\mathbf{u})$.

Proof. Induction on k. For $k = 1$, take $\mathbf{u} = \mathbf{v}_1$. Suppose the claim holds for all lists of length k, and let $\mathbf{v}_1, ..., \mathbf{v}_{k+1}$ be nonzero. Applying the inductive hypothesis to $\mathbf{v}_1, ..., \mathbf{v}_k$ gives $\mathbf{u}$ with $\mathbf{v}_1, ..., \mathbf{v}_k \in \operatorname{span}(\mathbf{u})$, and applying it to $\mathbf{v}_2, ..., \mathbf{v}_{k+1}$ gives $\mathbf{u}'$ with $\mathbf{v}_2, ..., \mathbf{v}_{k+1} \in \operatorname{span}(\mathbf{u}')$. The vector $\mathbf{v}_2$ lies on both lines and is nonzero, so $\operatorname{span}(\mathbf{u}) = \operatorname{span}(\mathbf{v}_2) = \operatorname{span}(\mathbf{u}')$. Hence all of $\mathbf{v}_1, ..., \mathbf{v}_{k+1}$ lie in $\operatorname{span}(\mathbf{u})$.

Solution on page 1087

Exercise 1.7.11. STANDARD Let V be a vector space over $\mathbb{F}_2$ and $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ a list in V. Let $Z(\mathcal{L})$ be the number of subsets $I \subseteq \{1, ..., k\}$ with $\sum_{i \in I} \mathbf{v}_i = \mathbf{0}$, where the empty set counts, with the empty sum $\mathbf{0}$.

(a) Let $\mathbf{w} \in V$. Show that $Z(\mathcal{L}, \mathbf{w}) = Z(\mathcal{L})$ if $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and $Z(\mathcal{L}, \mathbf{w}) = 2Z(\mathcal{L})$ if $\mathbf{w} \in \operatorname{span}(\mathcal{L})$.

(b) Deduce that $Z(\mathcal{L}) = 2^d$, where d is the number of entries that sifting deletes from $\mathcal{L}$ (the third proof of Theorem 1.6.10). In particular, $Z(\mathcal{L}) = 1$ if and only if $\mathcal{L}$ is linearly independent.

Solution on page 1088

Exercise 1.7.12. STANDARD

(a) Disprove: if $\mathcal{L}$ and $\mathcal{M}$ are linearly independent lists in V, then the concatenation $\mathcal{L}, \mathcal{M}$ is linearly independent. Use the smallest counterexample you can find.

(b) What do $\operatorname{span}(\mathcal{L})$ and $\operatorname{span}(\mathcal{M})$ have in common in your counterexample? Use the answer to formulate a corrected statement.

(c) Prove: if $\mathcal{L}$ and $\mathcal{M}$ are linearly independent, then $\mathcal{L}, \mathcal{M}$ is linearly independent if and only if $\operatorname{span}(\mathcal{L}) \cap \operatorname{span}(\mathcal{M}) = \{\mathbf{0}\}$.

Solution on page 1088

Exercise 1.7.13. ROUTINE Let U be a subspace of V, and let $\mathbf{v}, \mathbf{w} \in V$ with $\mathbf{v} \notin U$.

(a) Suppose $1 + 1 \neq 0$ in F. Show that $\mathbf{v} + \mathbf{w} \notin U$ or $\mathbf{v} - \mathbf{w} \notin U$.

(b) Show that (a) fails over $\mathbb{F}_2$.

Solution on page 1089

Exercise 1.7.14. STANDARD Show that there is no field F, and no scalar multiplication $F \times \mathbb{Z} \to \mathbb{Z}$, that make $\mathbb{Z}$ with its usual addition a vector space over F. (Exercise 1.2.46 treated the fields $\mathbb{F}_p$.)

Solution on page 1089

Exercise 1.7.15. ROUTINE Prove or disprove: every subspace of F^2 is the solution set $\{(x, y) \in F^2 \mid ax + by = 0\}$ of a single equation, for some $a, b \in F$. If the claim is false, find the subspaces for which it holds, and prove that it holds for them.

Solution on page 1090

Problems for Chapter 1

These problems combine the sections of Chapter 1. The first four are review prompts of the kind found on an exam review sheet: state the result, prove it without looking, and then answer a follow-up question that the text did not ask. The rest are ordered roughly by difficulty. Problems of kind "competition" are harder than anything the course expects.

Review: state and prove

Problem 1.1. STANDARD

(a) State the criterion for $\mathbb{Z}/n\mathbb{Z}$ to be a field, and prove it.

(b) One of the two proofs in Section 1.1.3 computes inverses. Use it to find 17^{-1} in $\mathbb{F}_{101}$, and then solve $17x + 3 = 5$ in $\mathbb{F}_{101}$.

Solution on page 1090

Problem 1.2. STANDARD Two of the vector space axioms, (A7) and (A8), involve both operations.

(a) Using only (A1)–(A7), and not (A8), prove that $a(\mathbf{v} - \mathbf{w}) = a\mathbf{v} - a\mathbf{w}$ for all $a \in F$ and $\mathbf{v}, \mathbf{w} \in V$.

(b) Prove that $(a - b)\mathbf{v} = a\mathbf{v} - b\mathbf{v}$ for all $a, b \in F$ and $\mathbf{v} \in V$. Which axioms does your proof use?

(c) Show that the identity in (b) can fail for a structure that satisfies every axiom except (A8).

Solution on page 1091

Problem 1.3. STANDARD Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ be a list in V.

(a) State the definition of linear independence and the test at zero (Theorem 1.5.5), and prove that they are equivalent.

(b) Suppose that *some* vector $\mathbf{w} \in \operatorname{span}(\mathcal{L})$ is represented by exactly one linear combination of $\mathcal{L}$. Show that $\mathcal{L}$ is linearly independent.

(c) Suppose instead that some $\mathbf{w} \in V$ is represented by *no* linear combination of $\mathcal{L}$. Can you conclude that $\mathcal{L}$ is independent? That it is dependent?

Solution on page 1092

Problem 1.4. STANDARD

(a) State the linear dependence lemma, including the statement about deleting an entry, and prove it.

(b) In the proof, j was the *largest* index with $a_j \neq 0$ in a nontrivial relation. Let i instead be the *smallest* such index. What does the relation say about $\mathbf{v}_i$?

(c) For the list $(1,0), (0,1), (0,2)$ in $\mathbb{R}^2$, find all j with $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, ..., \mathbf{v}_{j-1})$, and all i with $\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_{i+1}, ..., \mathbf{v}_3)$.

Solution on page 1093

Synthesis problems

Problem 1.5. ROUTINE Work in $\mathbb{F}_5^2$.

(a) List the lines $\operatorname{span}(\mathbf{v})$, $\mathbf{v} \neq \mathbf{0}$, of $\mathbb{F}_5^2$ by giving one nonzero vector on each. Which line contains $(2,3)$?

(b) Show that $\mathcal{B} = (1,2), (1,3)$ is a basis of $\mathbb{F}_5^2$, and find $[(x,y)]_{\mathcal{B}}$ and $[(2,3)]_{\mathcal{B}}$.

(c) How many bases of $\mathbb{F}_5^2$ have their first entry on the line $\operatorname{span}((1,2))$?

Solution on page 1093

Problem 1.6. ROUTINE In $F[x]_{\leq 3}$, let $U = \{p \mid p(1) = 0\}$ and $W = \{p \mid p(-1) = 0\}$.

(a) Suppose $1 + 1 \neq 0$ in F. Find a basis of $U \cap W$, and show that $U + W = F[x]_{\leq 3}$.

(b) Describe U, W, $U \cap W$, and $U + W$ when $F = \mathbb{F}_2$.

Solution on page 1094

Problem 1.7. ROUTINE In $\mathbb{R}^{\mathbb{R}}$, let $V = \operatorname{span}(1, \cos, \sin)$, where 1 is the constant function.

(a) Show that $1, \cos, \sin$ is a basis of V.

(b) Find the coordinates, with respect to this basis, of the functions $t \mapsto \cos(t + \pi/3)$, $t \mapsto \sin^2(t/2)$, and $t \mapsto (\cos(t/2) + \sin(t/2))^2$.

(c) Is the function $t \mapsto \cos^2 t$ in V?

Solution on page 1095

Problem 1.8. STANDARD Let K be a subfield of F with $K \neq F$, and let $n \geq 1$. Regard K^n as a subset of F^n.

(a) Show that K^n is a subspace of F^n regarded as a vector space over K (Proposition 1.2.19), but not a subspace of the F-vector space F^n.

(b) Show that the only subspace of the F-vector space F^n that is contained in K^n is $\{\mathbf{0}\}$.

(c) Let U be a subspace of the F-vector space F^n. Show that $U \cap K^n$ is a subspace of the K-vector space K^n. For $F = \mathbb{C}$, $K = \mathbb{R}$, and $n = 2$, find subspaces U of $\mathbb{C}^2$, other than $\{\mathbf{0}\}$ and $\mathbb{C}^2$, with $U \cap \mathbb{R}^2 = \{\mathbf{0}\}$ and with $U \cap \mathbb{R}^2 \neq \{\mathbf{0}\}$.

Solution on page 1095

Problem 1.9. STANDARD Let $P = (0, \infty)$ with the operations $x \oplus y = xy$ and $a \odot x = x^a$ for $a \in \mathbb{R}$. This is the real vector space obtained by transporting the structure of $\mathbb{R}$ along $\exp : \mathbb{R} \to (0, \infty)$ (Proposition 1.2.22), and you may take that for granted.

(a) What is the zero vector of P? Decide whether each of the lists $1, 2$, then $2, 4$, then $2, 3$ is linearly independent in P.

(b) Regard P as a vector space over $\mathbb{Q}$ (Proposition 1.2.19). Show that $2, 3$ is linearly independent over $\mathbb{Q}$, while $2, 3, 6$ and $2, 4$ are not.

(c) Show that $5 \notin \operatorname{span}(2, 3)$ over $\mathbb{Q}$.

Solution on page 1096

Problem 1.10. STANDARD Let U and W be subspaces of V, and let $\mathcal{L}$ be a list in V. Prove or disprove each statement.

(a) If $\mathbf{u} \in U$, $\mathbf{u} \notin W$, $\mathbf{w} \in W$, and $\mathbf{w} \notin U$, then $\mathbf{u}, \mathbf{w}$ is linearly independent.

(b) Under the same hypotheses, $\mathbf{u} + \mathbf{w}, \mathbf{u} - \mathbf{w}$ is linearly independent.

(c) The set of $\mathbf{v} \in V$ for which $\mathcal{L}, \mathbf{v}$ is linearly dependent is a subspace of V.

(d) If V is a complex vector space and $\mathcal{L}$ spans V, then $\mathcal{L}$ spans V regarded as a real vector space.

(e) If $\operatorname{span}(\mathcal{L}) \cap U = \{\mathbf{0}\}$, then $\mathcal{L}$ is linearly independent.

(f) If every entry of $\mathcal{L}$ lies in $U \cup W$, then $\operatorname{span}(\mathcal{L}) \subseteq U + W$.

Solution on page 1096

Problem 1.11. STANDARD Let U and W be subspaces of V with bases $\mathcal{L} = \mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathcal{M} = \mathbf{w}_1, \ldots, \mathbf{w}_m$, and sift the concatenation $\mathcal{L}, \mathcal{M}$ (the third proof of Theorem 1.6.10).

Claim. Every entry of $\mathcal{M}$ that sifting deletes lies in $U \cap W$.

Proof. Let $\mathbf{w}_j$ be deleted. It lies in W, since it is an entry of $\mathcal{M}$. It was deleted because it lies in the span of the entries before it, and these are the entries of $\mathcal{L}$, so $\mathbf{w}_j \in \operatorname{span}(\mathcal{L}) = U$. Hence $\mathbf{w}_j \in U \cap W$.

(a) Find the error, and show that the claim is false.

(b) Show that sifting deletes no entry of $\mathcal{L}$, and that the kept entries form a basis of $U + W$.

(c) Show that if $\mathbf{w}_j$ is deleted, then $\mathbf{w}_j - \mathbf{y}$ is a nonzero vector of $U \cap W$ for some $\mathbf{y} \in \operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_{j-1})$.

Solution on page 1097

Problem 1.12. HARD Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$, let $t = \sqrt[3]{2}$, and let

$$K = \operatorname{span}(1, t, t^2) = \{a + bt + ct^2 \mid a, b, c \in \mathbb{Q}\}.$$

You may use that no rational number has cube 2.

(a) Show that $1, t, t^2$ is linearly independent over $\mathbb{Q}$.

(b) Show that K is closed under multiplication.

(c) For $\alpha = a + bt + ct^2 \in K$, let $\beta = \big(a^2 - 2bc\big) + \big(2c^2 - ab\big)t + \big(b^2 - ac\big)t^2$. Show that $\alpha\beta = a^3 + 2b^3 + 4c^3 - 6abc$. Deduce that K is a subfield of $\mathbb{R}$. (Compare Exercise 1.1.53, where the span of $1, t$ alone was not.)

Solution on page 1097

Problem 1.13. HARD Let F be a finite field with q elements. Call a subspace of F^3 a *line* if it is $\operatorname{span}(\mathbf{v})$ for some $\mathbf{v} \neq \mathbf{0}$, and a *plane* if it is $\operatorname{span}(\mathbf{v}, \mathbf{w})$ for some linearly independent $\mathbf{v}, \mathbf{w}$. You may use Exercise 1.5.43: a linearly independent list of length k in a vector space over F has a span with exactly q^k elements, and F^n contains exactly $(q^n - 1)(q^n - q)\cdots\big(q^n - q^{k-1}\big)$ independent lists of length k.

(a) Show that every subspace of F^3 is $\{\mathbf{0}\}$, a line, a plane, or F^3, and that no subspace is of two of these kinds.

(b) Count the lines and the planes of F^3, and the subspaces of F^3. Check your count against $\mathbb{F}_2^3$, which has 16 subspaces.

Solution on page 1098

Problem 1.14. HARD Let F be a finite field with q elements, and let $n \geq 1$.

(a) Let U be a proper subspace of F^n. Show that U has at most q^{n-1} elements.

(b) Deduce that if $S \subseteq F^n$ has more than q^{n-1} elements, then every list whose entries run through the elements of S spans F^n.

(c) Show that (b) is sharp: find a set S with exactly q^{n-1} elements whose span is not F^n.

Solution on page 1099

Problem 1.15. HARD Let $n \geq 1$. In $\mathbb{R}^{\mathbb{R}}$, let $V_n = \operatorname{span}\big(1, t, t^2, \ldots, t^n\big)$, where t^k denotes the function $t \mapsto t^k$. For $0 \leq k \leq n$, let

$$b_k(t) = t^k(1 - t)^{n-k}.$$

(a) Show that $b_0, \ldots, b_n$ lie in V_n.

(b) Show that $b_0, \ldots, b_n$ is linearly independent. (Consider the smallest k with a nonzero coefficient, and let $t \to 0$.)

(c) Show that $b_0, \ldots, b_n$ spans V_n, so it is a basis of V_n.

(d) Find the coordinates of the constant function 1 and of $t \mapsto t$ in this basis.

Solution on page 1100

Problem 1.16. HARD Let F be a finite field. Write $m * 1 = 1 + \cdots + 1$ (m terms) for integers $m \geq 1$, and $0 * 1 = 0$. (Definition 1.1.28 writes $m \cdot 1$; we use $*$ because juxtaposition already means multiplication in F.)

(a) Show that $p = \operatorname{char} F$ is a prime. (You may use Exercise 1.1.48.)

(b) Let $K = \{0 * 1, 1 * 1, \ldots, (p-1) * 1\}$. Show that K has exactly p elements and is a subfield of F.

(c) Deduce that F has exactly p^n elements for some $n \geq 1$. In particular, there is no field with 6 elements.

Solution on page 1101

Problem 1.17. HARD Let X be a set and let $f_1, \ldots, f_n \in F^X$ with $n \geq 1$. Show that $f_1, \ldots, f_n$ is linearly independent if and only if there are points $t_1, \ldots, t_n \in X$ such that the only scalars $a_1, \ldots, a_n$ with

$$a_1 f_1(t_j) + \cdots + a_n f_n(t_j) = 0 \quad \text{for } j = 1, \ldots, n$$

are $a_1 = \cdots = a_n = 0$. That is, the independence of n functions can always be certified by evaluating at n well-chosen points.

Solution on page 1101

Problem 1.18. CHALLENGE Find all fields F such that every function $F \to F$ is the polynomial function (Definition 1.2.17) of some $p \in F[x]$. For each such field, show that every function $F \to F$ is the polynomial function of exactly one polynomial of degree less than $|F|$. You may use that a nonzero polynomial of degree d has at most d roots (Proposition 7.1.22).

Solution on page 1102

Problem 1.19. CHALLENGE Let $n \geq 1$. Find the largest k for which there is a list of k vectors in $\mathbb{F}_2^n$ in which every sublist of length at most 3 is linearly independent.

Solution on page 1103

Chapter 2

Linear maps and matrices

Linear maps, their matrices, matrix multiplication as composition, inverses, kernels and images, and the Fundamental Theorems.

2.1 Linear maps

In Section 1.6 we attached to each vector $\mathbf{v}$ of a space V with basis $\mathcal{B}$ its coordinate vector $[\mathbf{v}]_{\mathcal{B}} \in F^n$, and we saw that coordinates respect linear combinations: $[c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k]_{\mathcal{B}} = c_1[\mathbf{v}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{v}_k]_{\mathcal{B}}$ (Proposition 1.6.19). Forming a combination in V and then taking coordinates gives the same answer as taking coordinates first and forming the combination in F^n. A vector space has no structure besides addition and scalar multiplication, so the functions worth studying between vector spaces are the ones with this property. They are the linear maps, and from here on they are the main objects of the book.

2.1.1 Definition and first properties

Definition 2.1.1 (Linear map). Let V and W be vector spaces over the same field F. A function $T : V \to W$ is a **linear map**, or **linear transformation**, if it **respects linear combinations**: for every $k \geq 1$, all $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, and all $a_1, \ldots, a_k \in F$,

$$T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k).$$

A linear map from V to itself is also called a **linear operator**, or just an **operator**, on V.

Notation 2.1.2. $\mathrm{L}(V, W)$ denotes the set of linear maps $V \to W$, and $\mathrm{L}(V) = \mathrm{L}(V, V)$ the set of operators on V. Other texts write $T\mathbf{v}$ for $T(\mathbf{v})$; we keep the parentheses in this section. When the field matters, we say that T is F-**linear**.

Realize that the two sides of the defining equation are computed in different spaces. On the left, the linear combination is formed in V and then T is applied to the result. On the right, T is applied to each $\mathbf{v}_i$ first and the combination is formed in W, with the same coefficients. Linearity says that the order does not matter: applying T commutes with forming linear combinations.

Intuition. The definition asks about combinations of every length k, which is unpleasant to check directly. But a linear combination is built one operation at a time from scalar multiples and sums. So a function that respects scalar multiplication and addition separately should respect every linear combination: split off the last term $a_k\mathbf{v}_k$ using addition, pull out its coefficient using scalar multiplication, and

handle the shorter combination that remains in the same way. This is an induction on k. Conversely, $a\mathbf{v}$ and $\mathbf{u}+\mathbf{v}$ are themselves linear combinations, of lengths 1 and 2.

Theorem 2.1.3 (Checking linearity). Let V and W be vector spaces over F, and let $T : V \to W$ be a function. The following are equivalent.

(i) T is linear.

(ii) T is **additive** and **homogeneous**: $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(a\mathbf{v}) = aT(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $a \in F$.

(iii) For all $\mathbf{u}, \mathbf{v} \in V$ and all $a \in F$, $T(a\mathbf{u}+\mathbf{v}) = aT(\mathbf{u}) + T(\mathbf{v})$.

Proof. ROUTINE

Uses. Definition 2.1.1 (Linear map), Definition 1.2.1 (Vector space)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Sums and scalar multiples are linear combinations of length 2 and 1, so (i) ⇒ (iii) and (iii) ⇒ (ii) come down to choosing special cases. For (ii) ⇒ (i), induct on the length k.
Hint 2. For (iii) ⇒ (ii), first take $a = 1$ and $\mathbf{u} = \mathbf{v} = \mathbf{0}$ to learn that $T(\mathbf{0}) = \mathbf{0}$. Then $a = 1$ gives additivity and $\mathbf{v} = \mathbf{0}$ gives homogeneity.
Hint 3. For the induction step, write $a_1\mathbf{v}_1 + \cdots + a_{k+1}\mathbf{v}_{k+1}$ as $\mathbf{u} + a_{k+1}\mathbf{v}_{k+1}$ with $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. Apply additivity, then homogeneity to the last term and the induction hypothesis to $T(\mathbf{u})$.
Sketch. (i) ⇒ (iii) is the definition with $k = 2$ and coefficients $a, 1$. In (iii), the choice $a = 1$, $\mathbf{u} = \mathbf{v} = \mathbf{0}$ gives $T(\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, so $T(\mathbf{0}) = \mathbf{0}$; then $a = 1$ gives additivity and $\mathbf{v} = \mathbf{0}$ gives homogeneity. For (ii) ⇒ (i), induct on k: the case $k = 1$ is homogeneity, and additivity splits a combination of length $k+1$ into one of length k plus a single term.

We prove (i) ⇒ (iii) ⇒ (ii) ⇒ (i).

(i) ⇒ (iii). Let $\mathbf{u}, \mathbf{v} \in V$ and $a \in F$. The definition of linearity with $k = 2$, $\mathbf{v}_1 = \mathbf{u}$, $\mathbf{v}_2 = \mathbf{v}$, $a_1 = a$, and $a_2 = 1$ gives $T(a\mathbf{u} + 1\mathbf{v}) = aT(\mathbf{u}) + 1T(\mathbf{v})$. Since $1\mathbf{x} = \mathbf{x}$ for every vector $\mathbf{x}$, in V and in W, this is (iii).

(iii) ⇒ (ii). Taking $a = 1$ and $\mathbf{u} = \mathbf{v} = \mathbf{0}$ in (iii) gives

$$T(\mathbf{0}) = T(1 \cdot \mathbf{0} + \mathbf{0}) = 1 \cdot T(\mathbf{0}) + T(\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}).$$

Adding $-T(\mathbf{0})$ to both sides and regrouping, $\mathbf{0} = T(\mathbf{0})$. Now let $\mathbf{u}, \mathbf{v} \in V$ and $a \in F$. With $a = 1$, (iii) reads $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$, which is additivity. With $\mathbf{v} = \mathbf{0}$, it reads

$$T(a\mathbf{u}) = T(a\mathbf{u} + \mathbf{0}) = aT(\mathbf{u}) + T(\mathbf{0}) = aT(\mathbf{u}),$$

which is homogeneity.

(ii) ⇒ (i). We show by induction on $k \geq 1$ that $T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k)$ for all $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and $a_1, \ldots, a_k \in F$. For $k = 1$ this is homogeneity. Suppose it holds for some $k \geq 1$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1} \in V$ and $a_1, \ldots, a_{k+1} \in F$. Put $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. Then

$$\begin{aligned} T(a_1\mathbf{v}_1 + \cdots + a_{k+1}\mathbf{v}_{k+1}) &= T(\mathbf{u} + a_{k+1}\mathbf{v}_{k+1}) \\ &= T(\mathbf{u}) + T(a_{k+1}\mathbf{v}_{k+1}) \\ &= a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k) + a_{k+1}T(\mathbf{v}_{k+1}), \end{aligned}$$

as desired. The second step is additivity, and the third applies the induction hypothesis to $T(\mathbf{u})$ and homogeneity to the last term. □

From now on we verify linearity through (ii) or (iii), and we use (i) whenever we apply a linear map to a linear combination. Neither half of (ii) implies the other: complex conjugation is additive but not homogeneous over $\mathbb{C}$ (Non-example 2.1.14), and Exercise 2.1.38 gives a homogeneous map that is not additive.

Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives). Let $T \in \mathrm{L}(V, W)$. Then $T(\mathbf{0}_V) = \mathbf{0}_W$, and for all $\mathbf{u}, \mathbf{v} \in V$,

$$T(-\mathbf{v}) = -T(\mathbf{v}) \quad \text{and} \quad T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}).$$

Proof (via homogeneity). ROUTINE

Uses. Theorem 2.1.3 (Checking linearity), Proposition 1.2.32 (The scalar zero), Proposition 1.2.34 (Negatives and the scalar -1)
Hint 1. Write $\mathbf{0}_V$ as a scalar multiple of some vector, so that homogeneity applies.
Hint 2. $\mathbf{0}_V = 0 \cdot \mathbf{0}_V$, and $0\mathbf{w} = \mathbf{0}_W$ for every $\mathbf{w} \in W$. For negatives, $-\mathbf{v} = (-1)\mathbf{v}$.
Sketch. $T(\mathbf{0}) = T(0 \cdot \mathbf{0}) = 0 \cdot T(\mathbf{0}) = \mathbf{0}$, using $0\mathbf{x} = \mathbf{0}$ in V and then in W. Similarly $T(-\mathbf{v}) = T((-1)\mathbf{v}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$, and the last identity follows by additivity.

By Proposition 1.2.32, $0\mathbf{x} = \mathbf{0}$ for every vector $\mathbf{x}$, in V and in W. So by homogeneity,

$$T(\mathbf{0}_V) = T(0 \cdot \mathbf{0}_V) = 0 \cdot T(\mathbf{0}_V) = \mathbf{0}_W.$$

By Proposition 1.2.34, $-\mathbf{x} = (-1)\mathbf{x}$ in V and in W, so homogeneity gives $T(-\mathbf{v}) = T((-1)\mathbf{v}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$. Finally, by additivity, $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}+(-\mathbf{v})) = T(\mathbf{u}) + T(-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$. □

Proof (via additivity). ROUTINE

Uses. Theorem 2.1.3 (Checking linearity), Definition 1.2.1 (Vector space), Proposition 1.2.28 (Uniqueness of additive inverses)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Write $\mathbf{0}_V$ as a sum, so that additivity applies.
Hint 2. $T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$. Now cancel. For negatives, show that $T(-\mathbf{v})$ is an additive inverse of $T(\mathbf{v})$.
Sketch. $T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, and adding $-T(\mathbf{0})$ gives $\mathbf{0} = T(\mathbf{0})$. Then $T(\mathbf{v}) + T(-\mathbf{v}) = T(\mathbf{v}-\mathbf{v}) = T(\mathbf{0}) = \mathbf{0}$, so $T(-\mathbf{v}) = -T(\mathbf{v})$ by uniqueness of additive inverses.

By additivity, $T(\mathbf{0}_V) = T(\mathbf{0}_V + \mathbf{0}_V) = T(\mathbf{0}_V) + T(\mathbf{0}_V)$. Writing $\mathbf{w} = T(\mathbf{0}_V)$, we have $\mathbf{w} = \mathbf{w} + \mathbf{w}$, so

$$\mathbf{0}_W = \mathbf{w} + (-\mathbf{w}) = (\mathbf{w}+\mathbf{w}) + (-\mathbf{w}) = \mathbf{w} + (\mathbf{w} + (-\mathbf{w})) = \mathbf{w} + \mathbf{0}_W = \mathbf{w}.$$

Next, for $\mathbf{v} \in V$, additivity gives $T(\mathbf{v}) + T(-\mathbf{v}) = T(\mathbf{v} + (-\mathbf{v})) = T(\mathbf{0}_V) = \mathbf{0}_W$. So $T(-\mathbf{v})$ is an additive inverse of $T(\mathbf{v})$, and by Proposition 1.2.28 it equals $-T(\mathbf{v})$. The identity $T(\mathbf{u}-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ follows from additivity as in the first proof. □

The first proof used only homogeneity and the second only additivity. So a function satisfying either condition alone sends $\mathbf{0}$ to $\mathbf{0}$; we use this for additive maps in Exercise 2.1.36 and Exercise 2.1.37. In practice Proposition 2.1.4 is a quick test in the contrapositive: a function with $T(\mathbf{0}) \neq \mathbf{0}$ is not linear. It is only a necessary condition. The map $\mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto xy$, sends $\mathbf{0}$ to 0, but it sends $(1, 1)$ to 1 and $2(1, 1) = (2, 2)$ to $4 \neq 2 \cdot 1$, so it is not homogeneous.

Exercise 2.1.5. ROUTINE Let W be a vector space over F, and regard F as a vector space over itself.

(a) Show that for every $\mathbf{w} \in W$, the map $F \to W$, $x \mapsto x\mathbf{w}$, is linear.

(b) Show that every $T \in \mathrm{L}(F, W)$ has this form, with $\mathbf{w} = T(1)$. In particular, the linear maps $F \to F$ are exactly the maps $x \mapsto ax$ with $a \in F$.

Solution on page 1104

2.1.2 Examples

Many familiar operations are linear maps. In each example below, linearity is checked with Theorem 2.1.3.

Example 2.1.6 (Zero, identity, and scaling maps). Let V and W be vector spaces over F.

- The **zero map** $V \to W$, $\mathbf{v} \mapsto \mathbf{0}_W$, is linear, since both sides of (iii) in Theorem 2.1.3 are $\mathbf{0}_W$ (by Proposition 1.2.33). We denote it by 0; the context tells it apart from $0 \in F$ and $\mathbf{0}_W$.
- The **identity map** $\mathrm{id}_V : V \to V$, $\mathbf{v} \mapsto \mathbf{v}$, is linear, since both sides of (iii) are $a\mathbf{u} + \mathbf{v}$.
- For $\lambda \in F$, the **scaling map** $V \to V$, $\mathbf{v} \mapsto \lambda\mathbf{v}$, is linear:

$$\lambda(a\mathbf{u} + \mathbf{v}) = (\lambda a)\mathbf{u} + \lambda\mathbf{v} = (a\lambda)\mathbf{u} + \lambda\mathbf{v} = a(\lambda\mathbf{u}) + \lambda\mathbf{v}.$$

  The middle step is the commutativity of multiplication in F.

Example 2.1.7 (Maps $F^n \to F^m$ given by linear formulas). Let $A \in F^{m\times n}$ (Definition 1.2.7) have entries a_{ij}, and define $T_A : F^n \to F^m$ by

$$T_A(x_1, \ldots, x_n) = (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n).$$

That is, the i-th coordinate of $T_A(\mathbf{x})$ is $\sum_{j=1}^{n} a_{ij}x_j$. This map is linear. For $\mathbf{x}, \mathbf{y} \in F^n$ and $c \in F$, the i-th coordinate of $T_A(c\mathbf{x} + \mathbf{y})$ is

$$\sum_{j=1}^{n} a_{ij}(cx_j + y_j) = c\sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j,$$

which is the i-th coordinate of $cT_A(\mathbf{x}) + T_A(\mathbf{y})$. For example, $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (2x - y + 3z, x + 5z)$, is T_A with

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 5 \end{pmatrix}.$$

Over $\mathbb{F}_2$, the map $(x, y, z) \mapsto (x + y, y + z)$ is T_A with $A = \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right)$. The coordinate projections $(x_1, \ldots, x_n) \mapsto x_i$ are special cases too.

In Section 2.3 we write $T_A(\mathbf{x})$ as a matrix–vector product $A\mathbf{x}$ (Definition 2.3.8, Definition 2.3.15) and show that every linear map $F^n \to F^m$ is T_A for some A (Proposition 2.3.17).

Example 2.1.8 (Coordinates). Let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V. The coordinate map $V \to F^n$, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ (Definition 1.6.14), is linear; this is Proposition 1.6.19. In the other direction, for any list $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in V, basis or not, the map

$$F^n \to V, \quad (a_1, \ldots, a_n) \mapsto a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n,$$

is linear, because $\sum_i (ca_i + d_i)\mathbf{v}_i = c\sum_i a_i\mathbf{v}_i + \sum_i d_i\mathbf{v}_i$ by the vector space axioms. For the basis $\mathcal{B}$, the two maps undo each other: $[a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n]_{\mathcal{B}} = (a_1, \ldots, a_n)$ by the definition of coordinates, and $\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$ where $(c_1, \ldots, c_n) = [\mathbf{v}]_{\mathcal{B}}$. So once a basis is chosen, V and F^n correspond through a pair of linear maps; Section 4.5 develops this.

Example 2.1.9 (Maps of polynomials). Let $p = a_0 + a_1x + \cdots + a_nx^n \in F[x]$ (Definition 1.2.13).

- **Evaluation** at $t \in F$ is $\mathrm{ev}_t : F[x] \to F$, $\mathrm{ev}_t(p) = p(t) = a_0 + a_1t + \cdots + a_nt^n$.
- **Multiplication by** x is $M : F[x] \to F[x]$, $M(p) = xp = a_0x + a_1x^2 + \cdots + a_nx^{n+1}$.
- **Differentiation** is $D : F[x] \to F[x]$,

$$D(p) = a_1 + 2a_2x + 3a_3x^2 + \cdots + na_nx^{n-1},$$

  where ka_k means $a_k + \cdots + a_k$ (k terms). For $F = \mathbb{R}$ this is the derivative of calculus, and over any field we take the formula as the definition (the **formal derivative**). We also write p' for $D(p)$.

All three maps are linear, with the same proof. Let $q = b_0 + b_1x + \cdots + b_nx^n$ (padding with zero coefficients, we may use the same n for p and q) and $c \in F$. Polynomials are added and scaled coefficientwise, so $cp + q = \sum_k (ca_k + b_k)x^k$. Each of the three maps sends the coefficient of x^k to a fixed place and multiplies it by a factor that does not depend on the polynomial (t^k, 1, or k). For D,

$$D(cp+q)=\sum_{k=1}^{n}k(ca_k+b_k)x^{k-1}=c\sum_{k=1}^{n}ka_kx^{k-1}+\sum_{k=1}^{n}kb_kx^{k-1}=cD(p)+D(q),$$

where $k(ca_k+b_k)=c(ka_k)+kb_k$ because ky is a sum of k copies of y. The computations for ev_t and M are the same with t^k and x^{k+1} in place of kx^{k-1}.

Over a finite field the formal derivative behaves unlike the one from calculus. Over $\mathbb{F}_2$, for example, $D\big(x^2\big)=2x=0$, so a nonconstant polynomial can have derivative 0 (Exercise 2.1.34).

Example 2.1.10 (Integration). Let $C[0,1]$ be the real vector space of continuous functions $[0,1]\to\mathbb{R}$ (Section 1.2). The map

$$C[0,1]\to\mathbb{R},\quad f\mapsto\int_0^1 f(t)\,\mathrm{d}t,$$

is linear: $\int_0^1(cf+g)=c\int_0^1 f+\int_0^1 g$ is the linearity of the integral. Likewise $J:C[0,1]\to C[0,1]$, $(Jf)(x)=\int_0^x f(t)\,\mathrm{d}t$, is linear. Here Jf is continuous (even differentiable) by the fundamental theorem of calculus, and $J(cf+g)=cJf+Jg$ because the two functions agree at every $x\in[0,1]$.

Linear maps of the plane

Rotations, projections, and reflections of $\mathbb{R}^2$ that fix the origin are linear. In each case we derive a coordinate formula from the geometry. The formula has the shape of Example 2.1.7, so the map is linear.

Example 2.1.11 (Rotations, projections, and reflections). **Rotation.** Let $R_\theta:\mathbb{R}^2\to\mathbb{R}^2$ rotate the plane counterclockwise about $\mathbf{0}$ by the angle θ. Write a point in polar form, $(x,y)=(r\cos\varphi,r\sin\varphi)$. Rotation adds θ to the angle, so by the angle addition formulas,

$$\begin{aligned}R_\theta(x,y)&=(r\cos(\varphi+\theta),r\sin(\varphi+\theta))\\&=(r\cos\varphi\cos\theta-r\sin\varphi\sin\theta,r\sin\varphi\cos\theta+r\cos\varphi\sin\theta)\\&=(x\cos\theta-y\sin\theta,x\sin\theta+y\cos\theta).\end{aligned}$$

Projection. Let ℓ be the line through $\mathbf{0}$ in the direction of the unit vector $\mathbf{u}=(\cos\theta,\sin\theta)$. The projection $\mathrm{Pr}_\ell:\mathbb{R}^2\to\mathbb{R}^2$ sends $\mathbf{v}=(x,y)$ to the foot $t\mathbf{u}$ of the perpendicular from $\mathbf{v}$ to ℓ. Two vectors (a,b) and (c,d) are perpendicular when $ac+bd=0$, so t is determined by

$$(x-t\cos\theta)\cos\theta+(y-t\sin\theta)\sin\theta=0,\quad\text{that is,}\quad t=x\cos\theta+y\sin\theta,$$

using $\cos^2\theta+\sin^2\theta=1$. Hence

$$\begin{aligned}\mathrm{Pr}_\ell(x,y)&=(x\cos\theta+y\sin\theta)(\cos\theta,\sin\theta)\\&=\big(x\cos^2\theta+y\sin\theta\cos\theta,x\sin\theta\cos\theta+y\sin^2\theta\big).\end{aligned}$$

Reflection. The reflection S_ℓ across ℓ sends $\mathbf{v}$ to the point $S_\ell(\mathbf{v})$ for which $\mathrm{Pr}_\ell(\mathbf{v})$ is the midpoint of $\mathbf{v}$ and $S_\ell(\mathbf{v})$. So $S_\ell(\mathbf{v})=2\,\mathrm{Pr}_\ell(\mathbf{v})-\mathbf{v}$, and with the double angle formulas,

$$S_\ell(x,y)=(x\cos2\theta+y\sin2\theta,x\sin2\theta-y\cos2\theta).$$

All three formulas have the form of Example 2.1.7, so R_θ, Pr_ℓ, and S_ℓ are linear. For instance, if ℓ is the x-axis ($\theta=0$), then $\mathrm{Pr}_\ell(x,y)=(x,0)$ and $S_\ell(x,y)=(x,-y)$, and the rotation by a right angle is $R_{\pi/2}(x,y)=(-y,x)$. Orthogonal projections in any inner product space are the subject of Section 6.4.

Intuition (What a linear map does to the grid). Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be linear. Since $(x,y)=x\mathbf{e}_1+y\mathbf{e}_2$, linearity gives

$$T(x,y)=xT(\mathbf{e}_1)+yT(\mathbf{e}_2).$$

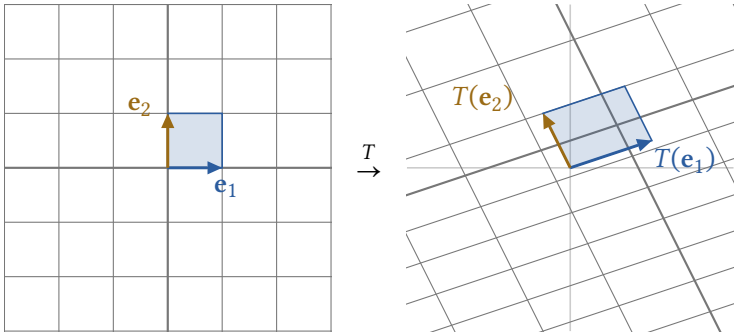
So T is determined by the two vectors $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ (Section 2.2 proves the analogue for every basis), and the grid point (x, y) goes to x steps along $T(\mathbf{e}_1)$ followed by y steps along $T(\mathbf{e}_2)$. The square grid becomes a grid of parallelograms spanned by $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$. To be precise about lines, apply T to the line through $\mathbf{p}$ in the direction $\mathbf{d}$:

$$T(\mathbf{p} + t\mathbf{d}) = T(\mathbf{p}) + tT(\mathbf{d}) \quad (t \in \mathbb{R}).$$

This is the line through $T(\mathbf{p})$ in the direction $T(\mathbf{d})$, or the single point $T(\mathbf{p})$ if $T(\mathbf{d}) = \mathbf{0}$. Reading off this formula:

- lines go to lines, or collapse to points;
- parallel lines, which share a direction $\mathbf{d}$, go to lines that share the direction $T(\mathbf{d})$ (they may coincide, or all collapse to points if $T(\mathbf{d}) = \mathbf{0}$);
- evenly spaced points $\mathbf{p}, \mathbf{p} + \mathbf{d}, \mathbf{p} + 2\mathbf{d}, \ldots$ go to evenly spaced points;
- the origin stays where it is.

Drag the tips of $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ below (or focus one and use the arrow keys) and watch the grid follow. The panel lists the four numbers in $T(x, y) = (ax + by, cx + dy)$ as an array A, the notation of Section 2.3; the handles labelled $A\mathbf{e}_1$ and $A\mathbf{e}_2$ are $T(\mathbf{e}_1) = (a, c)$ and $T(\mathbf{e}_2) = (b, d)$. The determinant readout belongs to Section 5.1 and can be ignored for now.



A linear map of the plane. Grid lines go to evenly spaced parallel lines, the unit square goes to the parallelogram spanned by $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$, and the origin stays fixed. (Interactive in the web version.)

2.1.3 Non-examples

The quickest way to see that a map is not linear is to find that it moves the origin (Proposition 2.1.4). The simplest maps that do this are the affine maps.

Non-example 2.1.12 (Affine maps). Let $a, b \in F$. The map $f : F \to F$, $f(x) = ax + b$, is linear if and only if $b = 0$. Indeed, if $b \neq 0$, then $f(0) = b \neq 0$, so f is not linear by Proposition 2.1.4; if $b = 0$, then $f(x) = ax$ is linear by Exercise 2.1.5. So the "linear functions" of school algebra, whose graphs are lines, are linear maps only when the line passes through the origin.

More generally, for $L \in \mathrm{L}(V, W)$ and $\mathbf{b} \in W$, the map $T : V \to W$, $T(\mathbf{v}) = L(\mathbf{v}) + \mathbf{b}$, is called **affine**. Since $T(\mathbf{0}) = \mathbf{b}$, it is linear only if $\mathbf{b} = \mathbf{0}$.

Intuition. The failure of an affine map is more specific than "$T(\mathbf{0}) \neq \mathbf{0}$". Compute both sides of the defining equation of linearity for $T = L + \mathbf{b}$: the left side picks up one copy of $\mathbf{b}$, and the right side picks up $c_1 + \cdots + c_k$ copies. So, when $\mathbf{b} \neq \mathbf{0}$, T respects exactly those combinations whose coefficients add up to 1.

Proposition 2.1.13 (Which combinations an affine map respects). Let $L \in \mathrm{L}(V, W)$, $\mathbf{b} \in W$, and $T(\mathbf{v}) = L(\mathbf{v}) + \mathbf{b}$. For all $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and $c_1, \ldots, c_k \in F$,

$$T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) - (c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k)) = (1 - c_1 - \cdots - c_k)\mathbf{b}.$$

In particular, if $\mathbf{b} \neq \mathbf{0}$, then $T\left(\sum_i c_i\mathbf{v}_i\right) = \sum_i c_iT(\mathbf{v}_i)$ holds if and only if $c_1 + \cdots + c_k = 1$. So T is linear if and only if $\mathbf{b} = \mathbf{0}$.

Proof. ROUTINE

Uses. Definition 2.1.1 (Linear map), Proposition 1.2.35 (Zero products)
Hint 1. Expand both $T\left(\sum_i c_i\mathbf{v}_i\right)$ and $\sum_i c_iT(\mathbf{v}_i)$ using the formula for T and the linearity of L.
Hint 2. The terms involving L cancel. What remains is $\mathbf{b} - (c_1 + \cdots + c_k)\mathbf{b}$. When is $s\mathbf{b} = \mathbf{0}$ for a nonzero $\mathbf{b}$?
Sketch. $T(\sum c_i\mathbf{v}_i) = \sum c_iL(\mathbf{v}_i) + \mathbf{b}$ and $\sum c_iT(\mathbf{v}_i) = \sum c_iL(\mathbf{v}_i) + (\sum c_i)\mathbf{b}$; subtract. If $\mathbf{b} \neq \mathbf{0}$, then $(1 - \sum c_i)\mathbf{b} = \mathbf{0}$ forces $\sum c_i = 1$. The combination $0 \cdot \mathbf{v}$ has coefficient sum 0, so T is not linear.

By the linearity of L,

$$T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = c_1L(\mathbf{v}_1) + \cdots + c_kL(\mathbf{v}_k) + \mathbf{b},$$

while, by distributivity in W,

$$c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) = c_1L(\mathbf{v}_1) + \cdots + c_kL(\mathbf{v}_k) + (c_1 + \cdots + c_k)\mathbf{b}.$$

Subtracting gives the displayed identity. If $\mathbf{b} \neq \mathbf{0}$, then by Proposition 1.2.35, $(1 - c_1 - \cdots - c_k)\mathbf{b} = \mathbf{0}$ if and only if $c_1 + \cdots + c_k = 1$. In particular, the combination $0 \cdot \mathbf{v}$ (with $k = 1$ and $c_1 = 0$) is not respected, so T is not linear. If $\mathbf{b} = \mathbf{0}$, then $T = L$ is linear. □

Combinations whose coefficients add up to 1 are called **affine combinations**. Examples are the points $(1 - t)\mathbf{u} + t\mathbf{v}$ of the line through $\mathbf{u}$ and $\mathbf{v}$, and in particular the midpoint $\frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$ when $1 + 1 \neq 0$. An affine map respects these, which is why an affine map of $\mathbb{R}^2$, like a linear one, still sends lines to lines or points, parallel lines to lines with a common direction, and evenly spaced points to evenly spaced points. What it does not respect is anything that refers to the origin, such as the scalar multiple $2\mathbf{v}$, whose single coefficient is 2. Compare the first two panels of Figure 2.1.1: the affine map produces the same grid as the linear one, translated. Affine maps and affine combinations are the subject of Section 4.7.

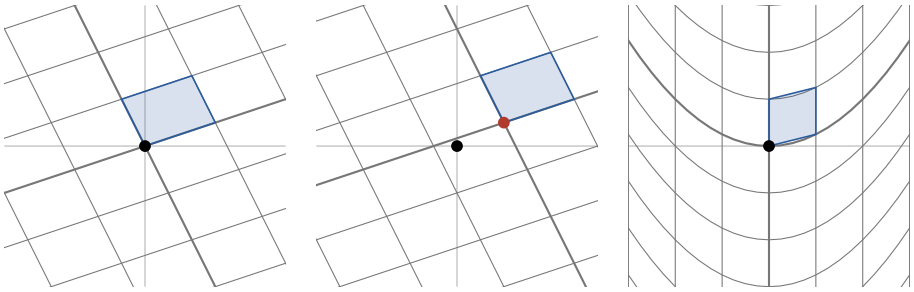


Figure 2.1.1. The square grid under three maps of $\mathbb{R}^2$; the faint lines are the original axes, and the shaded cell is the image of the unit square. Left: the linear map $T(x, y) = (1.5x - 0.5y, 0.5x + y)$. Middle: the affine map $\mathbf{v} \mapsto T(\mathbf{v}) + (1, 0.5)$, which moves the origin to the red point. Right: $(x, y) \mapsto \left(x, y + x^2/4\right)$, which fixes the origin but bends the horizontal grid lines, so it is not linear either.

Non-example 2.1.14 (Complex conjugation). Let $c : \mathbb{C} \to \mathbb{C}$ be complex conjugation, $c(a + bi) = \overline{a + bi} = a - bi$ for $a, b \in \mathbb{R}$. Regard $\mathbb{C}$ as a vector space over $\mathbb{R}$ (Section 1.2), with $a + bi$ corresponding to $(a, b) \in \mathbb{R}^2$. Then c is the map $(a, b) \mapsto (a, -b)$, the reflection in the real axis, which is $\mathbb{R}$-linear by Example 2.1.11.

Now regard $\mathbb{C}$ as a vector space over $\mathbb{C}$. Conjugation is still additive, since $\overline{z+w} = \overline{z} + \overline{w}$. But it is not homogeneous: $c(i \cdot 1) = \bar{i} = -i$, while $i \cdot c(1) = i$. So conjugation is $\mathbb{R}$-linear but not $\mathbb{C}$-linear. Whether a function is linear depends on the field of scalars, and not only on the sets and the function.

Formulas that are visibly nonlinear fail homogeneity at once. The map $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, sends $2 \cdot 1$ to $4 \neq 2 \cdot 1^2$, and $\mathbb{R} \to \mathbb{R}$, $x \mapsto \sin x$, sends $\pi = 2 \cdot \pi/2$ to $0 \neq 2\sin(\pi/2)$. As a rule of thumb, in a coordinate formula for a linear map, each output coordinate is a sum of constant multiples of the input coordinates: no constant terms, no products of coordinates, and no other functions applied to them. For maps $F^n \to F^m$, Proposition 2.3.17 makes this precise.

Exercise 2.1.15. ROUTINE Decide which of the following maps are linear.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x, y) = (y, x - 3y)$.

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x, y) = (x + 1, y)$.

(c) $T : \mathbb{R} \to \mathbb{R}^2$, $T(x) = (x, |x|)$.

(d) $T : \mathbb{F}_2^2 \to \mathbb{F}_2$, $T(x, y) = xy$.

Solution on page 1104

2.1.4 What linear maps preserve

Intuition. A vector space is a set with an addition and a scalar multiplication, and nothing else. Every notion we have defined so far (linear combination, span, linear relation, linear dependence) is defined from these two operations. A linear map respects both, so it carries each notion in V to the corresponding notion in W, at least in one direction. The direction matters: a linear map may send different vectors to the same vector, so it can destroy independence, but it never creates independence where there was none.

For a subset $A \subseteq V$ we write $T(A) = \{T(\mathbf{a}) \mid \mathbf{a} \in A\}$, and for a list $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ in V we write $T(\mathcal{L})$ for the list $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$ in W.

Proposition 2.1.16 (Linear maps preserve spans and linear relations). Let $T \in \mathrm{L}(V, W)$, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V.

(a) $\operatorname{im}\left(T|_{\operatorname{span}\mathcal{L}}\right) = \operatorname{span} T(\mathcal{L})$. In particular, if $\mathcal{L}$ spans V, then $T(\mathcal{L})$ spans $\operatorname{im} T$.

(b) If $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, then $a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k) = \mathbf{0}$. Consequently, if $\mathcal{L}$ is linearly dependent, then so is $T(\mathcal{L})$. Equivalently, if $T(\mathcal{L})$ is linearly independent, then so is $\mathcal{L}$.

Proof. ROUTINE

Uses. Definition 1.3.10 (Span), Definition 2.1.1 (Linear map), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For (a), prove the two inclusions separately. An element of $\operatorname{span}\mathcal{L}$ is a combination $\sum a_i\mathbf{v}_i$; what is its image?
Hint 2. For (b), apply T to both sides of the relation. For the statement about dependence, use the description of dependence by a nontrivial representation of $\mathbf{0}$ (Theorem 1.5.5).
Sketch. $T(\sum a_i\mathbf{v}_i) = \sum a_iT(\mathbf{v}_i)$ shows both that images of elements of $\operatorname{span}\mathcal{L}$ lie in $\operatorname{span} T(\mathcal{L})$, and that every element of $\operatorname{span} T(\mathcal{L})$ is such an image. Applying T to a nontrivial relation $\sum a_i\mathbf{v}_i = \mathbf{0}$ gives a relation among the $T(\mathbf{v}_i)$ with the same, not all zero, coefficients.

(a) If $k = 0$, then $\operatorname{span}\mathcal{L} = \{\mathbf{0}\}$ and $\operatorname{span} T(\mathcal{L}) = \{\mathbf{0}\}$, and $T(\{\mathbf{0}\}) = \{\mathbf{0}\}$ by Proposition 2.1.4. Let $k \geq 1$. We prove the two inclusions.

($\subseteq$) An element of $\operatorname{im}\left(T|_{\operatorname{span}\mathcal{L}}\right)$ has the form $T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k)$ with $a_i \in F$. By linearity it equals $a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k)$, which lies in $\operatorname{span} T(\mathcal{L})$.

($\supseteq$) An element of $\operatorname{span} T(\mathcal{L})$ has the form $a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k)$, which by linearity equals $T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k)$. Since $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \in \operatorname{span}\mathcal{L}$, this lies in $\operatorname{im}\left(T\restriction_{\operatorname{span}\mathcal{L}}\right)$.

If $\mathcal{L}$ spans V, then $\operatorname{im} T = \operatorname{im}\left(T\restriction_{\operatorname{span}\mathcal{L}}\right) = \operatorname{span} T(\mathcal{L})$, so $T(\mathcal{L})$ spans $\operatorname{im} T$.

(b) Applying T to both sides of $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ and using linearity and Proposition 2.1.4,

$$a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k) = T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = T(\mathbf{0}) = \mathbf{0}.$$

If $\mathcal{L}$ is linearly dependent, then by Theorem 1.5.5 there are scalars $a_1, \ldots, a_k$, not all 0, with $\sum_i a_i\mathbf{v}_i = \mathbf{0}$. By what we just showed, the same scalars give $\sum_i a_iT(\mathbf{v}_i) = \mathbf{0}$, a nontrivial representation of $\mathbf{0}$ by $T(\mathcal{L})$. So $T(\mathcal{L})$ is linearly dependent, again by Theorem 1.5.5. The last sentence of (b) is the contrapositive. □

The converse of the last statement fails. If ℓ is the x-axis, then Pr_ℓ sends the independent list $\mathbf{e}_1, \mathbf{e}_2$ to the dependent list $(1,0), (0,0)$, and the zero map sends every list to a list of zero vectors. The maps that preserve independence are identified in Section 2.6: they are the injective ones. Part (a) with $k = 1$ is the statement about lines from the grid picture: for $\mathbf{v} \neq \mathbf{0}$, the line $\operatorname{span}(\mathbf{v})$ through $\mathbf{0}$ goes to the line $\operatorname{span}(T(\mathbf{v}))$, or to $\{\mathbf{0}\}$ if $T(\mathbf{v}) = \mathbf{0}$.

2.1.5 Composition

Linear maps can be applied one after another. Recall that for functions $S : U \to V$ and $T : V \to W$, the **composition** $T \circ S : U \to W$ is $\mathbf{u} \mapsto T(S(\mathbf{u}))$ (Section A.1).

> **Proposition 2.1.17** (Compositions of linear maps are linear). Let U, V, and W be vector spaces over F. If $S \in \mathrm{L}(U, V)$ and $T \in \mathrm{L}(V, W)$, then $T \circ S \in \mathrm{L}(U, W)$.

Proof. ROUTINE

> *Uses.* Definition 2.1.1 (Linear map)
> *Hint 1.* Take a linear combination in U and follow it through S and then through T.
> *Hint 2.* Use the linearity of S to move the combination outside S, then the linearity of T to move it outside T.
> *Sketch.* $(T \circ S)(\sum a_i\mathbf{u}_i) = T(\sum a_iS(\mathbf{u}_i)) = \sum a_iT(S(\mathbf{u}_i))$.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_k \in U$ and $a_1, \ldots, a_k \in F$. Then

$$\begin{aligned}
(T \circ S)(a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k) &= T(S(a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k)) && \text{(definition of } \circ\text{)} \\
&= T(a_1S(\mathbf{u}_1) + \cdots + a_kS(\mathbf{u}_k)) && \text{(linearity of } S\text{)} \\
&= a_1T(S(\mathbf{u}_1)) + \cdots + a_kT(S(\mathbf{u}_k)) && \text{(linearity of } T\text{)} \\
&= a_1(T \circ S)(\mathbf{u}_1) + \cdots + a_k(T \circ S)(\mathbf{u}_k) && \text{(definition of } \circ\text{),}
\end{aligned}$$

as desired. □

> **Notation 2.1.18.** For linear maps we usually write TS for $T \circ S$. For $T \in \mathrm{L}(V)$ and an integer $k \geq 1$, $T^k = TT\cdots T$ (k factors), and $T^0 = \mathrm{id}_V$. Composition is associative (Proposition 2.1.24), so T^k needs no parentheses, and $T^jT^k = T^{j+k}$.

Example 2.1.19 (Composites).

(a) $D^2 = DD$ on $F[x]$ is the second formal derivative,

$$D^2(a_0 + a_1x + \cdots + a_nx^n) = 2a_2 + 6a_3x + \cdots + n(n-1)a_nx^{n-2}.$$

Each application of D sends constants to 0 and lowers the degree of a nonconstant polynomial by at least one (by exactly one over $\mathbb{R}$ or $\mathbb{C}$; over $\mathbb{F}_2$, $D(x^2 + x) = 1$). So $D^{n+1}(p) = 0$ whenever $\deg p \le n$.

(b) Order matters. Let $P(x, y) = (x, 0)$ be the projection onto the x-axis and $R(x, y) = (-y, x)$ the rotation by a right angle (Example 2.1.11). Then

$$(PR)(x, y) = P(-y, x) = (-y, 0), \quad (RP)(x, y) = R(x, 0) = (0, x).$$

So $PR \neq RP$: for instance $(PR)(\mathbf{e}_1) = \mathbf{0}$, while $(RP)(\mathbf{e}_1) = \mathbf{e}_2$.

Composition of linear maps is not commutative: even when TS and ST are both defined, they need not be equal. For matrices this becomes the noncommutativity of matrix multiplication (Section 2.4).

Exercise 2.1.20. ROUTINE Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^3$ be given by

$$S(x, y, z) = (x + y, y + z), \quad T(u, v) = (u - v, 2v, u).$$

Compute ST and TS.

Solution on page 1104

2.1.6 The vector space $\mathrm{L}(V, W)$

Functions with values in a vector space can be added and scaled value by value. For linear maps this makes $\mathrm{L}(V, W)$ a vector space, so everything we prove about vector spaces applies to spaces of linear maps.

Definition 2.1.21 (Pointwise operations). Let X be a set and W a vector space over F. For functions $f, g : X \to W$ and $c \in F$, define $f + g : X \to W$ and $cf : X \to W$ by

$$(f + g)(x) = f(x) + g(x), \quad (cf)(x) = cf(x) \quad (x \in X).$$

We write W^X for the set of all functions $X \to W$, with these operations.

Lemma 2.1.22 (Functions into a vector space). Let X be a set and W a vector space over F. Then W^X is a vector space over F. Its zero vector is the zero function $x \mapsto \mathbf{0}_W$, and the additive inverse of f is $x \mapsto -f(x)$.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), Definition 2.1.21 (Pointwise operations)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Each axiom is an equality between two functions $X \to W$. Two functions are equal when they agree at every point.
Hint 2. Evaluate both sides of an axiom at $x \in X$. You get the same axiom in W, applied to the vectors $f(x)$, $g(x)$, ...
Sketch. Each vector space axiom for W^X, evaluated at a point x, becomes the same axiom in W for the values at x. For example $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$.

The operations produce functions $X \to W$, so W^X is closed under them. Each vector space axiom asserts an equality between functions $X \to W$, and two such functions are equal when their values agree at every $x \in X$. At a point x, each axiom becomes the corresponding axiom of W. For example, for $f, g, h \in W^X$, $a, b \in F$, and $x \in X$:

$$\begin{aligned}
(f + g)(x) &= f(x) + g(x) = g(x) + f(x) = (g + f)(x), \\
((f + g) + h)(x) &= (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = (f + (g + h))(x), \\
(a(f + g))(x) &= a(f(x) + g(x)) = af(x) + ag(x) = (af + ag)(x),
\end{aligned}$$

$$((a+b)f)(x) = (a+b)f(x) = af(x) + bf(x) = (af+bf)(x),$$

and similarly $((ab)f)(x) = (ab)f(x) = a(bf(x)) = (a(bf))(x)$ and $(1f)(x) = 1f(x) = f(x)$. Let z be the zero function and $-f$ the function $x \mapsto -f(x)$. Then $(f+z)(x) = f(x) + \mathbf{0}_W = f(x)$ and $(f + (-f))(x) = f(x) + (-f(x)) = \mathbf{0}_W = z(x)$, so z is an additive identity and $-f$ is an additive inverse of f. □

For $W = F$ this is the vector space F^X of Section 1.2. Now $\mathrm{L}(V,W)$ is a subset of W^V.

Intuition. We could check the vector space axioms for $\mathrm{L}(V,W)$ one at a time, by evaluating at each $\mathbf{v}$. That would repeat the proof of Lemma 2.1.22 word for word. Passing through W^V separates the two things that are true. The axioms hold for all functions $V \to W$, linear or not; what is special about linear maps is only that the operations of W^V do not destroy linearity. The subspace criterion asks for exactly that.

Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space). Let V and W be vector spaces over F. If $S, T \in \mathrm{L}(V,W)$ and $c \in F$, then $S+T$ and cT are linear. Hence $\mathrm{L}(V,W)$ is a subspace of W^V, and in particular it is a vector space over F. Its zero vector is the zero map, and $-T$ is the map $\mathbf{v} \mapsto -T(\mathbf{v})$.

Proof. ROUTINE

Uses. Lemma 2.1.22 (Functions into a vector space), Theorem 1.4.3 (Subspace criterion), Theorem 2.1.3 (Checking linearity), Example 2.1.6 (Zero, identity, and scaling maps)
Hint 1. By Lemma 2.1.22 and the subspace criterion, it suffices to show that the zero map is linear and that $S+T$ and cT are linear.
Hint 2. Check (iii) of Theorem 2.1.3 for $S+T$: expand $(S+T)(a\mathbf{u}+\mathbf{v})$ by the definition of $S+T$, use linearity of S and of T, and regroup.
Hint 3. For cT you need to move c past a. Which field axiom allows that?
Sketch. The zero map is linear. For $S+T$, $(S+T)(a\mathbf{u}+\mathbf{v}) = aS(\mathbf{u}) + S(\mathbf{v}) + aT(\mathbf{u}) + T(\mathbf{v}) = a(S+T)(\mathbf{u}) + (S+T)(\mathbf{v})$. For cT, $c(aT(\mathbf{u}) + T(\mathbf{v})) = a(cT(\mathbf{u})) + cT(\mathbf{v})$ since $ca = ac$. So $\mathrm{L}(V,W)$ is a nonempty subset of the vector space W^V, closed under both operations.

By Lemma 2.1.22, W^V is a vector space, and $\mathrm{L}(V,W) \subseteq W^V$. We check the conditions of Theorem 1.4.3. The zero map is linear (Example 2.1.6), so $\mathrm{L}(V,W)$ is nonempty and contains the zero vector of W^V.

Let $S, T \in \mathrm{L}(V,W)$, $c \in F$, $\mathbf{u}, \mathbf{v} \in V$, and $a \in F$. Using the definition of $S+T$, then the linearity of S and T, then the vector space axioms of W to regroup,

$$\begin{aligned}(S+T)(a\mathbf{u}+\mathbf{v}) &= S(a\mathbf{u}+\mathbf{v}) + T(a\mathbf{u}+\mathbf{v}) \\ &= (aS(\mathbf{u}) + S(\mathbf{v})) + (aT(\mathbf{u}) + T(\mathbf{v})) \\ &= a(S(\mathbf{u}) + T(\mathbf{u})) + (S(\mathbf{v}) + T(\mathbf{v})) = a(S+T)(\mathbf{u}) + (S+T)(\mathbf{v}).\end{aligned}$$

Similarly,

$$\begin{aligned}(cT)(a\mathbf{u}+\mathbf{v}) &= c(aT(\mathbf{u}) + T(\mathbf{v})) = (ca)T(\mathbf{u}) + cT(\mathbf{v}) \\ &= (ac)T(\mathbf{u}) + cT(\mathbf{v}) = a(cT)(\mathbf{u}) + (cT)(\mathbf{v}),\end{aligned}$$

where the step $ca = ac$ is the commutativity of multiplication in F. By Theorem 2.1.3, $S+T$ and cT are linear. So $\mathrm{L}(V,W)$ is a subspace of W^V, hence a vector space, with the same zero vector and the same additive inverses as W^V. □

Linear maps now have three operations: addition, scalar multiplication, and composition. The next proposition records how they interact.

Proposition 2.1.24 (Algebra of composition). Let U, V, W, and X be vector spaces over F, and let $a \in F$.
(a) $(RS)T = R(ST)$ for $T \in \mathrm{L}(U,V)$, $S \in \mathrm{L}(V,W)$, and $R \in \mathrm{L}(W,X)$.
(b) $\mathrm{id}_W T = T = T\mathrm{id}_V$ for $T \in \mathrm{L}(V,W)$.
(c) $(S_1+S_2)T = S_1T + S_2T$ and $S(T_1+T_2) = ST_1 + ST_2$ for $T, T_1, T_2 \in \mathrm{L}(U,V)$ and $S, S_1, S_2 \in \mathrm{L}(V,W)$.
(d) $(aS)T = a(ST) = S(aT)$ for $T \in \mathrm{L}(U,V)$ and $S \in \mathrm{L}(V,W)$.

Proof. ROUTINE

> *Uses.* Proposition 2.1.23 (L(V,W) is a vector space), Definition 2.1.21 (Pointwise operations), Theorem 2.1.3 (Checking linearity)
> *Hint 1.* Two maps are equal when they agree at every vector. Evaluate both sides of each identity at $\mathbf{u}$.
> *Hint 2.* Most identities follow from the definitions alone. In $S(T_1+T_2)=ST_1+ST_2$ and $S(aT)=a(ST)$ you need a property of S.
> *Sketch.* Evaluate at $\mathbf{u}$. Associativity and the identity laws hold for all functions. $((S_1+S_2)T)(\mathbf{u})=S_1(T(\mathbf{u}))+S_2(T(\mathbf{u}))$ by the definition of S_1+S_2. $(S(T_1+T_2))(\mathbf{u})=S(T_1(\mathbf{u})+T_2(\mathbf{u}))$ splits by additivity of S, and $S(aT(\mathbf{u}))=aS(T(\mathbf{u}))$ by homogeneity of S.

In each case both sides are maps with the same domain and codomain, so it suffices to show that they agree at every vector $\mathbf{u}$ of the domain.

(a) $((RS)T)(\mathbf{u})=(RS)(T(\mathbf{u}))=R(S(T(\mathbf{u})))=R((ST)(\mathbf{u}))=(R(ST))(\mathbf{u})$.

(b) $(\mathrm{id}_W T)(\mathbf{u})=\mathrm{id}_W(T(\mathbf{u}))=T(\mathbf{u})$ and $(T\mathrm{id}_V)(\mathbf{u})=T(\mathrm{id}_V(\mathbf{u}))=T(\mathbf{u})$.

(c) By the definition of the sum of two maps,

$$((S_1+S_2)T)(\mathbf{u})=(S_1+S_2)(T(\mathbf{u}))=S_1(T(\mathbf{u}))+S_2(T(\mathbf{u}))=(S_1T+S_2T)(\mathbf{u}).$$

By the definition of T_1+T_2 and then additivity of S,

$$(S(T_1+T_2))(\mathbf{u})=S(T_1(\mathbf{u})+T_2(\mathbf{u}))=S(T_1(\mathbf{u}))+S(T_2(\mathbf{u}))=(ST_1+ST_2)(\mathbf{u}).$$

(d) $((aS)T)(\mathbf{u})=aS(T(\mathbf{u}))=(a(ST))(\mathbf{u})$ by the definition of aS, and by the definition of aT and homogeneity of S,

$$(S(aT))(\mathbf{u})=S(aT(\mathbf{u}))=aS(T(\mathbf{u}))=(a(ST))(\mathbf{u}).$$

□

Only the second identity in (c) and the last one in (d) used linearity, in both cases of the left factor S. For functions that are not linear they can fail: for $s(x)=x^2$ on $\mathbb{R}$, $(s\circ(\mathrm{id}_{\mathbb{R}}+\mathrm{id}_{\mathbb{R}}))(1)=s(2)=4$, while $(s\circ\mathrm{id}_{\mathbb{R}}+s\circ\mathrm{id}_{\mathbb{R}})(1)=2$. Parts (c) and (d) also say that composition is linear in each factor: for fixed $T\in\mathrm{L}(U,V)$, the map $\mathrm{L}(V,W)\to\mathrm{L}(U,W)$, $S\mapsto ST$, is linear, and for fixed $S\in\mathrm{L}(V,W)$, so is $\mathrm{L}(U,V)\to\mathrm{L}(U,W)$, $T\mapsto ST$.

Exercise 2.1.25. ROUTINE Let $T:F^2\to F^2, T(x,y)=(y,x-y)$. Show that $T^2+T=\mathrm{id}_{F^2}$, and deduce that $T\big(T+\mathrm{id}_{F^2}\big)=\big(T+\mathrm{id}_{F^2}\big)T=\mathrm{id}_{F^2}$.

Solution on page 1105

Assembling linear maps

The results of this section and the last give a way to prove linearity without computing.

Technique 2.1.26 (Assemble linear maps from linear pieces). To show that a map is linear, write it as a sum, scalar multiple, or composite of maps already known to be linear. By Proposition 2.1.23 and Proposition 2.1.17 the result is linear, with no computation.

Reach for it when a map is given by a formula built out of evaluations, derivatives, integrals, coordinates, projections, or other maps from the examples above. It also handles whole families at once. For instance, for $T\in\mathrm{L}(V)$ and $a_0,\dots,a_n\in F$, every $a_0\mathrm{id}_V+a_1T+\cdots+a_nT^n$ is linear. Conversely, a formula that involves a product of two inputs, a constant term, or a nonlinear function of a coordinate is a candidate for a counterexample; test $T(\mathbf{0})$ and $T(2\mathbf{v})$ first. The canonical example is the reflection $S_\ell=2\,\mathrm{Pr}_\ell-\mathrm{id}_{\mathbb{R}^2}$ (Example 2.1.27).

Used in. Exercise 2.1.29, Exercise 2.1.42, Exercise 2.1.47, Theorem 2.2.1, Theorem 2.3.20, Exercise 2.8.5, Proposition 4.6.7.

Example 2.1.27 (Maps assembled from linear pieces).

(a) By Example 2.1.11, $S_\ell(\mathbf{v}) = 2\operatorname{Pr}_\ell(\mathbf{v}) - \mathbf{v}$ for every $\mathbf{v}$. That is, $S_\ell = 2\operatorname{Pr}_\ell - \mathrm{id}_{\mathbb{R}^2}$ in $\mathrm{L}(\mathbb{R}^2)$. Since Pr_ℓ is linear, so is S_ℓ, and the double angle formulas are not needed for this.

(b) The map $\varphi : F[x] \to F$, $\varphi(p) = p(1) - 3p'(0)$, is linear, because $\varphi = \operatorname{ev}_1 - 3\operatorname{ev}_0 D$ with ev_t and D from Example 2.1.9.

(c) On the real vector space of infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$ (a subspace of $\mathbb{R}^{\mathbb{R}}$), differentiation $D : f \mapsto f'$ is linear, so $f \mapsto f'' - 3f' + 2f$ is linear, being $D^2 - 3D + 2\,\mathrm{id}$. This is why the differential equation $y'' - 3y' + 2y = 0$ is called linear: it asks for the vectors that a linear map sends to $\mathbf{0}$. By Section 2.6 such vectors always form a subspace.

Exercise 2.1.28. ROUTINE Fix $\mathbf{v}_0 \in V$. Show that the map $\mathrm{L}(V, W) \to W$, $T \mapsto T(\mathbf{v}_0)$, is linear.

Solution on page 1105

ASIDE: *Additive maps that are not linear* BEYOND 21-242

By Exercise 2.1.36 and Exercise 2.1.37, an additive map between vector spaces over $\mathbb{F}_p$ or over $\mathbb{Q}$ is automatically linear, because every scalar can be reached from 1 by addition, subtraction, and division by positive integers. Over $\mathbb{R}$ this is false, but no formula exhibits a counterexample. Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. It has a basis (a **Hamel basis**), by Zorn's lemma (Section 15.1). Pick a basis vector $\mathbf{b}$, and let $T : \mathbb{R} \to \mathbb{R}$ send each real number to its coordinate along $\mathbf{b}$. Then T is $\mathbb{Q}$-linear, hence additive, and it takes only rational values. If T were $\mathbb{R}$-linear, then $T(x) = xT(1)$ for all x (Exercise 2.1.5); since T takes only rational values, this forces $T(1) = 0$, and then $T = 0$, contradicting $T(\mathbf{b}) = 1$. Such a map is wildly discontinuous: an additive map $\mathbb{R} \to \mathbb{R}$ that is continuous is linear (Exercise 2.1.51).

Exercises

Exercise 2.1.29. ROUTINE Decide which of the following maps are linear. Here F is any field, $A \in F^{2\times 2}$ has entries a_{ij}, and $F^{2\times 2}$ is added and scaled entrywise.

(a) $F^{2\times 2} \to F$, $A \mapsto a_{11} + a_{22}$.

(b) $F^{2\times 2} \to F$, $A \mapsto a_{11}a_{22} - a_{12}a_{21}$.

(c) $F^{2\times 2} \to F^{2\times 2}$, the map that swaps the entries a_{12} and a_{21}.

(d) $F[x] \to F[x]$, $p \mapsto p + p'$.

(e) $\mathbb{C} \to \mathbb{C}$, $z \mapsto i\overline{z}$, with $\mathbb{C}$ regarded first as a vector space over $\mathbb{C}$, then as one over $\mathbb{R}$.

Solution on page 1105

Exercise 2.1.30. STANDARD How many functions $\mathbb{F}_2^2 \to \mathbb{F}_2$ are there? How many of them send $\mathbf{0}$ to 0? How many are linear? List the linear ones.

Solution on page 1106

Exercise 2.1.31. ROUTINE Let $T \in \mathrm{L}(V, W)$ and let $\mathcal{L}$ be a list in V. Prove or disprove each statement.

(a) If $\mathcal{L}$ spans V, then $T(\mathcal{L})$ spans W.

(b) If $T(\mathcal{L})$ spans W, then $\mathcal{L}$ spans V.

(c) If $T(\mathcal{L})$ spans W, then $\operatorname{im} T = W$.

Solution on page 1106

Exercise 2.1.32. ROUTINE Fix $q \in F[x]$. For $p = a_0 + a_1 x + \cdots + a_n x^n$, let $p(q) = a_0 + a_1 q + \cdots + a_n q^n \in F[x]$ be the result of substituting q for x. Show that the map $C_q : F[x] \to F[x]$, $p \mapsto p(q)$, is linear. (For $q = x + 1$, this is the shift $p(x) \mapsto p(x + 1)$.)

Solution on page 1106

Exercise 2.1.33. STANDARD Now substitute in the other order. For $q \in \mathbb{R}[x]$, let $L_q : \mathbb{R}[x] \to \mathbb{R}[x]$ be $p \mapsto q(p)$, the result of substituting p for x in q (notation of Exercise 2.1.32). Find all q for which L_q is linear.

Solution on page 1106

Exercise 2.1.34. ROUTINE Find all $p \in \mathbb{F}_2[x]$ with $D(p) = 0$, where D is the formal derivative. Do the same over $\mathbb{F}_3$ and over $\mathbb{R}$.

Solution on page 1107

Exercise 2.1.35. ROUTINE Find the error in the following argument.

Claim. Every additive map $T : \mathbb{R} \to \mathbb{R}$ is linear.

Proof. Let $c, x \in \mathbb{R}$. Then

$$T(cx) = T(x + x + \cdots + x) = T(x) + T(x) + \cdots + T(x) = cT(x),$$

so T is homogeneous. Since it is also additive, it is linear.

Solution on page 1107

Exercise 2.1.36. STANDARD Let V and W be vector spaces over F, and let $T : V \to W$ be additive. Prove or disprove that T is linear when

(a) $F = \mathbb{F}_p$ for a prime p;

(b) $F = \mathbb{C}$.

Solution on page 1107

Exercise 2.1.37. STANDARD Let V and W be vector spaces over $\mathbb{Q}$, and let $T : V \to W$ be additive. Show that T is linear.

Solution on page 1108

Exercise 2.1.38. STANDARD Let F be any field. Find a function $T : F^2 \to F$ that is homogeneous but not additive.

Solution on page 1108

Exercise 2.1.39. STANDARD Let $T : V \to W$ satisfy $T(\mathbf{0}) = \mathbf{0}$ and

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}) \quad \text{for all } \mathbf{u}, \mathbf{v} \in V \text{ and all } a, b \in F \text{ with } a + b = 1.$$

(a) Show that if $1 + 1 \neq 0$ in F, then T is linear.

(b) Show that the conclusion of (a) can fail over $\mathbb{F}_2$.

Solution on page 1108

Exercise 2.1.40. ROUTINE

(a) Find $S, T \in \mathrm{L}(\mathbb{R}^2)$ with $ST = 0$ but $TS \neq 0$.

(b) Find $T \in \mathrm{L}(\mathbb{R}^2)$ with $T \neq 0$ but $T^2 = 0$.

Solution on page 1109

Exercise 2.1.41. ROUTINE Find the error in the following argument.

Claim. If $P \in \mathrm{L}(V)$ satisfies $P^2 = P$, then $P = 0$ or $P = \mathrm{id}_V$.

Proof. Let $\mathbf{v} \in V$. Then $P(P(\mathbf{v}) - \mathbf{v}) = P^2(\mathbf{v}) - P(\mathbf{v}) = \mathbf{0}$. Since P is linear, either $P = 0$ or $P(\mathbf{v}) - \mathbf{v} = \mathbf{0}$. If $P \neq 0$, the second alternative holds for every $\mathbf{v}$, so $P = \mathrm{id}_V$.

Solution on page 1109

Exercise 2.1.42. STANDARD Let ℓ be the line through $\mathbf{0}$ at angle θ, and let $P = \mathrm{Pr}_\ell$ and $S = S_\ell$ be as in Example 2.1.11.

(a) Show from the formula for P that $P^2 = P$.

(b) Using only $S = 2P - \mathrm{id}$, (a), and Proposition 2.1.24, show that $S^2 = \mathrm{id}$ and $SP = PS = P$. Interpret these identities geometrically.

Solution on page 1109

Exercise 2.1.43. STANDARD For an angle α, let R_α be the rotation by α and S_α the reflection across the line through $\mathbf{0}$ at angle α (Example 2.1.11). Using the formulas, show that

$$R_\alpha R_\beta = R_{\alpha+\beta} \quad \text{and} \quad S_\alpha S_\beta = R_{2(\alpha-\beta)}.$$

Interpret the second identity geometrically.

Solution on page 1110

Exercise 2.1.44. ROUTINE Fix $\mathbf{v}_0 \in V$ and $\mathbf{w}_0 \in W$. Show that $\{T \in \mathrm{L}(V, W) \mid T(\mathbf{v}_0) = \mathbf{w}_0\}$ is a subspace of $\mathrm{L}(V, W)$ if and only if $\mathbf{w}_0 = \mathbf{0}$.

Solution on page 1110

Exercise 2.1.45. STANDARD Fix $A \in \mathrm{L}(V)$, and let $C(A) = \{T \in \mathrm{L}(V) \mid AT = TA\}$ be the set of operators that commute with A. Show that

(a) $C(A)$ is a subspace of $\mathrm{L}(V)$;

(b) if $S, T \in C(A)$, then $ST \in C(A)$;

(c) $a_0\mathrm{id}_V + a_1 A + \cdots + a_n A^n \in C(A)$ for all $a_0, \ldots, a_n \in F$.

Solution on page 1111

Exercise 2.1.46. STANDARD Let $M \in \mathrm{L}(F[x])$ be multiplication by x and D the formal derivative (Example 2.1.9).

(a) Show that $DM - MD = \mathrm{id}_{F[x]}$.

(b) Show that $DM^k - M^k D = kM^{k-1}$ for every $k \geq 1$, where kM^{k-1} means $(k \cdot 1)M^{k-1}$. What does this identity say when both sides are applied to a polynomial p?

Solution on page 1111

Exercise 2.1.47. STANDARD Regard $\mathbb{C}$ as a vector space over $\mathbb{R}$.

(a) Show that for every $\mathbb{R}$-linear $T : \mathbb{C} \to \mathbb{C}$ there are unique $\alpha, \beta \in \mathbb{C}$ with $T(z) = \alpha z + \beta \overline{z}$ for all $z \in \mathbb{C}$.

(b) Show that such a T is $\mathbb{C}$-linear if and only if $\beta = 0$, if and only if $T(i) = iT(1)$.

Solution on page 1111

Exercise 2.1.48. HARD Let $T \in \mathrm{L}(V)$ satisfy $T(\mathbf{v}) \in \operatorname{span}(\mathbf{v})$ for every $\mathbf{v} \in V$. Show that there is a $\lambda \in F$ with $T = \lambda \mathrm{id}_V$.

Solution on page 1112

Exercise 2.1.49. HARD A linear map $T : F[x] \to F[x]$ is a **derivation** if $T(pq) = T(p)q + pT(q)$ for all $p, q \in F[x]$.

(a) Show that the formal derivative D is a derivation (the product rule).

(b) Show that the derivations are exactly the maps $p \mapsto rD(p)$ for $r \in F[x]$, and that $r = T(x)$.

Solution on page 1112

Exercise 2.1.50. CHALLENGE Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$, and let $T \in \mathrm{L}(\mathbb{F}[x])$ commute with D, that is, $TD = DT$. Show that there are scalars $c_0, c_1, c_2, \ldots \in \mathbb{F}$ such that

$$T(p) = c_0 p + c_1 D(p) + c_2 D^2(p) + \cdots + c_n D^n(p) \quad \text{for every } p \text{ with } \deg p \le n.$$

(The terms $c_k D^k(p)$ with $k > \deg p$ are 0, so the right side does not depend on the choice of $n \ge \deg p$.) Conversely, show that every map of this form is linear and commutes with D.

Solution on page 1113

Exercise 2.1.51. HARD BEYOND 21-242 Let $T : \mathbb{R} \to \mathbb{R}$ be additive and continuous. Show that $T(x) = xT(1)$ for every $x \in \mathbb{R}$, so that T is linear.

Solution on page 1114

2.2 Linear maps are determined by a basis

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is linear, with $T(1,1) = (2,0)$ and $T(1,-1) = (0,4)$. Then T is already pinned down. For instance, $(3,1) = 2(1,1) + (1,-1)$, so linearity forces

$$T(3,1) = 2T(1,1) + T(1,-1) = (4,0) + (0,4) = (4,4),$$

and the same computation works at every vector, because every vector of $\mathbb{R}^2$ is a linear combination of $(1,1)$ and $(1,-1)$. Two questions remain. Is there a linear map with these two values at all? And if we wrote $(3,1)$ as a different combination, could the computation give a different answer? This section answers both questions for every basis of every vector space: a linear map may send the vectors of a basis anywhere, and once it has done so, it is determined. This is what will let us record a linear map by finitely many vectors, and in Section 2.3 by an array of scalars.

2.2.1 Existence and uniqueness

Intuition (Where the formula comes from). Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and suppose $T : V \to W$ is linear. Every $\mathbf{v} \in V$ is a linear combination $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, and linearity gives

$$T(\mathbf{v}) = a_1 T(\mathbf{v}_1) + \cdots + a_n T(\mathbf{v}_n).$$

So the n vectors $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ determine T. Now turn this around. Given vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n \in W$, a linear map with $T(\mathbf{v}_i) = \mathbf{w}_i$, if there is one, can only be

$$T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n.$$

Read this as a recipe: to find $T(\mathbf{v})$, write $\mathbf{v}$ in terms of the basis, then replace each $\mathbf{v}_i$ by $\mathbf{w}_i$. For the recipe to define a function $V \to W$, two things must hold. Every $\mathbf{v}$ must have a representation, or the recipe says nothing about $\mathbf{v}$; spanning provides this. And no $\mathbf{v}$ may have two different representations, or the recipe could produce two different answers; linear independence provides this. A basis provides both, and then it remains to check that the function we get is linear.

Theorem 2.2.1 (Linear maps are determined by a basis). Let V and W be vector spaces over F, let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and let $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be any list in W. Then there is exactly one linear map $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for $i = 1, \ldots, n$. It is given by

$$T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n \quad \text{for all } a_1, \ldots, a_n \in F. \tag{2.2.1}$$

That is, $T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$, where $(c_1, \ldots, c_n) = [\mathbf{v}]_{\mathcal{B}}$.

In other words, to specify a linear map out of V it suffices to say what it does to a basis of V, and every choice is allowed.

Proof, part 1 of 2: uniqueness. ROUTINE

Uses. Definition 1.3.10 (Span), Definition 2.1.1 (Linear map)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Suppose S and T both work. To compare them at an arbitrary $\mathbf{v} \in V$, write $\mathbf{v}$ in terms of the basis.
Hint 2. Apply the linearity of S and of T to $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Both give $a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n$.
Sketch. Since $\mathcal{B}$ spans V, every $\mathbf{v}$ is some $\sum_i a_i\mathbf{v}_i$. Linearity gives $S(\mathbf{v}) = \sum_i a_i\mathbf{w}_i = T(\mathbf{v})$. The same computation shows that any such map satisfies (2.2.1).

Suppose $S, T \in \mathrm{L}(V, W)$ both satisfy $S(\mathbf{v}_i) = \mathbf{w}_i = T(\mathbf{v}_i)$ for every i. We will show that $S(\mathbf{v}) = T(\mathbf{v})$ for every $\mathbf{v} \in V$. Let $\mathbf{v} \in V$. Since $\mathcal{B}$ spans V, there are scalars $a_1, \ldots, a_n$ with $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Then

$$\begin{aligned} S(\mathbf{v}) &= a_1S(\mathbf{v}_1) + \cdots + a_nS(\mathbf{v}_n) && \text{(linearity of } S\text{)} \\ &= a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n \\ &= a_1T(\mathbf{v}_1) + \cdots + a_nT(\mathbf{v}_n) = T(\mathbf{v}) && \text{(linearity of } T\text{)}. \end{aligned}$$

So $S = T$. The first two lines show more: every linear map that sends each $\mathbf{v}_i$ to $\mathbf{w}_i$ satisfies (2.2.1).

Intuition. The uniqueness proof did more than prove uniqueness: it found the only possible candidate, the map given by (2.2.1). This is a common way to organize an existence-and-uniqueness proof. First assume that a solution exists and work out what it must be, which proves uniqueness; then check that the forced candidate really is a solution, which proves existence. Here the check has three parts: (2.2.1) defines a function $V \to W$, this function sends $\mathbf{v}_i$ to $\mathbf{w}_i$, and it is linear. The uniqueness half used only that $\mathcal{B}$ spans V. Independence enters in the first part of the check.

Proof (directly), part 2 of 2: existence. ROUTINE

Uses. Definition 1.6.1 (Basis), Definition 1.5.1 (Linear independence), Definition 1.6.14 (Coordinates), Theorem 2.1.3 (Checking linearity)
Hint 1. Define $T(\mathbf{v})$ by (2.2.1), using the coordinates of $\mathbf{v}$. Why is this a single, well-defined vector of W?
Hint 2. What are the coordinates of $\mathbf{v}_i$ itself?
Hint 3. For linearity, if $\mathbf{u} = \sum_i b_i\mathbf{v}_i$ and $\mathbf{v} = \sum_i c_i\mathbf{v}_i$, then $a\mathbf{u} + \mathbf{v} = \sum_i (ab_i + c_i)\mathbf{v}_i$, and by uniqueness of representations this is *the* representation of $a\mathbf{u} + \mathbf{v}$.
Sketch. Each $\mathbf{v}$ has exactly one coordinate list $(c_1, \ldots, c_n)$ (at least one by spanning, at most one by independence), so $T(\mathbf{v}) = \sum_i c_i\mathbf{w}_i$ is a well-defined function. The coordinates of $\mathbf{v}_i$ are $\mathbf{e}_i$, so $T(\mathbf{v}_i) = \mathbf{w}_i$. The coordinates of $a\mathbf{u} + \mathbf{v}$ are $a[\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}$, which gives linearity.

T is a function $V \to W$. Let $\mathbf{v} \in V$. Since $\mathcal{B}$ is a basis, $\mathbf{v}$ can be written as a linear combination of $\mathcal{B}$ in exactly one way (Definition 1.6.1): in at least one way because $\mathcal{B}$ spans V, and in at most one way because $\mathcal{B}$ is linearly independent (Definition 1.5.1). Let $(c_1, \dots, c_n) = [\mathbf{v}]_{\mathcal{B}}$ be the coefficients (Definition 1.6.14), and define

$$T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n.$$

This assigns to each $\mathbf{v} \in V$ exactly one vector of W. If $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ is any representation, then $a_i = c_i$ for every i, so T satisfies (2.2.1).

$T(\mathbf{v}_i) = \mathbf{w}_i$. Since $\mathbf{v}_i = 0\mathbf{v}_1 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_n$, (2.2.1) gives $T(\mathbf{v}_i) = 0\mathbf{w}_1 + \cdots + 1\mathbf{w}_i + \cdots + 0\mathbf{w}_n = \mathbf{w}_i$.

T is linear. Let $\mathbf{u}, \mathbf{v} \in V$ and $a \in F$, and write $\mathbf{u} = \sum_i b_i\mathbf{v}_i$ and $\mathbf{v} = \sum_i c_i\mathbf{v}_i$. By the vector space axioms,

$$a\mathbf{u} + \mathbf{v} = \sum_{i=1}^{n} (ab_i + c_i)\mathbf{v}_i,$$

and since (2.2.1) holds for every representation,

$$T(a\mathbf{u} + \mathbf{v}) = \sum_{i=1}^{n} (ab_i + c_i)\mathbf{w}_i = a\sum_{i=1}^{n} b_i\mathbf{w}_i + \sum_{i=1}^{n} c_i\mathbf{w}_i = aT(\mathbf{u}) + T(\mathbf{v}).$$

By Theorem 2.1.3, T is linear. □

The recipe "take coordinates, then form the combination with the $\mathbf{w}_i$" is a composite of two maps that we already know to be linear. Seeing it that way gives a second proof of existence with no computation at all.

Proof (via coordinates), part 2 of 2: existence. ROUTINE

> *Uses.* Definition 1.6.14 (Coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Example 2.1.8 (Coordinates), Proposition 2.1.17 (Compositions of linear maps are linear)
> *Techniques.* Assemble linear maps from linear pieces (Technique 2.1.26)
> *Hint 1.* The coordinate map $V \to F^n$ is linear, and so is $F^n \to W$, $(a_1, \dots, a_n) \mapsto \sum_i a_i\mathbf{w}_i$.
> *Hint 2.* Compose them, and compute the composite at $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$.
> *Sketch.* Let $C(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ and $L(a_1, \dots, a_n) = \sum_i a_i\mathbf{w}_i$. Both are linear, so $T = LC$ is linear, and $T\left(\sum_i a_i\mathbf{v}_i\right) = L(a_1, \dots, a_n) = \sum_i a_i\mathbf{w}_i$.

Let $C : V \to F^n$ be the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$, which is linear by Proposition 1.6.19, and let $L : F^n \to W$ be the map $(a_1, \dots, a_n) \mapsto a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n$, which is linear by Example 2.1.8. By Proposition 2.1.17, $T = LC$ is linear. For scalars $a_1, \dots, a_n$, the definition of coordinates gives $[a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n]_{\mathcal{B}} = (a_1, \dots, a_n)$, so

$$T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = L(a_1, \dots, a_n) = a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n,$$

which is (2.2.1). In particular, taking $a_i = 1$ and the other coefficients 0, $T(\mathbf{v}_i) = \mathbf{w}_i$. □

In the second proof both halves of "basis" are hidden inside C: the coordinate vector $[\mathbf{v}]_{\mathcal{B}}$ is defined only because every vector has exactly one representation.

Nothing is assumed about $\mathbf{w}_1, \dots, \mathbf{w}_n$. They may be linearly dependent, or even all equal to $\mathbf{0}$, which gives the zero map. If $n = 0$, then $\mathcal{B}$ is the empty list, $V = \{\mathbf{0}\}$, and (2.2.1) with its empty sums says that the only linear map $\{\mathbf{0}\} \to W$ is the zero map.

Remark 2.2.2 (Where the two halves of "basis" were used). The proof of uniqueness used only that $\mathcal{B}$ spans V. The proof of existence used both halves: spanning so that the recipe applies to every $\mathbf{v}$, and independence so that it gives only one answer. Section 2.2.3 and Section 2.2.4 show what each half achieves on its own. A spanning list still gives uniqueness, but existence needs a condition on the $\mathbf{w}_i$. An independent list still gives existence, at least on its span, but if it does not span V, uniqueness fails (unless $W = \{\mathbf{0}\}$).

Geometrically, Theorem 2.2.1 says the following. A basis $\mathbf{b}_1, \mathbf{b}_2$ of $\mathbb{R}^2$ cuts the plane into a grid of parallelograms, and the grid point $a_1\mathbf{b}_1 + a_2\mathbf{b}_2$ is reached by a_1 steps along $\mathbf{b}_1$ and a_2 steps along $\mathbf{b}_2$. The linear

map with $T(\mathbf{b}_i) = \mathbf{w}_i$ sends it to the point reached by the same steps along $\mathbf{w}_1$ and $\mathbf{w}_2$ (Figure 2.2.1). This is the picture from Section 2.1, where $\mathbf{b}_1, \mathbf{b}_2$ was the standard basis.

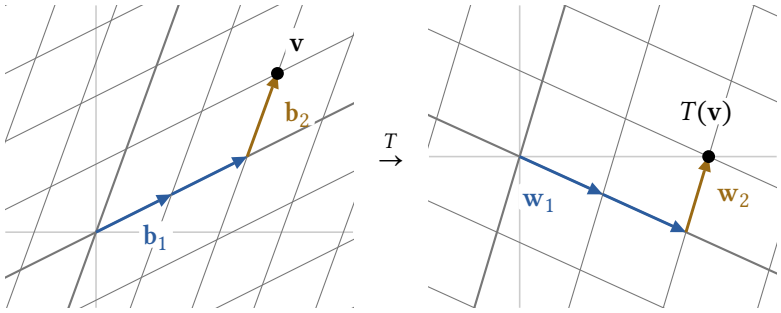


Figure 2.2.1. Left: the grid of $\mathcal{B}$-coordinates for the basis $\mathcal{B} = \mathbf{b}_1, \mathbf{b}_2$ of $\mathbb{R}^2$. The vector $\mathbf{v}$ with $[\mathbf{v}]_{\mathcal{B}} = (2, 1)$ is two steps along $\mathbf{b}_1$ and one along $\mathbf{b}_2$. Right: the linear map T with $T(\mathbf{b}_1) = \mathbf{w}_1$ and $T(\mathbf{b}_2) = \mathbf{w}_2$ carries the grid onto the grid built from $\mathbf{w}_1, \mathbf{w}_2$, and $T(\mathbf{v}) = 2\mathbf{w}_1 + \mathbf{w}_2$ is reached by the same steps.

In practice, Theorem 2.2.1 is used most often to build maps, by choosing a basis in which the desired map is easy to describe.

Example 2.2.3 (Building maps from a convenient basis). Let F be a field with $1 + 1 \neq 0$, and write $2 = 1 + 1$. In F^2 let $\mathbf{b}_1 = (1, 1)$ and $\mathbf{b}_2 = (1, -1)$. This is a basis of F^2. Indeed, $(x, y) = a\mathbf{b}_1 + b\mathbf{b}_2$ means $a + b = x$ and $a - b = y$. Adding and subtracting these equations gives $2a = x + y$ and $2b = x - y$, and since $2 \neq 0$,

$$a = \frac{x + y}{2}, \quad b = \frac{x - y}{2}.$$

Conversely these values satisfy both equations, so every (x, y) is a combination of $\mathbf{b}_1, \mathbf{b}_2$ in exactly one way.

By Theorem 2.2.1 there are unique linear maps $P, S : F^2 \to F^2$ with

$$P(\mathbf{b}_1) = \mathbf{b}_1, \quad P(\mathbf{b}_2) = \mathbf{0}, \quad S(\mathbf{b}_1) = \mathbf{b}_1, \quad S(\mathbf{b}_2) = -\mathbf{b}_2,$$

and (2.2.1) gives their formulas:

$$P(x, y) = \frac{x + y}{2}\mathbf{b}_1 = \left(\frac{x + y}{2}, \frac{x + y}{2}\right),$$
$$S(x, y) = \frac{x + y}{2}\mathbf{b}_1 - \frac{x - y}{2}\mathbf{b}_2 = (y, x).$$

Over $\mathbb{R}$, $\mathbf{b}_1$ spans the line ℓ given by $y = x$, and $\mathbf{b}_2$ is perpendicular to ℓ. So P fixes ℓ pointwise and kills the perpendicular direction: it is the projection Pr_ℓ. Similarly S fixes ℓ and flips the perpendicular direction: it is the reflection S_ℓ. The formulas agree with those of Example 2.1.11 for $\theta = \pi/4$, and this time no trigonometry was needed.

Over $\mathbb{F}_2$ the construction breaks down. There $\mathbf{b}_2 = (1, -1) = (1, 1) = \mathbf{b}_1$, so $\mathbf{b}_1, \mathbf{b}_2$ is not a basis, and indeed no linear map can send the same vector both to $\mathbf{b}_1$ and to $\mathbf{0}$.

> **Technique 2.2.4** (Construct a map by choosing it on a basis). To produce a linear map with some property, choose a basis of the domain that is adapted to the property, decide where each basis vector should go, and let Theorem 2.2.1 extend the choice to a linear map. Linearity then costs nothing, and the property itself can usually be checked on the basis (Technique 2.2.7). If a formula is wanted, write a general vector in the chosen basis and apply (2.2.1).
>
> Reach for it when the property is easy to state on special vectors (vectors to be fixed, sent to $\mathbf{0}$, swapped, or scaled), or when all you need is that some linear map with a given behaviour exists, as when building a

counterexample. The canonical example is the projection and reflection of Example 2.2.3. Another is that a subspace is the image of one linear map and the set of vectors sent to $\mathbf{0}$ by another (Exercise 2.2.35).
Used in. Proposition 2.2.11, Proposition 2.2.14, Exercise 2.2.28, Exercise 2.2.35, Exercise 2.2.40, Theorem 2.7.8, Theorem 2.7.4, Theorem 2.7.2, Exercise 2.7.29, Exercise 2.8.7, Problem 2.3, Problem 2.15, Proposition 3.5.9, Exercise 3.5.40, Proposition 4.1.19, Proposition 4.4.18, Exercise 4.4.42, Theorem 4.5.6, Theorem 4.5.8, Theorem 4.7.39, Problem 4.9, Problem 4.10, Corollary 8.4.7, Proposition 8.4.19, Exercise 7.2.58.

Exercise 2.2.5. ROUTINE Show that there is exactly one linear map $T : \mathbb{R}^2 \to \mathbb{R}$ with $T(1,1) = 1$ and $T(1,2) = 0$. Find $T(x,y)$ and $T(3,5)$.

Solution on page 1114

2.2.2 Checking on a basis

The proof of uniqueness in Theorem 2.2.1 used only that $\mathcal{B}$ spans V, so it proves more.

Corollary 2.2.6 (Agreeing on a spanning list). Let $S, T \in \mathrm{L}(V,W)$, and let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be a list that spans V. If $S(\mathbf{v}_i) = T(\mathbf{v}_i)$ for every i, then $S = T$. In particular, two linear maps that agree on a basis are equal.

Proof (by computation). ROUTINE
Uses. Definition 1.3.10 (Span), Definition 2.1.1 (Linear map)
Hint 1. Take any $\mathbf{v} \in V$ and write it as a combination of the spanning list.
Hint 2. This is the uniqueness half of the proof of Theorem 2.2.1, with a spanning list in place of a basis.
Sketch. Write $\mathbf{v} = \sum_i a_i \mathbf{v}_i$. Then $S(\mathbf{v}) = \sum_i a_i S(\mathbf{v}_i) = \sum_i a_i T(\mathbf{v}_i) = T(\mathbf{v})$.

Let $\mathbf{v} \in V$. Since $\mathbf{v}_1, \dots, \mathbf{v}_k$ spans V, we can write $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. By linearity of S and T,

$$S(\mathbf{v}) = \sum_{i=1}^{k} a_i S(\mathbf{v}_i) = \sum_{i=1}^{k} a_i T(\mathbf{v}_i) = T(\mathbf{v}).$$

As $\mathbf{v}$ was arbitrary, $S = T$. □

The second proof does not compute with a general vector at all. It looks at the set of vectors on which S and T agree, and shows that this set is a subspace containing the spanning list.

Proof (via a subspace). ROUTINE
Uses. Theorem 1.4.3 (Subspace criterion), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Hint 1. Consider $E = \{\mathbf{v} \in V \mid S(\mathbf{v}) = T(\mathbf{v})\}$. Which vectors do we know lie in E?
Hint 2. Show that E is a subspace. What is the smallest subspace containing $\mathbf{v}_1, \dots, \mathbf{v}_k$?
Sketch. $E = \{\mathbf{v} \mid S(\mathbf{v}) = T(\mathbf{v})\}$ contains $\mathbf{0}$ and is closed under addition and scalar multiplication, by linearity. It contains every $\mathbf{v}_i$, so it contains $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = V$.

Let $E = \{\mathbf{v} \in V \mid S(\mathbf{v}) = T(\mathbf{v})\}$. We check that E is a subspace of V with Theorem 1.4.3. It contains $\mathbf{0}$, since $S(\mathbf{0}) = \mathbf{0} = T(\mathbf{0})$ by Proposition 2.1.4. If $\mathbf{u}, \mathbf{v} \in E$ and $a \in F$, then

$$S(a\mathbf{u} + \mathbf{v}) = aS(\mathbf{u}) + S(\mathbf{v}) = aT(\mathbf{u}) + T(\mathbf{v}) = T(a\mathbf{u} + \mathbf{v}),$$

so $a\mathbf{u} + \mathbf{v} \in E$. By hypothesis, E contains $\mathbf{v}_1, \dots, \mathbf{v}_k$. Since the span of a list is the smallest subspace containing it (Proposition 1.4.14), $V = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) \subseteq E$. So $E = V$, that is, $S = T$. □

The second proof applies to more than equality. Whenever the vectors with some property form a subspace, the property holds on all of V as soon as it holds on a spanning list. (In the language of Section 2.6, E is the kernel of $S - T$.)

Technique 2.2.7 (Check it on a basis). To prove that two linear maps $S, T : V \to W$ are equal, it is enough to check $S(\mathbf{v}_i) = T(\mathbf{v}_i)$ for the vectors of a basis of V, or of any spanning list (Corollary 2.2.6). Many statements can be put in this form. An identity such as $P^2 = P$, $ST = TS$, or $T = a\,\mathrm{id} + bR$ is an

equation between linear maps, because sums, scalar multiples, and composites of linear maps are linear (Proposition 2.1.23, Proposition 2.1.17). A containment $T(V) \subseteq U$, for a subspace U of W, holds as soon as $T(\mathbf{v}_i) \in U$ for every vector of a spanning list, because $T(V)$ is the span of the $T(\mathbf{v}_i)$ (Proposition 2.1.16).

Reach for it when the basis vectors are easy to compute with, typically for a basis on which the maps involved act simply, and when a computation at a general vector would be messy. Before using it, make sure that both sides really are linear in the vector. A condition such as "$T(\mathbf{v})$ is a multiple of $\mathbf{v}$" is not an equation between linear maps, and it can hold on a basis without holding everywhere (Exercise 2.2.25). The canonical example is Corollary 2.2.6 itself.

Used in. Exercise 2.1.49, Exercise 2.2.9, Proposition 2.2.14, Exercise 2.2.23, Exercise 2.2.25, Exercise 2.2.28, Exercise 2.2.29, Exercise 2.2.30, Exercise 2.2.33, Exercise 2.2.34, Exercise 2.2.36, Exercise 2.2.39, Exercise 2.2.40, Exercise 2.3.38, Proposition 2.4.28, Exercise 2.4.49, Exercise 2.4.50, Proposition 2.6.18, Proposition 2.6.22, Theorem 2.7.8, Theorem 2.7.4, Theorem 2.7.2, Exercise 2.8.6, Exercise 2.8.8, Problem 2.3, Problem 2.4, Problem 2.5, Problem 2.6, Problem 2.7, Lemma 3.2.6, Proposition 3.5.9, Exercise 4.1.55, Theorem 4.5.8, Proposition 4.5.35, Proposition 6.1.22, Theorem 6.6.3, Corollary 7.1.12, Proposition 7.2.4, Proposition 7.2.27, Proposition 7.4.7, Proposition 7.4.26, Proposition 7.5.2, Exercise 7.6.48, Problem 7.9, Exercise 2.6.26, Exercise 3.5.40, Exercise 4.5.47.

Example 2.2.8 (Identities checked on a basis). Let P and S be the projection and reflection of Example 2.2.3. We claim that

$$P^2 = P, \quad S^2 = \mathrm{id}, \quad S = 2P - \mathrm{id}, \quad SP = PS = P.$$

Each side of each identity is a linear map $F^2 \to F^2$ (Proposition 2.1.17, Proposition 2.1.23), so by Corollary 2.2.6 it suffices to compare the two sides at $\mathbf{b}_1$ and at $\mathbf{b}_2$. Every map involved sends $\mathbf{b}_1$ to $\mathbf{b}_1$ (for $2P - \mathrm{id}$, because $2\mathbf{b}_1 - \mathbf{b}_1 = \mathbf{b}_1$), so both sides of each identity send $\mathbf{b}_1$ to $\mathbf{b}_1$. At $\mathbf{b}_2$,

$$\begin{aligned} P^2(\mathbf{b}_2) &= P(\mathbf{0}) = \mathbf{0} = P(\mathbf{b}_2), \\ S^2(\mathbf{b}_2) &= S(-\mathbf{b}_2) = \mathbf{b}_2, \\ (2P - \mathrm{id})(\mathbf{b}_2) &= \mathbf{0} - \mathbf{b}_2 = S(\mathbf{b}_2), \\ SP(\mathbf{b}_2) &= S(\mathbf{0}) = \mathbf{0} = P(\mathbf{b}_2), \\ PS(\mathbf{b}_2) &= P(-\mathbf{b}_2) = \mathbf{0} = P(\mathbf{b}_2). \end{aligned}$$

Exercise 2.1.42 proves the same identities over $\mathbb{R}$ from the coordinate formulas. In the basis $\mathbf{b}_1, \mathbf{b}_2$ each map does something simple to each basis vector, and the identities become one-line checks.

Exercise 2.2.9. ROUTINE Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and let $T \in \mathrm{L}(V)$ be the linear map with $T(\mathbf{v}_i) = \mathbf{v}_i + \mathbf{v}_1$ for every i (so $T(\mathbf{v}_1) = 2\mathbf{v}_1$). Show that $T^2 = 3T - 2\mathrm{id}_V$.

Solution on page 1114

2.2.3 Spanning lists and linear relations

Theorem 2.2.1 needs both halves of "basis". Here is what happens with a list that spans but is not independent.

Non-example 2.2.10 (A spanning list that is not independent). The list $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$ spans $\mathbb{R}^2$. Try to prescribe $T(\mathbf{e}_1) = \mathbf{w}_1$, $T(\mathbf{e}_2) = \mathbf{w}_2$, and $T(\mathbf{e}_1 + \mathbf{e}_2) = \mathbf{w}_3$ for a linear map $T : \mathbb{R}^2 \to W$.

Uniqueness survives, since its proof used only spanning: any such T is $T(x, y) = x\mathbf{w}_1 + y\mathbf{w}_2$. Existence fails unless $\mathbf{w}_3 = \mathbf{w}_1 + \mathbf{w}_2$, because a linear map satisfies $T(\mathbf{e}_1 + \mathbf{e}_2) = T(\mathbf{e}_1) + T(\mathbf{e}_2)$. In terms of the recipe: the vector $(1, 1)$ has the two representations $1 \cdot (\mathbf{e}_1 + \mathbf{e}_2)$ and $1 \cdot \mathbf{e}_1 + 1 \cdot \mathbf{e}_2$, and the recipe turns them into $\mathbf{w}_3$ and $\mathbf{w}_1 + \mathbf{w}_2$.

The obstruction is a linear relation. If $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, then every linear map satisfies $a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k) = \mathbf{0}$ (Proposition 2.1.16). So prescribed values must satisfy every linear relation that the vectors satisfy. For a spanning list, this is the only obstruction.

Proposition 2.2.11 (Prescribing values on a spanning list). Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list that spans V, and let $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$. There is a linear map $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for every i if and only if, for all $a_1, \ldots, a_k \in F$,

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \quad \Longrightarrow \quad a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k = \mathbf{0}. \tag{2.2.2}$$

In that case T is unique.

In words, the list $\mathbf{w}_1, \ldots, \mathbf{w}_k$ must satisfy every linear relation that $\mathcal{L}$ satisfies. If $\mathcal{L}$ is linearly independent, its only relation is the trivial one (Theorem 1.5.5), so the condition holds automatically, and we recover Theorem 2.2.1. There are two natural ways to prove the converse direction: repair the recipe directly, or throw away vectors until a basis remains.

Proof (directly). STANDARD

Uses. Proposition 2.1.16 (Linear maps preserve spans and linear relations), Corollary 2.2.6 (Agreeing on a spanning list), Theorem 2.1.3 (Checking linearity), Definition 1.3.10 (Span)
Hint 1. Necessity is Proposition 2.1.16. For sufficiency, define $T(\mathbf{v}) = \sum_i a_i\mathbf{w}_i$ for any representation $\mathbf{v} = \sum_i a_i\mathbf{v}_i$. Why does the choice of representation not matter?
Hint 2. If $\sum_i a_i\mathbf{v}_i = \sum_i b_i\mathbf{v}_i$, then $\sum_i(a_i - b_i)\mathbf{v}_i = \mathbf{0}$. Apply (2.2.2).
Hint 3. Once T is well defined, linearity and $T(\mathbf{v}_i) = \mathbf{w}_i$ follow because you may compute T from any representation you like.
Sketch. Necessity: apply T to a relation. Sufficiency: set $T\big(\sum_i a_i\mathbf{v}_i\big) = \sum_i a_i\mathbf{w}_i$. Two representations of the same vector differ by a relation of $\mathcal{L}$, which (2.2.2) turns into a relation of the $\mathbf{w}_i$, so the two values agree. Since any representation may be used, $T(\mathbf{v}_i) = \mathbf{w}_i$ and linearity follow as in Theorem 2.2.1. Uniqueness is Corollary 2.2.6.

Necessity. If T is linear with $T(\mathbf{v}_i) = \mathbf{w}_i$ and $\sum_i a_i\mathbf{v}_i = \mathbf{0}$, then $\sum_i a_i\mathbf{w}_i = \sum_i a_iT(\mathbf{v}_i) = \mathbf{0}$ by Proposition 2.1.16(b).

Sufficiency. Assume (2.2.2). Let $\mathbf{v} \in V$. Since $\mathcal{L}$ spans V, $\mathbf{v}$ has at least one representation $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. We claim that the vector $a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$ is the same for all representations of $\mathbf{v}$. Indeed, if also $\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k$, then subtracting,

$$(a_1 - b_1)\mathbf{v}_1 + \cdots + (a_k - b_k)\mathbf{v}_k = \mathbf{0},$$

so by (2.2.2), $\sum_i(a_i - b_i)\mathbf{w}_i = \mathbf{0}$, that is, $\sum_i a_i\mathbf{w}_i = \sum_i b_i\mathbf{w}_i$. So we may define $T : V \to W$ by

$$T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k,$$

computed from any representation. Using the representation $\mathbf{v}_i = 0\mathbf{v}_1 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_k$ gives $T(\mathbf{v}_i) = \mathbf{w}_i$. For linearity, let $\mathbf{u} = \sum_i a_i\mathbf{v}_i$, $\mathbf{v} = \sum_i b_i\mathbf{v}_i$, and $c \in F$. Then $c\mathbf{u} + \mathbf{v} = \sum_i(ca_i + b_i)\mathbf{v}_i$ is a representation of $c\mathbf{u} + \mathbf{v}$, so

$$T(c\mathbf{u} + \mathbf{v}) = \sum_{i=1}^{k}(ca_i + b_i)\mathbf{w}_i = cT(\mathbf{u}) + T(\mathbf{v}).$$

By Theorem 2.1.3, T is linear.

Uniqueness. Two such maps agree on the spanning list $\mathcal{L}$, so they are equal by Corollary 2.2.6. □

Proof (by reducing to a basis). STANDARD

Uses. Theorem 1.6.10 (Every spanning list contains a basis), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Some sublist of $\mathcal{L}$ is a basis of V (Theorem 1.6.10). Define T on that basis.

Hint 2. Each remaining $\mathbf{v}_j$ is a combination of the basis vectors. That expression is a linear relation among the vectors of $\mathcal{L}$; use (2.2.2) to compute $T(\mathbf{v}_j)$.
Sketch. Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be a sublist of $\mathcal{L}$ that is a basis of V, and let T send each $\mathbf{u}_r$ to its prescribed value. For any j, write $\mathbf{v}_j = \sum_r c_r \mathbf{u}_r$; this is a relation of $\mathcal{L}$, so the prescribed values satisfy it, and $T(\mathbf{v}_j) = \sum_r c_r T(\mathbf{u}_r)$ is the prescribed value $\mathbf{w}_j$.

Necessity and uniqueness are as in the first proof. For sufficiency, assume (2.2.2). By Theorem 1.6.10, some sublist $\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_n}$ of $\mathcal{L}$, with $i_1 < \cdots < i_n$, is a basis of V; write $\mathbf{u}_r = \mathbf{v}_{i_r}$. By Theorem 2.2.1 there is a linear map $T : V \to W$ with $T(\mathbf{u}_r) = \mathbf{w}_{i_r}$ for $r = 1, \dots, n$. We show that $T(\mathbf{v}_j) = \mathbf{w}_j$ for every j.

Fix j. Since $\mathbf{u}_1, \dots, \mathbf{u}_n$ spans V, there are scalars with $\mathbf{v}_j = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$. The equation

$$\mathbf{v}_j - c_1\mathbf{v}_{i_1} - \cdots - c_n\mathbf{v}_{i_n} = \mathbf{0}$$

is a linear relation among the vectors of $\mathcal{L}$ (collect the coefficient of each $\mathbf{v}_i$). By (2.2.2), the same combination of the $\mathbf{w}_i$ vanishes: $\mathbf{w}_j = c_1\mathbf{w}_{i_1} + \cdots + c_n\mathbf{w}_{i_n}$. Hence

$$T(\mathbf{v}_j) = c_1 T(\mathbf{u}_1) + \cdots + c_n T(\mathbf{u}_n) = c_1\mathbf{w}_{i_1} + \cdots + c_n\mathbf{w}_{i_n} = \mathbf{w}_j,$$

as desired. □

The first proof is the general method for defining a function by a formula that refers to a choice: show that the value does not depend on the choice. The second proof reduces to a case already handled, and it shows more than the statement does.

Remark 2.2.12 (Checking fewer relations). The second proof used (2.2.2) only for the $k - n$ relations $\mathbf{v}_j = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$ that express each vector of $\mathcal{L}$ outside the chosen basis in terms of the basis. So: if a sublist $\mathbf{u}_1, \dots, \mathbf{u}_n$ of $\mathcal{L}$ is a basis of V, and for each remaining $\mathbf{v}_j$ the prescribed values satisfy the relation expressing $\mathbf{v}_j$ in terms of $\mathbf{u}_1, \dots, \mathbf{u}_n$, then T exists. Every other relation of $\mathcal{L}$ is then satisfied automatically, by Proposition 2.1.16. In examples this is how existence is checked: one relation per redundant vector.

2.2.4 Prescribing values on any list

Now the other half: a list that is independent but does not span.

Non-example 2.2.13 (An independent list that does not span). The list $\mathbf{e}_1$ is linearly independent in $\mathbb{R}^2$ but does not span it. Prescribe $T(\mathbf{e}_1) = 1$ for a linear map $T : \mathbb{R}^2 \to \mathbb{R}$. Existence survives: $(x, y) \mapsto x$ works. Uniqueness fails: $(x, y) \mapsto x + cy$ works for every $c \in \mathbb{R}$. The recipe says nothing about $\mathbf{e}_2$, which has no representation in terms of $\mathbf{e}_1$, so $T(\mathbf{e}_2)$ can be anything.

In general, given vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ and $\mathbf{w}_1, \dots, \mathbf{w}_k \in W$, we can first define a map on $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ with Proposition 2.2.11, and then extend it to V by choosing values on vectors that complete a basis. The extension step needs such vectors to exist.

Proposition 2.2.14 (Prescribing values on any list). Let $\mathcal{L} = \mathbf{v}_1, \dots, \mathbf{v}_k$ be a list in V, and let $\mathbf{w}_1, \dots, \mathbf{w}_k \in W$. Suppose that V has a basis $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{x}_1, \dots, \mathbf{x}_r$ in which $\mathbf{u}_1, \dots, \mathbf{u}_m$ is a basis of $\operatorname{span}\mathcal{L}$.

(a) There is a linear map $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for every i if and only if (2.2.2) holds.

(b) If such a T exists, it is unique if and only if $\mathcal{L}$ spans V or $W = \{\mathbf{0}\}$.

Proof. STANDARD

Uses. Proposition 2.2.11 (Prescribing values on a spanning list), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Theorem 2.2.1 (Linear maps are determined by a basis), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
Hint 1. $\operatorname{span}\mathcal{L}$ is a vector space spanned by $\mathcal{L}$, so Proposition 2.2.11 gives a linear map $S : \operatorname{span}\mathcal{L} \to W$. Extend it with Theorem 2.2.1.
Hint 2. Let T agree with S on $\mathbf{u}_1, \dots, \mathbf{u}_m$ and send the $\mathbf{x}_j$ anywhere. Why does T agree with S on all of $\operatorname{span}\mathcal{L}$?

Hint 3. For (b), if $\mathcal{L}$ does not span V, then $r \geq 1$. Change the value at $\mathbf{x}_1$.
Sketch. Necessity in (a) is Proposition 2.1.16. Conversely, Proposition 2.2.11 applied to span $\mathcal{L}$ gives S with $S(\mathbf{v}_i) = \mathbf{w}_i$; let T agree with S on the $\mathbf{u}$'s and vanish on the $\mathbf{x}$'s. Then $T = S$ on $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = \operatorname{span} \mathcal{L}$. For (b), spanning gives uniqueness by Corollary 2.2.6; if $\mathcal{L}$ does not span, then $r \geq 1$, and changing $T(\mathbf{x}_1)$ to a nonzero vector gives a second map.

(a) If T exists, (2.2.2) holds by Proposition 2.1.16(b). Conversely, assume (2.2.2). The subspace span $\mathcal{L}$ is a vector space (Proposition 1.4.14), and $\mathcal{L}$ spans it. So Proposition 2.2.11, applied with span $\mathcal{L}$ in place of V, gives a linear map $S : \operatorname{span} \mathcal{L} \to W$ with $S(\mathbf{v}_i) = \mathbf{w}_i$ for every i. By Theorem 2.2.1 there is a linear map $T : V \to W$ with

$$T(\mathbf{u}_j) = S(\mathbf{u}_j) \quad (j = 1, \ldots, m), \quad T(\mathbf{x}_l) = \mathbf{0} \quad (l = 1, \ldots, r).$$

Every $\mathbf{v} \in \operatorname{span} \mathcal{L}$ is a combination $\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m$, and then $T(\mathbf{v}) = \sum_j c_j S(\mathbf{u}_j) = S(\mathbf{v})$. In particular $T(\mathbf{v}_i) = S(\mathbf{v}_i) = \mathbf{w}_i$.

(b) If $\mathcal{L}$ spans V, two such maps agree on a spanning list and are equal by Corollary 2.2.6. If $W = \{\mathbf{0}\}$, there is only one function $V \to W$ at all. Conversely, suppose that $\mathcal{L}$ does not span V and that W contains some $\mathbf{w} \neq \mathbf{0}$. Then $r \geq 1$: if $r = 0$, then $\mathbf{u}_1, \ldots, \mathbf{u}_m$ would be a basis of V, and $V = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = \operatorname{span} \mathcal{L}$. Let T be the map built in (a), and let T' be the linear map with $T'(\mathbf{u}_j) = S(\mathbf{u}_j)$ for all j, $T'(\mathbf{x}_1) = \mathbf{w}$, and $T'(\mathbf{x}_l) = \mathbf{0}$ for $l \geq 2$. The argument in (a) shows that $T'(\mathbf{v}_i) = \mathbf{w}_i$ for every i, but $T'(\mathbf{x}_1) = \mathbf{w} \neq \mathbf{0} = T(\mathbf{x}_1)$. So T is not unique. □

The hypothesis on V always holds when V is spanned by a finite list. By Theorem 1.6.10, span $\mathcal{L}$ has a basis taken from $\mathcal{L}$, and in Section 4.1 we will see that every linearly independent list in such a V extends to a basis (Theorem 4.1.16). In F^n one can also extend by hand, adding standard basis vectors one at a time. Without any finiteness assumption, the hypothesis still holds if one accepts the axiom of choice (Section 15.1).

Putting Proposition 2.2.11, Remark 2.2.12, and Proposition 2.2.14 together gives a procedure.

Algorithm 2.2.15 (Deciding whether a linear map with prescribed values exists)

Input Lists $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V and $\mathbf{w}_1, \ldots, \mathbf{w}_k$ in W.
Output Whether some $T \in \mathrm{L}(V, W)$ has $T(\mathbf{v}_i) = \mathbf{w}_i$ for every i, and whether it is unique.

1. Go through $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in order, and discard each vector that is a linear combination of the vectors kept before it. The kept vectors form a basis $\mathcal{B}$ of $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.
2. For each discarded $\mathbf{v}_j$, write $\mathbf{v}_j = \sum c_r \mathbf{u}_r$ with the $\mathbf{u}_r$ in $\mathcal{B}$, and check that the corresponding $\mathbf{w}$'s satisfy the same equation. If one check fails, no T exists. Otherwise a linear map with the prescribed values exists on $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.
3. Suppose every check passed. If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ spans V, then T exists and is unique. Otherwise extend $\mathcal{B}$ to a basis $\mathcal{B}, \mathbf{x}_1, \ldots, \mathbf{x}_r$ of V. Then T exists, its values at $\mathbf{x}_1, \ldots, \mathbf{x}_r$ can be chosen freely, and it is not unique unless $W = \{\mathbf{0}\}$.

In step 1, the kept vectors span the same subspace as the whole list, since each discarded vector is a combination of kept ones. They are linearly independent: otherwise, by the linear dependence lemma (Lemma 1.5.10), some kept vector would be a combination of the kept vectors before it, and it would have been discarded. Applied to $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, Remark 2.2.12 justifies step 2, and Proposition 2.2.14 justifies step 3.

Example 2.2.16 (Deciding existence and uniqueness).

(a) Is there a linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ with

$$T(1,2) = (1,1), \quad T(3,1) = (2,-1), \quad T(4,3) = (3,1)?$$

Keep $(1,2)$, and keep $(3,1)$, which is not a multiple of $(1,2)$. The two kept vectors form a basis of $\mathbb{R}^2$: the equations $a + 3b = x$ and $2a + b = y$ have the unique solution $b = (2x - y)/5$, $a = (3y - x)/5$. Then $(4,3) = (1,2) + (3,1)$ is discarded, and the check asks whether $(3,1) = (1,1) + (2,-1) = (3,0)$. It is not, so no such T exists.

(b) Change the last value to $T(4,3) = (3,0)$. Now the check passes, and the list spans $\mathbb{R}^2$, so exactly one T exists. With the coordinates from (a),

$$T(x,y) = \frac{3y - x}{5}(1,1) + \frac{2x - y}{5}(2,-1) = \left(\frac{3x + y}{5}, \frac{4y - 3x}{5}\right).$$

As a check, $T(1,2) = (1,1)$, $T(3,1) = (2,-1)$, and $T(4,3) = (3,0)$.

(c) Let F be any field. Is there a linear $T : F^3 \to F^2$ with

$$T(1,1,0) = (1,0), \quad T(0,1,1) = (0,1), \quad T(1,2,1) = (1,1)?$$

Keep $\mathbf{u}_1 = (1,1,0)$ and $\mathbf{u}_2 = (0,1,1)$, which is not a multiple of $\mathbf{u}_1$ (its first coordinate is 0 and its second is not). Discard $(1,2,1) = \mathbf{u}_1 + \mathbf{u}_2$; the check $(1,0) + (0,1) = (1,1)$ passes. The list does not span F^3: every vector $a\mathbf{u}_1 + b\mathbf{u}_2 = (a, a+b, b)$ of the span has middle coordinate equal to the sum of the other two, and $\mathbf{e}_1$ does not. Adding $\mathbf{e}_1$ gives a basis of F^3, because

$$(x,y,z) = a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{e}_1 = (a + c, a + b, b)$$

holds exactly when $b = z$, $a = y - z$, and $c = x - y + z$. So for each $\mathbf{c} \in F^2$ there is exactly one linear T with the prescribed values and $T(\mathbf{e}_1) = \mathbf{c}$, namely

$$T(x,y,z) = (y - z)(1,0) + z(0,1) + (x - y + z)\mathbf{c} = (y - z, z) + (x - y + z)\mathbf{c}.$$

Different choices of $\mathbf{c}$ give different maps, so T exists but is not unique.

2.2.5 Maps out of F^n

For $V = F^n$ with the standard basis, Theorem 2.2.1 takes a particularly concrete form.

Corollary 2.2.17 (Linear maps out of F^n). Let W be a vector space over F and $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$. There is exactly one linear map $T : F^n \to W$ with $T(\mathbf{e}_j) = \mathbf{w}_j$ for $j = 1, \dots, n$, namely

$$T(x_1, \dots, x_n) = x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n.$$

In particular, every $T \in \mathrm{L}(F^n, W)$ satisfies $T(\mathbf{x}) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n)$ for all $\mathbf{x} = (x_1, \dots, x_n) \in F^n$.

Proof. ROUTINE

Uses. Definition 1.6.8 (Standard bases), Theorem 2.2.1 (Linear maps are determined by a basis)
Hint 1. Apply Theorem 2.2.1 to the standard basis.
Hint 2. What are the coordinates of $(x_1, \dots, x_n)$ in the standard basis?
Sketch. $\mathbf{e}_1, \dots, \mathbf{e}_n$ is a basis of F^n, and $(x_1, \dots, x_n) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$. Apply (2.2.1).

The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ is a basis of F^n (Definition 1.6.8), and $(x_1, \dots, x_n) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$. So Theorem 2.2.1 gives exactly one linear T with $T(\mathbf{e}_j) = \mathbf{w}_j$, and (2.2.1) is the displayed formula. For an arbitrary $T \in \mathrm{L}(F^n, W)$, apply this with $\mathbf{w}_j = T(\mathbf{e}_j)$. □

Take $W = F^m$, and let a_{ij} be the i-th coordinate of $T(\mathbf{e}_j)$. Comparing i-th coordinates in $T(\mathbf{x}) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n)$ shows that the i-th coordinate of $T(\mathbf{x})$ is $a_{i1}x_1 + \cdots + a_{in}x_n$. So every linear map $F^n \to F^m$ is given by a linear formula as in Example 2.1.7, and the mn coefficients are the coordinates of $T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)$. Writing these n vectors side by side, as the columns of an $m \times n$ array, is the starting point of Section 2.3.

The corollary also explains the rotation formula of Example 2.1.11. Rotating $\mathbf{e}_1$ by θ gives $(\cos\theta, \sin\theta)$, and rotating $\mathbf{e}_2$ gives $(-\sin\theta, \cos\theta)$. Once we know that R_θ is linear, Corollary 2.2.17 gives $R_\theta(x, y) = x(\cos\theta, \sin\theta) + y(-\sin\theta, \cos\theta)$ without the angle addition formulas. (In Example 2.1.11 linearity was deduced from the formula, so this is a consistency check rather than an independent proof.)

Taken together, existence and uniqueness say that linear maps out of V correspond exactly to lists of vectors in W.

Corollary 2.2.18 (Linear maps correspond to lists of vectors). Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V. The map

$$T \mapsto (T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n))$$

is a bijection from $\mathrm{L}(V, W)$ onto the set of lists of length n in W. It respects the vector space operations: the list of $S + T$ is the entrywise sum of the lists of S and T, and the list of cT is c times the list of T.

Proof. ROUTINE

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 2.1.21 (Pointwise operations)
Hint 1. Surjectivity is the existence half of Theorem 2.2.1. What is injectivity?
Hint 2. For the operations, use the definition of $S + T$ and cT (Definition 2.1.21).
Sketch. Existence in Theorem 2.2.1 says every list is the list of some T; uniqueness says different maps have different lists. $(S + T)(\mathbf{v}_i) = S(\mathbf{v}_i) + T(\mathbf{v}_i)$ and $(cT)(\mathbf{v}_i) = cT(\mathbf{v}_i)$ by definition.

Given a list $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in W, the existence part of Theorem 2.2.1 gives $T \in \mathrm{L}(V, W)$ whose list is $\mathbf{w}_1, \ldots, \mathbf{w}_n$, so the map is surjective. If S and T have the same list, then $S = T$ by the uniqueness part, so the map is injective. Finally, by Definition 2.1.21, $(S + T)(\mathbf{v}_i) = S(\mathbf{v}_i) + T(\mathbf{v}_i)$ and $(cT)(\mathbf{v}_i) = cT(\mathbf{v}_i)$ for each i. □

Example 2.2.19 (Counting linear maps over $\mathbb{F}_p$). Let p be prime. The space $\mathbb{F}_p^m$ has p^m elements, so there are $(p^m)^n$ lists of length n in it. By Corollary 2.2.18, applied with the standard basis of $\mathbb{F}_p^n$,

$$\left|\mathrm{L}\left(\mathbb{F}_p^n, \mathbb{F}_p^m\right)\right| = p^{mn}.$$

For $n = 2$, $m = 1$, and $p = 2$ this gives the 4 linear maps $\mathbb{F}_2^2 \to \mathbb{F}_2$ found in Exercise 2.1.30. For comparison, there are $(p^m)^{p^n}$ functions $\mathbb{F}_p^n \to \mathbb{F}_p^m$ in all; for $\mathbb{F}_2^2 \to \mathbb{F}_2^2$ that is 256 functions, of which only 16 are linear.

2.2.6 Maps out of $F[x]$

No finite list spans $F[x]$ (Section 1.6), so Theorem 2.2.1 does not apply to it. But the monomials $1, x, x^2, \ldots$ behave like a basis: every polynomial is a combination of finitely many of them, and its coefficients are uniquely determined. The proof of Theorem 2.2.1 goes through unchanged.

Proposition 2.2.20 (Linear maps out of $F[x]$). Let W be a vector space over F, and let $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2, \ldots$ be any sequence of vectors in W. There is exactly one linear map $T : F[x] \to W$ with $T\left(x^k\right) = \mathbf{w}_k$ for every $k \geq 0$, namely

$$T\left(a_0 + a_1 x + \cdots + a_d x^d\right) = a_0\mathbf{w}_0 + a_1\mathbf{w}_1 + \cdots + a_d\mathbf{w}_d.$$

Proof. ROUTINE

Uses. Definition 1.2.13 (Polynomials), Notation 1.2.15, Theorem 2.1.3 (Checking linearity)
Hint 1. Imitate the proof of Theorem 2.2.1. What plays the role of the coordinates of a polynomial?
Hint 2. The coefficients of p are determined by p (Definition 1.2.13). Check that writing p with extra zero terms does not change the right side.
Sketch. Define $T(p) = \sum_k a_k\mathbf{w}_k$ over the coefficients a_k of p, a finite sum; extra zero terms contribute $0\mathbf{w}_k = \mathbf{0}$, so T is well defined. The operations on $F[x]$ are coefficientwise, which gives linearity, and x^k has a single coefficient 1, which gives $T\left(x^k\right) = \mathbf{w}_k$. A linear T' with $T'\left(x^k\right) = \mathbf{w}_k$ satisfies $T'(p) = \sum_k a_k T'\left(x^k\right) = T(p)$.

Let $p \in F[x]$ have coefficients $(a_0, a_1, \ldots)$, of which only finitely many are nonzero (Definition 1.2.13). Choose d with $a_k = 0$ for $k > d$, so that $p = a_0 + a_1 x + \cdots + a_d x^d$ (Notation 1.2.15), and define

$$T(p) = a_0\mathbf{w}_0 + a_1\mathbf{w}_1 + \cdots + a_d\mathbf{w}_d.$$

This does not depend on d, since increasing d only adds terms $0\mathbf{w}_k = \mathbf{0}$, and the coefficients are determined by p. So T is a function $F[x] \to W$.

Let $p, q \in F[x]$ have coefficients (a_k) and (b_k), let $c \in F$, and choose one d that works for both. The operations on $F[x]$ are coefficientwise, so $cp + q$ has coefficients $(ca_k + b_k)$, and

$$T(cp+q) = \sum_{k=0}^{d}(ca_k + b_k)\mathbf{w}_k = c\sum_{k=0}^{d} a_k\mathbf{w}_k + \sum_{k=0}^{d} b_k\mathbf{w}_k = cT(p) + T(q).$$

By Theorem 2.1.3, T is linear. The polynomial x^k has coefficient 1 in position k and 0 elsewhere, so $T\left(x^k\right) = \mathbf{w}_k$. Finally, if T' is linear and $T'\left(x^k\right) = \mathbf{w}_k$ for every k, then $T'(p) = \sum_{k=0}^{d} a_k T'\left(x^k\right) = \sum_{k=0}^{d} a_k\mathbf{w}_k = T(p)$, so $T' = T$. □

For example, the maps of Example 2.1.9 are the linear maps with $x^k \mapsto t^k$ (evaluation at t), $x^k \mapsto x^{k+1}$ (multiplication by x), and $x^k \mapsto kx^{k-1}$ (differentiation, with $1 \mapsto 0$), and the substitution map $p \mapsto p(q)$ of Exercise 2.1.32 is $x^k \mapsto q^k$. So Proposition 2.2.20 gives a second proof that these maps are linear. Its uniqueness half says that two linear maps out of $F[x]$ that agree on every monomial are equal, which is Technique 2.2.7 with the monomials in place of a basis; Exercise 2.1.49 used exactly this. The same argument works for any family of vectors such that every vector is a combination of them with uniquely determined coefficients, all but finitely many of which are 0. Such a family is a basis in the sense of Section 15.1.

Exercises

Exercise 2.2.21. ROUTINE

(a) Show that there is exactly one linear map $T : \mathbb{F}_5^2 \to \mathbb{F}_5^2$ with $T(1,2) = (1,0)$ and $T(2,1) = (0,1)$, and find $T(x,y)$.

(b) Show that there is no linear map $T : \mathbb{F}_3^2 \to \mathbb{F}_3^2$ with $T(1,2) = (1,0)$ and $T(2,1) = (0,1)$.

Solution on page 1115

Exercise 2.2.22. ROUTINE For each of the following, decide whether a linear map with the given values exists, and if so, whether it is unique.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(1,2) = (1,0)$ and $T(2,4) = (2,1)$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^2$ with $T(1,0,1) = (1,1)$, $T(0,1,1) = (2,0)$, and $T(1,1,2) = (3,1)$.

(c) $T : \mathbb{R}^2 \to \mathbb{R}$ with $T(1,1) = 2$, $T(1,-1) = 0$, and $T(3,1) = 4$.

(d) $T : \mathbb{F}_2^3 \to \mathbb{F}_2$ with $T(1,1,0) = T(0,1,1) = T(1,0,1) = 1$.

Solution on page 1115

Exercise 2.2.23. ROUTINE

(a) Show that the unique linear map $T : \mathbb{R}[x] \to \mathbb{R}$ with $T(x^n) = 1/(n+1)$ for every $n \geq 0$ is $p \mapsto \int_0^1 p(t)\,dt$.

(b) Let F be a field. Show that the unique linear map $U : F[x] \to F$ with $U(x^n) = n$ for every $n \geq 0$ (where n means $n \cdot 1 \in F$) is $p \mapsto p'(1)$. Compute $U\left(x^2 + x\right)$ over $\mathbb{R}$, over $\mathbb{F}_2$, and over $\mathbb{F}_3$.

Solution on page 1116

Exercise 2.2.24. ROUTINE Find the error in the following argument.

Claim. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span V, and let $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ be arbitrary. Then there is a linear map $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for every i.

Proof. For each $\mathbf{v} \in V$, choose scalars with $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, which is possible because the list spans V, and set $T(\mathbf{v}) = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$. Choosing $a_i = 1$ and $a_j = 0$ for $j \neq i$ shows $T(\mathbf{v}_i) = \mathbf{w}_i$. If $\mathbf{u} = \sum_i a_i\mathbf{v}_i$ and $\mathbf{v} = \sum_i b_i\mathbf{v}_i$, then $c\mathbf{u} + \mathbf{v} = \sum_i (ca_i + b_i)\mathbf{v}_i$, so $T(c\mathbf{u} + \mathbf{v}) = \sum_i (ca_i + b_i)\mathbf{w}_i = cT(\mathbf{u}) + T(\mathbf{v})$. So T is linear.

Solution on page 1116

Exercise 2.2.25. ROUTINE Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and let $T \in \mathrm{L}(V, W)$. Prove or disprove each statement.

(a) If $W = V$ and $T(\mathbf{v}_i) \in \operatorname{span}(\mathbf{v}_i)$ for every i, then $T(\mathbf{v}) \in \operatorname{span}(\mathbf{v})$ for every $\mathbf{v} \in V$.

(b) If U is a subspace of W and $T(\mathbf{v}_i) \in U$ for every i, then $T(V) \subseteq U$.

(c) If $T(\mathbf{v}_i) \neq \mathbf{0}$ for every i, then $T(\mathbf{v}) \neq \mathbf{0}$ for every nonzero $\mathbf{v} \in V$.

(d) If $T(\mathbf{v}_1) = T(\mathbf{v}_2) = \cdots = T(\mathbf{v}_n)$, then $T(V) \subseteq \operatorname{span}(T(\mathbf{v}_1))$.

Solution on page 1116

Exercise 2.2.26. STANDARD

(a) Find the error in the following argument.

Claim. There is a linear map $T : \mathbb{R}^2 \to \mathbb{R}$ with $T(1, 2) = 1$, $T(2, 1) = 1$, and $T(1, 1) = 1$.

Proof. Any two of the vectors $(1, 2)$, $(2, 1)$, $(1, 1)$ form a basis of $\mathbb{R}^2$. So by Theorem 2.2.1, for any two of the three prescribed values there is a linear map that takes them. Hence there is a linear map that takes all three.

(b) Find all $c \in \mathbb{R}$ for which some linear $T : \mathbb{R}^2 \to \mathbb{R}$ satisfies $T(1, 2) = T(2, 1) = 1$ and $T(1, 1) = c$.

Solution on page 1117

Exercise 2.2.27. STANDARD Regard $\mathbb{C}^2$ as a vector space over $\mathbb{C}$.

(a) Show that there is exactly one $\mathbb{C}$-linear map $T : \mathbb{C}^2 \to \mathbb{C}^2$ with $T(1, i) = (1, 0)$ and $T(1, -i) = (0, 1)$, and find $T(z_1, z_2)$.

(b) Show that there is no $\mathbb{C}$-linear map $T : \mathbb{C}^2 \to \mathbb{C}^2$ with $T(1, i) = (1, 0)$ and $T(i, -1) = (0, 1)$.

(c) Now regard $\mathbb{C}^2$ as a vector space over $\mathbb{R}$ (Proposition 1.2.19). Show that there is an $\mathbb{R}$-linear map $T : \mathbb{C}^2 \to \mathbb{C}^2$ with $T(1, i) = (1, 0)$ and $T(i, -1) = (0, 1)$.

Solution on page 1117

Exercise 2.2.28. STANDARD Let $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$. Show that there is exactly one linear map $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $T(\mathbf{u}) = \mathbf{u}$ for every $\mathbf{u} \in U$ and $T(1, 1, 1) = \mathbf{0}$. Find a formula for T, and show that $T^2 = T$.

Solution on page 1118

Exercise 2.2.29. STANDARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V with $n \geq 1$, and let $T \in \mathrm{L}(V)$ be the linear map with $T(\mathbf{v}_i) = \mathbf{v}_{i+1}$ for $i < n$ and $T(\mathbf{v}_n) = \mathbf{v}_1$. Let $\mathbf{s} = \mathbf{v}_1 + \cdots + \mathbf{v}_n$ and $N = \mathrm{id}_V + T + T^2 + \cdots + T^{n-1}$.

(a) Show that $T^n = \mathrm{id}_V$.

(b) Show that $N(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = (a_1 + \cdots + a_n)\mathbf{s}$ for all scalars a_i.

(c) Show that $TN = NT = N$.

Solution on page 1118

Exercise 2.2.30. STANDARD Let $R \in \mathrm{L}(\mathbb{R}^2)$ be the rotation $R(x, y) = (-y, x)$. Show that $T \in \mathrm{L}(\mathbb{R}^2)$ satisfies $TR = RT$ if and only if $T = a\,\mathrm{id} + bR$ for some $a, b \in \mathbb{R}$. Interpret the result using complex numbers.

Solution on page 1119

Exercise 2.2.31. STANDARD Let p be prime.

(a) How many $T \in \mathrm{L}(\mathbb{F}_p^2)$ satisfy $T(1, 1) \in \mathrm{span}((1, 1))$?

(b) How many $T \in \mathrm{L}(\mathbb{F}_p^2)$ satisfy both $T(1, 1) \in \mathrm{span}((1, 1))$ and $T(1, -1) \in \mathrm{span}((1, -1))$? (The answer depends on p.)

Solution on page 1119

Exercise 2.2.32. STANDARD Let $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a list in V.

(a) Suppose $W \neq \{\mathbf{0}\}$, and that for every list $\mathbf{w}_1, ..., \mathbf{w}_n$ in W there is a linear map $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for every i. Show that $\mathcal{L}$ is linearly independent.

(b) Conversely, show that if $\mathcal{L}$ is linearly independent, then for every vector space W and every $\mathbf{w}_1, ..., \mathbf{w}_n \in W$ there is a linear map $\mathrm{span}\,\mathcal{L} \to W$ sending each $\mathbf{v}_i$ to $\mathbf{w}_i$.

Solution on page 1120

Exercise 2.2.33. STANDARD Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V.

(a) Show that for each i there is exactly one linear map $\varphi_i : V \to F$ with $\varphi_i(\mathbf{v}_i) = 1$ and $\varphi_i(\mathbf{v}_j) = 0$ for $j \neq i$, and that $\varphi_i(\mathbf{v})$ is the i-th coordinate of $[\mathbf{v}]_{\mathcal{B}}$.

(b) Show that $\mathbf{v} = \varphi_1(\mathbf{v})\mathbf{v}_1 + \cdots + \varphi_n(\mathbf{v})\mathbf{v}_n$ for every $\mathbf{v} \in V$.

(c) Show that every $\varphi \in \mathrm{L}(V, F)$ equals $\varphi(\mathbf{v}_1)\varphi_1 + \cdots + \varphi(\mathbf{v}_n)\varphi_n$.

(The maps $\varphi_1, ..., \varphi_n$ form the *dual basis* of $\mathcal{B}$; see Section 13.1.)

Solution on page 1120

Exercise 2.2.34. STANDARD Let $\mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V and $\mathbf{w}_1, ..., \mathbf{w}_m$ a basis of W. For $1 \leq i \leq m$ and $1 \leq j \leq n$, let $E_{ij} \in \mathrm{L}(V, W)$ be the linear map with $E_{ij}(\mathbf{v}_j) = \mathbf{w}_i$ and $E_{ij}(\mathbf{v}_k) = \mathbf{0}$ for $k \neq j$. Show that the list of the mn maps

$$E_{11}, E_{21}, ..., E_{m1}, E_{12}, ..., E_{m2}, ..., E_{1n}, ..., E_{mn}$$

is a basis of $\mathrm{L}(V, W)$. (Once dimension is defined in Section 4.1, this says $\dim \mathrm{L}(V, W) = mn$.)

Solution on page 1120

Exercise 2.2.35. STANDARD Let U be a subspace of V, and suppose that V has a basis $\mathbf{u}_1, ..., \mathbf{u}_m, \mathbf{x}_1, ..., \mathbf{x}_r$ in which $\mathbf{u}_1, ..., \mathbf{u}_m$ is a basis of U. (By the results of Section 4.1, this holds whenever V is finite-dimensional.)

(a) Construct $P \in \mathrm{L}(V)$ with $P(V) = U$ and $P(\mathbf{u}) = \mathbf{u}$ for every $\mathbf{u} \in U$.

(b) Construct $S \in \mathrm{L}(V)$ with $\{\mathbf{v} \in V \mid S(\mathbf{v}) = \mathbf{0}\} = U$.

(c) Carry out (a) and (b) for $U = \mathrm{span}((1, 2, 3)) \subseteq \mathbb{R}^3$, and give formulas for P and S.

(In the language of Section 2.6, every such U is the image of a linear map and the kernel of another.)

Solution on page 1121

Exercise 2.2.36. HARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and let $T \in \mathrm{L}(V)$ satisfy $T(\mathbf{v}_1) = \mathbf{0}$ and $T(\mathbf{v}_k) \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$ for $2 \le k \le n$. Show that $T^n = 0$.

Solution on page 1121

Exercise 2.2.37. HARD Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in V and $\mathcal{M} = \mathbf{w}_1, \ldots, \mathbf{w}_k$ a list in W. Show that there is a bijective linear map $S : \operatorname{span}\mathcal{L} \to \operatorname{span}\mathcal{M}$ with $S(\mathbf{v}_i) = \mathbf{w}_i$ for every i if and only if $\mathcal{L}$ and $\mathcal{M}$ satisfy the same linear relations: for all $a_1, \ldots, a_k \in F$,

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \quad \Longleftrightarrow \quad a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k = \mathbf{0}.$$

Then decide whether such a map exists for $\mathcal{L} = (1,0), (0,1), (1,1)$ and $\mathcal{M} = (1,1), (1,-1), (2,0)$ in $\mathbb{R}^2$.

Solution on page 1122

Exercise 2.2.38. HARD Let $n \ge 2$, let $\mathbf{s} = \mathbf{e}_1 + \cdots + \mathbf{e}_n \in F^n$, and list the $n+1$ vectors $\mathbf{v}_1 = \mathbf{e}_1, \ldots, \mathbf{v}_n = \mathbf{e}_n, \mathbf{v}_{n+1} = \mathbf{s}$. For which permutations σ of $\{1, \ldots, n+1\}$ is there a linear map $T : F^n \to F^n$ with $T(\mathbf{v}_i) = \mathbf{v}_{\sigma(i)}$ for every i? The answer depends on whether $1 + 1 = 0$ in F.

Solution on page 1122

Exercise 2.2.39. HARD Find all linear maps $f : F[x] \to F[x]$ such that $f(pq) = f(p)f(q)$ for all $p, q \in F[x]$. (Products of polynomials are studied in Section 7.1. You may use that multiplication of polynomials is associative and commutative, distributes over addition, satisfies $c(pq) = (cp)q$ for $c \in F$, and has $x^j x^k = x^{j+k}$.)

Solution on page 1123

Exercise 2.2.40. CHALLENGE Let $n \ge 1$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1} \in V$ be such that any n of them, in their given order, form a basis of V.

(a) Show that $\mathbf{v}_{n+1} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ with every $a_i \neq 0$.

(b) Show that if $T \in \mathrm{L}(V)$ satisfies $T(\mathbf{v}_i) \in \operatorname{span}(\mathbf{v}_i)$ for $i = 1, \ldots, n+1$, then $T = \lambda \mathrm{id}_V$ for some $\lambda \in F$.

(c) Let $\mathbf{w}_1, \ldots, \mathbf{w}_{n+1} \in W$ be such that any n of them form a basis of W. Show that there are a linear map $T : V \to W$ and nonzero scalars $c_1, \ldots, c_{n+1}$ with $T(\mathbf{v}_i) = c_i\mathbf{w}_i$ for every i, and that any two such maps are scalar multiples of each other.

Solution on page 1124

2.3 Matrices of linear maps

By Theorem 2.2.1, a linear map $T : V \to W$ is determined by its values $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ on a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V, and these values can be any list in W (Corollary 2.2.18). If W has a basis $\mathbf{w}_1, \ldots, \mathbf{w}_m$ as well, each value $T(\mathbf{v}_j)$ is in turn determined by its coordinate vector in F^m. So T is recorded by mn scalars, and writing the n coordinate vectors side by side gives an $m \times n$ matrix, the matrix of T. Two things then have to be worked out: how to recover the values of T from its matrix, which will force the definition of the matrix–vector product, and how much the matrix depends on the chosen bases, which turns out to be a great deal.

Throughout this section, V and W are vector spaces over F with bases $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ of V and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_m$ of W, where $n, m \ge 1$, unless other bases are named. (We leave out the zero space, because matrices with no rows or no columns are not defined in Definition 1.2.7.)

2.3.1 The matrix of a linear map

Definition 2.3.1 (Matrix of a linear map). Let $T \in \mathrm{L}(V, W)$. The **matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$** is the $m \times n$ matrix

$$[T]_{\mathcal{C}\mathcal{B}} = \begin{pmatrix} [T(\mathbf{v}_1)]_{\mathcal{C}} & [T(\mathbf{v}_2)]_{\mathcal{C}} & \cdots & [T(\mathbf{v}_n)]_{\mathcal{C}} \end{pmatrix}$$

whose j-th column is the coordinate vector of $T(\mathbf{v}_j)$ with respect to $\mathcal{C}$. Equivalently, its entries a_{ij} are the scalars with

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{mj}\mathbf{w}_m \quad (j = 1, \ldots, n).$$

For an operator $T \in \mathrm{L}(V)$, the **matrix of T with respect to $\mathcal{B}$** is $[T]_{\mathcal{B}\mathcal{B}}$.

The matrix is well defined, since each $T(\mathbf{v}_j)$ has exactly one coordinate vector with respect to the basis $\mathcal{C}$ (Definition 1.6.14).

The matrix has one column for each vector of $\mathcal{B}$ and one row for each vector of $\mathcal{C}$, so its size is (length of $\mathcal{C}$) $\times$ (length of $\mathcal{B}$). Column j records where T sends the j-th basis vector of the domain, written in the coordinates of the target, and the entry a_{ij} is the coefficient of $\mathbf{w}_i$ in $T(\mathbf{v}_j)$. To keep the indices straight, imagine $\mathbf{v}_1, \ldots, \mathbf{v}_n$ written above the columns and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ written to the left of the rows. The target basis comes first in $[T]_{\mathcal{C}\mathcal{B}}$ so that the two copies of $\mathcal{B}$ sit next to each other in the formula $[T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$ of Theorem 2.3.9 below. We always write both subscripts, even when they are equal.

Notation 2.3.2 (The standard basis $\mathcal{E}$). We write $\mathcal{E}$ for the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_k$ of F^k (Definition 1.6.8), with k determined by the context. Since $\mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_k\mathbf{e}_k$ for $\mathbf{y} = (y_1, \ldots, y_k) \in F^k$, we have $[\mathbf{y}]_{\mathcal{E}} = \mathbf{y}$. That is, in the standard basis a vector of F^k is its own coordinate vector.

In particular, for a linear map $T : F^n \to F^m$, the matrix $[T]_{\mathcal{E}\mathcal{E}}$ with respect to the standard bases of F^n and F^m has the columns $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$.

Example 2.3.3 (Matrices with respect to standard bases).

(a) Let $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (2x - y + 3z, x + 5z)$, the map of Example 2.1.7. Then $T(\mathbf{e}_1) = (2, 1)$, $T(\mathbf{e}_2) = (-1, 0)$, and $T(\mathbf{e}_3) = (3, 5)$, so

$$[T]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 5 \end{pmatrix}.$$

This is the matrix A of Example 2.1.7: the coefficient of the j-th variable in the i-th coordinate of $T(x, y, z)$ is the entry in row i and column j. Proposition 2.3.17 shows that this always happens.

(b) The rotation R_θ of $\mathbb{R}^2$ by the angle θ (Example 2.1.11) sends $\mathbf{e}_1$ to $(\cos\theta, \sin\theta)$ and $\mathbf{e}_2$ to $(-\sin\theta, \cos\theta)$, so

$$[R_\theta]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \text{for example} \quad \left[R_{\pi/2}\right]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Example 2.3.4 (Differentiation). Let $V = F[x]_{\leq 2}$. Every element of V is $a + bx + cx^2$ for exactly one triple $(a, b, c) \in F^3$ (Definition 1.2.13, Definition 1.2.16), so $\mathcal{B} = 1, x, x^2$ is a basis of V and $\left[a + bx + cx^2\right]_{\mathcal{B}} = (a, b, c)$. Differentiation does not raise degrees, so D maps V into V, and we regard it as an operator $D \in \mathrm{L}(V)$ (Example 2.1.9). Since

$$D(1) = \mathbf{0}, \quad D(x) = 1, \quad D\left(x^2\right) = 2x,$$

whose coordinate vectors are $(0, 0, 0)$, $(1, 0, 0)$, and $(0, 2, 0)$,

$$[D]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Over $\mathbb{F}_2$ the entry 2 is 0. The domain and the target can also be different spaces. Regarded as a map $F[x]_{\leq 3} \to F[x]_{\leq 2}$, with the bases $\mathcal{B}' = 1, x, x^2, x^3$ and $\mathcal{B}$, differentiation has the 3×4 matrix

$$[D]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Strictly speaking these are two different linear maps, since their domains differ.

Definition 2.3.5 (Identity matrix). The $n \times n$ **identity matrix** I_n is the matrix with columns $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Its entry in row i and column j is 1 if $i = j$ and 0 if $i \neq j$. We write I when n is clear from the context.

Example 2.3.6 (The zero map and the identity map). The zero map $0 \in \mathrm{L}(V, W)$ sends each $\mathbf{v}_j$ to $\mathbf{0}$, whose coordinate vector is $\mathbf{0}$. So $[0]_{\mathcal{C}\mathcal{B}}$ is the zero matrix, for all bases. The identity map sends $\mathbf{v}_j$ to $\mathbf{v}_j = 0\mathbf{v}_1 + \cdots + 1\mathbf{v}_j + \cdots + 0\mathbf{v}_n$, whose $\mathcal{B}$-coordinate vector is $\mathbf{e}_j$. So

$$[\mathrm{id}_V]_{\mathcal{B}\mathcal{B}} = I_n \quad \text{for every basis } \mathcal{B} \text{ of } V.$$

Here the same basis is used on both sides. Section 2.3.5 shows that $[\mathrm{id}_V]_{\mathcal{C}\mathcal{B}}$ looks quite different when $\mathcal{C} \neq \mathcal{B}$.

Exercise 2.3.7. ROUTINE Let $T : F[x]_{\leq 1} \to F^2$, $T(p) = (p(0), p(1))$. Find $[T]_{\mathcal{E}\mathcal{B}}$ for the basis $\mathcal{B} = 1, x$, and $[T]_{\mathcal{E}\mathring{\mathcal{B}}}$ for the basis $\mathring{\mathcal{B}} = 1 - x, x$.

Solution on page 1125

2.3.2 The matrix–vector product

Suppose we know $A = [T]_{\mathcal{C}\mathcal{B}}$ and the coordinates $[\mathbf{v}]_{\mathcal{B}} = (x_1, \ldots, x_n)$ of a vector $\mathbf{v} \in V$, and we want the coordinates of $T(\mathbf{v})$.

Intuition (Computing $[T(\mathbf{v})]_{\mathcal{C}}$). Since $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$, linearity of T gives $T(\mathbf{v}) = x_1T(\mathbf{v}_1) + \cdots + x_nT(\mathbf{v}_n)$. Taking $\mathcal{C}$-coordinates is linear as well (Proposition 1.6.19), so

$$[T(\mathbf{v})]_{\mathcal{C}} = x_1[T(\mathbf{v}_1)]_{\mathcal{C}} + \cdots + x_n[T(\mathbf{v}_n)]_{\mathcal{C}}.$$

The right side is the linear combination of the columns of A whose coefficients are the entries of $[\mathbf{v}]_{\mathcal{B}}$. Realize that it no longer mentions T, the spaces V and W, or the bases: it is computed from the matrix A and the column $(x_1, \ldots, x_n)$ alone. So it is an operation on a matrix and a column, and we give it a name.

Definition 2.3.8 (Matrix–vector product). Let $A \in F^{m \times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n \in F^m$, and let $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$. The **product** of A and $\mathbf{x}$ is

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \in F^m.$$

It is defined only when the number of columns of A equals the number of entries of $\mathbf{x}$, and then $A\mathbf{x}$ has as many entries as A has rows.

In words, $A\mathbf{x}$ is the linear combination of the columns of A with coefficients $x_1, \ldots, x_n$. With this definition, the computation above becomes a theorem.

Theorem 2.3.9 (The matrix of a map computes the map). For every $T \in \mathrm{L}(V, W)$ and every $\mathbf{v} \in V$,

$$[T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}.$$

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 2.3.8 (Matrix–vector product), Definition 1.6.14 (Coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Definition 2.1.1 (Linear map)
Hint 1. Write $\mathbf{v}$ in terms of $\mathcal{B}$, using its coordinates $(x_1, \dots, x_n) = [\mathbf{v}]_{\mathcal{B}}$.
Hint 2. Apply T, then take $\mathcal{C}$-coordinates. Both steps respect linear combinations, and what comes out is a combination of the columns of $[T]_{\mathcal{C}\mathcal{B}}$.
Sketch. With $[\mathbf{v}]_{\mathcal{B}} = (x_1, \dots, x_n)$, linearity of T and of $[\cdot]_{\mathcal{C}}$ gives $[T(\mathbf{v})]_{\mathcal{C}} = \sum_j x_j \left[T(\mathbf{v}_j)\right]_{\mathcal{C}}$, a combination of the vectors $\left[T(\mathbf{v}_j)\right]_{\mathcal{C}}$. These are the columns of $[T]_{\mathcal{C}\mathcal{B}}$, so the combination is $[T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}$ by Definition 2.3.8.

Let $[\mathbf{v}]_{\mathcal{B}} = (x_1, \dots, x_n)$, so that $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$ (Definition 1.6.14). Then

$$\begin{aligned} [T(\mathbf{v})]_{\mathcal{C}} &= [x_1 T(\mathbf{v}_1) + \cdots + x_n T(\mathbf{v}_n)]_{\mathcal{C}} && \text{(linearity of } T\text{)} \\ &= x_1[T(\mathbf{v}_1)]_{\mathcal{C}} + \cdots + x_n[T(\mathbf{v}_n)]_{\mathcal{C}} && \text{(linearity of } [\cdot]_{\mathcal{C}}\text{)} \\ &= [T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} && \text{(definition of the product),} \end{aligned}$$

since the j-th column of $[T]_{\mathcal{C}\mathcal{B}}$ is $\left[T(\mathbf{v}_j)\right]_{\mathcal{C}}$ (Definition 2.3.1). □

So to evaluate T in coordinates, take the $\mathcal{B}$-coordinates of $\mathbf{v}$, multiply by the matrix, and read the result as $\mathcal{C}$-coordinates. Another way to say this: there are two routes from a vector $\mathbf{v} \in V$ to a column in F^m. One applies T and then takes $\mathcal{C}$-coordinates; the other takes $\mathcal{B}$-coordinates and then multiplies by $A = [T]_{\mathcal{C}\mathcal{B}}$. Theorem 2.3.9 says that the two routes always arrive at the same column (Figure 2.3.1). In terms of maps, $[\cdot]_{\mathcal{C}} \circ T = T_A \circ [\cdot]_{\mathcal{B}}$, where T_A is multiplication by A (Definition 2.3.15 below). The map T_A does to coordinate vectors what T does to vectors.

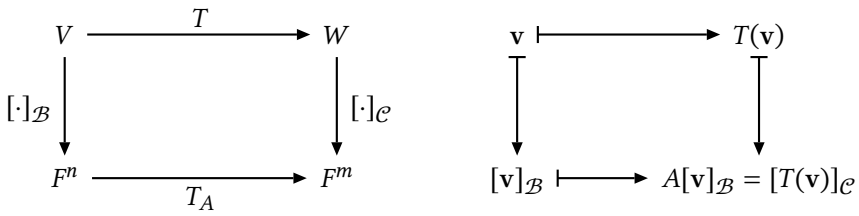


Figure 2.3.1. Two routes from V to F^m, where $A = [T]_{\mathcal{C}\mathcal{B}}$. Left: the spaces and maps. Right: what happens to a single vector $\mathbf{v}$. Going right and then down gives $[T(\mathbf{v})]_{\mathcal{C}}$, and going down and then right gives $A[\mathbf{v}]_{\mathcal{B}}$; Theorem 2.3.9 says that these agree.

You have probably met the product $A\mathbf{x}$ before, defined by a formula for its entries. The point here is that its definition is not a convention. Once we record vectors by their coordinates and linear maps by the columns $\left[T(\mathbf{v}_j)\right]_{\mathcal{C}}$, only one product makes Theorem 2.3.9 true. In fact, the special case of the standard bases already forces it.

Proposition 2.3.10 (The product is forced). Suppose that for all $m, n \geq 1$ we are given a rule that assigns to each $A \in F^{m\times n}$ and $\mathbf{x} \in F^n$ a vector $A \star \mathbf{x} \in F^m$, such that

$$[T]_{\mathcal{E}\mathcal{E}} \star \mathbf{x} = T(\mathbf{x}) \quad \text{for every } T \in \mathrm{L}(F^n, F^m) \text{ and every } \mathbf{x} \in F^n.$$

Then $A \star \mathbf{x} = A\mathbf{x}$ for all A and $\mathbf{x}$.

Since $[\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$ and $[T(\mathbf{x})]_{\mathcal{E}} = T(\mathbf{x})$, the hypothesis is exactly Theorem 2.3.9 for $V = F^n$ and $W = F^m$ with their standard bases, with $\star$ in place of the product.

Proof. ROUTINE

Uses. Corollary 2.2.17 (Linear maps out of F^n), Definition 2.3.1 (Matrix of a linear map), Notation 2.3.2 (The standard basis $\mathcal{E}$), Definition 2.3.8 (Matrix–vector product)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Given A and $\mathbf{x}$, find a linear map $T : F^n \to F^m$ whose matrix with respect to the standard bases is A.
Hint 2. By Corollary 2.2.17 there is a linear T with $T(\mathbf{e}_j) = \mathbf{a}_j$, the j-th column of A. That corollary also tells you what $T(\mathbf{x})$ is.
Sketch. Let T be the linear map with $T(\mathbf{e}_j) = \mathbf{a}_j$. Then $[T]_{\mathcal{E}\mathcal{E}} = A$ and $T(\mathbf{x}) = \sum_j x_j \mathbf{a}_j$ (Corollary 2.2.17), so $A \star \mathbf{x} = T(\mathbf{x}) = A\mathbf{x}$.

Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and let $\mathbf{x} = (x_1, \dots, x_n) \in F^n$. By Corollary 2.2.17 there is a linear map $T : F^n \to F^m$ with $T(\mathbf{e}_j) = \mathbf{a}_j$ for each j, and $T(\mathbf{x}) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$. The j-th column of $[T]_{\mathcal{E}\mathcal{E}}$ is $\big[T(\mathbf{e}_j)\big]_{\mathcal{E}} = \mathbf{a}_j$ (Notation 2.3.2), so $[T]_{\mathcal{E}\mathcal{E}} = A$. Hence

$$A \star \mathbf{x} = [T]_{\mathcal{E}\mathcal{E}} \star \mathbf{x} = T(\mathbf{x}) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}. \qquad \square$$

Matrix multiplication arises the same way in Section 2.4: it is the only product of matrices for which the matrix of a composite is the product of the matrices (Theorem 2.4.2).

The definition describes $A\mathbf{x}$ one column at a time. The familiar formula for its entries is a consequence.

Proposition 2.3.11 (Entries of $A\mathbf{x}$). Let $A = (a_{ij}) \in F^{m\times n}$ and $\mathbf{x} = (x_1, \dots, x_n) \in F^n$. For $i = 1, \dots, m$, the i-th entry of $A\mathbf{x}$ is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \sum_{j=1}^{n} a_{ij}x_j.$$

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 1.2.4 (The space F^n)
Hint 1. What is the i-th entry of the j-th column $\mathbf{a}_j$?
Hint 2. Vectors in F^m are added and scaled entry by entry.
Sketch. The i-th entry of $\mathbf{a}_j$ is a_{ij}, so the i-th entry of $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ is $x_1a_{i1} + \cdots + x_na_{in}$.

The j-th column $\mathbf{a}_j$ has i-th entry a_{ij}. Since the operations of F^m are entrywise (Definition 1.2.4), the i-th entry of $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ is $x_1a_{i1} + \cdots + x_na_{in} = a_{i1}x_1 + \cdots + a_{in}x_n$. $\square$

This is the row-times-column rule: the i-th entry of $A\mathbf{x}$ pairs row i of A with $\mathbf{x}$, multiplying entries in matching positions and adding. So there are two ways to read $A\mathbf{x}$. By columns, it is a linear combination of the columns of A, and this is the view to remember for proofs. By rows, it is a list of m numbers, each computed from one row of A. This is usually quicker by hand, and it is how Section 3.1 reads a system of equations.

Example 2.3.12 (A product computed two ways). Let $A = \begin{pmatrix} 1 & -2 & 0 \\ 3 & 1 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 3}$ and $\mathbf{x} = (2, 1, -1)$. By columns,

$$A\mathbf{x} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 1\begin{pmatrix} -2 \\ 1 \end{pmatrix} + (-1)\begin{pmatrix} 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 - 2 + 0 \\ 6 + 1 - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

By rows, the first entry is $1 \cdot 2 + (-2) \cdot 1 + 0 \cdot (-1) = 0$ and the second is $3 \cdot 2 + 1 \cdot 1 + 4 \cdot (-1) = 3$.

Proposition 2.3.13 (Linearity of the product). Let $A, A' \in F^{m\times n}$, $\mathbf{x}, \mathbf{y} \in F^n$, and $c \in F$, and let $\mathbf{a}_j$ be the j-th column of A.

(a) $A(c\mathbf{x} + \mathbf{y}) = cA\mathbf{x} + A\mathbf{y}$.
(b) $(cA + A')\mathbf{x} = cA\mathbf{x} + A'\mathbf{x}$.
(c) $A\mathbf{e}_j = \mathbf{a}_j$ for $j = 1, \dots, n$, and $I_n\mathbf{x} = \mathbf{x}$.

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 1.2.7 (Matrices), Definition 2.3.5 (Identity matrix), Definition 1.2.4 (The space F^n)

Hint 1. Write each side as a combination of columns using Definition 2.3.8, and regroup with the vector space axioms of F^m.
Hint 2. For (b), what are the columns of $cA + A'$? For (c), what are the entries of $\mathbf{e}_j$?
Sketch. (a) Expand $\sum_j (cx_j + y_j)\mathbf{a}_j$ and regroup. (b) Matrices are added and scaled entrywise, so the j-th column of $cA + A'$ is $c\mathbf{a}_j + \mathbf{a}'_j$; expand $\sum_j x_j(c\mathbf{a}_j + \mathbf{a}'_j)$ and regroup. (c) $\mathbf{e}_j$ has a single nonzero entry, a 1 in position j, and I_n has columns $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

(a) The vector $c\mathbf{x} + \mathbf{y}$ has entries $cx_j + y_j$ (Definition 1.2.4), so by Definition 2.3.8 and the vector space axioms of F^m,

$$A(c\mathbf{x} + \mathbf{y}) = \sum_{j=1}^{n} (cx_j + y_j)\mathbf{a}_j = c\sum_{j=1}^{n} x_j\mathbf{a}_j + \sum_{j=1}^{n} y_j\mathbf{a}_j = cA\mathbf{x} + A\mathbf{y}.$$

(b) Let $\mathbf{a}'_j$ be the j-th column of A'. Matrices are added and scaled entrywise (Definition 1.2.7), so the j-th column of $cA + A'$ is $c\mathbf{a}_j + \mathbf{a}'_j$. Hence

$$(cA + A')\mathbf{x} = \sum_{j=1}^{n} x_j(c\mathbf{a}_j + \mathbf{a}'_j) = c\sum_{j=1}^{n} x_j\mathbf{a}_j + \sum_{j=1}^{n} x_j\mathbf{a}'_j = cA\mathbf{x} + A'\mathbf{x}.$$

(c) The entries of $\mathbf{e}_j$ are 0 except for a 1 in position j, so $A\mathbf{e}_j = 0\mathbf{a}_1 + \cdots + 1\mathbf{a}_j + \cdots + 0\mathbf{a}_n = \mathbf{a}_j$. The columns of I_n are $\mathbf{e}_1, \ldots, \mathbf{e}_n$ (Definition 2.3.5), so $I_n\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n = \mathbf{x}$. □

Part (c) says that multiplying by $\mathbf{e}_j$ picks out the j-th column. It is the quickest way to get information about A out of the products $A\mathbf{x}$.

Exercise 2.3.14. ROUTINE Compute each product by columns and by rows.

(a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$ over $\mathbb{F}_5$.

(b) $\begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}\begin{pmatrix} i \\ 1 \end{pmatrix}$ over $\mathbb{C}$.

Solution on page 1125

2.3.3 Matrices as linear maps

Fixing the matrix and letting the vector vary turns the product into a function.

Definition 2.3.15 (The linear map of a matrix). For $A \in F^{m\times n}$, the map $T_A : F^n \to F^m$ is

$$T_A(\mathbf{x}) = A\mathbf{x}.$$

The map T_A is linear by Proposition 2.3.13(a). By Proposition 2.3.11, the i-th coordinate of $T_A(\mathbf{x})$ is $a_{i1}x_1 + \cdots + a_{in}x_n$, so T_A is the map given by a linear formula in Example 2.1.7, and the two definitions of T_A agree. So a matrix can be regarded either as a table of scalars or as a linear map. The next two results show that the two views carry the same information.

Proposition 2.3.16 (The matrix of T_A). For every $A \in F^{m\times n}$, $[T_A]_{\mathcal{E}\mathcal{E}} = A$.

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 2.3.15 (The linear map of a matrix), Proposition 2.3.13 (Linearity of the product), Notation 2.3.2 (The standard basis $\mathcal{E}$)
Hint 1. What is the j-th column of $[T_A]_{\mathcal{E}\mathcal{E}}$?
Hint 2. Compute $T_A(\mathbf{e}_j) = A\mathbf{e}_j$ with Proposition 2.3.13(c).
Sketch. The j-th column of $[T_A]_{\mathcal{E}\mathcal{E}}$ is $[A\mathbf{e}_j]_{\mathcal{E}} = A\mathbf{e}_j = \mathbf{a}_j$.

By Definition 2.3.1 and Notation 2.3.2, the j-th column of $[T_A]_{\mathcal{E}\mathcal{E}}$ is $[T_A(\mathbf{e}_j)]_{\mathcal{E}} = T_A(\mathbf{e}_j) = A\mathbf{e}_j$, which is the j-th column of A by Proposition 2.3.13(c). So the two matrices have the same columns. □

Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A). For every $T \in \mathrm{L}(F^n, F^m)$ there is exactly one $A \in F^{m\times n}$ with $T = T_A$, namely $A = [T]_{\mathcal{E}\mathcal{E}}$, the matrix whose columns are $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$.

Proof. ROUTINE

Uses. Theorem 2.3.9 (The matrix of a map computes the map), Notation 2.3.2 (The standard basis $\mathcal{E}$), Proposition 2.3.13 (Linearity of the product)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For existence, apply Theorem 2.3.9 with the standard bases on both sides. What are $[\mathbf{x}]_{\mathcal{E}}$ and $[T(\mathbf{x})]_{\mathcal{E}}$?
Hint 2. For uniqueness, suppose that $T_A = T_{A'}$, and evaluate both maps at $\mathbf{e}_j$.
Sketch. With $A = [T]_{\mathcal{E}\mathcal{E}}$, Theorem 2.3.9 gives $T(\mathbf{x}) = A[\mathbf{x}]_{\mathcal{E}} = A\mathbf{x}$, since standard coordinates do nothing. If $T_A = T_{A'}$, then $A\mathbf{e}_j = A'\mathbf{e}_j$, so A and A' have the same j-th column for every j.

Existence. Let $A = [T]_{\mathcal{E}\mathcal{E}}$. Its columns are $\big[T(\mathbf{e}_j)\big]_{\mathcal{E}} = T(\mathbf{e}_j)$ (Notation 2.3.2). For $\mathbf{x} \in F^n$, Theorem 2.3.9 with the standard bases of F^n and F^m gives

$$T(\mathbf{x}) = [T(\mathbf{x})]_{\mathcal{E}} = [T]_{\mathcal{E}\mathcal{E}}[\mathbf{x}]_{\mathcal{E}} = A\mathbf{x} = T_A(\mathbf{x}).$$

So $T = T_A$.

Uniqueness. Suppose $T_A = T_{A'}$ for some $A, A' \in F^{m\times n}$. By Proposition 2.3.13(c), for each j the j-th columns of A and A' are

$$A\mathbf{e}_j = T_A(\mathbf{e}_j) = T_{A'}(\mathbf{e}_j) = A'\mathbf{e}_j.$$

So A and A' have the same columns, and $A = A'$. □

Together with Proposition 2.3.16, this says that $A \mapsto T_A$ and $T \mapsto [T]_{\mathcal{E}\mathcal{E}}$ are inverse bijections between $F^{m\times n}$ and $\mathrm{L}(F^n, F^m)$. Two consequences are worth spelling out. First, the linear maps $F^n \to F^m$ are exactly the maps given by linear formulas, which makes precise the rule of thumb at the end of Section 2.1.3. Second, finding the matrix of a linear map $F^n \to F^m$ requires no formula at all: compute $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ and write them down as columns. From now on we often say "the matrix A" for the map T_A, and write $A\mathbf{x}$ rather than $T_A(\mathbf{x})$.

Exercise 2.3.18. ROUTINE For each linear map T, find the matrix A with $T = T_A$.

(a) $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (z, x, y)$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$, the projection $(x, y, z) \mapsto (x, y, 0)$ onto the xy-plane.

(c) $T : \mathbb{R}^3 \to \mathbb{R}$, $T(x, y, z) = x + y + z$.

(d) $T : \mathbb{F}_2^2 \to \mathbb{F}_2^2$, the linear map with $T(1, 0) = (1, 1)$ and $T(0, 1) = (0, 1)$. What is $T(1, 1)$?

Solution on page 1125

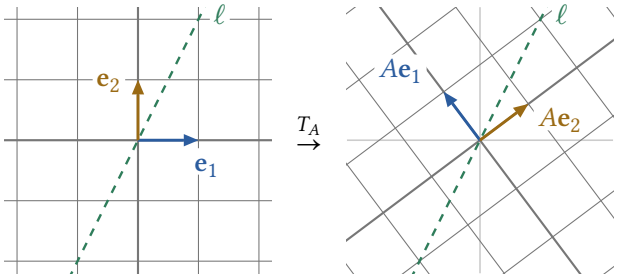
Example 2.3.19 (Projection and reflection). Let ℓ be the line $y = 2x$ in $\mathbb{R}^2$. It passes through $\mathbf{0}$ at the angle θ with $\cos\theta = 1/\sqrt{5}$ and $\sin\theta = 2/\sqrt{5}$, so $\cos^2\theta = 1/5$, $\sin\theta\cos\theta = 2/5$, $\sin^2\theta = 4/5$, $\cos 2\theta = -3/5$, and $\sin 2\theta = 4/5$. By the formulas of Example 2.1.11, the projection onto ℓ and the reflection across ℓ satisfy

$$\mathrm{Pr}_\ell(\mathbf{e}_1) = \left(\frac{1}{5}, \frac{2}{5}\right), \qquad \mathrm{Pr}_\ell(\mathbf{e}_2) = \left(\frac{2}{5}, \frac{4}{5}\right),$$
$$S_\ell(\mathbf{e}_1) = \left(-\frac{3}{5}, \frac{4}{5}\right), \qquad S_\ell(\mathbf{e}_2) = \left(\frac{4}{5}, \frac{3}{5}\right).$$

These are the columns of their matrices:

$$[\mathrm{Pr}_\ell]_{\mathcal{E}\mathcal{E}} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad [S_\ell]_{\mathcal{E}\mathcal{E}} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}.$$

The plane transformer below shows the reflection. Its panel displays the matrix A, and its two handles are the columns $A\mathbf{e}_1$ and $A\mathbf{e}_2$, that is, the images of $\mathbf{e}_1$ and $\mathbf{e}_2$. Dragging a handle changes one column of A and leaves the other alone, and the whole grid follows, because $A(x, y) = xA\mathbf{e}_1 + yA\mathbf{e}_2$.



The reflection across the line $y = 2x$. Its matrix has columns $A\mathbf{e}_1 = (-0.6, 0.8)$ and $A\mathbf{e}_2 = (0.8, 0.6)$. Type the entries $0.2, 0.4, 0.4, 0.8$ for the projection onto the same line, which flattens the grid onto it, or $0, -1, 1, 0$ for the rotation by a right angle. In each case the columns of the matrix are where $\mathbf{e}_1$ and $\mathbf{e}_2$ go. (Interactive in the web version.)

For general V and W the same holds once the bases are fixed: a linear map is determined by its matrix, and every matrix of the right size is the matrix of some linear map.

Theorem 2.3.20 (Linear maps correspond to matrices). Let $\mathcal{B}$ be a basis of V of length n and $\mathcal{C}$ a basis of W of length m.

(a) If $S, T \in \mathrm{L}(V, W)$ and $[S]_{\mathcal{C}\mathcal{B}} = [T]_{\mathcal{C}\mathcal{B}}$, then $S = T$.

(b) For every $A \in F^{m\times n}$ there is a $T \in \mathrm{L}(V, W)$ with $[T]_{\mathcal{C}\mathcal{B}} = A$.

That is, $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is a bijection from $\mathrm{L}(V, W)$ onto $F^{m\times n}$.

Intuition. For $V = F^n$ and $W = F^m$ with their standard bases, this is Proposition 2.3.17 together with Proposition 2.3.16. In general there are two natural strategies. One is to notice that the matrix is another way of writing the list $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$, which determines T by Section 2.2. The other is to move the problem into F^n and F^m with coordinates, where the special case applies, and move back. The first proof is shorter. The second exhibits every T as some T_A in disguise, which is the picture of Figure 2.3.1.

Proof (via lists). ROUTINE

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.14 (Coordinates), Definition 2.3.1 (Matrix of a linear map)

Hint 1. Two matrices are equal when their columns are equal. What does equality of the j-th columns of $[S]_{\mathcal{C}\mathcal{B}}$ and $[T]_{\mathcal{C}\mathcal{B}}$ say about $S(\mathbf{v}_j)$ and $T(\mathbf{v}_j)$?

Hint 2. For (b), the j-th column of A is the $\mathcal{C}$-coordinate vector of exactly one vector $\mathbf{y}_j \in W$. Send $\mathbf{v}_j$ there.

Sketch. (a) Equal columns mean $[S(\mathbf{v}_j)]_{\mathcal{C}} = [T(\mathbf{v}_j)]_{\mathcal{C}}$, so $S(\mathbf{v}_j) = T(\mathbf{v}_j)$ for all j, and $S = T$ by uniqueness in Theorem 2.2.1. (b) Let $\mathbf{y}_j = \sum_i a_{ij}\mathbf{w}_i$; the linear map with $T(\mathbf{v}_j) = \mathbf{y}_j$ has j-th column $[\mathbf{y}_j]_{\mathcal{C}} = \mathbf{a}_j$.

(a) Suppose $[S]_{\mathcal{C}\mathcal{B}} = [T]_{\mathcal{C}\mathcal{B}}$. Comparing j-th columns, $[S(\mathbf{v}_j)]_{\mathcal{C}} = [T(\mathbf{v}_j)]_{\mathcal{C}}$ for each j. A vector of W is determined by its coordinates: if $[\mathbf{w}]_{\mathcal{C}} = (c_1, \ldots, c_m)$, then $\mathbf{w} = c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m$ (Definition 1.6.14). So $S(\mathbf{v}_j) = T(\mathbf{v}_j)$ for every j, and $S = T$ by the uniqueness part of Theorem 2.2.1.

(b) Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the columns of A, and let $\mathbf{y}_j = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m \in W$, so that $[\mathbf{y}_j]_{\mathcal{C}} = \mathbf{a}_j$. By Theorem 2.2.1 there is a linear map $T : V \to W$ with $T(\mathbf{v}_j) = \mathbf{y}_j$ for each j. The j-th column of $[T]_{\mathcal{C}\mathcal{B}}$ is $[\mathbf{y}_j]_{\mathcal{C}} = \mathbf{a}_j$, so $[T]_{\mathcal{C}\mathcal{B}} = A$. □

Proof (via coordinates). ROUTINE

Uses. Theorem 2.3.9 (The matrix of a map computes the map), Example 2.1.8 (Coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Definition 2.3.15 (The linear map of a matrix), Proposition 2.1.17 (Compositions of linear maps are linear)

Techniques. Assemble linear maps from linear pieces (Technique 2.1.26)
Hint 1. Let $\Phi : F^m \to W$ be the map $(y_1, \dots, y_m) \mapsto y_1\mathbf{w}_1 + \cdots + y_m\mathbf{w}_m$. How are Φ and $[\cdot]_{\mathcal{C}}$ related?
Hint 2. For (a), use Theorem 2.3.9 to compare $[S(\mathbf{v})]_{\mathcal{C}}$ and $[T(\mathbf{v})]_{\mathcal{C}}$. For (b), try $T = \Phi \circ T_A \circ [\cdot]_{\mathcal{B}}$.
Sketch. Φ and $[\cdot]_{\mathcal{C}}$ undo each other. (a) Both $[S(\mathbf{v})]_{\mathcal{C}}$ and $[T(\mathbf{v})]_{\mathcal{C}}$ equal $A[\mathbf{v}]_{\mathcal{B}}$; apply Φ. (b) $T = \Phi \circ T_A \circ [\cdot]_{\mathcal{B}}$ is a composite of linear maps, and $\left[T(\mathbf{v}_j)\right]_{\mathcal{C}} = A\mathbf{e}_j = \mathbf{a}_j$.

Let $\Phi : F^m \to W$ be the linear map $(y_1, \dots, y_m) \mapsto y_1\mathbf{w}_1 + \cdots + y_m\mathbf{w}_m$ of Example 2.1.8. It undoes $\mathcal{C}$-coordinates in both orders: $\Phi\big([\mathbf{w}]_{\mathcal{C}}\big) = \mathbf{w}$ for $\mathbf{w} \in W$, and $[\Phi(\mathbf{y})]_{\mathcal{C}} = \mathbf{y}$ for $\mathbf{y} \in F^m$.

(a) Let $A = [S]_{\mathcal{CB}} = [T]_{\mathcal{CB}}$. For $\mathbf{v} \in V$, Theorem 2.3.9 gives

$$[S(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}},$$

and applying Φ gives $S(\mathbf{v}) = T(\mathbf{v})$.

(b) Define $T = \Phi \circ T_A \circ [\cdot]_{\mathcal{B}}$, that is, $T(\mathbf{v}) = \Phi(A[\mathbf{v}]_{\mathcal{B}})$. It is a composite of linear maps (Proposition 1.6.19, Definition 2.3.15, Example 2.1.8), so it is linear by Proposition 2.1.17. Since $\left[\mathbf{v}_j\right]_{\mathcal{B}} = \mathbf{e}_j$,

$$\left[T(\mathbf{v}_j)\right]_{\mathcal{C}} = \left[\Phi(A\mathbf{e}_j)\right]_{\mathcal{C}} = A\mathbf{e}_j = \mathbf{a}_j,$$

so the j-th column of $[T]_{\mathcal{CB}}$ is the j-th column of A. □

The second proof shows what an arbitrary linear map looks like through coordinates. By (a), every $T \in \mathrm{L}(V, W)$ is the map built in (b) from its own matrix: $T = \Phi \circ T_A \circ [\cdot]_{\mathcal{B}}$ with $A = [T]_{\mathcal{CB}}$. Read from right to left, this says: take coordinates, multiply by the matrix, and turn the resulting coordinates back into a vector. In Section 2.4 we will see that the bijection $T \mapsto [T]_{\mathcal{CB}}$ also respects addition and scalar multiplication and turns composition into matrix multiplication. So $\mathrm{L}(V, W)$ and $F^{m\times n}$ are isomorphic vector spaces (Section 4.5).

Exercise 2.3.21. ROUTINE Let $\mathcal{B} = 1, x$ and $\mathcal{C} = 1, 1 + x$, both bases of $F[x]_{\leq 1}$, and let $T \in \mathrm{L}\big(F[x]_{\leq 1}\big)$ be the operator with

$$[T]_{\mathcal{CB}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Find $T(a + bx)$, and find $[T]_{\mathcal{BB}}$.

Solution on page 1126

2.3.4 Different bases, different matrices

The matrix $[T]_{\mathcal{CB}}$ depends on three things: the map T, the basis $\mathcal{B}$, and the basis $\mathcal{C}$. Changing either basis, even by reordering it, can change the matrix, and one map can have matrices that look nothing alike.

Example 2.3.22 (Differentiation in four bases). Let $V = F[x]_{\leq 2}$, and let $D \in \mathrm{L}(V)$ be differentiation, as in Example 2.3.4.

(a) For $\mathcal{B} = 1, x, x^2$, we found

$$[D]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

(b) Reverse the order: $\mathcal{R} = x^2, x, 1$. Now $D(x^2) = 2x$, $D(x) = 1$, and $D(1) = \mathbf{0}$ have the $\mathcal{R}$-coordinates $(0, 2, 0)$, $(0, 0, 1)$, and $(0, 0, 0)$, so

$$[D]_{\mathcal{R}\mathcal{R}} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The vectors are the same as in (a), but a basis is a list, and its order matters.

(c) Suppose that $2 \neq 0$ in F, and let $\mathring{\mathcal{B}} = 1, x, \frac{1}{2}x^2$. This is a basis, since $a + bx + cx^2 = a \cdot 1 + bx + 2c \cdot \frac{1}{2}x^2$, with only these coefficients possible. Now $D\left(\frac{1}{2}x^2\right) = x$, so

$$[D]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Each basis vector goes to the one before it, and the first goes to $\mathbf{0}$.

(d) Keep the assumption $2 \neq 0$, and use different bases on the two sides: $\mathcal{B}' = x, \frac{1}{2}x^2, 1$ in the domain and $\mathcal{B} = 1, x, x^2$ in the target. Then $D(x) = 1$, $D\left(\frac{1}{2}x^2\right) = x$, and $D(1) = \mathbf{0}$ have the $\mathcal{B}$-coordinates $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{0}$, so

$$[D]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Intuition (One basis or two). The matrix in (d) is the matrix of the projection $(x_1, x_2, x_3) \mapsto (x_1, x_2, 0)$ of F^3, a map P with $P \circ P = P$. But D is nothing like that: $D(D(D(p))) = \mathbf{0}$ for every $p \in V$, since each application of D lowers the degree. Using different bases on the two sides gives so much freedom that the matrix forgets almost everything about D. What survives is how many independent values D takes, its rank (Section 4.2, and Theorem 3.5.27 for matrices).

For an operator $T \in \mathrm{L}(V)$, comparing $T(\mathbf{v})$ with $\mathbf{v}$ only makes sense in a single coordinate system, so we use the same basis on both sides, and then more survives. Let N be any of the three matrices in (a)–(c), with its basis $\mathcal{A}$. Applying Theorem 2.3.9 three times,

$$N\big(N\big(N[p]_{\mathcal{A}}\big)\big) = N\big(N[D(p)]_{\mathcal{A}}\big) = N[D(D(p))]_{\mathcal{A}} = [D(D(D(p)))]_{\mathcal{A}} = \mathbf{0}$$

for every p, and since every $\mathbf{x} \in F^3$ is $[p]_{\mathcal{A}}$ for some p, multiplying by N three times kills every vector. You can check this directly for each N. The matrix in (d) does not have this property: it fixes $\mathbf{e}_1$. Properties of an operator that are visible in its matrix in every basis are studied in Section 4.5, under the name of similarity.

Example 2.3.23 (Maps of the plane in adapted bases). Let ℓ be the line $y = 2x$, as in Example 2.3.19, and let $\mathcal{U} = \mathbf{u}_1, \mathbf{u}_2$ with $\mathbf{u}_1 = (1, 2)$ on ℓ and $\mathbf{u}_2 = (-2, 1)$ perpendicular to ℓ (since $1 \cdot (-2) + 2 \cdot 1 = 0$). This is a basis of $\mathbb{R}^2$: the equations $a - 2b = x$ and $2a + b = y$ have exactly one solution, $a = (x + 2y)/5$ and $b = (y - 2x)/5$.

The projection $P = \mathrm{Pr}_\ell$ fixes $\mathbf{u}_1$ and sends $\mathbf{u}_2$ to $\mathbf{0}$, and the reflection $S = S_\ell$ fixes $\mathbf{u}_1$ and sends $\mathbf{u}_2$ to $-\mathbf{u}_2$. (With the matrices of Example 2.3.19: $\frac{1}{5}(1 + 4, 2 + 8) = (1, 2)$ and $\frac{1}{5}(-2 + 2, -4 + 4) = (0, 0)$.) So

$$[P]_{\mathcal{UU}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad [S]_{\mathcal{UU}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

compared with $\frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $\frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$ in the standard basis. In the adapted basis, the matrices show what the maps do.

The rotation $R = R_{\pi/2}$ has $[R]_{\mathcal{EE}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. In the reversed standard basis $\mathbf{e}_2, \mathbf{e}_1$, the vector $R(\mathbf{e}_2) = -\mathbf{e}_1$ has coordinates $(0, -1)$ and $R(\mathbf{e}_1) = \mathbf{e}_2$ has coordinates $(1, 0)$, so the matrix becomes $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. That is the standard matrix of the clockwise rotation $R_{-\pi/2}$: reordering the basis reverses the orientation of the coordinate system. In the basis $\mathcal{U}$, on the other hand, $R(\mathbf{u}_1) = (-2, 1) = \mathbf{u}_2$ and $R(\mathbf{u}_2) = (-1, -2) = -\mathbf{u}_1$, so $[R]_{\mathcal{UU}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the same as in the standard basis.

Two lessons. First, a matrix alone does not determine a map; only the matrix together with the bases does (Theorem 2.3.20). For instance $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is $[S_\ell]_{\mathcal{UU}}$, and it is also the standard matrix of the reflection across the x-axis. Second, choosing a basis adapted to the map is how one understands it. If some basis consists of vectors that T sends to multiples of themselves, the matrix of T in that basis is diagonal. This is how Example 2.2.3 built P and S in the first place (Technique 2.2.4), and finding such bases is the subject of Section 7.4.

2.3.5 The identity map in two bases

By Example 2.3.6, $[\mathrm{id}_V]_{\mathcal{BB}} = I_n$ for every basis $\mathcal{B}$. With two different bases, the matrix of the identity map is a different matrix, and it can look like almost anything. What it does is given by Theorem 2.3.9: for any two bases $\mathcal{B}$ and $\mathcal{C}$ of V,

$$[\mathrm{id}_V]_{\mathcal{CB}}[\mathbf{v}]_{\mathcal{B}} = [\mathrm{id}_V(\mathbf{v})]_{\mathcal{C}} = [\mathbf{v}]_{\mathcal{C}} \quad \text{for every } \mathbf{v} \in V.$$

That is, the matrix of the identity converts $\mathcal{B}$-coordinates into $\mathcal{C}$-coordinates.

Example 2.3.24 (The identity in two bases). Let F be a field with $1 + 1 \neq 0$, and let $\mathcal{B} = (1, 1), (1, -1)$, a basis of F^2 (Example 2.2.3). The columns of $[\mathrm{id}]_{\mathcal{EB}}$ are the standard coordinates of $(1, 1)$ and $(1, -1)$, which are these vectors themselves. The columns of $[\mathrm{id}]_{\mathcal{BE}}$ are the $\mathcal{B}$-coordinates of $\mathbf{e}_1 = \frac{1}{2}(1, 1) + \frac{1}{2}(1, -1)$ and $\mathbf{e}_2 = \frac{1}{2}(1, 1) - \frac{1}{2}(1, -1)$. So

$$[\mathrm{id}]_{\mathcal{EB}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad [\mathrm{id}]_{\mathcal{BE}} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

For $\mathbf{v} = (3, 1) = 2(1, 1) + (1, -1)$ we have $[\mathbf{v}]_{\mathcal{B}} = (2, 1)$, and indeed the two matrices convert between the two coordinate vectors of $\mathbf{v}$:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = [\mathbf{v}]_{\mathcal{E}}, \quad \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = [\mathbf{v}]_{\mathcal{B}}.$$

The example generalizes in one direction. For any basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of F^n, $[\mathrm{id}]_{\mathcal{EB}}$ is the matrix whose columns are $\mathbf{b}_1, \ldots, \mathbf{b}_n$, because $[\mathbf{b}_j]_{\mathcal{E}} = \mathbf{b}_j$. The other direction, $[\mathrm{id}]_{\mathcal{BE}}$, requires solving for coordinates.

Remark 2.3.25 (A warning about the identity). The phrase "the matrix of the identity map" means nothing until both bases are specified: $[\mathrm{id}_V]_{\mathcal{CB}} = I_n$ exactly when $\mathcal{B} = \mathcal{C}$ (Exercise 2.3.26). In general, $[\mathrm{id}_V]_{\mathcal{CB}}$ is the matrix that changes $\mathcal{B}$-coordinates into $\mathcal{C}$-coordinates. These matrices are the subject of Section 4.5 (Definition 4.5.10). There, once matrix multiplication is available (Section 2.4), they compute $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}$ from $[T]_{\mathcal{CB}}$ for any other bases $\mathring{\mathcal{B}}$ and $\mathring{\mathcal{C}}$, and explain the different matrices of D and of the plane maps found above.

Exercise 2.3.26. ROUTINE Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_n$ be bases of V of the same length. Show that $[\mathrm{id}_V]_{\mathcal{C}\mathcal{B}} = I_n$ if and only if $\mathcal{B} = \mathcal{C}$, that is, $\mathbf{v}_j = \mathbf{w}_j$ for every j.

Solution on page 1126

2.3.6 Translating between maps and matrices

The results of this section give a dictionary between linear maps of spaces with chosen bases and matrices.

- A vector $\mathbf{v} \in V$ corresponds to its coordinate vector $[\mathbf{v}]_{\mathcal{B}} \in F^n$, and a map $T \in \mathrm{L}(V, W)$ to its matrix $[T]_{\mathcal{C}\mathcal{B}} \in F^{m \times n}$.
- The value $T(\mathbf{v})$ corresponds to the product $[T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}$ (Theorem 2.3.9), and the value $T(\mathbf{v}_j)$ on a basis vector to the j-th column.
- Two maps are equal if and only if their matrices are (Theorem 2.3.20).
- For $V = F^n$ and $W = F^m$ with the standard bases, T corresponds to the unique A with $T = T_A$ (Proposition 2.3.17), and id corresponds to I_n.

Technique 2.3.27 (Translate between maps and matrices). A linear map between spaces with chosen bases and its matrix carry the same information, so a question about one can be translated into a question about the other, using the dictionary above.

In one direction, to compute with a linear map, replace vectors by their coordinate vectors and the map by its matrix. Evaluating T becomes a matrix–vector product, and an equation $T(\mathbf{v}) = \mathbf{w}$ becomes the linear system $[T]_{\mathcal{C}\mathcal{B}}\mathbf{x} = [\mathbf{w}]_{\mathcal{C}}$ in the unknown coordinates $\mathbf{x}$ of $\mathbf{v}$ (Section 3.1). In the other direction, to prove a statement about a matrix A, read it as a statement about the linear map T_A, prove it with the theory of linear maps, which needs no indices, and translate back with the uniqueness in Proposition 2.3.17.

Reach for the first direction when you need numbers, and for the second when a matrix identity would otherwise need a computation with indices. A rule of thumb: compute with matrices, prove with maps. Before translating, fix the bases and keep track of which basis each coordinate vector refers to; most errors come from mixing them up. The canonical example is Example 2.3.28. Later instances: matrix multiplication is defined by translating composition (Section 2.4), the rank of a matrix is the rank of its map (Section 4.2), and the determinant and the eigenvalues of an operator are read off from its matrix in any basis (Section 5.4, Section 7.3).

Used in. Exercise 2.3.32, Exercise 2.3.35, Exercise 2.3.42, Theorem 2.4.2, Proposition 2.4.14, Proposition 2.5.20, Proposition 2.6.20, Corollary 2.7.16, Exercise 2.8.8, Exercise 2.8.13, Problem 2.2, Problem 2.13, Problem 3.12, Theorem 4.1.1, Proposition 4.1.31, Theorem 4.2.11, Proposition 4.2.15, Proposition 4.5.12, Proposition 4.5.16, Theorem 4.5.17, Theorem 4.5.25, Proposition 4.5.27, Proposition 5.4.20, Theorem 6.6.15, Proposition 6.6.18, Corollary 7.3.13, Theorem 7.4.2, Proposition 7.5.25, Proposition 8.1.7, Corollary 8.1.11, Theorem 8.4.4, Proposition B.1.18, Exercise B.1.41, Exercise B.1.42, Exercise 2.5.22, Exercise 6.6.46.

Example 2.3.28 (Translating in both directions).

(a) *From maps to matrices.* We find all $p \in \mathbb{Q}[x]_{\le 2}$ with $p(x+1) - p(x) = x$. The map $\Delta : \mathbb{Q}[x]_{\le 2} \to \mathbb{Q}[x]_{\le 2}$, $\Delta(p) = p(x+1) - p(x)$, is linear, being the difference of the substitution map of Exercise 2.1.32 and the identity (Technique 2.1.26). With $\mathcal{B} = 1, x, x^2$,

$$\Delta(1) = \mathbf{0}, \quad \Delta(x) = (x+1) - x = 1, \quad \Delta\big(x^2\big) = (x+1)^2 - x^2 = 1 + 2x,$$

so

$$[\Delta]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

By Theorem 2.3.9, $p = a + bx + cx^2$ satisfies $\Delta(p) = x$ if and only if $[\Delta]_{\mathcal{B}\mathcal{B}}(a, b, c) = [x]_{\mathcal{B}} = (0, 1, 0)$. By Proposition 2.3.11 this is the system $b + c = 0$, $2c = 1$, $0 = 0$, whose solutions are $c = 1/2$, $b = -1/2$, and a arbitrary. So the solutions are

$$p = a + \frac{x^2 - x}{2} \quad (a \in \mathbb{Q}).$$

This gives a formula for sums. For $q(t) = t(t-1)/2$ and every number t,

$$q(t+1) - q(t) = \frac{(t+1)t - t(t-1)}{2} = t,$$

so the sum $0 + 1 + \cdots + (n-1)$ telescopes:

$$\sum_{k=0}^{n-1} k = \sum_{k=0}^{n-1} (q(k+1) - q(k)) = q(n) - q(0) = \frac{n(n-1)}{2}.$$

(b) *From matrices to maps.* Let $A \in F^{n \times n}$, and suppose that for every $\mathbf{x} \in F^n$ the vector $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$. We show that $A = \lambda I_n$ for some $\lambda \in F$. The hypothesis says that $T_A(\mathbf{x}) \in \operatorname{span}(\mathbf{x})$ for every $\mathbf{x}$. By Exercise 2.1.48, $T_A = \lambda\,\mathrm{id}$ for some $\lambda \in F$. Now $(\lambda I_n)\mathbf{x} = \lambda(I_n\mathbf{x}) = \lambda\mathbf{x}$ by Proposition 2.3.13, so $\lambda\,\mathrm{id} = T_{\lambda I_n}$. Hence $T_A = T_{\lambda I_n}$, and $A = \lambda I_n$ by the uniqueness in Proposition 2.3.17. A direct proof with entries is possible, but it would repeat the argument of Exercise 2.1.48 in coordinates.

Exercises

Exercise 2.3.29. ROUTINE Let $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x, y) = (x + y, x - y, 2x)$, and consider the bases $\mathcal{B} = (1, 1), (1, -1)$ of $\mathbb{R}^2$ and $\mathcal{C} = (1, 0, 0), (1, 1, 0), (1, 1, 1)$ of $\mathbb{R}^3$.

(a) Find $[T]_{\mathcal{C}\mathcal{B}}$.

(b) Check Theorem 2.3.9 for $\mathbf{v} = (3, 1)$ by computing both sides.

Solution on page 1126

Exercise 2.3.30. ROUTINE Find the error in the following computation, and correct it.

Let $\mathcal{B} = (1, 1), (1, -1)$, a basis of $\mathbb{R}^2$, and let $\mathbf{v} = (3, 1)$. The matrix $[\mathrm{id}]_{\mathcal{E}\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ relates the two bases, so the $\mathcal{B}$-coordinates of $\mathbf{v}$ are

$$[\mathbf{v}]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{E}\mathcal{B}}\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

Solution on page 1127

Exercise 2.3.31. ROUTINE Find the error in the following argument.

Claim. Let F be any field. If $A \in F^{2\times 2}$ satisfies $A(1,1) = (1,1)$ and $A(1,-1) = (1,-1)$, then $A = I_2$.

Proof. The maps T_A and $\mathrm{id} = T_{I_2}$ agree on the vectors $(1,1)$ and $(1,-1)$, which form a basis of F^2. Two linear maps that agree on a basis are equal, so $T_A = T_{I_2}$, and $A = I_2$ by Proposition 2.3.17.

Solution on page 1127

Exercise 2.3.32. STANDARD Let $T : F[x]_{\leq 3} \to F[x]_{\leq 3}$, $T(p) = p(x+1)$, which is linear by Exercise 2.1.32, and let $\mathcal{B} = 1, x, x^2, x^3$.

(a) Find $[T]_{\mathcal{B}\mathcal{B}}$ for $F = \mathbb{Q}$ and for $F = \mathbb{F}_2$.

(b) Over $\mathbb{F}_2$, use the matrix to find all $p \in \mathbb{F}_2[x]_{\leq 3}$ with $p(x+1) = p(x)$.

(c) Do the same over $\mathbb{Q}$.

Solution on page 1127

Exercise 2.3.33. STANDARD Let $T : F[x]_{\leq 2} \to F^3$, $T(p) = (p(0), p(1), p(2))$, where $2 = 1 + 1$.

(a) Find $[T]_{\mathcal{E}\mathcal{B}}$ for $\mathcal{B} = 1, x, x^2$.

(b) Suppose $2 \neq 0$ in F. Find a basis $\mathring{\mathcal{B}} = p_0, p_1, p_2$ of $F[x]_{\leq 2}$ with $[T]_{\mathcal{E}\mathring{\mathcal{B}}} = I_3$, and prove that it is a basis.

(c) With $\mathring{\mathcal{B}}$ as in (b), show that $[q]_{\mathring{\mathcal{B}}} = (q(0), q(1), q(2))$ for every $q \in F[x]_{\leq 2}$.

(d) Show that over $\mathbb{F}_2$ no basis $\mathring{\mathcal{B}}$ with $[T]_{\mathcal{E}\mathring{\mathcal{B}}} = I_3$ exists.

Solution on page 1128

Exercise 2.3.34. STANDARD Let $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x+y, 2x+2y)$.

(a) Find bases $\mathcal{B}$ and $\mathcal{C}$ of $\mathbb{R}^2$ with $[T]_{\mathcal{C}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(b) Show that there is no basis $\mathcal{B}$ of $\mathbb{R}^2$ with $[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(c) Find a basis $\mathcal{B}$ of $\mathbb{R}^2$ for which $[T]_{\mathcal{B}\mathcal{B}}$ is diagonal.

Solution on page 1129

Exercise 2.3.35. STANDARD Regard $\mathbb{C}$ as a vector space over $\mathbb{R}$ with the basis $\mathcal{B} = 1, i$.

(a) For $z = a + bi$ with $a, b \in \mathbb{R}$, let $M_z : \mathbb{C} \to \mathbb{C}$, $M_z(w) = zw$. Find $[M_z]_{\mathcal{B}\mathcal{B}}$.

(b) Find the matrix of complex conjugation $w \mapsto \overline{w}$ with respect to $\mathcal{B}$.

(c) Show that a matrix $A \in \mathbb{R}^{2\times 2}$ is $[T]_{\mathcal{B}\mathcal{B}}$ for some $\mathbb{C}$-linear $T : \mathbb{C} \to \mathbb{C}$ if and only if $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ for some $a, b \in \mathbb{R}$.

Solution on page 1129

Exercise 2.3.36. STANDARD Let V and W be vector spaces over $\mathbb{R}$ with bases. Prove or disprove each statement.

(a) If $S, T \in \mathrm{L}(V,W)$ and $[S]_{\mathcal{C}\mathcal{B}} = [T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}$ for some bases $\mathcal{B}, \mathring{\mathcal{B}}$ of V and $\mathcal{C}, \mathring{\mathcal{C}}$ of W, then $S = T$.

(b) If $T \in \mathrm{L}(V)$ and $[T]_{\mathcal{B}\mathcal{B}}$ is the zero matrix for some basis $\mathcal{B}$, then $T = 0$.

(c) If $T \in \mathrm{L}(V)$ and $[T]_{\mathcal{C}\mathcal{B}} = I_n$ for some bases $\mathcal{B}, \mathcal{C}$ of V, then $T = \mathrm{id}_V$.

(d) If $T \in \mathrm{L}(V)$ and $[T]_{\mathcal{B}\mathcal{B}}$ is diagonal for some basis $\mathcal{B}$, then $[T]_{\mathcal{C}\mathcal{C}}$ is diagonal for every basis $\mathcal{C}$ of V.

(e) There are bases $\mathcal{B}, \mathcal{C}$ of $\mathbb{R}^2$ with $[\mathrm{id}]_{\mathcal{C}\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Solution on page 1130

Exercise 2.3.37. STANDARD Let $T \in \mathrm{L}(V, W)$ and $A = [T]_{\mathcal{C}\mathcal{B}}$. Describe how A changes when the bases are modified as follows, and prove your answer. (Each modified list is again a basis; check this too.)

(a) Two vectors $\mathbf{v}_j$ and $\mathbf{v}_k$ of $\mathcal{B}$ are swapped.

(b) $\mathbf{v}_j$ is replaced by $c\mathbf{v}_j$, where $c \neq 0$.

(c) Two vectors $\mathbf{w}_i$ and $\mathbf{w}_k$ of $\mathcal{C}$ are swapped.

(d) $\mathbf{w}_i$ is replaced by $c\mathbf{w}_i$, where $c \neq 0$.

(e) $\mathbf{w}_k$ is replaced by $\mathbf{w}_k + c\mathbf{w}_i$, where $i \neq k$ and $c \in F$.

Solution on page 1130

Exercise 2.3.38. STANDARD A matrix $A \in F^{n\times n}$ is **upper triangular** if $a_{ij} = 0$ whenever $i > j$. Let $T \in \mathrm{L}(V)$, and for $k = 1, \ldots, n$ let $V_k = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. Show that the following are equivalent.

(i) $[T]_{\mathcal{B}\mathcal{B}}$ is upper triangular.

(ii) $T(\mathbf{v}_k) \in V_k$ for $k = 1, \ldots, n$.

(iii) $T(V_k) \subseteq V_k$ for $k = 1, \ldots, n$.

Solution on page 1131

Exercise 2.3.39. HARD Let $N \in F^{4\times 4}$ be the matrix with 1 in positions $(1,2)$, $(2,3)$, $(3,4)$ and 0 elsewhere, and let D be differentiation on $F[x]_{\leq 3}$. For each of $F = \mathbb{Q}$, $\mathbb{F}_2$, $\mathbb{F}_3$, $\mathbb{F}_5$, decide whether there is a basis $\mathcal{B}$ of $F[x]_{\leq 3}$ with $[D]_{\mathcal{B}\mathcal{B}} = N$.

Solution on page 1131

Exercise 2.3.40. HARD Let $T \in \mathrm{L}(V)$, where V has a basis of length $n \geq 1$. Show that $[T]_{\mathcal{B}\mathcal{B}}$ is the same matrix for every basis $\mathcal{B}$ of V of length n if and only if $T = \lambda \mathrm{id}_V$ for some $\lambda \in F$. (In Section 4.1 we will see that every basis of V has length n.)

Solution on page 1132

Exercise 2.3.41. HARD *Lights out on a line.* A row of $n \geq 2$ lamps is controlled by n buttons. Pressing button i switches lamp i and its neighbours (lamps $i-1$ and $i+1$, when they exist) from on to off or from off to on. Record a configuration of the lamps as $\mathbf{s} \in \mathbb{F}_2^n$, with $s_i = 1$ when lamp i is on.

(a) Let $A \in \mathbb{F}_2^{n\times n}$ have $a_{ij} = 1$ when $|i - j| \leq 1$ and $a_{ij} = 0$ otherwise. Show that pressing buttons in any order changes the configuration $\mathbf{s}$ to $\mathbf{s} + A\mathbf{x}$, where $x_j \in \mathbb{F}_2$ is the number of times button j was pressed, modulo 2.

(b) Show that every configuration can be switched off completely if and only if n is not of the form $3k + 2$.

Solution on page 1133

Exercise 2.3.42. STANDARD Using the matrix of $\Delta(p) = p(x+1) - p(x)$ on $\mathbb{Q}[x]_{\leq 3}$ with respect to $1, x, x^2, x^3$, find all $p \in \mathbb{Q}[x]_{\leq 3}$ with $p(x+1) - p(x) = x^2$. Deduce a formula for $0^2 + 1^2 + \cdots + (n-1)^2$, as in Example 2.3.28.

Solution on page 1134

Exercise 2.3.43. STANDARD For $\mathbf{y} \in F^k$ let $s(\mathbf{y}) = y_1 + \cdots + y_k$ be the sum of its entries, and let $A \in F^{m\times n}$.

(a) Show that $s(A\mathbf{x}) = s(\mathbf{x})$ for every $\mathbf{x} \in F^n$ if and only if the entries of each column of A add up to 1.

(b) Show that $A(1, \ldots, 1) = c(1, \ldots, 1)$ if and only if the entries of each row of A add up to c.

(c) Let $F = \mathbb{R}$, and suppose that all entries of A are nonnegative and each column adds up to 1. Show that if the entries of $\mathbf{x}$ are nonnegative and add up to 1, then the entries of $A\mathbf{x}$ are also nonnegative and add up to 1.

Solution on page 1134

Exercise 2.3.44. STANDARD BEYOND 21-242 The matrix of an operator on $F[x]$, which has no finite basis, can still be defined. For $T \in \mathrm{L}(F[x])$, let $a_{ij} \in F$ $(i, j \geq 0)$ be the coefficients with $T(x^j) = \sum_i a_{ij} x^i$. Call the array $(a_{ij})_{i,j\geq 0}$ the matrix of T.

(a) Show that each column of the matrix has only finitely many nonzero entries, and that every array $(a_{ij})_{i,j\geq 0}$ with this property is the matrix of exactly one $T \in \mathrm{L}(F[x])$.

(b) Find the matrices of differentiation D and of multiplication by x.

(c) Show that if p has coefficients $(c_0, c_1, \ldots)$, then $T(p)$ has coefficients $\left(\sum_j a_{ij} c_j\right)_{i\geq 0}$, where each sum has only finitely many nonzero terms.

(d) Give an operator whose matrix has a row with infinitely many nonzero entries.

Solution on page 1135

Exercise 2.3.45. ROUTINE Let $T : \mathbb{C}^2 \to \mathbb{C}^2$, $T(z, w) = (z + iw, iz - w)$.

(a) Find $[T]_{\mathcal{E}\mathcal{E}}$.

(b) Show that $\mathcal{B} = (1, i), (1, -i)$ is a basis of $\mathbb{C}^2$, and find $[T]_{\mathcal{B}\mathcal{B}}$. What does it tell you about $T \circ T$?

Solution on page 1135

Exercise 2.3.46. STANDARD For $\varphi \in \mathbb{R}$, let $\mathbf{u}_1 = (\cos\varphi, \sin\varphi)$ and $\mathbf{u}_2 = (-\sin\varphi, \cos\varphi)$, and let $\mathcal{U} = \mathbf{u}_1, \mathbf{u}_2$. Let R_θ be the rotation by θ and S_α the reflection across the line through $\mathbf{0}$ at angle α (Example 2.1.11).

(a) Show that $a\mathbf{u}_1 + b\mathbf{u}_2 = R_\varphi(a, b)$, that $\mathcal{U}$ is a basis of $\mathbb{R}^2$, and that the vector $(r\cos\beta, r\sin\beta)$ has $\mathcal{U}$-coordinates $(r\cos(\beta - \varphi), r\sin(\beta - \varphi))$.

(b) Show that $[R_\theta]_{\mathcal{U}\mathcal{U}} = [R_\theta]_{\mathcal{E}\mathcal{E}}$.

(c) Show that $[S_\alpha]_{\mathcal{U}\mathcal{U}} = \left[S_{\alpha-\varphi}\right]_{\mathcal{E}\mathcal{E}}$, and interpret this.

Solution on page 1136

Exercise 2.3.47. HARD Let $A \in \mathbb{C}^{n\times n}$ satisfy

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}| \quad \text{for every } i = 1, \ldots, n.$$

Show that if $A\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$.

Solution on page 1136

Exercise 2.3.48. CHALLENGE Let $n \geq 1$ and let $T : F[x]_{\leq n} \to F[x]_{\leq n}$, $T(f) = f(x + 1)$. Let $J \in F^{(n+1)\times(n+1)}$ have 1 in each diagonal position (k, k) and in each position $(k, k + 1)$ just above the diagonal, and 0 elsewhere. (Use the usual algebra of polynomials freely, including that substituting $x + 1$ respects products and that substituting twice in a row gives $f(x + 2)$; these facts are proved in Section 7.1.)

(a) Over $\mathbb{Q}$, let $b_0 = 1$ and $b_k = x(x-1)\cdots(x-k+1)/k!$ for $k \geq 1$. Show that $\mathcal{B} = b_0, b_1, \ldots, b_n$ is a basis of $\mathbb{Q}[x]_{\leq n}$ with $[T]_{\mathcal{B}\mathcal{B}} = J$.

(b) Let p be a prime. Show that over $\mathbb{F}_p$ there is a basis $\mathcal{B}$ of $\mathbb{F}_p[x]_{\leq n}$ with $[T]_{\mathcal{B}\mathcal{B}} = J$ if and only if $n < p$.

Solution on page 1137

2.4 Matrix multiplication

Once bases are fixed, a linear map and its matrix carry the same information (Theorem 2.3.20). Linear maps can be composed (Proposition 2.1.17), so matrices should be able to multiply, and the product of two matrices should be the matrix of the composite. This section takes that as the definition of the product. The formula for the entries, the algebraic laws, and the failures of commutativity and cancellation all follow from it, and each of them is the shadow of a fact about composition.

Throughout, U, V, and W are vector spaces over F with bases $\mathcal{A} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ of U, $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_m$ of V, and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_\ell$ of W, where $n, m, \ell \geq 1$. As in Notation 2.1.18, we write TS for the composite $T \circ S$.

2.4.1 The product of two matrices

Intuition (Computing $[TS]$ from $[T]$ and $[S]$). Let $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$, so that $TS \in \mathrm{L}(U,W)$. The j-th column of $[TS]_{\mathcal{C}\mathcal{A}}$ is the coordinate vector of $(TS)(\mathbf{u}_j) = T\big(S(\mathbf{u}_j)\big)$. By Theorem 2.3.9, applied to T at the vector $S(\mathbf{u}_j)$,

$$\big[T(S(\mathbf{u}_j))\big]_{\mathcal{C}} = [T]_{\mathcal{C}\mathcal{B}}\big[S(\mathbf{u}_j)\big]_{\mathcal{B}},$$

and $\big[S(\mathbf{u}_j)\big]_{\mathcal{B}}$ is the j-th column of $[S]_{\mathcal{B}\mathcal{A}}$. So the j-th column of $[TS]_{\mathcal{C}\mathcal{A}}$ is $[T]_{\mathcal{C}\mathcal{B}}$ times the j-th column of $[S]_{\mathcal{B}\mathcal{A}}$. As with the matrix–vector product in Section 2.3.2, the answer no longer mentions S, T, the spaces, or the bases. It is computed from the two matrices alone, so it is an operation on matrices, and we give it a name.

Definition 2.4.1 (Matrix product). Let $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$, and let $\mathbf{b}_1, \ldots, \mathbf{b}_n \in F^m$ be the columns of B. The **product** AB is the $\ell \times n$ matrix whose j-th column is $A\mathbf{b}_j$:

$$AB = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{pmatrix}.$$

It is defined only when the number of columns of A equals the number of rows of B. By Proposition 2.3.11, applied to the j-th column, the entry of AB in row i and column j is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}.$$

The sizes behave like dominoes: an $\ell \times m$ matrix times an $m \times n$ matrix is an $\ell \times n$ matrix, and the two inner sizes must match. A column $\mathbf{x} \in F^m$ is an $m \times 1$ matrix, and the product of A with it in the sense of Definition 2.4.1 has the single column $A\mathbf{x}$. So the matrix–vector product of Definition 2.3.8 is the special case $n = 1$, and the notation $A\mathbf{x}$ is unambiguous.

Theorem 2.4.2 (The matrix of a composite).

(a) For $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$,

$$[TS]_{\mathcal{C}\mathcal{A}} = [T]_{\mathcal{C}\mathcal{B}}[S]_{\mathcal{B}\mathcal{A}}.$$

(b) For $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$, $T_{AB} = T_A T_B$. That is, $(AB)\mathbf{x} = A(B\mathbf{x})$ for every $\mathbf{x} \in F^n$.

Proof. ROUTINE

> *Uses.* Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.9 (The matrix of a map computes the map), Definition 2.4.1 (Matrix product), Proposition 2.3.16 (The matrix of T_A), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Proposition 2.1.17 (Compositions of linear maps are linear)
> *Techniques.* Translate between maps and matrices (Technique 2.3.27)
> *Hint 1.* Two matrices of the same size are equal when their columns are. What is the j-th column of each side of (a)?
> *Hint 2.* The j-th column of $[TS]_{\mathcal{C}\mathcal{A}}$ is $\big[T\big(S(\mathbf{u}_j)\big)\big]_{\mathcal{C}}$. Apply Theorem 2.3.9 to T at the vector $S(\mathbf{u}_j)$.
> *Hint 3.* For (b), apply (a) with the standard bases, and then use that a linear map $F^n \to F^\ell$ is determined by its standard matrix.
> *Sketch.* (a) The j-th column of $[TS]_{\mathcal{C}\mathcal{A}}$ is $\big[T\big(S(\mathbf{u}_j)\big)\big]_{\mathcal{C}} = [T]_{\mathcal{C}\mathcal{B}}\big[S(\mathbf{u}_j)\big]_{\mathcal{B}}$, which is $[T]_{\mathcal{C}\mathcal{B}}$ times the j-th column of $[S]_{\mathcal{B}\mathcal{A}}$, the j-th column of the product. (b) By (a) with standard bases, $[T_A T_B]_{\mathcal{E}\mathcal{E}} = AB$, so $T_A T_B = T_{AB}$ by Proposition 2.3.17.

(a) Both sides are $\ell \times n$ matrices. By Definition 2.3.1, the j-th column of $[TS]_{\mathcal{C}\mathcal{A}}$ is $\big[(TS)(\mathbf{u}_j)\big]_{\mathcal{C}}$. By the definition of composition and Theorem 2.3.9,

$$\begin{aligned}\big[(TS)(\mathbf{u}_j)\big]_{\mathcal{C}} &= \big[T\big(S(\mathbf{u}_j)\big)\big]_{\mathcal{C}} \\ &= [T]_{\mathcal{C}\mathcal{B}}\big[S(\mathbf{u}_j)\big]_{\mathcal{B}}.\end{aligned}$$

The vector $\big[S(\mathbf{u}_j)\big]_{\mathcal{B}}$ is the j-th column of $[S]_{\mathcal{B}\mathcal{A}}$ (Definition 2.3.1), so by Definition 2.4.1 this is the j-th column of $[T]_{\mathcal{C}\mathcal{B}}[S]_{\mathcal{B}\mathcal{A}}$. The two matrices have the same columns, so they are equal.

(b) The map $T_A T_B : F^n \to F^\ell$ is linear (Proposition 2.1.17). By (a) with the standard bases of F^n, F^m, and F^ℓ, and by Proposition 2.3.16,

$$[T_A T_B]_{\mathcal{E}\mathcal{E}} = [T_A]_{\mathcal{E}\mathcal{E}}[T_B]_{\mathcal{E}\mathcal{E}} = AB.$$

By Proposition 2.3.17, every linear map $F^n \to F^\ell$ is multiplication by its standard matrix, so $T_A T_B = T_{AB}$. Evaluating at $\mathbf{x}$ gives $A(B\mathbf{x}) = (AB)\mathbf{x}$. □

The subscripts in (a) behave like the sizes: in $[T]_{\mathcal{C}\mathcal{B}}[S]_{\mathcal{B}\mathcal{A}}$ the two inner bases agree and drop out, leaving $[TS]_{\mathcal{C}\mathcal{A}}$. Also notice the order. TS means "first S, then T", and the matrix of the map applied first stands on the right, as S does in $T(S(\mathbf{u}))$. This is the price of writing functions to the left of their arguments, and it is why products of matrices are read from right to left.

As with the matrix–vector product (Proposition 2.3.10), the definition is not a convention. Even the special case of standard bases leaves no other choice.

> **Proposition 2.4.3** (The product is forced). Suppose that for all $\ell, m, n \geq 1$ we are given a rule that assigns to each $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$ a matrix $A \star B \in F^{\ell\times n}$, such that for all $S \in \mathrm{L}(F^n, F^m)$ and $T \in \mathrm{L}\big(F^m, F^\ell\big)$,
>
> $$[TS]_{\mathcal{E}\mathcal{E}} = [T]_{\mathcal{E}\mathcal{E}} \star [S]_{\mathcal{E}\mathcal{E}}.$$
>
> Then $A \star B = AB$ for all A and B.

Proof. ROUTINE

> *Uses.* Proposition 2.3.16 (The matrix of T_A), Theorem 2.4.2 (The matrix of a composite)
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* Given A and B, find linear maps whose standard matrices are A and B.
> *Hint 2.* Take $T = T_A$ and $S = T_B$, and compute $[T_A T_B]_{\mathcal{E}\mathcal{E}}$ in two ways.
> *Sketch.* With $T = T_A$ and $S = T_B$, the hypothesis gives $A \star B = [T_A T_B]_{\mathcal{E}\mathcal{E}}$, and Theorem 2.4.2 gives $[T_A T_B]_{\mathcal{E}\mathcal{E}} = AB$.

Let $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$. By Proposition 2.3.16, $[T_A]_{\mathcal{E}\mathcal{E}} = A$ and $[T_B]_{\mathcal{E}\mathcal{E}} = B$. Applying the hypothesis to $T = T_A$ and $S = T_B$, and then Theorem 2.4.2,

$$\begin{aligned}A \star B &= [T_A]_{\mathcal{E}\mathcal{E}} \star [T_B]_{\mathcal{E}\mathcal{E}} = [T_A T_B]_{\mathcal{E}\mathcal{E}} \\ &= [T_A]_{\mathcal{E}\mathcal{E}}[T_B]_{\mathcal{E}\mathcal{E}} = AB.\end{aligned}$$

□

Figure 2.4.1 shows Theorem 2.4.2(b) in the plane. The grid in the last panel of each row is the image of the square grid under the composite, and the two colored arrows are the columns of the product.

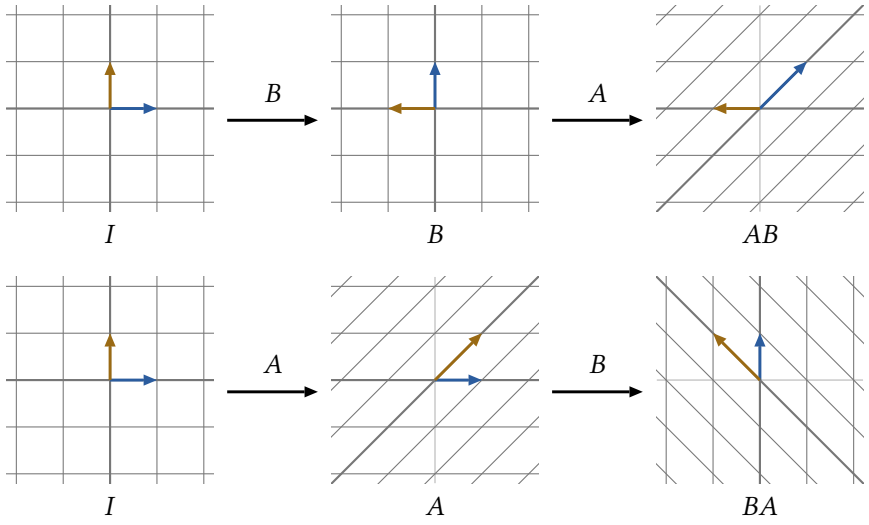


Figure 2.4.1. Composition in the plane, with $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (a shear) and $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (rotation by a right angle). Blue and ochre arrows are the images of $\mathbf{e}_1$ and $\mathbf{e}_2$. Top row: first B, then A. The result is $T_A T_B = T_{AB}$, and its arrows are the columns $(1, 1)$ and $(-1, 0)$ of $AB = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$. Bottom row: the same two maps in the other order, giving $BA = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$. The final grids differ, so $AB \neq BA$.

Example 2.4.4 (A product by columns). Let $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}$ and $B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ -1 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 2}$. The columns of AB are $A(2, 1, -1)$ and $A(1, 0, 4)$, which by Definition 2.3.8 are

$$2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \end{pmatrix},$$

so $AB = \begin{pmatrix} -1 & 9 \\ 2 & 4 \end{pmatrix}$. The product BA is also defined, but it is a 3×3 matrix:

$$BA = \begin{pmatrix} 2 & 1 & 5 \\ 1 & -1 & 2 \\ -1 & 13 & 2 \end{pmatrix},$$

whose columns are $B\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $B\begin{pmatrix} -1 \\ 3 \end{pmatrix}$, and $B\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Even the sizes of AB and BA differ.

Example 2.4.5 (Composites through their matrices).

(a) Let $D \in \mathrm{L}\big(F[x]_{\leq 2}\big)$ be differentiation and $\mathcal{B} = 1, x, x^2$, so that $N = [D]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ (Example 2.3.4). By Theorem 2.4.2, $\big[D^2\big]_{\mathcal{B}\mathcal{B}} = NN$ and $\big[D^3\big]_{\mathcal{B}\mathcal{B}} = NNN$, and indeed

$$NN = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad NNN = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The first says $D^2\big(a + bx + cx^2\big) = 2c$. The second is the matrix form of $D^3 = 0$, which we could only observe one vector at a time in Intuition (One basis or two).

(b) Rotating by β and then by α is rotating by $\alpha + \beta$, so $R_\alpha R_\beta = R_{\alpha+\beta}$ (Exercise 2.1.43). By Theorem 2.4.2 and Example 2.3.3,

$$\begin{aligned} &\big[R_{\alpha+\beta}\big]_{\mathcal{E}\mathcal{E}} \\ &= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}. \end{aligned}$$

The first column of the left side is $(\cos(\alpha+\beta), \sin(\alpha+\beta))$, and by the column view of the product, the first column of the right side is

$$\begin{aligned} &\cos\beta\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} + \sin\beta\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{pmatrix}. \end{aligned}$$

These are the angle addition formulas for $\cos(\alpha+\beta)$ and $\sin(\alpha+\beta)$. In Example 2.1.11 we used those formulas to show that rotations are linear, so here they are recovered rather than proved. But if one accepts from geometry that rotations are linear, then their matrices come from Corollary 2.2.17 without any trigonometric identity, as noted in Section 2.2.5, and this computation is a proof of the angle addition formulas.

(c) Let $\mathcal{B} = (1,1), (1,-1)$, a basis of F^2 when $1 + 1 \neq 0$. In Example 2.3.24 we found $[\mathrm{id}]_{\mathcal{E}\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $[\mathrm{id}]_{\mathcal{B}\mathcal{E}} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Since id id = id, Theorem 2.4.2 gives

$$[\mathrm{id}]_{\mathcal{E}\mathcal{B}}[\mathrm{id}]_{\mathcal{B}\mathcal{E}} = [\mathrm{id}]_{\mathcal{E}\mathcal{E}} = I_2,$$

and indeed $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, which becomes I_2 after multiplying by $\frac{1}{2}$. Converting standard coordinates into $\mathcal{B}$-coordinates and back changes nothing. Matrices whose products are I in this way are the invertible matrices of Section 2.5, and those of the form $[\mathrm{id}]_{\mathcal{C}\mathcal{B}}$ are studied in Section 4.5.

Exercise 2.4.6. ROUTINE

(a) Over $\mathbb{F}_5$, let $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 \\ 0 & 3 \\ 1 & 1 \end{pmatrix}$. Compute AB and BA. Which of $A^2 = AA$ and $B^2 = BB$ are defined?

(b) Over $\mathbb{C}$, compute $\begin{pmatrix} 1 & i \\ 0 & 2 \end{pmatrix}\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$ and $\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} 1 & i \\ 0 & 2 \end{pmatrix}$.

Solution on page 1138

2.4.2 Four views of the product

The definition describes AB one column at a time, but a product can be read in several other ways, and choosing the right one often makes a problem easy. We first fix some notation.

Notation 2.4.7 (Entries, rows, and columns). For a matrix M given by an expression, such as AB, we write $(M)_{ij}$ or M_{ij} for its entry in row i and column j, and

$$M_{i,\cdot} = \text{row } i \text{ of } M, \quad M_{\cdot,j} = \text{column } j \text{ of } M.$$

A row of an $\ell \times m$ matrix is a $1 \times m$ matrix, and a column is an $\ell \times 1$ matrix, that is, an element of F^{ℓ}. We identify a 1×1 matrix (c) with the scalar c. The **Kronecker delta** is

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j, \end{cases}$$

so that $I_n = (\delta_{ij})$ and $\mathbf{e}_j = (\delta_{1j}, \ldots, \delta_{nj})$.

Proposition 2.4.8 (Four views of the product). Let $A \in F^{\ell \times m}$ and $B \in F^{m \times n}$, and let $1 \leq i \leq \ell$ and $1 \leq j \leq n$.

(a) *Entries.* The (i, j) entry of AB is row i of A times column j of B:

$$(AB)_{ij} = A_{i,\cdot} B_{\cdot,j} = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

(b) *Columns.* Column j of AB is a combination of the columns of A, with coefficients from column j of B:

$$\begin{aligned} (AB)_{\cdot,j} &= AB_{\cdot,j} \\ &= b_{1j} A_{\cdot,1} + \cdots + b_{mj} A_{\cdot,m}. \end{aligned}$$

(c) *Rows.* Row i of AB is a combination of the rows of B, with coefficients from row i of A:

$$\begin{aligned} (AB)_{i,\cdot} &= A_{i,\cdot} B \\ &= a_{i1} B_{1,\cdot} + \cdots + a_{im} B_{m,\cdot}. \end{aligned}$$

(d) *Columns times rows.* AB is a sum of m products of a column with a row:

$$AB = A_{\cdot,1} B_{1,\cdot} + \cdots + A_{\cdot,m} B_{m,\cdot}.$$

Proof. ROUTINE

Uses. Definition 2.4.1 (Matrix product), Definition 2.3.8 (Matrix–vector product), Definition 1.2.7 (Matrices)
Hint 1. (a) and (b) are restatements of Definition 2.4.1. For (c) and (d), compare entries using (a).
Hint 2. In (d), $A_{\cdot,k} B_{k,\cdot}$ is an $\ell \times 1$ matrix times a $1 \times n$ matrix. Its inner size is 1, so the sum in (a) has a single term. What is its (i, j) entry?
Sketch. (a) is the entry formula of Definition 2.4.1, read for a $1 \times m$ times $m \times 1$ product. (b) is the definition together with Definition 2.3.8. In (c) and (d), the (i, j) entry of each side is $\sum_k a_{ik} b_{kj}$: the matrix $A_{\cdot,k} B_{k,\cdot}$ has (i, j) entry $a_{ik} b_{kj}$, and the operations on matrices are entrywise.

(a) The formula $(AB)_{ij} = \sum_k a_{ik} b_{kj}$ is part of Definition 2.4.1. Applied to the $1 \times m$ matrix $A_{i,\cdot}$ and the $m \times 1$ matrix $B_{\cdot,j}$, the same formula says that $A_{i,\cdot} B_{\cdot,j}$ is the 1×1 matrix with entry $\sum_k a_{ik} b_{kj}$.

(b) By Definition 2.4.1, column j of AB is $A\mathbf{b}_j$ with $\mathbf{b}_j = B_{\cdot,j} = (b_{1j}, \ldots, b_{mj})$, and by Definition 2.3.8, $A\mathbf{b}_j = b_{1j} A_{\cdot,1} + \cdots + b_{mj} A_{\cdot,m}$.

(c) By (a), applied to $A_{i,\cdot} \in F^{1 \times m}$ and B, the $1 \times n$ matrix $A_{i,\cdot} B$ has j-th entry $\sum_k a_{ik} b_{kj}$. Matrices are added and scaled entrywise (Definition 1.2.7), so the j-th entry of $\sum_k a_{ik} B_{k,\cdot}$ is also $\sum_k a_{ik} b_{kj}$. By (a), both equal the j-th entry of $(AB)_{i,\cdot}$.

(d) By (a), applied with inner size 1, the $\ell \times n$ matrix $A_{\cdot,k}B_{k,\cdot}$ has (i,j) entry $a_{ik}b_{kj}$. Adding these matrices entrywise over $k = 1, \ldots, m$ gives the matrix with (i,j) entry $\sum_k a_{ik}b_{kj} = (AB)_{ij}$. □

Each view is the right one for some job.

- The **entry view** (a) is the quickest way to compute a product by hand, and the one to use in proofs about individual entries, such as the shape of a product of triangular matrices (Proposition 2.4.36).
- The **column view** (b) answers questions about columns. Since every column of AB is a linear combination of the columns of A, the span of the columns of AB is contained in the span of the columns of A (Proposition 1.4.14): multiplying A on the right cannot produce new directions. This is the first step towards $\operatorname{rank}(AB) \le \operatorname{rank} A$ (Section 4.3). It also says that multiplying on the right recombines columns, which is how column operations become matrix products (Section 3.2).
- The **row view** (c) is the mirror image: multiplying B on the left recombines its rows. This is why row operations, and Gauss–Jordan elimination as a whole, are left multiplications (Section 3.2, Section 3.3).
- The **column–row view** (d) writes AB as a sum of m simple pieces, each built from one column and one row. It is the view to use when A has few columns, for block multiplication (Theorem 2.4.34), and for decompositions of a matrix into simple pieces (Section 4.3, Section 9.1).

Example 2.4.9 (One product, four ways). Let $A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 1 & 0 \\ 1 & -2 & 4 \end{pmatrix}$ over $\mathbb{R}$.

Entries. $(AB)_{11} = 1 \cdot 3 + 2 \cdot 1 = 5$, $(AB)_{12} = 1 - 4 = -3$, and so on, giving $AB = \begin{pmatrix} 5 & -3 & 8 \\ -3 & -1 & 0 \end{pmatrix}$.

Columns. $(AB)_{\cdot,1} = 3\begin{pmatrix} 1 \\ -1 \end{pmatrix} + 1\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$, and similarly for the other two columns.

Rows. Row 1 of A is $(1 \;\; 2)$, so

$$\begin{aligned}(AB)_{1,\cdot} &= 1(3 \;\; 1 \;\; 0) + 2(1 \;\; -2 \;\; 4) \\ &= (5 \;\; -3 \;\; 8).\end{aligned}$$

Row 2 of A is $(-1 \;\; 0)$, so row 2 of AB is -1 times row 1 of B, that is, $(-3 \;\; -1 \;\; 0)$. The row view gave this row with no computation at all.

Columns times rows.

$$\begin{aligned}AB = &\begin{pmatrix} 1 \\ -1 \end{pmatrix}(3 \;\; 1 \;\; 0) \\ &+\begin{pmatrix} 2 \\ 0 \end{pmatrix}(1 \;\; -2 \;\; 4).\end{aligned}$$

The two terms are $\begin{pmatrix} 3 & 1 & 0 \\ -3 & -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 2 & -4 & 8 \\ 0 & 0 & 0 \end{pmatrix}$, and their sum is AB.

Example 2.4.10 (Diagonal matrices). A square matrix D is **diagonal** if $d_{ij} = 0$ whenever $i \ne j$. We write $D = \operatorname{diag}(d_1, \ldots, d_n)$ for the diagonal matrix with diagonal entries $d_1, \ldots, d_n$. Row i of D is $d_i \mathbf{e}_i$ written as a row, so by the row view, row i of DA is $d_i A_{i,\cdot}$: multiplying by D on the left scales row i by d_i. Column j of D is $d_j \mathbf{e}_j$, so by the column view, column j of AD is $d_j A_{\cdot,j}$: multiplying on the right scales columns. In particular

$$\begin{aligned}&\operatorname{diag}(c_1, \ldots, c_n)\operatorname{diag}(d_1, \ldots, d_n) \\ &\quad = \operatorname{diag}(c_1 d_1, \ldots, c_n d_n),\end{aligned}$$

so diagonal matrices multiply entry by entry and commute with each other.

Exercise 2.4.11. ROUTINE Let $A \in F^{\ell \times m}$ and $B \in F^{m \times n}$.

(a) Show that if column j of B is zero, then so is column j of AB, and that if row i of A is zero, then so is row i of AB.

(b) Show that if $B_{:,3} = B_{:,1} + 2B_{:,2}$, then $(AB)_{:,3} = (AB)_{:,1} + 2(AB)_{:,2}$.

(c) Give an example in which AB has a zero column but B has none.

Solution on page 1138

2.4.3 Sums and scalar multiples

Matrices of the same size can also be added and scaled entrywise (Definition 1.2.7), and linear maps can be added and scaled pointwise (Proposition 2.1.23). These operations correspond to each other as well.

Proposition 2.4.12 (Matrices of sums and scalar multiples). Let $S, T \in \mathrm{L}(V, W)$ and $a \in F$. Then

$$[S+T]_{\mathcal{C}\mathcal{B}} = [S]_{\mathcal{C}\mathcal{B}} + [T]_{\mathcal{C}\mathcal{B}},$$
$$[aT]_{\mathcal{C}\mathcal{B}} = a[T]_{\mathcal{C}\mathcal{B}}.$$

So $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is a linear bijection from $\mathrm{L}(V, W)$ onto $F^{\ell\times m}$.

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 2.1.21 (Pointwise operations), Proposition 1.6.19 (Coordinates respect linear combinations), Definition 1.2.7 (Matrices), Theorem 2.3.20 (Linear maps correspond to matrices)
Hint 1. Compare the j-th columns of the two sides.
Hint 2. $(S+T)(\mathbf{v}_j) = S(\mathbf{v}_j) + T(\mathbf{v}_j)$ by the definition of $S+T$. Taking coordinates is linear.
Sketch. The j-th column of $[S+T]_{\mathcal{C}\mathcal{B}}$ is $\big[S(\mathbf{v}_j) + T(\mathbf{v}_j)\big]_{\mathcal{C}} = \big[S(\mathbf{v}_j)\big]_{\mathcal{C}} + \big[T(\mathbf{v}_j)\big]_{\mathcal{C}}$, the sum of the j-th columns of $[S]$ and $[T]$. Scalars are the same. Bijectivity is Theorem 2.3.20.

By Definition 2.3.1, Definition 2.1.21, and the linearity of $[\cdot]_{\mathcal{C}}$ (Proposition 1.6.19), the j-th column of $[S+T]_{\mathcal{C}\mathcal{B}}$ is

$$\begin{aligned}\big[(S+T)(\mathbf{v}_j)\big]_{\mathcal{C}} &= \big[S(\mathbf{v}_j) + T(\mathbf{v}_j)\big]_{\mathcal{C}} \\ &= \big[S(\mathbf{v}_j)\big]_{\mathcal{C}} + \big[T(\mathbf{v}_j)\big]_{\mathcal{C}}.\end{aligned}$$

Matrices are added entrywise (Definition 1.2.7), so the j-th column of $[S]_{\mathcal{C}\mathcal{B}} + [T]_{\mathcal{C}\mathcal{B}}$ is the sum of their j-th columns, which is the same vector. Likewise the j-th column of $[aT]_{\mathcal{C}\mathcal{B}}$ is $\big[aT(\mathbf{v}_j)\big]_{\mathcal{C}} = a\big[T(\mathbf{v}_j)\big]_{\mathcal{C}}$, the j-th column of $a[T]_{\mathcal{C}\mathcal{B}}$. So $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is linear, and it is a bijection by Theorem 2.3.20. □

So the dictionary of Section 2.3.6 has three more entries: sums, scalar multiples, and composites of maps correspond to sums, scalar multiples, and products of matrices. In the language of Section 4.5, the proposition says that $\mathrm{L}(V, W)$ and $F^{\ell\times m}$ are isomorphic vector spaces, and dimension counting there gives $\dim \mathrm{L}(V, W) = \ell m$.

Example 2.4.13 (The difference operator). Let $T \in \mathrm{L}\big(F[x]_{\leq 2}\big)$ be the substitution $T(p) = p(x+1)$, and $\mathcal{B} = 1, x, x^2$. From $T(1) = 1$, $T(x) = 1 + x$, and $T\big(x^2\big) = 1 + 2x + x^2$,

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

The difference operator of Example 2.3.28 is $\Delta = T - \mathrm{id}$, so by Proposition 2.4.12 and Example 2.3.6,

$$[\Delta]_{\mathcal{B}\mathcal{B}} = [T]_{\mathcal{B}\mathcal{B}} - I_3 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

which is the matrix computed there column by column.

2.4.4 The algebraic laws

Proposition 2.4.14 (Algebra of matrices). Let $A, A' \in F^{\ell\times m}$, $B, B' \in F^{m\times n}$, $C \in F^{n\times p}$, and $a \in F$.
(a) *Associativity.* $(AB)C = A(BC)$.
(b) *Identities.* $I_\ell A = A = AI_m$.
(c) *Distributivity.* $A(B+B') = AB + AB'$ and $(A+A')B = AB + A'B$.
(d) *Scalars.* $(aA)B = a(AB) = A(aB)$.

Intuition (The laws come from composition). Read a product of matrices as a sequence of maps, applied from right to left:

$$F^p \xrightarrow{T_C} F^n \xrightarrow{T_B} F^m \xrightarrow{T_A} F^\ell.$$

Both $(AB)C$ and $A(BC)$ are the matrix of "first T_C, then T_B, then T_A". The parentheses only record which two of the three steps were combined first, and the composite of three functions does not depend on that: both are $\mathbf{x} \mapsto T_A(T_B(T_C(\mathbf{x})))$. So associativity of matrix multiplication should follow from associativity of composition (Proposition 2.1.24) with no computation, once we can translate in both directions. That translation is Technique 2.3.27, and the dictionary we need is now complete: $T_{AB} = T_A T_B$ (Theorem 2.4.2), $T_{aA+A'} = aT_A + T_{A'}$ and $T_{I_\ell} = \mathrm{id}$ (Proposition 2.3.13), and A is determined by T_A (Proposition 2.3.17).

Proof (via composition). ROUTINE

Uses. Theorem 2.4.2 (The matrix of a composite), Proposition 2.1.24 (Algebra of composition), Proposition 2.3.13 (Linearity of the product), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Two matrices of the same size are equal when the maps T_X they define are equal (Proposition 2.3.17). So compute $T_{(AB)C}$ and $T_{A(BC)}$.
Hint 2. $T_{(AB)C} = T_{AB}T_C = (T_AT_B)T_C$. Now use associativity of composition and go back.
Hint 3. For (b), (c), and (d) you also need $T_{I_\ell} = \mathrm{id}$, $T_{B+B'} = T_B + T_{B'}$, and $T_{aA} = aT_A$, all from Proposition 2.3.13.
Sketch. Translate each side into a composite of maps T_X with Theorem 2.4.2 and Proposition 2.3.13, apply the matching identity of Proposition 2.1.24, and translate back. For instance $T_{(AB)C} = (T_AT_B)T_C = T_A(T_BT_C) = T_{A(BC)}$, so $(AB)C = A(BC)$ by the uniqueness in Proposition 2.3.17.

By Proposition 2.3.17, two matrices $X, Y \in F^{r\times s}$ with $T_X = T_Y$ are equal. So in each case it suffices to show that the two sides define the same linear map. We use Theorem 2.4.2(b), $T_{XY} = T_XT_Y$, and Proposition 2.3.13, which says that $T_{cX+X'} = cT_X + T_{X'}$ and that $T_{I_r} = \mathrm{id}_{F^r}$.

(a) By Theorem 2.4.2 and part (a) of Proposition 2.1.24,

$$\begin{aligned} T_{(AB)C} &= T_{AB}T_C = (T_AT_B)T_C \\ &= T_A(T_BT_C) = T_AT_{BC} = T_{A(BC)}. \end{aligned}$$

(b) $T_{I_\ell A} = T_{I_\ell}T_A = \mathrm{id}\, T_A = T_A$, and $T_{AI_m} = T_A\, \mathrm{id} = T_A$, by part (b) of Proposition 2.1.24.

(c) By part (c) of Proposition 2.1.24,

$$\begin{aligned} T_{A(B+B')} &= T_A(T_B + T_{B'}) = T_AT_B + T_AT_{B'} \\ &= T_{AB} + T_{AB'} = T_{AB+AB'}, \end{aligned}$$

and in the same way $T_{(A+A')B} = (T_A + T_{A'})T_B$ equals $T_{AB+A'B}$.

(d) By part (d) of Proposition 2.1.24,

$$\begin{aligned} T_{(aA)B} &= (aT_A)T_B = a(T_AT_B) \\ &= T_A(aT_B) = T_{A(aB)}, \end{aligned}$$

and the middle expression is $aT_{AB} = T_{a(AB)}$. □

The first proof never looks at an entry. The second proof works with entries only, and it is the prototype of a move we will use constantly.

> **Technique 2.4.15** (Compute entry by entry). To prove an identity between matrices, show that the (i, j) entries of the two sides agree for all i and j. Write each entry of a product with Proposition 2.4.8(a), using a new summation index for every product, then rearrange: pull constant factors out of sums, exchange the order of two finite sums, $\sum_k \sum_l = \sum_l \sum_k$, and collapse sums containing a Kronecker delta, $\sum_k \delta_{ik} x_k = x_i$.
>
> Reach for it when the statement is about the entries themselves (transposes, traces, triangular or block shapes), or when there is no linear map in sight, for example for matrices whose entries are integers, polynomials, or other matrices. Two things go wrong in practice. Reusing a summation index in a nested sum silently identifies two different indices. And moving a factor past another needs commutativity of the entries, which holds for scalars but not for blocks (Section 2.4.8). The canonical example is the second proof of associativity below.
>
> *Used in.* Proposition 2.4.14, Proposition 2.4.23, Proposition 2.4.28, Theorem 2.4.34, Proposition 2.4.36, Proposition 2.4.39, Exercise 2.8.6, Exercise 2.8.11, Problem 2.4, Problem 2.12, Theorem 5.5.9, Proposition 6.3.27, Proposition 7.1.3, Theorem 7.6.28, Proposition 8.3.25, Exercise 2.6.46, Exercise 6.3.60, Exercise 8.2.45.

Proof (entry by entry). ROUTINE

> *Uses.* Proposition 2.4.8 (Four views of the product), Definition 1.1.1 (Field), Notation 2.4.7 (Entries, rows, and columns)
> *Techniques.* Compute entry by entry (Technique 2.4.15)
> *Hint 1.* Write $((AB)C)_{ij}$ as a double sum, using the entry formula twice with two different summation indices.
> *Hint 2.* $((AB)C)_{ij} = \sum_l \bigl(\sum_k a_{ik} b_{kl}\bigr) c_{lj}$. Distribute, exchange the two sums, and factor out a_{ik}.
> *Sketch.* Both $((AB)C)_{ij}$ and $(A(BC))_{ij}$ equal $\sum_k \sum_l a_{ik} b_{kl} c_{lj}$, after distributing and exchanging the finite sums. The other laws follow in the same way from the field axioms, with $(I_\ell)_{ik} = \delta_{ik}$ for (b).

(a) For $1 \le i \le \ell$ and $1 \le j \le p$, by Proposition 2.4.8(a) twice,

$$\begin{aligned}((AB)C)_{ij} &= \sum_{l=1}^{n} (AB)_{il} c_{lj} = \sum_{l=1}^{n} \left(\sum_{k=1}^{m} a_{ik} b_{kl} \right) c_{lj} \\ &= \sum_{l=1}^{n} \sum_{k=1}^{m} a_{ik} b_{kl} c_{lj} = \sum_{k=1}^{m} \sum_{l=1}^{n} a_{ik} b_{kl} c_{lj} \\ &= \sum_{k=1}^{m} a_{ik} \left(\sum_{l=1}^{n} b_{kl} c_{lj} \right) = \sum_{k=1}^{m} a_{ik} (BC)_{kj} \\ &= (A(BC))_{ij}.\end{aligned}$$

The steps used distributivity and associativity in F and the reordering of a finite sum.

(b) $(I_\ell A)_{ij} = \sum_k \delta_{ik} a_{kj} = a_{ij}$, since only the term $k = i$ survives, and similarly $(AI_m)_{ij} = \sum_k a_{ik} \delta_{kj} = a_{ij}$.

(c) By distributivity in F,

$$\begin{aligned}(A(B + B'))_{ij} &= \sum_k a_{ik} \bigl(b_{kj} + b'_{kj} \bigr) \\ &= \sum_k a_{ik} b_{kj} + \sum_k a_{ik} b'_{kj} \\ &= (AB + AB')_{ij},\end{aligned}$$

and the other law is the same with the factors on the other side.

(d) Using commutativity of F to move a to the front,

$$((aA)B)_{ij} = \sum_k (aa_{ik})b_{kj} = a\sum_k a_{ik}b_{kj}$$
$$= \sum_k a_{ik}\big(ab_{kj}\big) = (A(aB))_{ij},$$

and the middle expression is $(a(AB))_{ij}$. □

The two proofs teach different things. The first explains *why* the laws hold: matrices multiply like composites because they are composites in disguise, and composition of functions is associative for trivial reasons. All the work went into the dictionary between maps and matrices. The second is self-contained and mechanical. Look at what it used in (a): the laws of addition, associativity of multiplication, and distributivity, but nothing about inverses or commutativity of multiplication. So the same computation proves associativity for matrices with entries in $\mathbb{Z}$, in $F[x]$, or in any other number system with these laws, even one whose multiplication is not commutative, where there may be no vector spaces and linear maps to translate into. This is what makes block multiplication associative (Section 2.4.8), since blocks are matrices whose "entries" are matrices.

By associativity, products of three or more matrices need no parentheses.

Notation 2.4.16 (Powers of a matrix). For $A \in F^{n\times n}$ and an integer $k \geq 1$, $A^k = AA\cdots A$ (k factors), and $A^0 = I_n$. Then $A^jA^k = A^{j+k}$, and if $T \in \mathrm{L}(V)$ and $\mathcal{B}$ is a basis of V, then $\left[T^k\right]_{\mathcal{B}\mathcal{B}} = \left([T]_{\mathcal{B}\mathcal{B}}\right)^k$ by Theorem 2.4.2 and induction on k. Together with Proposition 2.4.12,

$$\left[a_0\,\mathrm{id} + a_1T + \cdots + a_kT^k\right]_{\mathcal{B}\mathcal{B}}$$
$$= a_0I_n + a_1A + \cdots + a_kA^k,$$

where $A = [T]_{\mathcal{B}\mathcal{B}}$.

Exercise 2.4.17. ROUTINE Let $A, B \in F^{n\times n}$. Show that $(A+B)^2 = A^2 + AB + BA + B^2$, and deduce that $(A+B)^2 = A^2 + 2AB + B^2$ if and only if $AB = BA$. (Here $2AB = AB + AB$.)

Solution on page 1139

2.4.5 What fails

Matrix multiplication shares associativity and distributivity with the multiplication of numbers, but three other familiar properties fail. In each case the failure is already visible for composition of maps.

Example 2.4.18 ($AB \neq BA$). Let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

In terms of maps, T_A sends $\mathbf{e}_2 \mapsto \mathbf{e}_1 \mapsto \mathbf{0}$ and T_B sends $\mathbf{e}_1 \mapsto \mathbf{e}_2 \mapsto \mathbf{0}$. Following $\mathbf{e}_1$ through T_AT_B returns it to $\mathbf{e}_1$, but following it through T_BT_A kills it. For a geometric example, let P be the projection onto the x-axis and R the rotation by a right angle, so that $PR \neq RP$ by Example 2.1.19. In matrices,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Figure 2.4.1 is a third example. When AB is defined, BA may not be, and if both are defined they may have different sizes (Example 2.4.4). So "$AB = BA$" is a strong condition, which we express by saying that A and B **commute**.

Example 2.4.19 (Zero divisors and the failure of cancellation). With A as above and $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,

$$A^2 = AP = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad PA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

So a product of nonzero matrices can be zero, and even a nonzero square can be zero. In the language of Proposition 1.1.17, $F^{2\times 2}$ has **zero divisors**. Note also that $AP = 0$ while $PA \neq 0$.

Cancellation fails with them: $AP = A0$ with $A \neq 0$, but $P \neq 0$. In general, $AB = AC$ means $A(B - C) = 0$ (Proposition 2.4.14(c)), so cancelling A is legitimate exactly when $AX = 0$ forces $X = 0$. For A that are left-invertible (Section 2.5) this holds; for the A above it does not, since T_A kills $\mathbf{e}_1$, and any X whose columns lie in $\operatorname{span}(\mathbf{e}_1)$ satisfies $AX = 0$.

These examples are 2×2, which is the smallest size where anything can go wrong: $F^{1\times 1}$ is just F, where multiplication is commutative and there are no zero divisors. For $n \geq 2$, put the 2×2 examples in the top left corner of an $n \times n$ matrix that is zero elsewhere. By block multiplication (Theorem 2.4.34) the products are computed in that corner, so $F^{n\times n}$ has the same failures. Finding such examples is an instance of Technique 1.4.18: the first place to look is 2×2 matrices whose entries are 0 and 1.

Some matrices do commute. A scalar matrix cI_n commutes with every $A \in F^{n\times n}$, since $(cI_n)A = cA = A(cI_n)$ by Proposition 2.4.14. Any A commutes with its own powers, since $AA^k = A^{k+1} = A^kA$, and hence with every $a_0I_n + a_1A + \cdots + a_kA^k$. Diagonal matrices commute with each other (Example 2.4.10), and so do the rotation matrices $[R_\alpha]$ and $[R_\beta]$, since both products equal $[R_{\alpha+\beta}]$. Which matrices commute with a given one is a recurring question; see Exercise 2.4.42 and Exercise 2.4.43.

Exercise 2.4.20. STANDARD Find $A, B \in \mathbb{R}^{2\times 2}$, none of whose eight entries is 0, with $AB = 0$ but $BA \neq 0$.

Solution on page 1139

2.4.6 The transpose

The row view and the column view are mirror images of each other. The transpose makes the mirror precise.

Definition 2.4.21 (Transpose). The **transpose** of $A = (a_{ij}) \in F^{m\times n}$ is the matrix $A^{\mathsf{T}} \in F^{n\times m}$ with

$$\left(A^{\mathsf{T}}\right)_{ij} = a_{ji} \quad (1 \leq i \leq n, 1 \leq j \leq m).$$

That is, row i of A^{T} is column i of A written as a row, and column j of A^{T} is row j of A written as a column. A square matrix A is **symmetric** if $A^{\mathsf{T}} = A$ and **skew-symmetric** if $A^{\mathsf{T}} = -A$.

For example,

$$\begin{pmatrix} 1 & 0 & -2 \\ 3 & 5 & 4 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & 3 \\ 0 & 5 \\ -2 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 2 & 7 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 1 & 2 \\ 2 & 7 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix}.$$

The second matrix is symmetric and the third skew-symmetric; symmetric matrices form the subspace of Example 1.4.11. A vector $\mathbf{x} \in F^n$ is a column, so $\mathbf{x}^{\mathsf{T}} = (x_1 \ \cdots \ x_n)$ is a row, and for $\mathbf{x}, \mathbf{y} \in F^n$ the 1×1 product

$$\mathbf{x}^{\mathsf{T}}\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

is a scalar. In the other order, for $\mathbf{u} \in F^m$ and $\mathbf{v} \in F^n$, the column–row product $\mathbf{u}\mathbf{v}^{\mathsf{T}} \in F^{m\times n}$ has (i,j) entry u_iv_j, as in Proposition 2.4.8(d).

Proposition 2.4.22 (Transpose is a linear involution). For $A, B \in F^{m\times n}$ and $a \in F$,

$$\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A, \quad I_n^{\mathsf{T}} = I_n,$$
$$(aA + B)^{\mathsf{T}} = aA^{\mathsf{T}} + B^{\mathsf{T}}.$$

So $A \mapsto A^{\mathsf{T}}$ is a linear bijection $F^{m\times n} \to F^{n\times m}$, whose inverse is transposition $F^{n\times m} \to F^{m\times n}$.

Proof. ROUTINE

Uses. Definition 2.4.21 (Transpose), Definition 1.2.7 (Matrices), Definition 2.3.5 (Identity matrix)
Hint 1. Compare entries. The (i,j) entry of $\left(A^{\mathsf{T}}\right)^{\mathsf{T}}$ is the (j,i) entry of A^{T}.
Hint 2. For the last claim, transposing twice gives back the matrix you started with, whichever of the two directions you apply first.
Sketch. $\left(\left(A^{\mathsf{T}}\right)^{\mathsf{T}}\right)_{ij} = \left(A^{\mathsf{T}}\right)_{ji} = a_{ij}$, $\left((aA+B)^{\mathsf{T}}\right)_{ij} = aa_{ji} + b_{ji}$, and $\delta_{ji} = \delta_{ij}$. The two transposition maps $F^{m\times n} \to F^{n\times m}$ and back are inverse to each other.

By Definition 2.4.21, $\left(\left(A^{\mathsf{T}}\right)^{\mathsf{T}}\right)_{ij} = \left(A^{\mathsf{T}}\right)_{ji} = a_{ij}$, and since matrices are added and scaled entrywise,

$$\begin{aligned}\left((aA+B)^{\mathsf{T}}\right)_{ij} &= (aA+B)_{ji} = aa_{ji} + b_{ji} \\ &= \left(aA^{\mathsf{T}} + B^{\mathsf{T}}\right)_{ij}.\end{aligned}$$

The (i,j) entry of I_n^{T} is $\delta_{ji} = \delta_{ij}$. So transposition is linear, and applying it twice gives the identity map. Hence it is a bijection from $F^{m\times n}$ to $F^{n\times m}$, with transposition $F^{n\times m} \to F^{m\times n}$ as its inverse. □

Intuition (Why the order reverses). Before guessing a formula for $(AB)^{\mathsf{T}}$, check sizes. If $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$, then $(AB)^{\mathsf{T}}$ is $n \times \ell$, while $A^{\mathsf{T}} \in F^{m\times\ell}$ and $B^{\mathsf{T}} \in F^{n\times m}$. So $A^{\mathsf{T}}B^{\mathsf{T}}$ is not even defined unless $\ell = n$, and the only product of the two transposes with the right size is $B^{\mathsf{T}}A^{\mathsf{T}}$. The same reversal will appear for inverses of composites (Section 2.5) and for adjoints (Section 6.6).

Proposition 2.4.23 (Transpose of a product). For $A \in F^{\ell\times m}$ and $B \in F^{m\times n}$,

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}.$$

Proof (entry by entry). ROUTINE

Uses. Definition 2.4.21 (Transpose), Proposition 2.4.8 (Four views of the product), Definition 1.1.1 (Field)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Compute the (i,j) entry of each side with Definition 2.4.21 and the entry formula.
Hint 2. $\left((AB)^{\mathsf{T}}\right)_{ij} = (AB)_{ji} = \sum_k a_{jk}b_{ki}$. Write each factor as an entry of a transpose, and reorder each product.
Sketch. The (i,j) entry of $(AB)^{\mathsf{T}}$ is $\sum_k a_{jk}b_{ki}$. Since $a_{jk}b_{ki} = b_{ki}a_{jk} = \left(B^{\mathsf{T}}\right)_{ik}\left(A^{\mathsf{T}}\right)_{kj}$, this sum is the (i,j) entry of $B^{\mathsf{T}}A^{\mathsf{T}}$.

Both sides are $n \times \ell$. Let $1 \le i \le n$ and $1 \le j \le \ell$. By Definition 2.4.21 and Proposition 2.4.8(a), then commutativity of multiplication in F, then Definition 2.4.21 and Proposition 2.4.8(a) again,

$$\left((AB)^{\mathsf{T}}\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{m} a_{jk}b_{ki} = \sum_{k=1}^{m} b_{ki}a_{jk}$$
$$= \sum_{k=1}^{m} \left(B^{\mathsf{T}}\right)_{ik}\left(A^{\mathsf{T}}\right)_{kj} = \left(B^{\mathsf{T}}A^{\mathsf{T}}\right)_{ij}.$$

□

The first proof used commutativity of F at one step, and that step is where the order of the factors flips. The second proof explains the reversal through the views of the product: the row view of AB is the column view of $B^{\mathsf{T}}A^{\mathsf{T}}$ in a mirror.

Proof (rows become columns). ROUTINE

> *Uses.* Proposition 2.4.8 (Four views of the product), Proposition 2.4.22 (Transpose is a linear involution), Definition 2.4.21 (Transpose), Definition 2.3.8 (Matrix–vector product)
> *Hint 1.* Column i of $(AB)^{\mathsf{T}}$ is row i of AB, written as a column. What is row i of AB in the row view?
> *Hint 2.* Column i of $B^{\mathsf{T}}A^{\mathsf{T}}$ is B^{T} times column i of A^{T}. Expand this with the column view, and compare with the transpose of the row-view expansion.
> *Sketch.* By the row view, row i of AB is $\sum_k a_{ik}B_{k,\cdot}$; transposing (a linear operation) gives $\sum_k a_{ik}\left(B^{\mathsf{T}}\right)_{\cdot,k}$. By the column view, column i of $B^{\mathsf{T}}A^{\mathsf{T}}$ is $B^{\mathsf{T}}\left(A^{\mathsf{T}}\right)_{\cdot,i} = \sum_k a_{ik}\left(B^{\mathsf{T}}\right)_{\cdot,k}$, since $\left(A^{\mathsf{T}}\right)_{ki} = a_{ik}$. The two matrices have the same columns.

By Definition 2.4.21, column i of $(AB)^{\mathsf{T}}$ is the transpose of row i of AB. By the row view (Proposition 2.4.8(c)) and the linearity of transposition (Proposition 2.4.22),

$$\left((AB)_{i,\cdot}\right)^{\mathsf{T}} = \left(\sum_{k=1}^{m} a_{ik}B_{k,\cdot}\right)^{\mathsf{T}} = \sum_{k=1}^{m} a_{ik}\left(B_{k,\cdot}\right)^{\mathsf{T}}$$
$$= \sum_{k=1}^{m} a_{ik}\left(B^{\mathsf{T}}\right)_{\cdot,k},$$

since the transpose of row k of B is column k of B^{T}. On the other side, column i of A^{T} is the transpose of row i of A, with entries $a_{i1}, \ldots, a_{im}$. By the column view (Proposition 2.4.8(b)), column i of $B^{\mathsf{T}}A^{\mathsf{T}}$ is

$$B^{\mathsf{T}}\left(A^{\mathsf{T}}\right)_{\cdot,i} = a_{i1}\left(B^{\mathsf{T}}\right)_{\cdot,1} + \cdots + a_{im}\left(B^{\mathsf{T}}\right)_{\cdot,m}.$$

So $(AB)^{\mathsf{T}}$ and $B^{\mathsf{T}}A^{\mathsf{T}}$ have the same columns. □

By induction, $(A_1A_2\cdots A_k)^{\mathsf{T}} = A_k^{\mathsf{T}}\cdots A_2^{\mathsf{T}}A_1^{\mathsf{T}}$, and in particular $\left(A^k\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^k$. Proposition 2.4.23 is the engine of Technique 3.2.27, which turns facts about rows and left multiplication into facts about columns and right multiplication. The transpose also has a meaning in terms of linear maps, which needs more structure than a matrix: with an inner product it is the adjoint (Section 6.6), and in general it is the dual map (Section 13.1).

The column–row products $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ of standard basis vectors are the simplest matrices of all, and we will use them repeatedly.

Lemma 2.4.24 (Matrix units). For $1 \le i \le m$ and $1 \le j \le n$, the matrix $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} \in F^{m\times n}$ has a 1 in position (i, j) and 0 everywhere else.

(a) Every $A \in F^{m\times n}$ is $A = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$, and these are the only coefficients that work. So the mn matrices $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ form a basis of $F^{m\times n}$.
(b) For $A \in F^{m\times n}$: $A\mathbf{e}_j = A_{\cdot,j}$, $\mathbf{e}_i^{\mathsf{T}}A = A_{i,\cdot}$, and $\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_j = a_{ij}$.
(c) $\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right)\left(\mathbf{e}_k\mathbf{e}_l^{\mathsf{T}}\right) = \delta_{jk}\mathbf{e}_i\mathbf{e}_l^{\mathsf{T}}$, for matrices of compatible sizes.

Proof. ROUTINE

> *Uses.* Proposition 2.4.8 (Four views of the product), Proposition 2.4.14 (Algebra of matrices), Proposition 2.3.13 (Linearity of the product), Definition 1.6.1 (Basis), Notation 2.4.7 (Entries, rows, and columns)
> *Hint 1.* The (k, l) entry of the column–row product $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is u_kv_l. For $\mathbf{u} = \mathbf{e}_i$ and $\mathbf{v} = \mathbf{e}_j$ that is $\delta_{ki}\delta_{lj}$.

Hint 2. For (c), regroup with associativity: $\mathbf{e}_i\big(\mathbf{e}_j^{\mathsf{T}}\mathbf{e}_k\big)\mathbf{e}_l^{\mathsf{T}}$, and the middle factor is a 1×1 matrix.
Sketch. The (k,l) entry of $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ is $\delta_{ki}\delta_{lj}$. (a) The (k,l) entry of $\sum c_{ij}\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ is c_{kl}, so it equals A exactly when $c_{kl}=a_{kl}$. (b) Column and row views. (c) Associativity and $\mathbf{e}_j^{\mathsf{T}}\mathbf{e}_k=\delta_{jk}$.

By Proposition 2.4.8(a) with inner size 1, the (k,l) entry of $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ is $\delta_{ki}\delta_{lj}$, which is 1 if $(k,l)=(i,j)$ and 0 otherwise.

(a) Since matrices are added and scaled entrywise, the (k,l) entry of $\sum_{i,j} c_{ij}\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ is $\sum_{i,j} c_{ij}\delta_{ki}\delta_{lj}=c_{kl}$. So this sum equals A if and only if $c_{kl}=a_{kl}$ for all k,l. Every matrix is therefore a linear combination of the $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ in exactly one way, which is what it means for them to form a basis (Definition 1.6.1), with the entries of A as its coordinates.

(b) $A\mathbf{e}_j=A_{\cdot,j}$ is Proposition 2.3.13(c). By the row view (Proposition 2.4.8(c)), $\mathbf{e}_i^{\mathsf{T}}A=\sum_k \delta_{ik}A_{k,\cdot}=A_{i,\cdot}$. Then $\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_j=A_{i,\cdot}\mathbf{e}_j$ is the j-th entry of row i, which is a_{ij}.

(c) By associativity (Proposition 2.4.14) and (b), $\mathbf{e}_j^{\mathsf{T}}\mathbf{e}_k$ is the (j,k) entry of I, which is δ_{jk}. So

$$\big(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\big)\big(\mathbf{e}_k\mathbf{e}_l^{\mathsf{T}}\big)=\mathbf{e}_i\big(\mathbf{e}_j^{\mathsf{T}}\mathbf{e}_k\big)\mathbf{e}_l^{\mathsf{T}}=\delta_{jk}\mathbf{e}_i\mathbf{e}_l^{\mathsf{T}}.$$

□

Part (c) gives the noncommutativity of Example 2.4.18 in its purest form: for $i\neq j$, $\big(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\big)\big(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\big)=\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$, while $\big(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\big)\big(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\big)=\mathbf{e}_j\mathbf{e}_j^{\mathsf{T}}$. The matrices A and B of that example are $\mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}$ and $\mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}$.

Exercise 2.4.25. ROUTINE

(a) Show that $A^{\mathsf{T}}A$ and AA^{T} are symmetric for every $A\in F^{m\times n}$.

(b) Let $A,B\in F^{n\times n}$ be symmetric. Show that $AB+BA$ is symmetric, that $AB-BA$ is skew-symmetric, and that AB is symmetric if and only if $AB=BA$.

Solution on page 1139

2.4.7 The trace

Matrix multiplication is not commutative, but there is a number attached to a square matrix that cannot tell AB from BA.

Definition 2.4.26 (Trace). The **trace** of $A=\big(a_{ij}\big)\in F^{n\times n}$ is the sum of its diagonal entries,

$$\operatorname{tr}A=a_{11}+a_{22}+\cdots+a_{nn}=\sum_{i=1}^{n}a_{ii}.$$

For example, $\operatorname{tr}\left(\begin{smallmatrix}1&2\\3&4\end{smallmatrix}\right)=5$ and $\operatorname{tr}I_n=n*1$, which is n over $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ and is 0 over $\mathbb{F}_p$ when $p\mid n$. In this notation, the trace-zero matrices of Example 1.4.11 are those with $\operatorname{tr}A=0$.

Proposition 2.4.27 (The trace is linear). For $A,B\in F^{n\times n}$ and $a\in F$, $\operatorname{tr}(aA+B)=a\operatorname{tr}A+\operatorname{tr}B$ and $\operatorname{tr}A^{\mathsf{T}}=\operatorname{tr}A$.

Proof. ROUTINE

Uses. Definition 2.4.26 (Trace), Definition 1.2.7 (Matrices), Definition 2.4.21 (Transpose)
Hint 1. The diagonal entries of $aA+B$ are $aa_{ii}+b_{ii}$.
Hint 2. Transposing does not move the diagonal entries: $\big(A^{\mathsf{T}}\big)_{ii}=a_{ii}$.
Sketch. The diagonal entries of $aA+B$ are $aa_{ii}+b_{ii}$; add them up. Transposing fixes the diagonal: $\big(A^{\mathsf{T}}\big)_{ii}=a_{ii}$.

Matrices are added and scaled entrywise, so

$$\operatorname{tr}(aA+B)=\sum_i(aa_{ii}+b_{ii})$$

$$= a \sum_i a_{ii} + \sum_i b_{ii} = a \operatorname{tr} A + \operatorname{tr} B.$$

By Definition 2.4.21, $\left(A^{\mathsf{T}}\right)_{ii} = a_{ii}$, so $\operatorname{tr} A^{\mathsf{T}} = \operatorname{tr} A$. □

The trace is not multiplicative: for $A = \operatorname{diag}(1,0)$ and $B = \operatorname{diag}(0,1)$ we have $AB = 0$, so $\operatorname{tr}(AB) = 0$, but $\operatorname{tr} A \operatorname{tr} B = 1$. What it does satisfy is the following.

Proposition 2.4.28 (Cyclicity of the trace). For $A \in F^{m\times n}$ and $B \in F^{n\times m}$,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Intuition. Here AB is $m \times m$ and BA is $n \times n$, so the two products may not even have the same size, let alone be equal. But the diagonal entries of AB are $(AB)_{ii} = \sum_k a_{ik}b_{ki}$, so $\operatorname{tr}(AB)$ is the sum of $a_{ik}b_{ki}$ over all pairs (i,k): every entry of A multiplied by the entry of B in the mirror position. That description is symmetric in A and B. There is also an approach that avoids indices. For fixed A, both sides are linear in B, so by Technique 2.2.7 it is enough to check the identity for the matrix units of Lemma 2.4.24.

Proof (entry by entry). ROUTINE

Uses. Definition 2.4.26 (Trace), Proposition 2.4.8 (Four views of the product), Definition 1.1.1 (Field)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Write $\operatorname{tr}(AB) = \sum_i (AB)_{ii}$ and expand each diagonal entry with the entry formula.
Hint 2. You get a double sum $\sum_i \sum_k a_{ik}b_{ki}$. Exchange the order of summation and recognize the diagonal entries of BA.
Sketch. Both $\operatorname{tr}(AB)$ and $\operatorname{tr}(BA)$ expand to the double sum of $a_{ik}b_{ki}$ over all $i \le m$ and $k \le n$; the two expansions differ only in the order of summation.

By Definition 2.4.26 and Proposition 2.4.8(a),

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} a_{ik}b_{ki} \\ &= \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ki}a_{ik} \\ &= \sum_{k=1}^{n} (BA)_{kk} = \operatorname{tr}(BA),\end{aligned}$$

where we exchanged the two finite sums and used commutativity of multiplication in F. □

Proof (check on matrix units). STANDARD

Uses. Lemma 2.4.24 (Matrix units), Proposition 2.4.27 (The trace is linear), Proposition 2.4.14 (Algebra of matrices), Corollary 2.2.6 (Agreeing on a spanning list), Definition 2.4.26 (Trace)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. Fix A and define $\varphi, \psi : F^{n\times m} \to F$ by $\varphi(X) = \operatorname{tr}(AX)$ and $\psi(X) = \operatorname{tr}(XA)$. Why are these linear?
Hint 2. By Corollary 2.2.6, it is enough that φ and ψ agree on the matrix units $\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}$ $(1 \le k \le n,\ 1 \le i \le m)$.
Hint 3. $A\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}$ is zero except for column i, which is $A_{:,k}$. Which diagonal entry survives? Do the same for $\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}A$ with rows.
Sketch. $\varphi(X) = \operatorname{tr}(AX)$ and $\psi(X) = \operatorname{tr}(XA)$ are linear in X by distributivity and Proposition 2.4.27. $A\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right) = A_{:,k}\mathbf{e}_i^{\mathsf{T}}$ has only column i nonzero, so its trace is its (i,i) entry a_{ik}; $\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right)A = \mathbf{e}_k A_{i,:}$ has only row k nonzero, with (k,k) entry a_{ik}. They agree on a spanning list, so $\varphi = \psi$; take $X = B$.

Fix $A \in F^{m\times n}$, and define $\varphi, \psi : F^{n\times m} \to F$ by $\varphi(X) = \operatorname{tr}(AX)$ and $\psi(X) = \operatorname{tr}(XA)$. Both are linear: for instance $\varphi(cX + Y) = \operatorname{tr}(cAX + AY) = c\varphi(X) + \varphi(Y)$ by Proposition 2.4.14 and Proposition 2.4.27. By Lemma 2.4.24(a), the matrices $\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}$ with $1 \le k \le n$ and $1 \le i \le m$ span $F^{n\times m}$, so by Corollary 2.2.6 it suffices to show $\varphi\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right) = \psi\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right)$.

By associativity and Lemma 2.4.24(b), $A\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right) = (A\mathbf{e}_k)\mathbf{e}_i^{\mathsf{T}} = A_{:,k}\mathbf{e}_i^{\mathsf{T}}$. This $m \times m$ matrix has (r,s) entry $a_{rk}\delta_{si}$, so its only nonzero column is column i, and its only diagonal entry that can be nonzero is the (i,i) entry a_{ik}. Hence $\varphi\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right) = a_{ik}$. Similarly, $\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right)A = \mathbf{e}_k\left(\mathbf{e}_i^{\mathsf{T}}A\right) = \mathbf{e}_k A_{i,:}$ is the $n \times n$ matrix with (r,s) entry $\delta_{rk}a_{is}$,

whose only possibly nonzero diagonal entry is the (k,k) entry a_{ik}. So $\psi\left(\mathbf{e}_k\mathbf{e}_i^{\mathsf{T}}\right) = a_{ik}$ as well. Hence $\varphi = \psi$, and $\varphi(B) = \psi(B)$ is the claim. □

The first proof is shorter. The second shows what linearity in B buys: only mn very simple products need to be computed, and each is a picture (one nonzero column, one nonzero row) rather than a sum.

Cyclicity extends to longer products by grouping: $\operatorname{tr}(ABC) = \operatorname{tr}(A(BC)) = \operatorname{tr}((BC)A) = \operatorname{tr}(BCA)$, and likewise $\operatorname{tr}(BCA) = \operatorname{tr}(CAB)$. A product inside a trace can be rotated, but not rearranged arbitrarily (Exercise 2.4.32). This makes the trace a test for identities in which the order of factors matters.

Technique 2.4.29 (Cyclicity of trace). The trace is linear, and it does not see the order of two factors: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, so $\operatorname{tr}(A_1A_2\cdots A_k)$ is unchanged by rotating the factors cyclically. Take the trace of both sides of a claimed identity. If the traces differ, the identity is false; and quantities such as $AB - BA$, which are invisible to the trace, can be eliminated.

Reach for it when commutators $AB - BA$ appear (they always have trace 0), when a product is conjugated as in $S^{-1}AS$ (once inverses are available, $\operatorname{tr}\left(S^{-1}AS\right) = \operatorname{tr}\left(ASS^{-1}\right) = \operatorname{tr} A$, which makes the trace a similarity invariant, Section 4.5), and when an inner product of matrices is written as $\operatorname{tr}\left(B^{\mathsf{T}}A\right)$ (Section 6.1). Only cyclic rotations are allowed: $\operatorname{tr}(ABC)$ and $\operatorname{tr}(BAC)$ differ in general. And the trace detects nothing about a matrix of trace 0; it is a single number, so it can rule things out but rarely proves them. The canonical example is Proposition 2.4.30.

Used in. Proposition 2.4.30, Exercise 2.4.50, Exercise 2.4.52, Exercise 2.4.53, Exercise 2.4.54, Exercise 2.5.44, Exercise 2.8.12, Problem 2.9, Problem 2.16, Problem 2.17, Exercise 4.3.38, Exercise 4.3.43, Proposition 4.5.27, Exercise 4.5.47, Exercise 6.4.51, Exercise 6.5.38, Problem 6.11, Problem 6.18, Proposition 7.5.25, Corollary 7.5.26, Exercise 7.5.41, Exercise 7.5.44, Exercise 7.6.56, Exercise 8.1.33, Exercise 8.1.37, Exercise 8.1.38, Exercise 8.3.34, Exercise 6.3.48.

Proposition 2.4.30 ($AB - BA \neq I$). Let $n \geq 1$, and suppose that $n * 1 \neq 0$ in F; for instance, F has characteristic 0 (Definition 1.1.28). Then $AB - BA \neq I_n$ for all $A, B \in F^{n\times n}$.

Proof. ROUTINE

Uses. Proposition 2.4.27 (The trace is linear), Proposition 2.4.28 (Cyclicity of the trace), Definition 2.4.26 (Trace)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Take the trace of both sides of $AB - BA = I_n$.
Hint 2. By linearity and cyclicity, $\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. What is $\operatorname{tr} I_n$?
Sketch. $\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$, but $\operatorname{tr} I_n = n * 1 \neq 0$.

By Proposition 2.4.27 and Proposition 2.4.28, $\operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. But $\operatorname{tr} I_n = 1 + \cdots + 1 = n * 1$, which is nonzero by hypothesis. So $AB - BA$ and I_n have different traces and cannot be equal. □

In terms of maps: on the infinite-dimensional space $F[x]$, differentiation D and multiplication by x satisfy $DM - MD = \mathrm{id}$, by the product rule (Exercise 2.4.51). Proposition 2.4.30 says that this identity has no solution by $n \times n$ matrices, that is, by operators on F^n, when $n * 1 \neq 0$. In quantum mechanics the same relation holds between the position and momentum operators, and this is why they cannot act on a finite-dimensional space.

Remark 2.4.31 (The hypothesis $n * 1 \neq 0$ is needed). The matrices of Example 2.4.18 satisfy

$$AB - BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Over $\mathbb{F}_2$, where $-1 = 1$, this is I_2. There is no contradiction with Proposition 2.4.30, because $2 * 1 = 0$ in $\mathbb{F}_2$, and indeed $\operatorname{tr} I_2 = 0$ there. In general, over $\mathbb{F}_p$ there are $p \times p$ matrices with $AB - BA = I_p$ (Exercise 2.4.52).

Exercise 2.4.32. STANDARD Find the error in the following argument, and show that its conclusion is false.

Claim. For all $A, B, C \in F^{n\times n}$, $\operatorname{tr}(ABC) = \operatorname{tr}(BAC)$.

Proof. By associativity and Proposition 2.4.28,

$$\begin{aligned}\operatorname{tr}(ABC) &= \operatorname{tr}((AB)C)\\ &= \operatorname{tr}((BA)C) = \operatorname{tr}(BAC).\end{aligned}$$

Solution on page 1140

2.4.8 Block matrices

Large matrices often have structure that is visible only when we group their entries into rectangular pieces, such as a block of zeros or a copy of the identity. Multiplication respects this grouping.

Definition 2.4.33 (Block matrix). Let $A \in F^{\ell\times m}$. Split the row indices $1, \dots, \ell$ into r consecutive groups of sizes $\ell_1, \dots, \ell_r$, so $\ell_1 + \cdots + \ell_r = \ell$, and the column indices into s consecutive groups of sizes $m_1, \dots, m_s$. The **block** $A_{IK} \in F^{\ell_I \times m_K}$ is the submatrix of A formed by the entries in the I-th group of rows and the K-th group of columns. We then write A as an $r \times s$ array of blocks,

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{pmatrix},$$

and call this a **block matrix**, partitioned with row sizes $\ell_1, \dots, \ell_r$ and column sizes $m_1, \dots, m_s$. A block of zeros is written 0.

For example, with row sizes $1, 2$ and column sizes $2, 2$,

$$\left(\begin{array}{cc|cc} 1 & 2 & 0 & 3 \\ \hline 0 & 1 & 4 & 5 \\ 2 & 0 & 6 & 7 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where, for instance, $A_{11} = \begin{pmatrix} 1 & 2 \end{pmatrix}$ and $A_{22} = \begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$. A square block matrix with square diagonal blocks $A_{11}, \dots, A_{rr}$ is **block diagonal** if all other blocks are zero, and **block upper triangular** if $A_{IK} = 0$ whenever $I > K$.

Theorem 2.4.34 (Block multiplication). Let $A \in F^{\ell\times m}$ be partitioned with row sizes $\ell_1, \dots, \ell_r$ and column sizes $m_1, \dots, m_s$, and let $B \in F^{m\times n}$ be partitioned with row sizes $m_1, \dots, m_s$ (the column sizes of A) and column sizes $n_1, \dots, n_t$. Partition AB with row sizes $\ell_1, \dots, \ell_r$ and column sizes $n_1, \dots, n_t$. Then

$$(AB)_{IJ} = \sum_{K=1}^{s} A_{IK} B_{KJ}$$

for $1 \le I \le r$ and $1 \le J \le t$. That is, block matrices are multiplied by the entry formula, treating blocks as entries and keeping each product in its order.

Proof (entry by entry). STANDARD

Uses. Definition 2.4.33 (Block matrix), Proposition 2.4.8 (Four views of the product)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Fix an entry of AB in block (I, J), and write it with the entry formula as a sum over $k = 1, \dots, m$.
Hint 2. Split that sum according to the group of columns of A (equivalently, rows of B) that k lies in. Each piece is an entry of some $A_{IK}B_{KJ}$.

Sketch. For row i in group I and column j in group J, $(AB)_{ij} = \sum_k a_{ik}b_{kj}$. Grouping the indices k by the K-th group of columns of A, the K-th partial sum is exactly the corresponding entry of $A_{IK}B_{KJ}$, by the entry formula for that product. Adding over K gives the entry of $\sum_K A_{IK}B_{KJ}$.

Let i be the p-th row index in the I-th group of rows, and j the q-th column index in the J-th group of columns. Write G_K for the K-th group of indices in $\{1, \dots, m\}$, which has m_K elements. For k the u-th element of G_K, the entries a_{ik} and b_{kj} are the entries $(A_{IK})_{pu}$ and $\big(B_{KJ}\big)_{uq}$ of the blocks. Splitting the sum over k into the groups $G_1, \dots, G_s$,

$$\begin{aligned}(AB)_{ij} &= \sum_{k=1}^{m} a_{ik}b_{kj} = \sum_{K=1}^{s} \sum_{k\in G_K} a_{ik}b_{kj} \\ &= \sum_{K=1}^{s} \sum_{u=1}^{m_K} (A_{IK})_{pu}\big(B_{KJ}\big)_{uq} \\ &= \sum_{K=1}^{s} \big(A_{IK}B_{KJ}\big)_{pq}.\end{aligned}$$

The left side is the (p, q) entry of $(AB)_{IJ}$ and the right side is the (p, q) entry of $\sum_K A_{IK}B_{KJ}$. □

The second proof reduces the theorem to three special cases, each of which is one of the views of Proposition 2.4.8.

Proof (from the views). STANDARD

Uses. Proposition 2.4.8 (Four views of the product), Definition 2.4.33 (Block matrix)
Hint 1. Handle three special cases first: only the columns of B are split; only the rows of A are split; only the inner index is split, $A = (A_1 \ \cdots \ A_s)$ and B is $B_1, \dots, B_s$ stacked.
Hint 2. The first two cases are the column and row views. For the third, group the terms of the column–row view.
Hint 3. For the general case, block (I, J) of AB is (rows of group I of A) times (columns of group J of B) by the first two cases; now apply the third.
Sketch. By the column view, the columns of AB in group J are A times the columns of B in group J; by the row view, the rows of AB in group I are the rows of A in group I times B. So $(AB)_{IJ} = A_I B^J$, with A_I the I-th block row of A and B^J the J-th block column of B. By the column–row view, $A_I B^J = \sum_k (A_I)_{\cdot,k}\big(B^J\big)_{k,\cdot}$, and grouping the k by blocks gives $\sum_K A_{IK}B_{KJ}$.

Let A_I be the I-th block row of A, that is, the $\ell_I \times m$ matrix formed by the rows of A in the I-th group, and let B^J be the J-th block column of B, the $m \times n_J$ matrix formed by its columns in the J-th group.

By the row view (Proposition 2.4.8(c)), each row of AB in the I-th group is the corresponding row of A times B, so these rows form the matrix $A_I B$. By the column view (Proposition 2.4.8(b)), applied to $A_I B$, its columns in the J-th group are A_I times the corresponding columns of B, which form $A_I B^J$. Hence $(AB)_{IJ} = A_I B^J$.

By the column–row view (Proposition 2.4.8(d)), $A_I B^J = \sum_{k=1}^{m} (A_I)_{\cdot,k}\big(B^J\big)_{k,\cdot}$. Group the terms by the groups $G_1, \dots, G_s$ of the index k. The columns of A_I with $k \in G_K$ are the columns of A_{IK}, and the rows of B^J with $k \in G_K$ are the rows of B_{KJ}, in the same order. So by the column–row view again, applied to $A_{IK}B_{KJ}$,

$$\sum_{k\in G_K} (A_I)_{\cdot,k}\big(B^J\big)_{k,\cdot} = A_{IK}B_{KJ},$$

and summing over K gives $(AB)_{IJ} = \sum_K A_{IK}B_{KJ}$. □

In the second proof the three special cases are what to remember: blocks of columns of B produce blocks of columns of AB, blocks of rows of A produce blocks of rows of AB, and splitting the inner index turns the product into a sum. The partitions must match: the column sizes of A must be the row sizes of B, or the products $A_{IK}B_{KJ}$ are not defined. And the factors cannot be swapped, since blocks are matrices: $A_{IK}B_{KJ}$ is not $B_{KJ}A_{IK}$.

For 2×2 block matrices the theorem says that the (I, J) block of AB is $A_{I1}B_{1J} + A_{I2}B_{2J}$, exactly as for the entries of a product of 2×2 matrices. For block diagonal matrices it gives

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} A_1B_1 & 0 \\ 0 & A_2B_2 \end{pmatrix},$$

and for block upper triangular matrices

$$\begin{pmatrix} A_1 & X \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} B_1 & Y \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} A_1B_1 & A_1Y + XB_2 \\ 0 & A_2B_2 \end{pmatrix}.$$

So block diagonal matrices multiply block by block, and in particular $\operatorname{diag}(A_1, A_2)^k = \operatorname{diag}\left(A_1^k, A_2^k\right)$, where $\operatorname{diag}(A_1, A_2)$ denotes the block diagonal matrix.

Technique 2.4.35 (Block decomposition). Split matrices into blocks that match the structure of the problem, and compute with the blocks as if they were entries (Theorem 2.4.34). Typical splittings: off a single first row and column, to set up an induction on the size; along a basis that is split into two parts, for example a basis of a subspace extended to a basis of the whole space; or around a block of zeros or a copy of I.

Reach for it when a matrix has large zero or identity blocks, when an argument should reduce from size n to size $n-1$, or when a statement concerns a map that preserves a subspace. Two checks: the partitions must match, and products of blocks must keep their order. The canonical example is Proposition 2.4.36. Later it computes determinants of block triangular matrices (Section 5.4), compares AB with BA (Section 7.3), and organizes the generalized eigenspace decomposition and the Jordan form (Section 10.2, Section 10.3).

Used in. Proposition 2.4.36, Exercise 2.4.55, Exercise 2.5.40, Exercise 2.8.9, Problem 2.13, Problem 2.16, Theorem 3.5.27, Problem 3.4, Problem 3.14, Theorem 4.7.42, Lemma 4.8.2, Exercise 4.8.13, Problem 4.12, Theorem 5.4.26, Corollary 5.4.27, Exercise 5.4.43, Exercise 5.5.47, Exercise 5.6.14, Proposition 7.3.22, Problem 7.2, Problem 7.6, Theorem 8.3.23, Exercise 2.4.37, Exercise 2.4.52, Exercise 3.5.46, Exercise 4.5.53, Exercise 6.3.63.

Recall from Exercise 2.3.38 that $A \in F^{n\times n}$ is **upper triangular** if $a_{ij} = 0$ whenever $i > j$.

Proposition 2.4.36 (Products of upper triangular matrices). If $A, B \in F^{n\times n}$ are upper triangular, then so is AB, and its diagonal entries are $(AB)_{ii} = a_{ii}b_{ii}$.

Intuition. There are two natural strategies. The entry formula $(AB)_{ij} = \sum_k a_{ik}b_{kj}$ suggests looking at which terms can be nonzero. Alternatively, an upper triangular matrix of size n is a 1×1 corner, a row, a zero column, and an upper triangular matrix of size $n-1$. That suggests block multiplication and induction on n. The second strategy is longer here, but it is the template for every later induction on the size of a triangular matrix (Section 5.3, Section 7.5).

Proof (by blocks, induction on n). STANDARD

Uses. Theorem 2.4.34 (Block multiplication), Definition 2.4.33 (Block matrix)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Partition with row and column sizes $1, n-1$. What do the blocks of an upper triangular matrix look like?
Hint 2. Write $A = \begin{pmatrix} a_{11} & \mathbf{a}^{\mathsf{T}} \\ \mathbf{0} & A' \end{pmatrix}$ and similarly for B, with A' and B' upper triangular of size $n-1$. Multiply by blocks and apply the induction hypothesis to $A'B'$.

Sketch. Induct on n; $n = 1$ is clear. Split off the first row and column: $A = \begin{pmatrix} a_{11} & \mathbf{a}^{\mathsf{T}} \\ 0 & A' \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & \mathbf{b}^{\mathsf{T}} \\ 0 & B' \end{pmatrix}$ with A', B' upper triangular. By block multiplication, AB has corner $a_{11}b_{11}$, a zero column below it, and lower right block $A'B'$, which is upper triangular with diagonal entries $a_{ii}b_{ii}$ ($i \geq 2$) by induction.

We induct on n. For $n = 1$ every matrix is upper triangular and $(AB)_{11} = a_{11}b_{11}$.

Let $n \geq 2$ and assume the claim for size $n - 1$. Partition A and B with row sizes and column sizes $1, n-1$. Since A is upper triangular, its lower left block is the zero column $\mathbf{0} \in F^{n-1}$, and its lower right block A', which has entries a_{ij} with $i, j \geq 2$, is upper triangular. So

$$A = \begin{pmatrix} a_{11} & \mathbf{a}^{\mathsf{T}} \\ \mathbf{0} & A' \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \mathbf{b}^{\mathsf{T}} \\ \mathbf{0} & B' \end{pmatrix},$$

for some $\mathbf{a}, \mathbf{b} \in F^{n-1}$ and upper triangular $A', B' \in F^{(n-1)\times(n-1)}$. By Theorem 2.4.34, the blocks of AB are

$$a_{11}b_{11} + \mathbf{a}^{\mathsf{T}}\mathbf{0} = a_{11}b_{11}, \qquad a_{11}\mathbf{b}^{\mathsf{T}} + \mathbf{a}^{\mathsf{T}}B',$$
$$\mathbf{0}b_{11} + A'\mathbf{0} = \mathbf{0}, \qquad \mathbf{0}\mathbf{b}^{\mathsf{T}} + A'B' = A'B',$$

that is,

$$AB = \begin{pmatrix} a_{11}b_{11} & a_{11}\mathbf{b}^{\mathsf{T}} + \mathbf{a}^{\mathsf{T}}B' \\ \mathbf{0} & A'B' \end{pmatrix}.$$

By the induction hypothesis, $A'B'$ is upper triangular with diagonal entries $a_{ii}b_{ii}$ for $i = 2, \ldots, n$. So every entry of AB below the diagonal is either in the zero column or below the diagonal of $A'B'$, hence zero, and the diagonal entries of AB are $a_{11}b_{11}$ and those of $A'B'$. □

Proof (entry by entry). ROUTINE

Uses. Proposition 2.4.8 (Four views of the product)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. In $(AB)_{ij} = \sum_k a_{ik}b_{kj}$, when can the term $a_{ik}b_{kj}$ be nonzero?
Hint 2. It needs $i \leq k$ and $k \leq j$.
Sketch. $a_{ik}b_{kj} \neq 0$ forces $i \leq k \leq j$. If $i > j$ there is no such k, and if $i = j$ only $k = i$ remains.

In $(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, the factor a_{ik} is zero unless $i \leq k$, and b_{kj} is zero unless $k \leq j$. So only terms with $i \leq k \leq j$ can be nonzero. If $i > j$ there are none, and $(AB)_{ij} = 0$. If $i = j$, the only such term is $k = i$, so $(AB)_{ii} = a_{ii}b_{ii}$. □

A third proof translates to maps: by Exercise 2.3.38, $[T]_{\mathcal{B}\mathcal{B}}$ is upper triangular exactly when T maps each $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ into itself, and a composite of two maps with this property has it too. That proof is the one that generalizes to triangularization in Section 7.5.

Exercise 2.4.37. ROUTINE Let $X \in F^{p\times q}$, and let $M = \begin{pmatrix} I_p & X \\ 0 & I_q \end{pmatrix}$. Show that $M^k = \begin{pmatrix} I_p & kX \\ 0 & I_q \end{pmatrix}$ for every $k \geq 1$, where $kX = X + \cdots + X$ (k terms), and that $\begin{pmatrix} I_p & X \\ 0 & I_q \end{pmatrix}\begin{pmatrix} I_p & -X \\ 0 & I_q \end{pmatrix} = I_{p+q}$.

Solution on page 1140

2.4.9 Counting walks BEYOND 21-242

Applications to graphs are not part of 21-242.

Powers of a matrix count paths. This is the simplest instance of a theme that runs through Markov chains (Section 11.1) and spectral graph theory (Section 16.5). (Another application of the matrix product, homogeneous coordinates, which turn translations of F^n into matrix products, is in Section 4.7.)

Definition 2.4.38 (Adjacency matrix). A **directed graph** G on the vertices $1, \ldots, n$ is a set of ordered pairs (i, j) of vertices, called **edges**. A **walk of length** k from i to j is a sequence of vertices $i =$

$v_0, v_1, \dots, v_k = j$ such that (v_{t-1}, v_t) is an edge for each $t = 1, \dots, k$. The **adjacency matrix** of G is the $n \times n$ matrix A over $\mathbb{Q}$ with $a_{ij} = 1$ if (i, j) is an edge and $a_{ij} = 0$ otherwise. An undirected graph is a directed graph in which (j, i) is an edge whenever (i, j) is; its adjacency matrix is symmetric.

Proposition 2.4.39 (Powers count walks). Let A be the adjacency matrix of a directed graph G on $1, \dots, n$. For every $k \geq 0$, $\left(A^k\right)_{ij}$ is the number of walks of length k from i to j. In particular, $\operatorname{tr}\left(A^k\right)$ is the number of closed walks of length k, counted with their starting vertex.

Proof. STANDARD

Uses. Proposition 2.4.8 (Four views of the product), Definition 2.4.38 (Adjacency matrix), Notation 2.4.16 (Powers of a matrix)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Induct on k. For $k = 0$, $A^0 = I_n$, and a walk of length 0 from i to j exists only when $i = j$.
Hint 2. A walk of length $k + 1$ from i to j is a walk of length k from i to some vertex l, followed by the edge (l, j). Sort the walks by l, and compare with $\left(A^{k+1}\right)_{ij} = \sum_l \left(A^k\right)_{il} a_{lj}$.
Sketch. Induct on k. Grouping the walks of length $k + 1$ from i to j by their next-to-last vertex l, there are (walks of length k from i to l) $\times\, a_{lj}$ of them for each l, and the sum over l is $\sum_l \left(A^k\right)_{il} a_{lj} = \left(A^{k+1}\right)_{ij}$.

We induct on k. For $k = 0$, $\left(A^0\right)_{ij} = \delta_{ij}$, and there is exactly one walk of length 0 from i to i (the sequence $v_0 = i$) and none from i to $j \neq i$.

Suppose the claim holds for k. A walk $v_0, \dots, v_{k+1}$ of length $k + 1$ from i to j consists of the walk $v_0, \dots, v_k$ of length k from i to $l = v_k$, followed by the edge (l, j); conversely, any walk of length k from i to l can be extended in this way exactly when (l, j) is an edge. So the number of walks of length $k + 1$ from i to j whose next-to-last vertex is l equals the number of walks of length k from i to l times a_{lj}, which is $\left(A^k\right)_{il} a_{lj}$ by the induction hypothesis. Summing over l, the number of walks of length $k + 1$ from i to j is

$$\sum_{l=1}^{n} \left(A^k\right)_{il} a_{lj} = \left(A^{k+1}\right)_{ij}.$$

The last claim follows by adding the diagonal entries. □

Example 2.4.40 (Walks in a small graph).

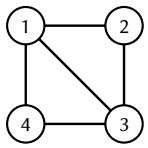


Figure 2.4.2. The undirected graph with edges $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$, $\{4, 1\}$, and $\{1, 3\}$.

The graph of Figure 2.4.2 has

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

$$A^2 = \begin{pmatrix} 3 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 1 & 3 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}.$$

For instance $(A^2)_{13} = 2$ counts the walks $1,2,3$ and $1,4,3$. In an undirected graph, $(A^2)_{ii}$ is the number of neighbours of i, and $(A^2)_{ij}$ is the number of common neighbours of i and j. The diagonal entries of A^3 are $4,2,4,2$, so $\operatorname{tr}(A^3) = 12$. A closed walk of length 3 in a graph without loops goes around a triangle, and each triangle gives 6 closed walks (3 starting vertices, 2 directions). So the graph has $12/6 = 2$ triangles, namely $\{1,2,3\}$ and $\{1,3,4\}$.

Exercises

Exercise 2.4.41. ROUTINE Let $D : F[x]_{\le 3} \to F[x]_{\le 2}$ be differentiation and $M : F[x]_{\le 2} \to F[x]_{\le 3}$ multiplication by x, with the bases $\mathcal{B}_2 = 1, x, x^2$ and $\mathcal{B}_3 = 1, x, x^2, x^3$.

(a) Find $[D]_{\mathcal{B}_2\mathcal{B}_3}$ and $[M]_{\mathcal{B}_3\mathcal{B}_2}$.

(b) Compute $[DM]_{\mathcal{B}_2\mathcal{B}_2}$ and $[MD]_{\mathcal{B}_3\mathcal{B}_3}$ as products of matrices, and check the answers against the formulas $DM(p) = (xp)'$ and $MD(p) = xp'$.

(c) Compare $\operatorname{tr}(DM)$ and $\operatorname{tr}(MD)$, meaning the traces of these two matrices.

Solution on page 1140

Exercise 2.4.42. STANDARD Let $D = \operatorname{diag}(d_1, \ldots, d_n) \in F^{n\times n}$.

(a) Show that $(DX)_{ij} = d_i x_{ij}$ and $(XD)_{ij} = x_{ij} d_j$ for $X \in F^{n\times n}$.

(b) Suppose $d_1, \ldots, d_n$ are distinct. Show that X commutes with D if and only if X is diagonal.

(c) Find all $X \in F^{3\times 3}$ that commute with $\operatorname{diag}(1,1,0)$.

Solution on page 1141

Exercise 2.4.43. STANDARD Let $n \ge 2$. Show that $A \in F^{n\times n}$ commutes with every upper triangular matrix in $F^{n\times n}$ if and only if $A = cI_n$ for some $c \in F$.

Solution on page 1142

Exercise 2.4.44. STANDARD Prove or disprove each statement about matrices in $F^{n\times n}$, $n \ge 2$.

(a) If $A^2 = I_n$, then $A = I_n$ or $A = -I_n$.

(b) If A and B are symmetric, then AB is symmetric.

(c) $(AB)^2 = A^2B^2$.

(d) If $F = \mathbb{R}$ and $A^{\mathsf{T}}A = 0$, then $A = 0$. What if $F = \mathbb{C}$ or $F = \mathbb{F}_2$?

(e) If A is upper triangular and $A^2 = 0$, then every diagonal entry of A is 0.

Solution on page 1142

Exercise 2.4.45. STANDARD

(a) Show that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$ satisfies $A^2 = 0$ if and only if $a + d = 0$ and $a^2 + bc = 0$.

(b) Let p be prime. How many $A \in \mathbb{F}_p^{2\times 2}$ satisfy $A^2 = 0$?

Solution on page 1143

Exercise 2.4.46. STANDARD

(a) For $\lambda \in F$, let $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Show that $J^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$ for every $k \geq 1$, where $k\lambda^{k-1}$ means $(k * 1)\lambda^{k-1}$.

(b) Deduce that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^p = I_2$ over $\mathbb{F}_p$.

(c) Find $A \in \mathbb{F}_2^{2\times 2}$ and $A' \in \mathbb{R}^{2\times 2}$ with $A^3 = I_2$ and $A'^3 = I_2$, neither equal to I_2.

Solution on page 1143

Exercise 2.4.47. STANDARD Let $f_0 = 0$, $f_1 = 1$, and $f_{k+1} = f_k + f_{k-1}$ for $k \geq 1$ (the Fibonacci numbers), and let $Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$.

(a) Show that $Q^k = \begin{pmatrix} f_{k+1} & f_k \\ f_k & f_{k-1} \end{pmatrix}$ for every $k \geq 1$.

(b) Deduce that $f_{m+n} = f_{m+1}f_n + f_m f_{n-1}$ for all $m, n \geq 1$, and in particular $f_{2n} = f_n(f_{n+1} + f_{n-1})$.

Solution on page 1144

Exercise 2.4.48. ROUTINE Find the error in the following computation, which claims that $(AB)^\mathsf{T} = A^\mathsf{T}B^\mathsf{T}$ for $A, B \in F^{n\times n}$:

$$\begin{aligned}\left((AB)^\mathsf{T}\right)_{ij} = (AB)_{ji} &= \sum_k a_{jk}b_{ki} \\ &= \sum_k \left(A^\mathsf{T}\right)_{kj}\left(B^\mathsf{T}\right)_{ik} = \left(A^\mathsf{T}B^\mathsf{T}\right)_{ij}.\end{aligned}$$

Solution on page 1144

Exercise 2.4.49. STANDARD Let $A \in F^{m\times n}$.

(a) Show that $\operatorname{tr}\left(A\mathbf{e}_j\mathbf{e}_i^\mathsf{T}\right) = a_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.

(b) Show that if $\operatorname{tr}(AX) = 0$ for every $X \in F^{n\times m}$, then $A = 0$.

(c) Show that every linear map $\varphi : F^{n\times m} \to F$ has the form $\varphi(X) = \operatorname{tr}(BX)$ for exactly one $B \in F^{m\times n}$.

Solution on page 1145

Exercise 2.4.50. HARD Let $\varphi : F^{n\times n} \to F$ be linear with $\varphi(AB) = \varphi(BA)$ for all $A, B \in F^{n\times n}$. Show that $\varphi = c\operatorname{tr}$ for some $c \in F$. (So the trace is, up to a scalar, the only linear functional with the property of Proposition 2.4.28.)

Solution on page 1145

Exercise 2.4.51. STANDARD Let $D, M \in \mathrm{L}(F[x])$ be differentiation and multiplication by x (Example 2.1.9).

(a) Show that $DM - MD = \mathrm{id}_{F[x]}$.

(b) Explain why (a) does not contradict Proposition 2.4.30, even over $\mathbb{Q}$.

(c) Show that if U is a subspace of $F[x]$ spanned by finitely many polynomials and $M(U) \subseteq U$, then $U = \{\mathbf{0}\}$.

Solution on page 1145

Exercise 2.4.52. HARD Let p be prime.

(a) Let $A, B \in \mathbb{F}_p^{p\times p}$ be the matrices with $A\mathbf{e}_1 = \mathbf{0}$, $A\mathbf{e}_{k+1} = k\mathbf{e}_k$ for $1 \leq k \leq p-1$, and $B\mathbf{e}_k = \mathbf{e}_{k+1}$ for $k \leq p-1$, $B\mathbf{e}_p = \mathbf{0}$. Show that $AB - BA = I_p$.

(b) Show that there are $X, Y \in \mathbb{F}_p^{n\times n}$ with $XY - YX = I_n$ if and only if p divides n.

Solution on page 1146

Exercise 2.4.53. HARD Let $A, B \in F^{n\times n}$, and suppose that $C = AB - BA$ commutes with A. Show that $\operatorname{tr}\left(C^k\right) = 0$ for every $k \geq 1$.

Solution on page 1146

Exercise 2.4.54. CHALLENGE Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric. Show that $\operatorname{tr}(ABAB) \leq \operatorname{tr}\left(A^2B^2\right)$, with equality if and only if $AB = BA$.

Solution on page 1147

Exercise 2.4.55. STANDARD Let $A \in F^{p\times p}$, $X \in F^{p\times q}$, and $D \in F^{q\times q}$, and let $N = \begin{pmatrix} A & X \\ 0 & D \end{pmatrix}$.

(a) Show that $N^k = \begin{pmatrix} A^k & X_k \\ 0 & D^k \end{pmatrix}$ for every $k \geq 1$, where $X_k = \sum_{i=0}^{k-1} A^i X D^{k-1-i}$.

(b) Deduce that if $A^k = 0$ and $D^k = 0$, then $N^{2k} = 0$.

Solution on page 1147

Exercise 2.4.56. ROUTINE BEYOND 21-242 Let G be the directed graph on $1, 2, 3$ with edges $(1, 2)$, $(2, 3)$, $(3, 1)$, and $(1, 1)$.

(a) Write down its adjacency matrix A and compute A^2 and A^3.

(b) How many walks of length 3 go from 1 to 1? List them. How many closed walks of length 3 are there in all?

Solution on page 1148

Exercise 2.4.57. HARD BEYOND 21-242 Let $n \geq 2$, let $J \in \mathbb{Q}^{n\times n}$ be the matrix all of whose entries are 1, and let $A = J - I_n$, the adjacency matrix of the complete graph on n vertices (every two distinct vertices are joined).

(a) Show that $J^2 = nJ$.

(b) Show that $A^k = a_k J + (-1)^k I_n$ for $k \geq 1$, where $a_k = \left((n-1)^k - (-1)^k\right)/n$.

(c) How many walks of length k are there between two distinct vertices, and how many closed walks of length k are there in all?

Solution on page 1148

2.5 Invertibility

To solve $T(\mathbf{v}) = \mathbf{w}$ is to undo T. In F, multiplication by $a \neq 0$ is undone by multiplication by a^{-1}, and $a^{-1}a = aa^{-1} = 1$, so it does not matter on which side a^{-1} stands. For linear maps the two sides come apart, and they do different jobs. Let $T \in \mathrm{L}(V, W)$ and $S \in \mathrm{L}(W, V)$. If $ST = \mathrm{id}_V$, then every solution of $T(\mathbf{v}) = \mathbf{w}$ satisfies

$$\mathbf{v} = S(T(\mathbf{v})) = S(\mathbf{w}),$$

so there is at most one solution. If $TS = \mathrm{id}_W$, then $T(S(\mathbf{w})) = \mathbf{w}$, so $S(\mathbf{w})$ is a solution. A map that undoes T on one side therefore answers one of the two questions about $T(\mathbf{v}) = \mathbf{w}$, uniqueness or existence, and the inverses of this section are the maps that answer both.

Throughout, U, V, and W are vector spaces over F, and as in Notation 2.1.18 we write TS for $T \circ S$.

2.5.1 Left and right inverses

Definition 2.5.1 (Left inverse). Let $T \in \mathrm{L}(V, W)$. A linear map $S \in \mathrm{L}(W, V)$ is a **left inverse** of T if $ST = \mathrm{id}_V$, that is, $S(T(\mathbf{v})) = \mathbf{v}$ for every $\mathbf{v} \in V$. If T has a left inverse, T is **left-invertible**.

Likewise, for $A \in F^{m\times n}$, a matrix $B \in F^{n\times m}$ is a **left inverse** of A if $BA = I_n$, and A is **left-invertible** if it has a left inverse.

Definition 2.5.2 (Right inverse). Let $T \in \mathrm{L}(V, W)$. A linear map $S \in \mathrm{L}(W, V)$ is a **right inverse** of T if $TS = \mathrm{id}_W$, that is, $T(S(\mathbf{w})) = \mathbf{w}$ for every $\mathbf{w} \in W$. If T has a right inverse, T is **right-invertible**.

Likewise, for $A \in F^{m\times n}$, a matrix $C \in F^{n\times m}$ is a **right inverse** of A if $AC = I_m$, and A is **right-invertible** if it has a right inverse.

"Left" and "right" refer to the side on which S stands in the product that equals the identity. The two notions are mirror images: S is a left inverse of T exactly when T is a right inverse of S. Note the sizes. A one-sided inverse of an $m \times n$ matrix is $n \times m$, so that BA and AC are both defined and square, but they are identity matrices of different sizes, I_n and I_m. Also, we insist that one-sided inverses be linear; Remark 2.5.8 explains what changes without this.

The definitions for matrices are the definitions for maps, applied to T_A. Indeed $T_B T_A = T_{BA}$ by Theorem 2.4.2, and $T_{I_n} = \mathrm{id}_{F^n}$ by Proposition 2.3.13. A matrix is determined by its map (Proposition 2.3.17), so $BA = I_n$ if and only if $T_B T_A = \mathrm{id}_{F^n}$. Thus B is a left inverse of A exactly when T_B is a left inverse of T_A. Since every linear map $F^m \to F^n$ is T_B for some B, it follows that A is left-invertible exactly when T_A is. The same holds on the right, and Proposition 2.5.20 extends all of this to matrices of maps in arbitrary bases.

Example 2.5.3 (The x-axis in the plane). Let $T : F \to F^2$, $T(x) = (x, 0)$, whose standard matrix is the column $\begin{pmatrix}1\\0\end{pmatrix}$. For $c \in F$, let

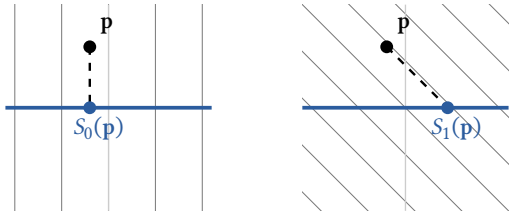
$$S_c : F^2 \to F, \quad S_c(x, y) = x + cy.$$

Each S_c is linear, and $S_c(T(x)) = S_c(x, 0) = x$, so every S_c is a left inverse of T. There are no others. By Corollary 2.2.17, a linear map $S : F^2 \to F$ is $S(x, y) = ax + by$ with $a = S(\mathbf{e}_1)$ and $b = S(\mathbf{e}_2)$, and $S(T(1)) = S(\mathbf{e}_1) = a$, so $ST = \mathrm{id}_F$ forces $a = 1$. In matrix terms, the left inverses of $\begin{pmatrix}1\\0\end{pmatrix}$ are the rows $\begin{pmatrix}1 & c\end{pmatrix}$ with $c \in F$.

On the other hand, T has no right inverse. If $TS = \mathrm{id}_{F^2}$ for some $S \in \mathrm{L}\big(F^2, F\big)$, then

$$(0, 1) = T(S(0, 1)) = (S(0, 1), 0),$$

and comparing second coordinates gives $1 = 0$, which is false. The obstruction is that $(0, 1)$ is not a value of T at all.



$S_0(x, y) = x$ $\qquad$ $S_1(x, y) = x + y$

Figure 2.5.1. Two left inverses of $T(x) = (x, 0)$ over $\mathbb{R}$. The values of T form the x-axis (blue). S_c is constant along each line $x + cy = t$ (grey) and sends the whole line to its intersection $(t, 0)$ with the axis, identified with t. Both maps fix the axis pointwise, as a left inverse must, but they treat the point $\mathbf{p}$ differently.

Intuition (Choices off the image). A left inverse S of T must send $T(\mathbf{v})$ back to $\mathbf{v}$, so its values on the set $T(V)$ of values of T are forced. Off that set, S is constrained only by linearity, and in Example 2.5.3 this leaves a free choice: S_c collapses the plane onto the x-axis along the direction $(-c, 1)$, and every direction other than that of the axis itself works (Figure 2.5.1). A right inverse has the opposite problem. It must produce, for every $\mathbf{w} \in W$, a vector $S(\mathbf{w})$ with $T(S(\mathbf{w})) = \mathbf{w}$, and this is impossible as soon as some $\mathbf{w}$ is not a value of T. So left inverses are about T losing no information, and right inverses about T reaching everything.

To make this precise, recall that a function $f : X \to Y$ is *injective* if $f(x) = f(x')$ implies $x = x'$, and *surjective* if every $y \in Y$ equals $f(x)$ for some $x \in X$ (Section A.1). These properties of linear maps are the subject of Section 2.6.

Proposition 2.5.4 (Left inverses give uniqueness). Suppose $T \in \mathrm{L}(V, W)$ has a left inverse S. If $T(\mathbf{v}) = \mathbf{w}$, then $\mathbf{v} = S(\mathbf{w})$. In particular, T is injective.

Likewise, if $A \in F^{m\times n}$ has a left inverse B and $A\mathbf{x} = \mathbf{y}$, then $\mathbf{x} = B\mathbf{y}$, and T_A is injective.

Proof. ROUTINE

Uses. Definition 2.5.1 (Left inverse)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Apply S to both sides of $T(\mathbf{v}) = \mathbf{w}$.
Hint 2. For injectivity, suppose $T(\mathbf{v}) = T(\mathbf{v}')$. Both $\mathbf{v}$ and $\mathbf{v}'$ solve $T(\mathbf{x}) = \mathbf{w}$ for one and the same $\mathbf{w}$.
Sketch. Applying S to $T(\mathbf{v}) = \mathbf{w}$ gives $\mathbf{v} = S(T(\mathbf{v})) = S(\mathbf{w})$. If $T(\mathbf{v}) = T(\mathbf{v}') = \mathbf{w}$, then $\mathbf{v}$ and $\mathbf{v}'$ both equal $S(\mathbf{w})$. The matrix statement is the same argument with $BA = I_n$.

Suppose $T(\mathbf{v}) = \mathbf{w}$. Applying S to both sides and using $ST = \mathrm{id}_V$,

$$\mathbf{v} = \mathrm{id}_V(\mathbf{v}) = S(T(\mathbf{v})) = S(\mathbf{w}).$$

Now let $\mathbf{v}, \mathbf{v}' \in V$ with $T(\mathbf{v}) = T(\mathbf{v}')$, and call this vector $\mathbf{w}$. By what we just showed, $\mathbf{v} = S(\mathbf{w}) = \mathbf{v}'$. So T is injective.

For the matrix statement, if $A\mathbf{x} = \mathbf{y}$, then $\mathbf{x} = I_n\mathbf{x} = (BA)\mathbf{x} = B(A\mathbf{x}) = B\mathbf{y}$ by Theorem 2.4.2, and injectivity of T_A follows as before. □

Proposition 2.5.5 (Right inverses give existence). Suppose $T \in \mathrm{L}(V, W)$ has a right inverse S. Then for every $\mathbf{w} \in W$, the vector $\mathbf{v} = S(\mathbf{w})$ satisfies $T(\mathbf{v}) = \mathbf{w}$. In particular, T is surjective.

Likewise, if $A \in F^{m\times n}$ has a right inverse C, then $\mathbf{x} = C\mathbf{y}$ satisfies $A\mathbf{x} = \mathbf{y}$ for every $\mathbf{y} \in F^m$, and T_A is surjective.

Proof. ROUTINE

Uses. Definition 2.5.2 (Right inverse)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Fix $\mathbf{w} \in W$. Which vector could possibly map to $\mathbf{w}$, if all you know about T is $TS = \mathrm{id}_W$?
Hint 2. Compute $T(S(\mathbf{w}))$.
Sketch. $T(S(\mathbf{w})) = (TS)(\mathbf{w}) = \mathrm{id}_W(\mathbf{w}) = \mathbf{w}$, so every $\mathbf{w}$ is a value of T. For matrices, $A(C\mathbf{y}) = (AC)\mathbf{y} = I_m\mathbf{y} = \mathbf{y}$.

Let $\mathbf{w} \in W$ and $\mathbf{v} = S(\mathbf{w})$. Then

$$T(\mathbf{v}) = T(S(\mathbf{w})) = (TS)(\mathbf{w}) = \mathrm{id}_W(\mathbf{w}) = \mathbf{w}.$$

As $\mathbf{w}$ was arbitrary, every vector of W is a value of T, so T is surjective. For matrices, $A(C\mathbf{y}) = (AC)\mathbf{y} = I_m\mathbf{y} = \mathbf{y}$ by Theorem 2.4.2 and Proposition 2.3.13. □

Neither proposition says more than it claims. A left inverse produces the only possible solution, but that candidate need not be a solution. With $A = \begin{pmatrix}1\\0\end{pmatrix}$, its left inverse $B = \begin{pmatrix}1 & 0\end{pmatrix}$ from Example 2.5.3, and $\mathbf{y} = (0, 1)$, the candidate is $B\mathbf{y} = 0$, but $Ax = (x, 0) \neq \mathbf{y}$ for every x: the system has no solution at all. A right inverse

produces a solution, but not necessarily the only one. The row $A = (1 \;\; 0)$ has the right inverses $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}1\\1\end{pmatrix}$, and applied to $y = 1$ they give the two different solutions $(1,0)$ and $(1,1)$ of $A\mathbf{x} = 1$. So "at most one solution" and "at least one solution" really are separate properties, detected by inverses on separate sides.

The uniqueness half has a useful reformulation for products.

Corollary 2.5.6 (One-sided cancellation). Let $T \in \mathrm{L}(V,W)$.
(a) If T is left-invertible, then $TR_1 = TR_2$ implies $R_1 = R_2$, for all $R_1, R_2 \in \mathrm{L}(U,V)$.
(b) If T is right-invertible, then $R_1T = R_2T$ implies $R_1 = R_2$, for all $R_1, R_2 \in \mathrm{L}(W,U)$.

Likewise, if $A \in F^{m\times n}$ is left-invertible, then $AX = AY$ implies $X = Y$ for $X, Y \in F^{n\times p}$, and if A is right-invertible, then $XA = YA$ implies $X = Y$ for $X, Y \in F^{p\times m}$.

Proof. ROUTINE

Uses. Definition 2.5.1 (Left inverse), Definition 2.5.2 (Right inverse), Proposition 2.1.24 (Algebra of composition), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Multiply both sides of $TR_1 = TR_2$ by a left inverse S of T, on the left.
Hint 2. Regroup with associativity: $S(TR_i) = (ST)R_i = R_i$.
Sketch. With $ST = \mathrm{id}_V$: $R_1 = (ST)R_1 = S(TR_1) = S(TR_2) = (ST)R_2 = R_2$. With $TS = \mathrm{id}_W$: $R_1 = R_1(TS) = (R_1T)S = (R_2T)S = R_2$. Matrices are the same with Proposition 2.4.14.

(a) Let S be a left inverse of T. By Proposition 2.1.24,

$$\begin{aligned} R_1 &= \mathrm{id}_V R_1 = (ST)R_1 = S(TR_1) \\ &= S(TR_2) = (ST)R_2 = \mathrm{id}_V R_2 = R_2. \end{aligned}$$

(b) Let S be a right inverse of T. Then

$$\begin{aligned} R_1 &= R_1\mathrm{id}_W = R_1(TS) = (R_1T)S \\ &= (R_2T)S = R_2(TS) = R_2. \end{aligned}$$

The matrix statements are the same computations, with $BA = I_n$ or $AC = I_m$ and associativity and identities from Proposition 2.4.14. □

This answers the question left open in Example 2.4.19: a left-invertible matrix can be cancelled on the left, and in particular $AX = 0$ forces $X = 0$. The side matters. $A = \begin{pmatrix}1\\0\end{pmatrix}$ is left-invertible, yet $(0 \;\; 1)A = 0 = (0 \;\; 0)A$; cancelling A on the right would need a right inverse.

Remark 2.5.7 (The converses). Every injective linear map between finite-dimensional spaces has a left inverse, and every surjective one has a right inverse. For T_A these statements are among the Fundamental Theorems (Theorem 2.7.2, Theorem 2.7.4), which are stated in Section 2.7, where one-sided inverses are also built directly from a basis, and completed in Section 3.5, where they are constructed by row reduction. The converses are harder than the two propositions because a left inverse must be defined, linearly, on all of W, while injectivity of T only says what it has to do on $T(V)$.

Remark 2.5.8 (Why one-sided inverses must be linear). Every injective function $f : X \to Y$ with X nonempty has a left inverse function, and every surjective function has a right inverse function (Section A.1). For a linear map we ask for more, a *linear* one-sided inverse. Without linearity there are many more left inverses: for $T(x) = (x, 0)$ from $\mathbb{R}$ to $\mathbb{R}^2$, the function $g(x,y) = x + y^2$ satisfies $g(T(x)) = x$ for all x, but g is not linear, since $g(0,2) = 4 \neq 2g(0,1)$. For two-sided inverses the distinction disappears: the inverse function of a bijective linear map is automatically linear (Proposition 2.7.7).

Exercise 2.5.9. ROUTINE Let $A \in F^{m\times n}$.

(a) Show that if some column of A is zero, then A has no left inverse.

(b) Show that if some row of A is zero, then A has no right inverse.

Solution on page 1149

2.5.2 Invertible maps and matrices

In Example 2.5.3 the map T had many left inverses and no right inverse. That is no accident. A left inverse S_l and a right inverse S_r of T are both maps $W \to V$, and to compare them we want a single expression that involves both. The composite $S_l T S_r$ is one: grouping it as $S_l(TS_r)$ uses that S_r is a right inverse, and grouping it as $(S_lT)S_r$ uses that S_l is a left inverse.

Theorem 2.5.10 (Left and right inverses agree). Let $T \in \mathrm{L}(V,W)$. If $S_l \in \mathrm{L}(W,V)$ is a left inverse of T and $S_r \in \mathrm{L}(W,V)$ is a right inverse of T, then $S_l = S_r$.

Consequently, if T has both a left inverse and a right inverse, then it has exactly one left inverse and exactly one right inverse, these are equal, and this map is both a left and a right inverse of T.

The same holds for matrices: if $A \in F^{m\times n}$ and $B, C \in F^{n\times m}$ satisfy $BA = I_n$ and $AC = I_m$, then $B = C$.

Proof. ROUTINE

Uses. Definition 2.5.1 (Left inverse), Definition 2.5.2 (Right inverse), Proposition 2.1.24 (Algebra of composition), Proposition 2.4.14 (Algebra of matrices)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Find one composite that involves S_l, T, and S_r, and evaluate it in two ways.
Hint 2. Group S_lTS_r as $S_l(TS_r)$, and then as $(S_lT)S_r$.
Hint 3. For the second paragraph: if R is another left inverse, apply the first paragraph to R and S_r.
Sketch. By associativity, $S_l = S_l\mathrm{id}_W = S_l(TS_r) = (S_lT)S_r = \mathrm{id}_V S_r = S_r$. Every left inverse therefore equals the fixed right inverse S_r, so there is only one left inverse; symmetrically there is only one right inverse, and the two coincide. Matrices: $B = B(AC) = (BA)C = C$.

By Proposition 2.1.24,

$$\begin{aligned} S_l &= S_l\mathrm{id}_W && \text{(identity)} \\ &= S_l(TS_r) && \text{(as } S_r \text{ is a right inverse)} \\ &= (S_lT)S_r && \text{(associativity)} \\ &= \mathrm{id}_V S_r && \text{(as } S_l \text{ is a left inverse)} \\ &= S_r && \text{(identity).} \end{aligned}$$

Now suppose T has a left inverse S_l and a right inverse S_r. If R is any left inverse, the first part applied to R and S_r gives $R = S_r$. So every left inverse equals S_r, and T has exactly one left inverse, namely S_r. In the same way every right inverse equals S_l, so S_l is the only right inverse. Finally, the common map $S_l = S_r$ is a left inverse (as S_l) and a right inverse (as S_r).

For matrices, by Proposition 2.4.14,

$$B = BI_m = B(AC) = (BA)C = I_nC = C,$$

and the rest follows as for maps. □

This is Technique 1.1.8 once more. The expression S_lTS_r plays the role that $0 + 0'$ played for the zero of a field (Theorem 1.1.9): it is evaluated twice, once using the property of each candidate. Read contrapositively, the theorem explains Example 2.5.3.

Corollary 2.5.11 (Many one-sided inverses). If $T \in \mathrm{L}(V, W)$ has two different left inverses, then T has no right inverse. If T has two different right inverses, then T has no left inverse. The same holds for matrices.

Proof. ROUTINE

Uses. Theorem 2.5.10 (Left and right inverses agree)
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. Prove the contrapositive.
Hint 2. If T has a right inverse and a left inverse, how many left inverses can it have?
Sketch. If T had a right inverse S_r, then by Theorem 2.5.10 every left inverse would equal S_r, so any two left inverses would be equal. The second statement is symmetric.

Suppose S and S' are left inverses of T and T has a right inverse S_r. By Theorem 2.5.10, $S = S_r$ and $S' = S_r$, so $S = S'$. Hence if T has two different left inverses, it has no right inverse. The statement about right inverses is proved the same way, and the matrix statements follow from the matrix part of Theorem 2.5.10. □

In Example 2.5.3, $S_0(x, y) = x$ and $S_1(x, y) = x + y$ are different left inverses of T, since $S_0(0, 1) = 0 \neq 1 = S_1(0, 1)$. So Corollary 2.5.11 gives a second proof that T has no right inverse, one that never needed to find a vector outside $T(F)$.

Definition 2.5.12 (Invertible). A linear map $T \in \mathrm{L}(V, W)$ is **invertible** if there is a linear map $S \in \mathrm{L}(W, V)$ that is both a left and a right inverse of T:

$$ST = \mathrm{id}_V \quad \text{and} \quad TS = \mathrm{id}_W.$$

Such an S is a **two-sided inverse** of T. By Theorem 2.5.10, T is invertible if and only if it is both left-invertible and right-invertible, and in that case it has exactly one two-sided inverse, called **the inverse** of T and written T^{-1}.

Likewise, $A \in F^{m\times n}$ is **invertible** if there is $B \in F^{n\times m}$ with $BA = I_n$ and $AB = I_m$. Again B is unique; it is **the inverse** of A, written A^{-1}, and A is invertible if and only if it is both left-invertible and right-invertible.

So an invertible map satisfies $T^{-1}T = \mathrm{id}_V$ and $TT^{-1} = \mathrm{id}_W$. By Proposition 2.5.4 and Proposition 2.5.5 it is injective and surjective, so for every $\mathbf{w} \in W$ the equation $T(\mathbf{v}) = \mathbf{w}$ has exactly one solution, $\mathbf{v} = T^{-1}(\mathbf{w})$. In particular T^{-1} is the inverse function of the bijection T. Conversely, every bijective linear map is invertible, because its inverse function turns out to be linear (Proposition 2.7.7). By the discussion after Definition 2.5.2, a matrix A is invertible exactly when T_A is, and then $(T_A)^{-1} = T_{A^{-1}}$; this is also part (c) of Proposition 2.5.20 below.

Remark 2.5.13 (Invertible matrices are square). The definition does not ask A to be square, and nothing so far rules out an invertible $m \times n$ matrix with $m \neq n$: Theorem 2.5.10 shows that a left inverse and a right inverse coincide, not that $m = n$. In fact every invertible matrix is square (Corollary 3.5.15). The proof uses rank (Section 3.5), although over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ the trace gives a short argument (Exercise 2.5.44). Nothing below depends on it.

Example 2.5.14 (Invertible maps and matrices).

(a) id_V is its own inverse, and $I_n^{-1} = I_n$. For $c \in F$ with $c \neq 0$, $c\mathrm{id}_V$ is invertible with inverse $c^{-1}\mathrm{id}_V$, since $\big(c^{-1}\mathrm{id}_V\big)(c\mathrm{id}_V) = \big(c^{-1}c\big)\mathrm{id}_V = \mathrm{id}_V$ in both orders (Proposition 2.1.24(d)).

(b) Rotating by θ and then by $-\theta$ in either order is rotating by 0, so $R_\theta R_{-\theta} = R_{-\theta}R_\theta = \mathrm{id}_{\mathbb{R}^2}$ (Exercise 2.1.43), and $R_\theta^{-1} = R_{-\theta}$. In matrices,

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

A reflection satisfies $S_\ell^2 = \mathrm{id}_{\mathbb{R}^2}$ (Exercise 2.1.42), so it is its own inverse: reflecting twice returns every point.

(c) For a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of V, the coordinate map $V \to F^n$, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$, is invertible, and its inverse is $(a_1, \ldots, a_n) \mapsto a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n$ (Example 2.1.8).

(d) If $d_1, \ldots, d_n \in F$ are all nonzero, then $\mathrm{diag}(d_1, \ldots, d_n)$ is invertible with inverse $\mathrm{diag}\big(d_1^{-1}, \ldots, d_n^{-1}\big)$, because diagonal matrices multiply entry by entry (Example 2.4.10). If some $d_i = 0$, then column i is zero, and Exercise 2.5.9 shows that the matrix has no left inverse.

(e) Suppose $1 + 1 \neq 0$ in F, and write $2 = 1 + 1$. Let $T : F[x] \to F[x]$ be $\sum_k a_k x^k \mapsto \sum_k 2^k a_k x^k$, the substitution $p(x) \mapsto p(2x)$. The map $S : \sum_k a_k x^k \mapsto \sum_k 2^{-k} a_k x^k$ is linear for the same reason as T, and both ST and TS multiply the coefficient of each x^k by $2^k 2^{-k} = 1$. So $T^{-1} = S$, that is, $T^{-1}(p) = p(x/2)$.

(f) The map $T(x, y) = (y, x - y)$ of Exercise 2.1.25 satisfies $T(T + \mathrm{id}) = (T + \mathrm{id})T = \mathrm{id}$, so $T^{-1} = T + \mathrm{id}$, that is, $T^{-1}(x, y) = (x + y, x)$.

Non-example 2.5.15 (Maps that are not invertible). The projection Pr_ℓ onto a line ℓ through $\mathbf{0}$ in $\mathbb{R}^2$ (Example 2.1.11) is neither left- nor right-invertible: it sends every vector perpendicular to ℓ to $\mathbf{0}$, so it is not injective, and its values all lie on ℓ, so it is not surjective. The zero map $V \to W$ is not left-invertible unless $V = \{\mathbf{0}\}$, and not right-invertible unless $W = \{\mathbf{0}\}$. Differentiation D on $F[x]$ is not left-invertible, since $D(1) = 0 = D(0)$; whether it is right-invertible is taken up in Example 2.5.24.

2.5.3 The algebra of inverses

To show that a map is invertible, we never have to solve for its inverse. It is enough to name a candidate and check it.

> **Technique 2.5.16** (Guess the inverse, then check both products). To show that T is invertible and to identify T^{-1}, produce a candidate S and verify $ST = \mathrm{id}$ and $TS = \mathrm{id}$. Then T is invertible and $S = T^{-1}$, by the uniqueness in Theorem 2.5.10. How the candidate was found plays no role in the proof.
>
> Candidates come from undoing: to undo a rotation by θ, rotate by $-\theta$, and to undo a composite, undo its steps in reverse order. A polynomial identity also produces one: if $T^2 + T = \mathrm{id}$, then $T(T + \mathrm{id}) = \mathrm{id} = (T + \mathrm{id})T$, and if $N^k = 0$, then $(\mathrm{id} - N)\big(\mathrm{id} + N + \cdots + N^{k-1}\big) = \mathrm{id} - N^k = \mathrm{id}$, which is the geometric series $1/(1 - x) = 1 + x + x^2 + \cdots$ cut off (Exercise 2.5.39). Heuristics of this kind may be as informal as you like, because the verification is what proves the result.
>
> Check both products, unless you already know that T has an inverse on the other side. One product shows only that S is a one-sided inverse, and one-sided inverses need not be two-sided (Example 2.5.24). The canonical example is Proposition 2.5.17.
>
> *Used in.* Proposition 2.5.17, Proposition 2.5.26, Proposition 2.5.28, Exercise 2.5.38, Exercise 2.5.39, Exercise 2.5.40, Exercise 2.5.47, Exercise 2.7.31, Exercise 2.8.9, Exercise 2.8.10, Exercise 2.8.14, Problem 2.5, Problem 2.10, Exercise 4.3.37, Proposition 6.6.18, Exercise 2.5.19, Exercise 2.5.22, Exercise 2.5.35, Exercise 8.1.35.

Intuition (Socks and shoes). To undo "first put on socks, then shoes", take off the shoes first, then the socks. Likewise, to undo TS, which is "first S, then T", first undo T and then undo S. So the candidate for $(TS)^{-1}$ is $S^{-1}T^{-1}$, with the order reversed, as it was for transposes in Proposition 2.4.23.

Proposition 2.5.17 (Algebra of inverses).

(a) If $T \in \mathrm{L}(V,W)$ is invertible, then so is T^{-1}, and $\left(T^{-1}\right)^{-1} = T$.

(b) If $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$ are invertible, then so is $TS \in \mathrm{L}(U,W)$, and

$$(TS)^{-1} = S^{-1}T^{-1}.$$

(c) If $T \in \mathrm{L}(V,W)$ is invertible and $a \in F$ is nonzero, then aT is invertible, and $(aT)^{-1} = a^{-1}T^{-1}$.

(d) Statements (a)–(c) hold for matrices of compatible sizes. Moreover, if $A \in F^{m\times n}$ is invertible, then so is A^{T}, and $\left(A^{\mathsf{T}}\right)^{-1} = \left(A^{-1}\right)^{\mathsf{T}}$.

Proof. ROUTINE

Uses. Definition 2.5.12 (Invertible), Theorem 2.5.10 (Left and right inverses agree), Proposition 2.1.24 (Algebra of composition), Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.23 (Transpose of a product), Proposition 2.4.22 (Transpose is a linear involution)

Techniques. Guess the inverse, then check both products (Technique 2.5.16)

Hint 1. In each part the formula names the candidate. Multiply it by the given map on both sides and simplify.

Hint 2. For (b), regroup $(TS)\left(S^{-1}T^{-1}\right)$ as $T\left(SS^{-1}\right)T^{-1}$.

Hint 3. For (d), transpose the equations $A^{-1}A = I_n$ and $AA^{-1} = I_m$ with Proposition 2.4.23.

Sketch. (a) The equations $T^{-1}T = \mathrm{id}_V$ and $TT^{-1} = \mathrm{id}_W$ say that T is a two-sided inverse of T^{-1}. (b) $(TS)\left(S^{-1}T^{-1}\right) = T\left(SS^{-1}\right)T^{-1} = TT^{-1} = \mathrm{id}_W$, and similarly in the other order. (c) $\left(a^{-1}T^{-1}\right)(aT) = \left(a^{-1}a\right)T^{-1}T = \mathrm{id}_V$, and similarly. (d) $A^{\mathsf{T}}\left(A^{-1}\right)^{\mathsf{T}} = \left(A^{-1}A\right)^{\mathsf{T}} = I_n$ and $\left(A^{-1}\right)^{\mathsf{T}}A^{\mathsf{T}} = \left(AA^{-1}\right)^{\mathsf{T}} = I_m$.

In each part we exhibit a two-sided inverse; by Definition 2.5.12 the map or matrix is then invertible, and by Theorem 2.5.10 the exhibited inverse is *the* inverse.

(a) The equations $T^{-1}T = \mathrm{id}_V$ and $TT^{-1} = \mathrm{id}_W$, read with T^{-1} as the map being inverted, say that T is a left and a right inverse of T^{-1}. So T^{-1} is invertible with inverse T.

(b) By associativity and the identity laws (Proposition 2.1.24),

$$\begin{aligned}(TS)\left(S^{-1}T^{-1}\right) &= T\left(SS^{-1}\right)T^{-1} = T\mathrm{id}_V T^{-1}\\ &= TT^{-1} = \mathrm{id}_W,\\ \left(S^{-1}T^{-1}\right)(TS) &= S^{-1}\left(T^{-1}T\right)S = S^{-1}\mathrm{id}_V S\\ &= S^{-1}S = \mathrm{id}_U.\end{aligned}$$

(c) By Proposition 2.1.24(d), $\left(a^{-1}T^{-1}\right)(aT) = \left(a^{-1}a\right)\left(T^{-1}T\right) = \mathrm{id}_V$, and likewise $(aT)\left(a^{-1}T^{-1}\right) = \mathrm{id}_W$.

(d) For matrices, the computations in (a)–(c) work verbatim, with Proposition 2.4.14 in place of Proposition 2.1.24 and identity matrices in place of identity maps. For the transpose, $A^{-1} \in F^{n\times m}$, with $A^{-1}A = I_n$ and $AA^{-1} = I_m$. By Proposition 2.4.23 and Proposition 2.4.22,

$$\begin{aligned}A^{\mathsf{T}}\left(A^{-1}\right)^{\mathsf{T}} &= \left(A^{-1}A\right)^{\mathsf{T}} = I_n^{\mathsf{T}} = I_n,\\ \left(A^{-1}\right)^{\mathsf{T}}A^{\mathsf{T}} &= \left(AA^{-1}\right)^{\mathsf{T}} = I_m^{\mathsf{T}} = I_m,\end{aligned}$$

so $\left(A^{-1}\right)^{\mathsf{T}}$ is a two-sided inverse of A^{T}. □

By induction on k, part (b) gives $(T_1T_2\cdots T_k)^{-1} = T_k^{-1}\cdots T_2^{-1}T_1^{-1}$ for invertible $T_1,\ldots,T_k$ that can be composed, and in particular $\left(T^k\right)^{-1} = \left(T^{-1}\right)^k$ for invertible $T \in \mathrm{L}(V)$. We write T^{-k} for this map, and A^{-k} for $\left(A^{-1}\right)^k$.

Part (b) also runs backwards: if two of S, T, and TS are invertible, the third can be written as a product of invertible maps.

Corollary 2.5.18 (Two out of three). Let $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$. If two of S, T, and TS are invertible, then so is the third. The same holds for matrices.

Proof. ROUTINE

Uses. Proposition 2.5.17 (Algebra of inverses), Proposition 2.1.24 (Algebra of composition)
Hint 1. If S and T are invertible, this is Proposition 2.5.17(b). In the other two cases, express the unknown factor as a product of two invertible maps.
Hint 2. $S = T^{-1}(TS)$ and $T = (TS)S^{-1}$.
Sketch. The case where S and T are invertible is Proposition 2.5.17(b). If T and TS are invertible, then $S = T^{-1}(TS)$ is a product of invertible maps; if S and TS are, then $T = (TS)S^{-1}$ is. Apply Proposition 2.5.17(a) and (b).

If S and T are invertible, then TS is invertible by Proposition 2.5.17(b).

If T and TS are invertible, then by associativity $T^{-1}(TS) = \big(T^{-1}T\big)S = S$. The map T^{-1} is invertible by Proposition 2.5.17(a), so S is a composite of two invertible maps, which is invertible by Proposition 2.5.17(b). Explicitly, $S^{-1} = (TS)^{-1}T$.

If S and TS are invertible, then $(TS)S^{-1} = T\big(SS^{-1}\big) = T$ is invertible in the same way, with $T^{-1} = S(TS)^{-1}$. The matrix case is identical. □

If only TS is known to be invertible, neither factor need be (Example 2.5.27).

Exercise 2.5.19. ROUTINE Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ in $\mathbb{Q}^{2\times 2}$.

(a) Check that $A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$.

(b) Compute $(AB)^{-1}$, $A^{-1}B^{-1}$, and $B^{-1}A^{-1}$. Which two agree?

Solution on page 1149

Inverses also translate between maps and matrices, in every pair of bases. The proof is a direct application of Technique 2.3.27.

Proposition 2.5.20 (Inverses and matrices). Let $\mathcal{B}$ be a basis of V of length n and $\mathcal{C}$ a basis of W of length m, and let $T \in \mathrm{L}(V,W)$ and $S \in \mathrm{L}(W,V)$.

(a) S is a left inverse of T if and only if $[S]_{\mathcal{B}\mathcal{C}}$ is a left inverse of $[T]_{\mathcal{C}\mathcal{B}}$, and S is a right inverse of T if and only if $[S]_{\mathcal{B}\mathcal{C}}$ is a right inverse of $[T]_{\mathcal{C}\mathcal{B}}$.

(b) T is left-invertible, right-invertible, or invertible if and only if $[T]_{\mathcal{C}\mathcal{B}}$ is. If T is invertible, then

$$\big[T^{-1}\big]_{\mathcal{B}\mathcal{C}} = \big([T]_{\mathcal{C}\mathcal{B}}\big)^{-1}.$$

(c) In particular, $A \in F^{m\times n}$ is left-invertible, right-invertible, or invertible if and only if T_A is, and if A is invertible, then $(T_A)^{-1} = T_{A^{-1}}$.

Proof. STANDARD

Uses. Theorem 2.4.2 (The matrix of a composite), Example 2.3.6 (The zero map and the identity map), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.3.16 (The matrix of T_A), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Definition 2.5.12 (Invertible)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. By Theorem 2.4.2, $[S]_{\mathcal{B}\mathcal{C}}[T]_{\mathcal{C}\mathcal{B}} = [ST]_{\mathcal{B}\mathcal{B}}$. What is the matrix of id_V in the bases $\mathcal{B}, \mathcal{B}$?
Hint 2. Two maps $V \to V$ are equal exactly when their matrices in $\mathcal{B}, \mathcal{B}$ are equal (Theorem 2.3.20(a)).
Hint 3. For the "if" in (b), you are given a matrix. Use Theorem 2.3.20(b) to find a map with that matrix, then apply (a).
Sketch. (a) $[S]_{\mathcal{B}\mathcal{C}}[T]_{\mathcal{C}\mathcal{B}} = [ST]_{\mathcal{B}\mathcal{B}}$ and $I_n = [\mathrm{id}_V]_{\mathcal{B}\mathcal{B}}$, and maps $V \to V$ are determined by their matrices, so $ST = \mathrm{id}_V$ iff the matrix product is I_n; the right side is the same with TS and I_m. (b) "Only if" is (a). For "if", realize a one-sided inverse X of $[T]$ as $[S]$ by Theorem 2.3.20(b) and apply (a). The inverse formula is (a) plus uniqueness. (c) is (b) for the standard bases, where $[T_A]_{\mathcal{E}\mathcal{E}} = A$.

(a) By Theorem 2.4.2, $[S]_{\mathcal{BC}}[T]_{\mathcal{CB}} = [ST]_{\mathcal{BB}}$, and by Example 2.3.6, $[\mathrm{id}_V]_{\mathcal{BB}} = I_n$. Both ST and id_V lie in $\mathrm{L}(V)$, and by Theorem 2.3.20(a) (with the basis $\mathcal{B}$ on both sides) they are equal if and only if their matrices are. So

$$\begin{aligned} ST = \mathrm{id}_V &\iff [ST]_{\mathcal{BB}} = [\mathrm{id}_V]_{\mathcal{BB}} \\ &\iff [S]_{\mathcal{BC}}[T]_{\mathcal{CB}} = I_n. \end{aligned}$$

In the same way, $TS = \mathrm{id}_W$ if and only if $[T]_{\mathcal{CB}}[S]_{\mathcal{BC}} = [TS]_{\mathcal{CC}} = [\mathrm{id}_W]_{\mathcal{CC}} = I_m$.

(b) If T has a left inverse S, then $[S]_{\mathcal{BC}}$ is a left inverse of $[T]_{\mathcal{CB}}$ by (a). Conversely, suppose $X \in F^{n\times m}$ is a left inverse of $[T]_{\mathcal{CB}}$. By Theorem 2.3.20(b) there is $S \in \mathrm{L}(W, V)$ with $[S]_{\mathcal{BC}} = X$, and by (a) this S is a left inverse of T. The same argument works for right inverses, and invertibility means having both (Definition 2.5.12).

If T is invertible, then by (a) the matrix $\left[T^{-1}\right]_{\mathcal{BC}}$ is both a left and a right inverse of $[T]_{\mathcal{CB}}$. By the uniqueness in Definition 2.5.12, it is $\left([T]_{\mathcal{CB}}\right)^{-1}$.

(c) Apply (b) to $T_A \in \mathrm{L}(F^n, F^m)$ with the standard bases, where $[T_A]_{\mathcal{EE}} = A$ (Proposition 2.3.16). If A is invertible, (b) gives $\left[(T_A)^{-1}\right]_{\mathcal{EE}} = A^{-1}$, so $(T_A)^{-1} = T_{A^{-1}}$ by Proposition 2.3.17. □

Example 2.5.21 (Change of coordinates is invertible). Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. The identity map id_V is its own inverse, so by Proposition 2.5.20(b), applied with the basis $\mathcal{B}$ on the domain and $\mathcal{C}$ on the target,

$$\left([\mathrm{id}_V]_{\mathcal{CB}}\right)^{-1} = [\mathrm{id}_V]_{\mathcal{BC}}.$$

The matrix converting $\mathcal{B}$-coordinates into $\mathcal{C}$-coordinates is invertible, and its inverse converts back. Example 2.4.5(c) checked one instance by multiplying out. These matrices are the subject of Section 4.5. Once we know that invertible matrices are square (Corollary 3.5.15), this also shows that any two bases of the same space have the same length, which Section 4.1 proves by other means.

Exercise 2.5.22. ROUTINE Let $T \in \mathrm{L}\left(F[x]_{\le 2}\right)$ be the shift $T(p) = p(x+1)$, whose matrix in $\mathcal{B} = 1, x, x^2$ is $[T]_{\mathcal{BB}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ (Example 2.4.13). Let $S \in \mathrm{L}\left(F[x]_{\le 2}\right)$ be the shift $S(p) = p(x-1)$. Compute $[S]_{\mathcal{BB}}$, and use Proposition 2.5.20 to show that T is invertible with $T^{-1} = S$.

Solution on page 1150

2.5.4 One-sided inverses that are not two-sided

Here are the main sources of maps and matrices that have an inverse on one side but not on the other. In each example below there are many inverses on the one side, and by Corollary 2.5.11 having two of them already rules out an inverse on the other side.

Matrices that are not square

Example 2.5.23 (One-sided inverses of a tall matrix). Let $p, q \ge 1$ and $X \in F^{q\times p}$, and consider the $(p+q)\times p$ matrix

$$A = \begin{pmatrix} I_p \\ X \end{pmatrix}.$$

It sends $\mathbf{x} \in F^p$ to $(\mathbf{x}, X\mathbf{x})$, the vector $\mathbf{x}$ followed by q more entries. For every $Y \in F^{p\times q}$, block multiplication (Theorem 2.4.34) gives

$$\begin{pmatrix} I_p - YX & Y \end{pmatrix}\begin{pmatrix} I_p \\ X \end{pmatrix}$$

$$= \big(I_p - YX\big) + YX = I_p,$$

so $B_Y = \begin{pmatrix} I_p - YX & Y \end{pmatrix}$ is a left inverse of A. These are all of them: if $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$ with $B_1 \in F^{p\times p}$ and $B_2 \in F^{p\times q}$ satisfies $BA = B_1 + B_2X = I_p$, then $B_1 = I_p - B_2X$, so $B = B_{B_2}$. Different Y give different B_Y, since Y is the right block of B_Y, so A has more than one left inverse, and therefore no right inverse (Corollary 2.5.11).

For instance, take $p = 2$, $q = 1$, and $X = \begin{pmatrix} 2 & -1 \end{pmatrix}$. With $Y = \begin{pmatrix} s \\ t \end{pmatrix}$, the left inverses of

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -1 \end{pmatrix}$$

are the matrices

$$B_{s,t} = \begin{pmatrix} 1-2s & s & s \\ -2t & 1+t & t \end{pmatrix} \quad (s,t \in F).$$

The simplest, $B_{0,0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, reads off the first two entries. The others also use the third entry $2x_1 - x_2$, which is redundant information, so it can be mixed in with any weight.

Transposing turns this into right inverses of wide matrices: $(BA)^{\mathsf{T}} = A^{\mathsf{T}}B^{\mathsf{T}}$, so B is a left inverse of A exactly when B^{T} is a right inverse of A^{T} (Exercise 2.5.45). The pattern is general. A left-invertible $m \times n$ matrix has $n \le m$, and a right-invertible one has $n \ge m$ (Theorem 3.5.12, Theorem 3.5.13). So a matrix that is not square can have an inverse on at most one side.

Differentiation and integration

Example 2.5.24 (Differentiation has a right inverse but no left inverse). Let F be a field of characteristic 0 (Definition 1.1.28), such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, so that $k * 1 = 1 + \cdots + 1$ (k terms) is nonzero for every $k \ge 1$; we write $1/k$ for its inverse. Let $D : F[x] \to F[x]$ be differentiation (Example 2.1.9), and let $J : F[x] \to F[x]$ be the antiderivative with constant term 0,

$$J\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1}.$$

Like D, the map J moves each coefficient to a fixed place and multiplies it by a fixed scalar, so it is linear.

- $DJ = \mathrm{id}_{F[x]}$. Both sides are linear, and $D\Big(J\big(x^k\big)\Big) = D\big(x^{k+1}/(k+1)\big) = x^k$ for every $k \ge 0$. The monomials span $F[x]$, so $DJ = \mathrm{id}$ by Corollary 2.2.6. So J is a right inverse of D, and D is a left inverse of J. This is the fundamental theorem of calculus, in its polynomial form.
- $JD \ne \mathrm{id}_{F[x]}$, since $J(D(1)) = J(0) = 0$. In general $J(D(p)) = p - a_0$: differentiating loses the constant term, and J cannot recover it.
- D has no left inverse, since $D(1) = 0 = D(0)$ shows that D is not injective (Proposition 2.5.4). J has no right inverse, since every value of J has constant term 0, so 1 is not a value of J (Proposition 2.5.5).
- D has many right inverses. If $E \in \mathrm{L}(F[x])$ takes only constant values, then $DE = 0$ and $D(J + E) = \mathrm{id}$. For instance, for $t \in F$ the map $J_t(p) = J(p) - (J(p))(t)$, where the scalar $(J(p))(t)$ is subtracted as a constant polynomial, is a right inverse of D. It sends p to its antiderivative that vanishes at t, and $J_0 = J$.

The same happens between finite-dimensional spaces of different dimensions. For $n \ge 1$, differentiation $D : F[x]_{\le n} \to F[x]_{\le n-1}$ and the antiderivative $J : F[x]_{\le n-1} \to F[x]_{\le n}$ still satisfy $DJ = \mathrm{id}$ and $JD \ne \mathrm{id}$. For $n = 2$, in the bases $1, x, x^2$ and $1, x$,

$$[D][J] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix} = I_2,$$

$$[J][D] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \neq I_3.$$

The finite-dimensional version is an instance of the previous subsection: a 2×3 matrix with a right inverse. The version on $F[x]$ is new. There D and J map a space to *itself*, and still $DJ = \mathrm{id}$ while $JD \neq \mathrm{id}$. For a linear map from a finite-dimensional space to itself this cannot happen: a one-sided inverse of a square matrix is automatically two-sided (Theorem 3.5.23), and correspondingly an operator on a finite-dimensional space is injective if and only if it is surjective (Corollary 4.2.18). $F[x]$ escapes because no finite list spans it, so D can be onto without being one-to-one: it shifts the monomials down by one step, and the bottom one, 1, falls off. Over $\mathbb{F}_p$ even the right inverse is lost (Exercise 2.5.42).

Remark 2.5.25 (One-sided inverses in infinite dimensions). BEYOND 21-242 The converses of Remark 2.5.7 hold in every vector space: an injective linear map has a left inverse and a surjective one has a right inverse. Both proofs need a basis of W (in the injective case, one that extends a basis of $T(V)$), and in infinite dimensions bases are known to exist only by Zorn's lemma (Section 15.1).

Products

If S and T are invertible, so is TS (Proposition 2.5.17). The converse fails, but a one-sided version survives.

Proposition 2.5.26 (One-sided inverses of composites). Let $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$.

(a) If S and T are left-invertible, so is TS. If S and T are right-invertible, so is TS.
(b) If TS is left-invertible, then so is S. If TS is right-invertible, then so is T.
(c) In particular, if TS is invertible, then $(TS)^{-1}T$ is a left inverse of S, and $S(TS)^{-1}$ is a right inverse of T.

The same holds for matrices: for instance, if $A \in F^{m\times n}$ and $B \in F^{n\times m}$ and AB is invertible, then $B(AB)^{-1}$ is a right inverse of A and $(AB)^{-1}A$ is a left inverse of B.

Proof. ROUTINE

Uses. Definition 2.5.1 (Left inverse), Definition 2.5.2 (Right inverse), Proposition 2.1.24 (Algebra of composition), Proposition 2.4.14 (Algebra of matrices)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. For (a), the candidate is the composite of the one-sided inverses in reverse order, as in Proposition 2.5.17(b).
Hint 2. For (b), if $R(TS) = \mathrm{id}_U$, regroup: $(RT)S = \mathrm{id}_U$.
Sketch. (a) If $S'S = \mathrm{id}_U$ and $T'T = \mathrm{id}_V$, then $(S'T')(TS) = S'(T'T)S = S'S = \mathrm{id}_U$; right inverses are symmetric. (b) $R(TS) = \mathrm{id}_U$ gives $(RT)S = \mathrm{id}_U$, and $(TS)R = \mathrm{id}_W$ gives $T(SR) = \mathrm{id}_W$. (c) is (b) with $R = (TS)^{-1}$.

(a) Let $S'S = \mathrm{id}_U$ and $T'T = \mathrm{id}_V$. By Proposition 2.1.24,

$$(S'T')(TS) = S'(T'T)S = S'\mathrm{id}_V S = S'S = \mathrm{id}_U,$$

so $S'T'$ is a left inverse of TS. If $SS'' = \mathrm{id}_V$ and $TT'' = \mathrm{id}_W$, then $(TS)(S''T'') = T(SS'')T'' = TT'' = \mathrm{id}_W$, so $S''T''$ is a right inverse of TS.

(b) If $R(TS) = \mathrm{id}_U$, then $(RT)S = \mathrm{id}_U$, so RT is a left inverse of S. If $(TS)R = \mathrm{id}_W$, then $T(SR) = \mathrm{id}_W$, so SR is a right inverse of T.

(c) Apply (b) with $R = (TS)^{-1}$, which is both a left and a right inverse of TS. The matrix statements are the same computations with Proposition 2.4.14. □

Example 2.5.27 (An invertible product of non-invertible factors). Let $A = \begin{pmatrix} 1 & 0 \end{pmatrix} \in F^{1\times 2}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in F^{2\times 1}$. Then $AB = (1) = I_1$ is invertible. But A is not: it has the two right inverses $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, so by

Corollary 2.5.11 it has no left inverse. Similarly B has the two left inverses $(1\ \ 0)$ and $(1\ \ 1)$, so it has no right inverse. In the other order, $BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has a zero row and is not even right-invertible (Exercise 2.5.9).

For a 2×2 example of the same kind, take $A = (I_2\ \ 0) \in F^{2\times 3}$ and $B = \begin{pmatrix} I_2 \\ 0 \end{pmatrix} \in F^{3\times 2}$, so that $AB = I_2$ by block multiplication. For an example with maps from a space to itself, take $DJ = \mathrm{id}_{F[x]}$ from Example 2.5.24, where neither D nor J is invertible.

These examples use matrices that are not square, or a space that is not finite-dimensional, and they have to. For square matrices, AB invertible implies that A and B are both invertible (Theorem 3.5.23).

2.5.5 Inverting 2×2 matrices

For 2×2 matrices there is a formula for the inverse, and a single number decides whether it exists. The candidate comes from the following computation. With $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, swap the diagonal entries and negate the other two:

$$\begin{aligned} &\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ &= \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} \\ &= \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \end{aligned}$$

So the second matrix is an inverse of A up to the scalar $ad - bc$.

Proposition 2.5.28 (Inverse of a 2×2 matrix). Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, and let $\Delta = ad - bc$.

(a) If $\Delta \neq 0$, then A is invertible, and

$$A^{-1} = \frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(b) If $\Delta = 0$, then A has neither a left inverse nor a right inverse.

Proof. ROUTINE

Uses. Definition 2.5.12 (Invertible), Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.8 (Four views of the product)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Let $A' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Compute AA' and $A'A$.
Hint 2. Both products are ΔI_2. For (a), divide by Δ.
Hint 3. For (b), suppose $BA = I_2$. Then $A' = (BA)A' = B(AA')$. What is AA' when $\Delta = 0$, and what does that say about A', and then about A?
Sketch. With $A' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, the entry formula gives $AA' = A'A = \Delta I_2$. (a) $\Delta^{-1}A'$ is a two-sided inverse. (b) If $\Delta = 0$, then $AA' = A'A = 0$. A left inverse B would give $A' = BAA' = 0$, so $a = b = c = d = 0$ and $A = 0$, which has no left inverse. A right inverse C gives $A' = A'AC = 0$ in the same way.

Let $A' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. By the entry formula,

$$AA' = \begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} = \Delta I_2,$$

$$A'A = \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{pmatrix} = \Delta I_2.$$

(a) If $\Delta \neq 0$, then by Proposition 2.4.14(d),

$$A\left(\Delta^{-1}A'\right) = \Delta^{-1}(AA') = \Delta^{-1}\Delta I_2 = I_2,$$

and likewise $(\Delta^{-1}A')A = I_2$. So $\Delta^{-1}A'$ is a two-sided inverse of A.

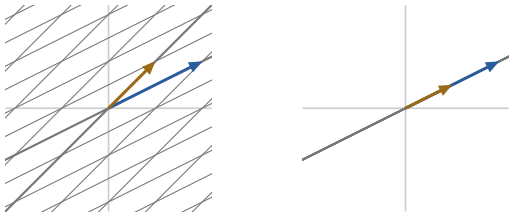
(b) Suppose $\Delta = 0$, so that $AA' = A'A = 0$. If B is a left inverse of A, then

$$A' = I_2A' = (BA)A' = B(AA') = B0 = 0.$$

So $a = b = c = d = 0$ and $A = 0$. But the zero matrix has no left inverse, since $B0 = 0 \neq I_2$, a contradiction. If C is a right inverse of A, then $A' = A'(AC) = (A'A)C = 0$, and we reach the same contradiction. □

The number $\Delta = ad - bc$ is the **determinant** of A, studied in Chapter 5, and A' is the 2×2 case of the adjugate matrix, which inverts matrices of every size (Theorem 5.5.9). For a rotation, $\Delta = \cos^2\theta + \sin^2\theta = 1$ and the formula returns $[R_{-\theta}]$, as in Example 2.5.14(b).

Geometrically, $|\Delta|$ is the area of the parallelogram spanned by the columns of A, which is the image of the unit square under T_A (Section 5.1). When $\Delta = 0$ the columns are parallel, and T_A flattens the plane onto a line through $\mathbf{0}$, or onto $\{\mathbf{0}\}$. A flattened plane cannot be unflattened: two points that land on the same point cannot both be recovered. Drag the columns below until they line up, and watch the area reach 0.



The map of $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, with $\Delta = 1$. Move $A\mathbf{e}_2$ onto the line through $A\mathbf{e}_1$ (for example to $(1, 1/2)$) and the grid collapses onto that line; the signed area $ad - bc$ is then 0, and by Proposition 2.5.28 the matrix has no inverse on either side. (Interactive in the web version.)

Example 2.5.29 (One matrix over several fields). Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, with $\Delta = 4 - 6 = -2$. Whether A is invertible depends on the field in which we read its entries.

- Over $\mathbb{Q}$, $\Delta = -2 \neq 0$ and $A^{-1} = -\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.
- Over $\mathbb{F}_5$, $\Delta = -2 = 3$ and $3^{-1} = 2$, so $A^{-1} = 2\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -4 \\ -6 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}$. As a check, $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 11 & 5 \\ 25 & 11 \end{pmatrix}$, which is I_2 modulo 5.
- Over $\mathbb{F}_3$, $\Delta = -2 = 1$ and $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, so $A^{-1} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
- Over $\mathbb{F}_2$, $\Delta = 0$, so A is not invertible. Indeed $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ there, with a zero column.

So a matrix with integer entries can be invertible over $\mathbb{Q}$ and over some $\mathbb{F}_p$ but not others; here it fails exactly for $p = 2$, the only prime dividing $\Delta = -2$.

Exercises

Exercise 2.5.30. ROUTINE Let $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \in \mathbb{Q}^{3\times 1}$.

(a) Find all left inverses of $\mathbf{a}$.

(b) Show that $\mathbf{a}$ has no right inverse, in two ways: directly from Definition 2.5.2, and from (a).

Solution on page 1151

Exercise 2.5.31. ROUTINE

(a) Over $\mathbb{C}$, find the inverse of $\begin{pmatrix} 1 & i \\ i & 2 \end{pmatrix}$.

(b) Over $\mathbb{F}_7$, find the inverse of $\begin{pmatrix} 3 & 5 \\ 2 & 6 \end{pmatrix}$, and check your answer.

(c) For which $t \in \mathbb{R}$ is $\begin{pmatrix} t & 1 \\ 4 & t \end{pmatrix}$ invertible? For which $t \in \mathbb{F}_5$?

Solution on page 1151

Exercise 2.5.32. ROUTINE Let $A \in F^{n\times n}$ with $n \geq 1$.

(a) Show that if A is invertible and $AB = 0$ for some $B \in F^{n\times p}$, then $B = 0$.

(b) Show that if $A^k = 0$ for some $k \geq 1$, then A is not invertible.

(c) Show that if A is invertible and $A^2 = A$, then $A = I_n$.

(d) Show that if $A^k = I_n$ for some $k \geq 1$, then A is invertible with $A^{-1} = A^{k-1}$.

Solution on page 1151

Exercise 2.5.33. ROUTINE Let $T \in \mathrm{L}(V, W)$ and $S \in \mathrm{L}(W, V)$.

(a) Suppose $ST = \mathrm{id}_V$ and T is surjective. Show that $TS = \mathrm{id}_W$.

(b) Suppose $TS = \mathrm{id}_W$ and T is injective. Show that $ST = \mathrm{id}_V$.

(c) Deduce that a left-invertible linear map is invertible if and only if it is surjective, and a right-invertible one is invertible if and only if it is injective.

Solution on page 1152

Exercise 2.5.34. STANDARD Find the error in the following argument, and show that its conclusion is false. Then add a hypothesis on A under which the conclusion holds.

Claim. Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$. If $AB = I_m$, then $BA = I_n$.

Proof. Since $AB = I_m$, the matrix B is the inverse of A, that is, $B = A^{-1}$. Hence $BA = A^{-1}A = I_n$.

Solution on page 1152

Exercise 2.5.35. STANDARD Prove or disprove each statement, where $m, n \geq 1$.

(a) If $A, B \in F^{n\times n}$ are invertible, then $A + B$ is invertible.

(b) If $A \in F^{n\times n}$ and A^2 is invertible, then A is invertible.

(c) If $A \in F^{n\times n}$ is invertible and symmetric, then A^{-1} is symmetric.

(d) If $A \in F^{m\times n}$ and $A^{\mathsf{T}}A$ is invertible, then A is left-invertible.

(e) If $A \in F^{m\times n}$ is left-invertible, then $A^{\mathsf{T}}A$ is invertible.

Solution on page 1152

Exercise 2.5.36. STANDARD Let p be a prime and $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{F}_p^{3\times 2}$.

(a) How many left inverses does A have?

(b) Does A have a right inverse?

(c) How many right inverses does A^{T} have?

Solution on page 1153

Exercise 2.5.37. STANDARD Let $T \in \mathrm{L}(V,W)$, and let S be a left inverse of T. Show that T is invertible if and only if S is left-invertible, and that in this case $T^{-1} = S$.

Solution on page 1153

Exercise 2.5.38. STANDARD Let $T \in \mathrm{L}(V)$, and suppose that $a_k T^k + \cdots + a_1 T + a_0 \mathrm{id}_V = 0$ for some $a_0, \ldots, a_k \in F$ with $a_0 \neq 0$.

(a) Show that T is invertible, and that $T^{-1} = -a_0^{-1}\left(a_k T^{k-1} + \cdots + a_2 T + a_1 \mathrm{id}_V\right)$.

(b) Let $A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Verify that $A^2 - A + I_2 = 0$, find A^{-1} from (a), and show that $A^3 = -I_2$.

Solution on page 1154

Exercise 2.5.39. STANDARD

(a) Let $N \in \mathrm{L}(V)$ with $N^k = 0$ for some $k \geq 1$. Show that $\mathrm{id}_V - N$ and $\mathrm{id}_V + N$ are invertible, with

$$(\mathrm{id}_V - N)^{-1} = \mathrm{id}_V + N + N^2 + \cdots + N^{k-1}.$$

(b) Use (a) to invert $\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$ for $a, b, c \in F$.

(c) Show that $\mathrm{id} + D$ is invertible on $F[x]_{\leq 3}$, where D is differentiation, and compute $(\mathrm{id} + D)^{-1}\left(x^3\right)$.

Solution on page 1154

Exercise 2.5.40. STANDARD Let $A \in F^{p\times p}$ and $D \in F^{q\times q}$ be invertible, and let $X \in F^{p\times q}$.

(a) Show that $M = \begin{pmatrix} A & X \\ 0 & D \end{pmatrix}$ is invertible, with

$$M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}XD^{-1} \\ 0 & D^{-1} \end{pmatrix}.$$

(b) Use (a) to invert $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$.

Solution on page 1155

Exercise 2.5.41. STANDARD Let $M : F[x] \to F[x]$ be multiplication by x, $M(p) = xp$, and let $L : F[x] \to F[x]$ be

$$\begin{aligned} &L(a_0 + a_1 x + \cdots + a_n x^n) \\ &= a_1 + a_2 x + \cdots + a_n x^{n-1}. \end{aligned}$$

(a) Show that $LM = \mathrm{id}_{F[x]}$, compute ML, and show that M has no right inverse and L no left inverse.

(b) Find all left inverses of M.

Solution on page 1156

Exercise 2.5.42. STANDARD

(a) On $\mathbb{Q}[x]$, construct a right inverse R of differentiation D such that $(R(p))(1) = 0$ for every $p \in \mathbb{Q}[x]$.

(b) Let p be a prime. Show that x^{p-1} is not a value of $D : \mathbb{F}_p[x] \to \mathbb{F}_p[x]$, and conclude that over $\mathbb{F}_p$, D has no right inverse.

Solution on page 1156

Exercise 2.5.43. STANDARD Let p be a prime. Use Proposition 2.5.28 to show that exactly $\left(p^2-1\right)\left(p^2-p\right)$ matrices in $\mathbb{F}_p^{2\times 2}$ are invertible. List them for $p=2$.

Solution on page 1157

Exercise 2.5.44. STANDARD

(a) Let $A \in F^{m\times n}$ be invertible. Show that $m * 1 = n * 1$ in F. Deduce that over a field of characteristic 0, such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, every invertible matrix is square.

(b) More generally, let $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $AB = I_m$, and suppose $m * 1 \neq n * 1$ in F. Show that $BA \neq I_n$.

(c) Explain why this argument does not show that invertible matrices over $\mathbb{F}_2$ are square.

Solution on page 1158

Exercise 2.5.45. ROUTINE Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$. Show that B is a left inverse of A if and only if B^{T} is a right inverse of A^{T}. Deduce that A is left-invertible if and only if A^{T} is right-invertible, and that A is invertible if and only if A^{T} is.

Solution on page 1158

Exercise 2.5.46. HARD

(a) Show that every $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$ satisfies

$$A^2 - (a+d)A + (ad-bc)I_2 = 0.$$

(b) Use (a) and Exercise 2.5.38 to give a second proof of Proposition 2.5.28(a).

(c) Let p be an odd prime. How many $A \in \mathbb{F}_p^{2\times 2}$ satisfy $A^{-1} = A$?

Solution on page 1158

Exercise 2.5.47. HARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$.

(a) Suppose that $I_m - AB$ is invertible. Show that $I_n - BA$ is invertible, with

$$(I_n - BA)^{-1} = I_n + B(I_m - AB)^{-1}A.$$

(b) Deduce: if $\mathbf{u}, \mathbf{v} \in F^n$ and $1 + \mathbf{v}^{\mathsf{T}}\mathbf{u} \neq 0$, then $I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ is invertible, with inverse $I_n - \left(1+\mathbf{v}^{\mathsf{T}}\mathbf{u}\right)^{-1}\mathbf{u}\mathbf{v}^{\mathsf{T}}$.

Solution on page 1159

Exercise 2.5.48. HARD Let $S, T \in \mathrm{L}(V)$ satisfy $ST = S + T$.

(a) Show that $(\mathrm{id} - S)(\mathrm{id} - T) = \mathrm{id}$.

(b) Suppose that $\mathrm{id} - S$ is invertible. Show that $ST = TS$. (When $V = F^n$, this hypothesis follows from (a) by Theorem 3.5.23.)

(c) Find $S, T \in \mathrm{L}(\mathbb{Q}[x])$ with $ST = S + T$ but $ST \neq TS$.

Solution on page 1160

Exercise 2.5.49. CHALLENGE Let $A, B \in F^{n\times n}$ with $n \geq 1$ and $AB = I_n$. You may use that any n^2+1 matrices in $F^{n\times n}$ are linearly dependent, because $F^{n\times n}$ has dimension n^2 (Section 4.1).

(a) Show that $A^r B^k = B^{k-r}$ whenever $k \geq r \geq 0$.

(b) Show that B has a right inverse that is a linear combination of $I_n, B, B^2, \ldots$.

(c) Conclude that $BA = I_n$.

Solution on page 1161

2.6 Kernel and image

In Section 2.5, a left inverse of T showed that the equation $T(\mathbf{v}) = \mathbf{w}$ has at most one solution, and a right inverse showed that it has at least one. This section studies the two properties themselves. As properties of functions they are injectivity and surjectivity, and each is a statement about every $\mathbf{w} \in W$ at once. For a linear map, linearity packs each of them into a single subspace: the kernel, which consists of the vectors sent to $\mathbf{0}$, and the image, which consists of the vectors that are reached. In particular, whether $T(\mathbf{v}) = \mathbf{w}$ has at most one solution for every $\mathbf{w}$ is decided by the one equation $T(\mathbf{v}) = \mathbf{0}$.

Throughout, U, V, and W are vector spaces over F, and $T \in \mathrm{L}(V, W)$ unless said otherwise.

2.6.1 Injective, surjective, and bijective maps

The three notions below make sense for any function $f : X \to Y$ between sets, and Section A.1 treats them in that generality. For $y \in Y$, an element $x \in X$ with $f(x) = y$ is a **preimage** of y. The definitions count preimages.

Definition 2.6.1 (Injective). A function $f : X \to Y$ is **injective** if every $y \in Y$ has at most one preimage. Equivalently, for all $x, x' \in X$, if $f(x) = f(x')$, then $x = x'$.

Definition 2.6.2 (Surjective). A function $f : X \to Y$ is **surjective** if every $y \in Y$ has at least one preimage, that is, if for every $y \in Y$ there is some $x \in X$ with $f(x) = y$.

Definition 2.6.3 (Bijective). A function $f : X \to Y$ is **bijective** if every $y \in Y$ has exactly one preimage, that is, if f is both injective and surjective.

The two forms of injectivity say the same thing: if $f(x) = f(x') = y$, then x and x' are both preimages of y, and "at most one" says that they are equal. The second form is the one we use in proofs. Its contrapositive reads: different inputs have different values. Other common names are *one-to-one* for injective, *onto* for surjective, and *one-to-one correspondence* for a bijection.

For a linear map T, the preimages of $\mathbf{w}$ are the solutions of the equation $T(\mathbf{v}) = \mathbf{w}$. So T is injective if and only if this equation has at most one solution for every $\mathbf{w} \in W$, surjective if and only if it has at least one, and bijective if and only if it has exactly one. In this language, Proposition 2.5.4 and Proposition 2.5.5 say that a left-invertible linear map is injective and a right-invertible one is surjective, so an invertible linear map is bijective. Conversely, a bijective linear map is invertible, because its inverse function turns out to be linear (Proposition 2.7.7). The converses for one-sided inverses are harder (Remark 2.5.7).

Example 2.6.4 (Injective, surjective, both, neither).

(a) $T : F \to F^2$, $T(x) = (x, 0)$, is injective: if $(x, 0) = (x', 0)$, then $x = x'$. It is not surjective, since $(0, 1)$ has no preimage (Example 2.5.3).

(b) $P : F^2 \to F$, $P(x, y) = x$, is surjective, since $x = P(x, 0)$ for every $x \in F$. It is not injective, since $P(0, 1) = 0 = P(0, 0)$.

(c) The rotation R_θ of $\mathbb{R}^2$ is bijective, since it is invertible with inverse $R_{-\theta}$ (Example 2.5.14).

(d) The projection Pr_ℓ onto a line ℓ through $\mathbf{0}$ in $\mathbb{R}^2$ is neither injective nor surjective (Non-example 2.5.15).

(e) On $F[x]$, multiplication by x, $M(p) = xp$ (Example 2.1.9), is injective but not surjective. If $p = \sum_k a_k x^k$, then $xp = \sum_k a_k x^{k+1}$ has the coefficients of p, moved up by one place, so $xp = xq$ forces $p = q$. Every value of M has constant term 0, so 1 has no preimage. In the other direction, if F has characteristic 0, then differentiation D is surjective but not injective (Example 2.5.24).

In (e), an operator on $F[x]$ is injective without being surjective, and another is surjective without being injective. For an operator on a finite-dimensional space this cannot happen (Corollary 4.2.18).

2.6.2 Kernel and image

Checking injectivity straight from Definition 2.6.1 means comparing $T(\mathbf{v})$ with $T(\mathbf{v}')$ for every pair of vectors. For a linear map there is a shortcut, since $T(\mathbf{v}) = T(\mathbf{v}')$ if and only if $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$. Every coincidence of values is therefore recorded by a vector that T sends to $\mathbf{0}$. Surjectivity, on the other hand, asks which vectors T reaches.

Definition 2.6.5 (Kernel). The **kernel** of $T \in \mathrm{L}(V, W)$ is the set of vectors that T sends to $\mathbf{0}$,

$$\ker T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}_W\} \subseteq V.$$

Definition 2.6.6 (Image). The **image** of $T \in \mathrm{L}(V, W)$ is the set of values of T,

$$\mathrm{im}\, T = \{\mathbf{w} \in W \mid \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\} = T(V) \subseteq W.$$

The kernel is also called the **null space** of T, and the image its **range**. Note where each one lives: $\ker T$ is a subset of the domain V, and $\mathrm{im}\, T$ is a subset of the target W. Directly from Definition 2.6.2, T is surjective if and only if $\mathrm{im}\, T = W$. The kernel always contains $\mathbf{0}_V$, since $T(\mathbf{0}_V) = \mathbf{0}_W$ (Proposition 2.1.4).

Example 2.6.7 (First kernels and images).

(a) The zero map $V \to W$ has kernel V and image $\{\mathbf{0}\}$. The identity id_V has kernel $\{\mathbf{0}\}$ and image V.

(b) For $T(x) = (x, 0)$ from Example 2.6.4(a), $\ker T = \{0\}$ and $\mathrm{im}\, T = \{(x, 0) \mid x \in F\} = \mathrm{span}(\mathbf{e}_1)$. For $P(x, y) = x$ from part (b) of that example, $\ker P = \mathrm{span}(\mathbf{e}_2)$ and $\mathrm{im}\, P = F$.

(c) Let $Q : F^3 \to F^3$, $Q(x, y, z) = (x, y, 0)$. Then $Q(x, y, z) = \mathbf{0}$ if and only if $x = y = 0$, so $\ker Q = \mathrm{span}(\mathbf{e}_3)$, the z-axis. Every value of Q has third entry 0, and every such vector is a value, since $Q(x, y, 0) = (x, y, 0)$. So $\mathrm{im}\, Q = \mathrm{span}(\mathbf{e}_1, \mathbf{e}_2)$, the xy-plane.

(d) Let $\ell = \mathrm{span}(\mathbf{u})$ with $\mathbf{u} = (c, s) = (\cos\theta, \sin\theta)$. By Example 2.1.11, $\mathrm{Pr}_\ell(x, y) = (xc + ys)\mathbf{u}$. Since $\mathbf{u} \neq \mathbf{0}$, $\mathrm{Pr}_\ell(x, y) = \mathbf{0}$ if and only if $xc + ys = 0$ (Proposition 1.2.35). Using $c^2 + s^2 = 1$, every $(x, y) \in \mathbb{R}^2$ satisfies

$$(x, y) = (xc + ys)(c, s) + (yc - xs)(-s, c),$$

so the solutions of $xc + ys = 0$ are exactly the multiples of $(-s, c)$. Hence $\ker \mathrm{Pr}_\ell = \mathrm{span}((-s, c))$, the line through $\mathbf{0}$ perpendicular to ℓ. Every value of Pr_ℓ lies on ℓ, and every $t\mathbf{u} \in \ell$ is a value, since $\mathrm{Pr}_\ell(t\mathbf{u}) = t(c^2 + s^2)\mathbf{u} = t\mathbf{u}$. So $\mathrm{im}\, \mathrm{Pr}_\ell = \ell$.

Intuition (What the kernel and the image measure). The kernel measures how far T is from being injective, and the image measures how close it comes to being surjective. Two vectors with the same value differ

by an element of the kernel, so a large kernel means that T glues many vectors together and loses information. Theorem 2.6.11 below shows that T loses nothing exactly when the kernel is as small as it can be, namely $\{\mathbf{0}\}$. The image is what T reaches, and T is surjective exactly when the image is as large as it can be, namely W. In Example 2.6.7(d), Pr_ℓ forgets the component perpendicular to ℓ and reaches only ℓ.

Both sets are subspaces (Proposition 2.6.8), so in finite dimensions their sizes can be measured by dimension. Rank–nullity (Section 4.2) shows that the two measurements are tied together: the more T loses, the less it reaches.

Proposition 2.6.8 (Kernels and images are subspaces). For every $T \in \mathrm{L}(V, W)$, $\ker T$ is a subspace of V and $\operatorname{im} T$ is a subspace of W.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Definition 2.1.1 (Linear map), Proposition 1.2.33 (The zero vector)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Use part (III) of the subspace criterion (Theorem 1.4.3): check that each set contains the zero vector and is closed under addition and scalar multiplication.
Hint 2. For the kernel, apply T to $\mathbf{u}+\mathbf{v}$ and to $a\mathbf{u}$, where $T(\mathbf{u}) = T(\mathbf{v}) = \mathbf{0}$. For the image, write $T(\mathbf{u}) + T(\mathbf{v})$ and $aT(\mathbf{u})$ as single values of T.
Sketch. Both sets contain $\mathbf{0}$ because $T(\mathbf{0}) = \mathbf{0}$. For $\mathbf{u}, \mathbf{v} \in \ker T$, linearity gives $T(\mathbf{u}+\mathbf{v}) = \mathbf{0}+\mathbf{0} = \mathbf{0}$ and $T(a\mathbf{u}) = a\mathbf{0} = \mathbf{0}$. For values, $T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u}+\mathbf{v})$ and $aT(\mathbf{u}) = T(a\mathbf{u})$.

The kernel. By Proposition 2.1.4, $T(\mathbf{0}_V) = \mathbf{0}_W$, so $\mathbf{0}_V \in \ker T$. Let $\mathbf{u}, \mathbf{v} \in \ker T$ and $a \in F$. Since T is linear,

$$T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0}+\mathbf{0} = \mathbf{0}, \quad T(a\mathbf{u}) = aT(\mathbf{u}) = a\mathbf{0} = \mathbf{0},$$

where the last step is Proposition 1.2.33. So $\mathbf{u}+\mathbf{v}$ and $a\mathbf{u}$ lie in $\ker T$, and $\ker T$ is a subspace of V by Theorem 1.4.3.

The image. Again $\mathbf{0}_W = T(\mathbf{0}_V) \in \operatorname{im} T$. Let $\mathbf{w}, \mathbf{w}' \in \operatorname{im} T$ and $a \in F$, say $\mathbf{w} = T(\mathbf{u})$ and $\mathbf{w}' = T(\mathbf{v})$ with $\mathbf{u}, \mathbf{v} \in V$. Since T is linear,

$$\mathbf{w}+\mathbf{w}' = T(\mathbf{u}+\mathbf{v}), \quad a\mathbf{w} = T(a\mathbf{u}),$$

so both vectors lie in $\operatorname{im} T$, and $\operatorname{im} T$ is a subspace of W by Theorem 1.4.3. □

The proof used only that $\{\mathbf{0}\}$ is a subspace of W and that V is a subspace of itself. The same argument shows that the image $\operatorname{im}(T|_X)$ of a subspace X of V is a subspace of W, and that the set of $\mathbf{v}$ with $T(\mathbf{v}) \in U$ is a subspace of V for every subspace U of W (Exercise 2.6.51). Linearity cannot be dropped.

Non-example 2.6.9 (Without linearity).

(a) The function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$, sends exactly the points of the two coordinate axes to 0. This set contains $(1, 0)$ and $(0, 1)$ but not their sum.

(b) The values of $g : \mathbb{R} \to \mathbb{R}^2$, $g(t) = (t, t^2)$, form the parabola $y = x^2$. It contains $(1, 1)$ but not $2(1, 1) = (2, 2)$, since $2 \neq 2^2$.

(c) An affine map $\mathbf{v} \mapsto T(\mathbf{v}) + \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$ (Non-example 2.1.12) sends $\mathbf{0}$ to $\mathbf{b}$, so the set of vectors it sends to $\mathbf{0}$ does not contain $\mathbf{0}$.

Exercise 2.6.10. ROUTINE Find the kernel and the image of each map, and decide whether it is injective and whether it is surjective.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(x, y) = (x - y, 2y - 2x)$.

(b) $S : F^2 \to F^3$, $S(x, y) = (x, x + y, y)$.

(c) $\mathrm{ev}_1 : F[x]_{\leq 1} \to F$, $\mathrm{ev}_1(p) = p(1)$.

Solution on page 1161

2.6.3 Injectivity and the kernel

Intuition. Suppose $T(\mathbf{v}) = T(\mathbf{v}')$. Then $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$, so the difference $\mathbf{v} - \mathbf{v}'$ lies in $\ker T$. Conversely, adding a vector $\mathbf{z} \in \ker T$ does not change values: $T(\mathbf{v} + \mathbf{z}) = T(\mathbf{v}) + \mathbf{0} = T(\mathbf{v})$. So T gives two vectors the same value exactly when they differ by an element of the kernel. A coincidence anywhere in V is a coincidence with $\mathbf{0}$, moved over, and the coincidences with $\mathbf{0}$ are recorded by the nonzero vectors of $\ker T$.

Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$). A linear map $T \in \mathrm{L}(V, W)$ is injective if and only if $\ker T = \{\mathbf{0}\}$.

Proof. ROUTINE

Uses. Definition 2.6.1 (Injective), Definition 2.6.5 (Kernel), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For "only if", recall that $\mathbf{0} \in \ker T$ always. If $\mathbf{v} \in \ker T$, compare $T(\mathbf{v})$ with $T(\mathbf{0})$.
Hint 2. For "if", suppose $T(\mathbf{v}) = T(\mathbf{v}')$ and apply T to $\mathbf{v} - \mathbf{v}'$.
Sketch. If T is injective and $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$, then $\mathbf{v} = \mathbf{0}$. If $\ker T = \{\mathbf{0}\}$ and $T(\mathbf{v}) = T(\mathbf{v}')$, then $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$ by linearity, so $\mathbf{v} - \mathbf{v}' = \mathbf{0}$.

Only if. Suppose T is injective. We always have $\mathbf{0} \in \ker T$ (Proposition 2.1.4). If $\mathbf{v} \in \ker T$, then $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$, and injectivity gives $\mathbf{v} = \mathbf{0}$. So $\ker T = \{\mathbf{0}\}$.

If. Suppose $\ker T = \{\mathbf{0}\}$, and let $\mathbf{v}, \mathbf{v}' \in V$ with $T(\mathbf{v}) = T(\mathbf{v}')$. By Proposition 2.1.4,

$$T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}') = \mathbf{0},$$

so $\mathbf{v} - \mathbf{v}' \in \ker T = \{\mathbf{0}\}$. Hence $\mathbf{v} = \mathbf{v}'$, and T is injective. □

Linearity is essential here. The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, sends only 0 to 0, but $f(1) = f(-1)$.

Technique 2.6.12 (Test injectivity at $\mathbf{0}$). To prove that a linear map T is injective, take a vector $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{0}$ and show that $\mathbf{v} = \mathbf{0}$ (Theorem 2.6.11). This replaces a statement about all pairs of vectors by a statement about one equation, and $T(\mathbf{v}) = \mathbf{0}$ is usually the easiest equation to analyze. To prove that T is *not* injective, exhibit a single nonzero vector of $\ker T$.

The same move is how injectivity gets used. A relation $a_1 T(\mathbf{v}_1) + \cdots + a_k T(\mathbf{v}_k) = \mathbf{0}$ can be pulled back through T: by linearity it says $T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = \mathbf{0}$, and if T is injective, then $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$.

Reach for it whenever injectivity of a linear map is to be proved or used. Check linearity first: for other functions the test proves nothing, as $x \mapsto x^2$ shows. The canonical example is Example 2.6.13.
Used in. Proposition 2.6.14, Exercise 2.6.16, Proposition 2.6.22, Exercise 2.8.7, Exercise 2.8.10, Problem 2.1, Problem 2.15, Problem 2.18, Problem 2.19, Exercise 4.4.35, Proposition 6.5.13, Proposition 6.5.21, Corollary 8.4.6, Exercise 2.6.37.

Example 2.6.13 (Testing at $\mathbf{0}$).

(a) Let $T : F[x] \to F[x]$, $T(p) = p + D(p)$, where D is differentiation. We claim that T is injective over every field. Differentiation lowers the degree of every nonzero polynomial: if $\deg p = d \geq 1$, then $D(p)$ has degree at most $d - 1$ by the formula in Example 2.1.9, and if p is a nonzero constant, then $D(p) = \mathbf{0}$, whose degree is $-\infty$. So for $p \neq \mathbf{0}$ the degrees of p and $D(p)$ differ, and Proposition 1.2.18(a) gives $\deg(p + D(p)) = \deg p \geq 0$. In particular $T(p) \neq \mathbf{0}$. Hence $\ker T = \{\mathbf{0}\}$, and T is injective by Theorem 2.6.11. Differentiation alone is not injective; adding the identity repairs this.

(b) Let $S : F^2 \to F^2$, $S(x, y) = (x + y, x - y)$. If $S(x, y) = \mathbf{0}$, then $x + y = 0$ and $x - y = 0$. Adding and subtracting the two equations gives $2x = 0$ and $2y = 0$, where $2 = 1 + 1$. If $2 \neq 0$ in F, then $x = y = 0$, so S is injective. Over $\mathbb{F}_2$, $S(1, 1) = (1 + 1, 1 - 1) = \mathbf{0}$, so $(1, 1)$ is a nonzero vector of $\ker S$, and S is not injective.

Proposition 2.1.16(b) says that every linear map sends linearly dependent lists to linearly dependent lists. It may also send independent lists to dependent ones, as Pr_ℓ does with $\mathbf{e}_1, \mathbf{e}_2$. The injective maps are exactly the ones that never do.

Proposition 2.6.14 (Injective maps preserve independence). For $T \in \mathrm{L}(V, W)$, the following are equivalent.

(I) T is injective.

(II) For every linearly independent list $\mathcal{L} = \mathbf{v}_1, ..., \mathbf{v}_k$ in V, the list $T(\mathcal{L}) = T(\mathbf{v}_1), ..., T(\mathbf{v}_k)$ is linearly independent.

Proof. ROUTINE

Uses. Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.5.15 (Short lists, zeros, and repeats), Definition 2.1.1 (Linear map)
Techniques. Test injectivity at $\mathbf{0}$ (Technique 2.6.12), Contrapositive (Technique 1.5.4)
Hint 1. For (I) $\Rightarrow$ (II), start from a relation $\sum_i a_i T(\mathbf{v}_i) = \mathbf{0}$ and pull it back through T.
Hint 2. For (II) $\Rightarrow$ (I), work with the kernel. Which list could you feed to (II) to learn something about a vector $\mathbf{v} \in \ker T$?
Hint 3. A single nonzero vector forms a linearly independent list, and the list $\mathbf{0}$ is linearly dependent.
Sketch. (I) $\Rightarrow$ (II): $\sum_i a_i T(\mathbf{v}_i) = \mathbf{0}$ gives $T\left(\sum_i a_i \mathbf{v}_i\right) = \mathbf{0}$, so $\sum_i a_i \mathbf{v}_i = \mathbf{0}$, and all $a_i = 0$. (II) $\Rightarrow$ (I): if $\ker T$ contained some $\mathbf{v} \neq \mathbf{0}$, the list $\mathbf{v}$ would be independent while its image, the list $\mathbf{0}$, is not.

(I) $\Rightarrow$ (II). Suppose T is injective, let $\mathcal{L}$ be linearly independent, and suppose that $a_1 T(\mathbf{v}_1) + \cdots + a_k T(\mathbf{v}_k) = \mathbf{0}$. By linearity, $T(a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k) = \mathbf{0}$, so $a_1 \mathbf{v}_1 + \cdots + a_k \mathbf{v}_k \in \ker T = \{\mathbf{0}\}$ by Theorem 2.6.11. As $\mathcal{L}$ is linearly independent, $a_1 = \cdots = a_k = 0$ (Theorem 1.5.5). So $\mathbf{0}$ has only the trivial representation by $T(\mathcal{L})$, and $T(\mathcal{L})$ is linearly independent by Theorem 1.5.5.

(II) $\Rightarrow$ (I). We prove the contrapositive. Suppose T is not injective. By Theorem 2.6.11 there is a vector $\mathbf{v} \neq \mathbf{0}$ in $\ker T$. The list $\mathbf{v}$ is linearly independent, but its image is the list $T(\mathbf{v}) = \mathbf{0}$, which is linearly dependent (Proposition 1.5.15). So (II) fails. □

Composites

Under composition, kernels can only grow and images can only shrink.

Proposition 2.6.15 (Kernels and images of composites). Let $S \in \mathrm{L}(U, V)$ and $T \in \mathrm{L}(V, W)$.

(a) $\ker S \subseteq \ker(TS)$, with equality if T is injective.

(b) $\mathrm{im}(TS) = \mathrm{im}(T|_{\mathrm{im}\, S}) \subseteq \mathrm{im}\, T$, with equality if S is surjective.

(c) If TS is injective, then S is injective. If TS is surjective, then T is surjective.

Proof. ROUTINE

Uses. Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For (a): if $S(\mathbf{u}) = \mathbf{0}$, what is $T(S(\mathbf{u}))$? For the equality, if $T(S(\mathbf{u})) = \mathbf{0}$, which vector lies in $\ker T$?
Hint 2. For (b), unwind the definitions: a value of TS is T applied to a value of S.

Hint 3. (c) follows from (a) and (b): $\ker S \subseteq \ker(TS)$ and $\operatorname{im}(TS) \subseteq \operatorname{im} T$.
Sketch. (a) $T(S(\mathbf{u})) = T(\mathbf{0}) = \mathbf{0}$ for $\mathbf{u} \in \ker S$; if T is injective and $T(S(\mathbf{u})) = \mathbf{0}$, then $S(\mathbf{u}) \in \ker T = \{\mathbf{0}\}$. (b) Both $\operatorname{im}(TS)$ and $\operatorname{im}(T\restriction_{\operatorname{im} S})$ are the set of all $T(S(\mathbf{u}))$, which lie in $\operatorname{im} T$; if $\operatorname{im} S = V$, then $T\restriction_{\operatorname{im} S} = T$. (c) $\ker S \subseteq \ker(TS) = \{\mathbf{0}\}$, and $W = \operatorname{im}(TS) \subseteq \operatorname{im} T$.

(a) If $\mathbf{u} \in \ker S$, then $(TS)(\mathbf{u}) = T(\mathbf{0}) = \mathbf{0}$ by Proposition 2.1.4, so $\mathbf{u} \in \ker(TS)$. If T is injective and $\mathbf{u} \in \ker(TS)$, then $S(\mathbf{u}) \in \ker T = \{\mathbf{0}\}$ by Theorem 2.6.11, so $\mathbf{u} \in \ker S$.

(b) By definition, $\operatorname{im}(TS) = \{T(S(\mathbf{u})) \mid \mathbf{u} \in U\}$, and

$$\operatorname{im}(T\restriction_{\operatorname{im} S}) = \{T(\mathbf{v}) \mid \mathbf{v} \in \operatorname{im} S\} = \{T(S(\mathbf{u})) \mid \mathbf{u} \in U\}$$

as well. Each $T(S(\mathbf{u}))$ is a value of T, so this set is contained in $\operatorname{im} T$. If S is surjective, then $\operatorname{im} S = V$, so $T\restriction_{\operatorname{im} S} = T$ and the two images agree.

(c) If TS is injective, then by (a) and Theorem 2.6.11, $\ker S \subseteq \ker(TS) = \{\mathbf{0}\}$. So $\ker S = \{\mathbf{0}\}$, and S is injective. If TS is surjective, then by (b), $W = \operatorname{im}(TS) \subseteq \operatorname{im} T \subseteq W$, so $\operatorname{im} T = W$. □

Part (c) is the counterpart of Proposition 2.5.26(b) for injectivity and surjectivity, and it cannot be improved. With $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the composite $T_A T_B = T_{AB} = \mathrm{id}_F$ is bijective, but $T_A(x, y) = x$ is not injective and $T_B(x) = (x, 0)$ is not surjective (Example 2.5.27).

Exercise 2.6.16. ROUTINE

(a) Show that $T : F^2 \to F^3$, $T(x, y) = (x, x + y, x + 2y)$, is injective over every field.

(b) Show that $T : F[x] \to F[x]$, $T(p) = xp + p$, is injective. (Look at the coefficient of the highest power of x in $T(p)$.)

(c) Let $T \in \mathrm{L}(V)$ with $T^2 = 0$. Show that T is not injective unless $V = \{\mathbf{0}\}$.

Solution on page 1162

2.6.4 Null space and column space

A matrix $A \in F^{m \times n}$ gives the linear map $T_A : F^n \to F^m$, $\mathbf{x} \mapsto A\mathbf{x}$ (Definition 2.3.15), and the kernel and image of T_A have names of their own.

Definition 2.6.17 (Null space and column space). Let $A \in F^{m \times n}$. The **null space** of A is

$$\ker T_A = \{\mathbf{x} \in F^n \mid A\mathbf{x} = \mathbf{0}\},$$

a subspace of F^n, and the **column space** of A is

$$\operatorname{im} T_A = \{A\mathbf{x} \mid \mathbf{x} \in F^n\},$$

a subspace of F^m. We sometimes abbreviate them to $\ker A$ and $\operatorname{im} A$.

By Proposition 3.1.8, the null space of A is the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$, and the system $A\mathbf{x} = \mathbf{y}$ has a solution exactly when $\mathbf{y}$ lies in the column space. The name "column space" comes from the following description.

Proposition 2.6.18 (The column space is spanned by the columns). Let $A \in F^{m \times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Then

$$\operatorname{im} T_A = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

Proof (from the definition of $A\mathbf{x}$). ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 1.3.10 (Span), Definition 2.6.6 (Image)
Hint 1. What is $A\mathbf{x}$, by definition, in terms of the columns of A?

Hint 2. As $\mathbf{x}$ runs over F^n, its entries $x_1, \dots, x_n$ run over all lists of n scalars.
Sketch. $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$, and as $\mathbf{x}$ runs over F^n these are all the linear combinations of the columns.

By Definition 2.3.8, $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ for $\mathbf{x} = (x_1, \dots, x_n)$. Every list of scalars $x_1, \dots, x_n$ is the list of entries of some $\mathbf{x} \in F^n$, so the values of T_A are exactly the linear combinations of $\mathbf{a}_1, \dots, \mathbf{a}_n$. By Definition 1.3.10, the set of these is $\operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$. □

Proof (from a spanning list). ROUTINE

Uses. Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 2.3.13 (Linearity of the product), Proposition 1.6.7 (Standard bases)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ spans F^n. What does Proposition 2.1.16(a) say about the image of T_A?
Hint 2. $A\mathbf{e}_j = \mathbf{a}_j$ by Proposition 2.3.13(c).
Sketch. Since $\mathbf{e}_1, \dots, \mathbf{e}_n$ spans F^n, $\operatorname{im} T_A = \operatorname{span}(A\mathbf{e}_1, \dots, A\mathbf{e}_n) = \operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$.

The standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ spans F^n (Proposition 1.6.7). By Proposition 2.1.16(a) and Proposition 2.3.13(c),

$$\operatorname{im} T_A = \operatorname{span}(A\mathbf{e}_1, \dots, A\mathbf{e}_n) = \operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_n).$$ □

The first proof is a reading of the definition of the product. The second uses nothing about matrices except $A\mathbf{e}_j = \mathbf{a}_j$, so it proves more: for every $T \in \mathrm{L}(V, W)$ and every list $\mathcal{L}$ that spans V,

$$\operatorname{im} T = \operatorname{span} T(\mathcal{L}). \tag{2.6.1}$$

That is, the image of a linear map is spanned by the values on any spanning list. Section 2.6.5 builds on this.

Example 2.6.19 (Computing a null space and a column space).

(a) Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix}$, over any field F.

Null space. A vector $\mathbf{x}$ lies in $\ker T_A$ if and only if $x_1 + 2x_2 + 3x_3 = 0$ and $2x_1 + 4x_2 + 7x_3 = 0$. If $\mathbf{x}$ satisfies both, then subtracting twice the first equation from the second gives $x_3 = 0$, and then the first gives $x_1 = -2x_2$. Conversely, $(-2t, t, 0)$ satisfies both equations for every $t \in F$. So $\ker T_A = \operatorname{span}((-2, 1, 0))$.

Column space. By Proposition 2.6.18, $\operatorname{im} T_A = \operatorname{span}((1, 2), (2, 4), (3, 7))$. The second column is twice the first, so it can be dropped without changing the span (Corollary 1.3.20(c)). The remaining two columns span all of F^2, since

$$7(1, 2) - 2(3, 7) = (1, 0), \quad -3(1, 2) + (3, 7) = (0, 1),$$

and a subspace containing $\mathbf{e}_1$ and $\mathbf{e}_2$ contains $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) = F^2$ (Proposition 1.4.14). So T_A is surjective but not injective: every system $A\mathbf{x} = \mathbf{y}$ has solutions, and Section 2.6.7 shows that the solutions of each one form a translate of the line $\ker T_A$.

(b) Let $A = \begin{pmatrix} 1 & 2i \\ i & -2 \end{pmatrix} \in \mathbb{C}^{2\times 2}$. Its columns are $\mathbf{a}_1 = (1, i)$ and $\mathbf{a}_2 = (2i, -2) = 2i\mathbf{a}_1$, so $\operatorname{im} T_A = \operatorname{span}(\mathbf{a}_1, 2i\mathbf{a}_1) = \operatorname{span}((1, i))$, a line in $\mathbb{C}^2$. Also, $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = (x_1 + 2ix_2)\mathbf{a}_1$, which is $\mathbf{0}$ if and only if $x_1 + 2ix_2 = 0$, since $\mathbf{a}_1 \neq \mathbf{0}$ (Proposition 1.2.35). So $\ker T_A = \{(-2it, t) \mid t \in \mathbb{C}\} = \operatorname{span}((-2i, 1))$, and T_A is neither injective nor surjective.

For larger matrices both computations are done by Gauss–Jordan elimination. Theorem 3.4.9 writes every vector of the null space as a linear combination of the *basic solutions* of A, one for each free variable, and Section 4.2 reads off from the reduced matrix which columns of A suffice to span the column space.

In the span explorer below, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are the columns of

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

so their span is the column space of A, the plane $\{(a,b,c) \mid c = a+b\}$ in $\mathbb{R}^3$. The explorer finds the linear relation $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$, which says that $(1,1,-1) \in \ker T_A$. For a target $\mathbf{w}$ in the plane it shows coefficients x_1, x_2, x_3 with $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{w}$, and (x_1, x_2, x_3) is a solution of $A\mathbf{x} = \mathbf{w}$. Move the target off the plane and the system becomes inconsistent. Drag $\mathbf{v}_3$ off the plane and the columns span $\mathbb{R}^3$: then T_A is surjective, and the explorer finds no relation, so T_A is also injective (Corollary 2.6.24).

$$\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}, \quad (1,-1,0) = \mathbf{v}_1 - \mathbf{v}_2,$$
$$\operatorname{im} T_A = \{(a,b,c) \in \mathbb{R}^3 \mid c = a+b\}.$$

The columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of A span the plane $c = a + b$, the column space of A. The relation $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$ gives the vector $(1,1,-1)$ of $\ker T_A$, and the target $\mathbf{w} = (1,-1,0) = \mathbf{v}_1 - \mathbf{v}_2$ lies in the column space, so $A\mathbf{x} = \mathbf{w}$ is consistent. (Interactive in the web version.)

Kernels and images through a matrix

By Proposition 2.3.16, $[T_A]_{\mathcal{E}\mathcal{E}} = A$, so the null space and the column space of A are the kernel and the image of the linear map whose matrix in the standard bases is A. In the other direction, the kernel and the image of any linear map between spaces with bases can be computed from any matrix of the map.

Proposition 2.6.20 (Kernel and image in coordinates). Let $\mathcal{B}$ be a basis of V of length n and $\mathcal{C}$ a basis of W of length m, let $T \in \mathrm{L}(V,W)$, and let $A = [T]_{\mathcal{C}\mathcal{B}}$.

(a) For $\mathbf{v} \in V$: $\mathbf{v} \in \ker T$ if and only if $[\mathbf{v}]_{\mathcal{B}} \in \ker T_A$.
(b) For $\mathbf{w} \in W$: $\mathbf{w} \in \operatorname{im} T$ if and only if $[\mathbf{w}]_{\mathcal{C}} \in \operatorname{im} T_A$.
(c) T is injective if and only if T_A is, and T is surjective if and only if T_A is.

In particular, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ maps $\ker T$ bijectively onto $\ker T_A$, and $\mathbf{w} \mapsto [\mathbf{w}]_{\mathcal{C}}$ maps $\operatorname{im} T$ bijectively onto $\operatorname{im} T_A$.

Proof (by chasing coordinates). STANDARD

Uses. Theorem 2.3.9 (The matrix of a map computes the map), Proposition 1.6.16 (The coordinate map is a bijection), Proposition 1.6.19 (Coordinates respect linear combinations), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. The bridge between T and T_A is $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$ (Theorem 2.3.9).
Hint 2. The coordinate maps are bijections that send $\mathbf{0}$ to $\mathbf{0}$. So $[\mathbf{w}]_{\mathcal{C}} = \mathbf{0}$ if and only if $\mathbf{w} = \mathbf{0}$, and every $\mathbf{x} \in F^n$ is $[\mathbf{v}]_{\mathcal{B}}$ for some $\mathbf{v}$.
Sketch. (a) $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$, which is $\mathbf{0}$ iff $T(\mathbf{v}) = \mathbf{0}$. (b) If $\mathbf{w} = T(\mathbf{v})$, then $[\mathbf{w}]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$. If $[\mathbf{w}]_{\mathcal{C}} = A\mathbf{x}$, write $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$; then $[\mathbf{w}]_{\mathcal{C}} = [T(\mathbf{v})]_{\mathcal{C}}$, so $\mathbf{w} = T(\mathbf{v})$. The bijections follow because the coordinate maps are bijections, and (c) follows from the bijections, which match $\{\mathbf{0}\}$ with $\{\mathbf{0}\}$ and W with F^m.

By Theorem 2.3.9, $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$ for every $\mathbf{v} \in V$. The coordinate maps $V \to F^n$ and $W \to F^m$ are bijections (Proposition 1.6.16) that send $\mathbf{0}$ to $\mathbf{0}$ (Proposition 1.6.19).

(a) $[\mathbf{v}]_{\mathcal{B}} \in \ker T_A$ means $[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$. Since the coordinate map of W is injective and $[\mathbf{0}]_{\mathcal{C}} = \mathbf{0}$, this holds if and only if $T(\mathbf{v}) = \mathbf{0}$.

(b) If $\mathbf{w} = T(\mathbf{v})$, then $[\mathbf{w}]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}} \in \operatorname{im} T_A$. Conversely, suppose that $[\mathbf{w}]_{\mathcal{C}} = A\mathbf{x}$ for some $\mathbf{x} \in F^n$. Since the coordinate map of V is surjective, $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ for some $\mathbf{v} \in V$, and then $[\mathbf{w}]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$. Since the coordinate map of W is injective, $\mathbf{w} = T(\mathbf{v}) \in \operatorname{im} T$.

The bijections. By (a), the coordinate map of V, which is injective, sends $\ker T$ into $\ker T_A$. Every $\mathbf{x} \in \ker T_A$ is $[\mathbf{v}]_{\mathcal{B}}$ for some $\mathbf{v} \in V$, and this $\mathbf{v}$ lies in $\ker T$ by (a). So $\ker T \to \ker T_A$, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$, is a bijection. The same argument with (b) and the coordinate map of W shows that $\operatorname{im} T \to \operatorname{im} T_A$, $\mathbf{w} \mapsto [\mathbf{w}]_{\mathcal{C}}$, is a bijection.

(c) The bijection $\ker T \to \ker T_A$ sends $\mathbf{0}$ to $\mathbf{0}$, so $\ker T = \{\mathbf{0}\}$ if and only if $\ker T_A = \{\mathbf{0}\}$. By Theorem 2.6.11, T is injective if and only if T_A is. If $\operatorname{im} T = W$, then $\operatorname{im} T_A$ is the image of all of W under the coordinate map

of W, which is F^m. Conversely, if $\operatorname{im} T_A = F^m$, then $[\mathbf{w}]_{\mathcal{C}} \in \operatorname{im} T_A$ for every $\mathbf{w} \in W$, so every $\mathbf{w}$ lies in $\operatorname{im} T$ by (b). □

Proof (through the square of maps). STANDARD

> *Uses.* Theorem 2.3.9 (The matrix of a map computes the map), Example 2.5.14 (Invertible maps and matrices), Proposition 2.6.15 (Kernels and images of composites), Proposition 2.1.17 (Compositions of linear maps are linear)
> *Techniques.* Translate between maps and matrices (Technique 2.3.27)
> *Hint 1.* Write $C_{\mathcal{B}} : V \to F^n$ and $C_{\mathcal{C}} : W \to F^m$ for the coordinate maps. Rewrite Theorem 2.3.9 as an equation between composites.
> *Hint 2.* $C_{\mathcal{C}} T = T_A C_{\mathcal{B}}$, and both coordinate maps are invertible. Apply Proposition 2.6.15 to each side, using that $C_{\mathcal{C}}$ is injective and $C_{\mathcal{B}}$ is surjective.
> *Sketch.* With $C_{\mathcal{C}} T = T_A C_{\mathcal{B}}$: since $C_{\mathcal{C}}$ is injective, $\ker T = \ker(C_{\mathcal{C}} T) = \ker(T_A C_{\mathcal{B}})$, which is (a). Since $C_{\mathcal{B}}$ is surjective, $\operatorname{im}(C_{\mathcal{C}}|_{\operatorname{im} T}) = \operatorname{im}(C_{\mathcal{C}} T) = \operatorname{im}(T_A C_{\mathcal{B}}) = \operatorname{im} T_A$, which with injectivity of $C_{\mathcal{C}}$ is (b). For (c), $T_A = C_{\mathcal{C}} T C_{\mathcal{B}}^{-1}$ and $T = C_{\mathcal{C}}^{-1} T_A C_{\mathcal{B}}$, and composites of injective (surjective) maps are injective (surjective).

Let $C_{\mathcal{B}} : V \to F^n$ and $C_{\mathcal{C}} : W \to F^m$ be the coordinate maps. They are linear and invertible (Example 2.5.14(c)), hence bijective, and Theorem 2.3.9 says that

$$C_{\mathcal{C}} \circ T = T_A \circ C_{\mathcal{B}},$$

that is, the two routes around the square of Figure 2.3.1 agree.

(a) Since $C_{\mathcal{C}}$ is injective, Proposition 2.6.15(a) gives $\ker T = \ker(C_{\mathcal{C}} T) = \ker(T_A C_{\mathcal{B}})$. A vector $\mathbf{v}$ lies in $\ker(T_A C_{\mathcal{B}})$ exactly when $C_{\mathcal{B}}(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ lies in $\ker T_A$.

(b) Since $C_{\mathcal{B}}$ is surjective, Proposition 2.6.15(b) gives

$$\operatorname{im}(C_{\mathcal{C}}|_{\operatorname{im} T}) = \operatorname{im}(C_{\mathcal{C}} T) = \operatorname{im}(T_A C_{\mathcal{B}}) = \operatorname{im} T_A.$$

So $C_{\mathcal{C}}$ maps $\operatorname{im} T$ onto $\operatorname{im} T_A$, and it is injective. In particular, if $[\mathbf{w}]_{\mathcal{C}} \in \operatorname{im} T_A$, then $[\mathbf{w}]_{\mathcal{C}} = [\mathbf{w}']_{\mathcal{C}}$ for some $\mathbf{w}' \in \operatorname{im} T$, and $\mathbf{w} = \mathbf{w}' \in \operatorname{im} T$. The converse is immediate.

(c) By Proposition 2.6.15(a) and (b), a composite of injective linear maps is injective (its kernel is the kernel of the first map) and a composite of surjective linear maps is surjective (its image is the image of the last map). Now $T_A = C_{\mathcal{C}} T C_{\mathcal{B}}^{-1}$ and $T = C_{\mathcal{C}}^{-1} T_A C_{\mathcal{B}}$, and the coordinate maps and their inverses are bijective. So if T is injective or surjective, so is T_A, and conversely. □

The first proof keeps track of individual vectors. The second works with whole maps: it records $[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$ as the equation $C_{\mathcal{C}} T = T_A C_{\mathcal{B}}$ between linear maps and lets Proposition 2.6.15 do the rest. Any property that is preserved by composing with invertible maps on both sides transfers between T and T_A in this way, and Section 4.5 uses the same idea for change of basis.

Example 2.6.21 (Differentiation in coordinates). Let $D : F[x]_{\le 3} \to F[x]_{\le 2}$ be differentiation. By Example 2.3.4, its matrix in the bases $1, x, x^2, x^3$ and $1, x, x^2$ is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad \text{so} \quad A\mathbf{x} = (x_2, 2x_3, 3x_4).$$

If $2 \neq 0$ and $3 \neq 0$ in F (for instance if F has characteristic 0), then $A\mathbf{x} = \mathbf{0}$ if and only if $x_2 = x_3 = x_4 = 0$, so $\ker T_A = \operatorname{span}(\mathbf{e}_1)$. The columns $\mathbf{0}, \mathbf{e}_1, 2\mathbf{e}_2, 3\mathbf{e}_3$ span F^3. By Proposition 2.6.20, $\ker D$ consists of the polynomials with coordinates in $\operatorname{span}(\mathbf{e}_1)$, the constants, and D is surjective.

Over $\mathbb{F}_3$, the entry 3 is 0, so $A\mathbf{x} = (x_2, 2x_3, 0)$. Now $\ker T_A = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_4)$ and $\operatorname{im} T_A = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$. Translating back, $\ker D = \operatorname{span}(1, x^3)$ and $\operatorname{im} D = \operatorname{span}(1, x) = \mathbb{F}_3[x]_{\le 1}$. Indeed $D(x^3) = 3x^2 = 0$, and no polynomial of degree at most 3 over $\mathbb{F}_3$ has derivative x^2.

2.6.5 Testing on a basis

Intuition. By Theorem 2.2.1, a linear map is determined by its values on a basis $\mathcal{B}$ of V, and those values can be chosen freely. So every property of T must be visible in the list $T(\mathcal{B})$. Injectivity and surjectivity are visible in the simplest possible way. They become linear independence and spanning, the two halves of "basis" (Proposition 1.6.2), carried over to W.

Proposition 2.6.22 (Testing injectivity and surjectivity on a basis). Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V, let $T \in \mathrm{L}(V, W)$, and write $T(\mathcal{B}) = T(\mathbf{v}_1), ..., T(\mathbf{v}_n)$.

(a) T is injective if and only if $T(\mathcal{B})$ is linearly independent.
(b) T is surjective if and only if $T(\mathcal{B})$ spans W.
(c) T is bijective if and only if $T(\mathcal{B})$ is a basis of W.
(d) In every case $T(\mathcal{B})$ spans $\operatorname{im} T$, and it is a basis of $\operatorname{im} T$ if and only if T is injective.

Proof. ROUTINE

Uses. Proposition 2.6.14 (Injective maps preserve independence), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Theorem 1.5.5 (Independence is decided at zero), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 1.6.2 (Bases are independent spanning lists), Definition 2.6.3 (Bijective)
Techniques. Test injectivity at $\mathbf{0}$ (Technique 2.6.12), Check it on a basis (Technique 2.2.7)
Hint 1. For (a), one direction is Proposition 2.6.14. For the other, take $\mathbf{v} \in \ker T$, expand it in the basis $\mathcal{B}$, and apply T.
Hint 2. For (b) and (d), express $\operatorname{im} T$ in terms of $T(\mathcal{B})$ with Proposition 2.1.16(a).
Hint 3. (c) is (a) and (b) together.
Sketch. (a) An injective map preserves independence. Conversely, if $\mathbf{v} = \sum_i a_i \mathbf{v}_i \in \ker T$, then $\sum_i a_i T(\mathbf{v}_i) = \mathbf{0}$, so all $a_i = 0$ and $\mathbf{v} = \mathbf{0}$. (b) $\operatorname{im} T = \operatorname{span} T(\mathcal{B})$ since $\mathcal{B}$ spans V, and this equals W iff $T(\mathcal{B})$ spans W. (c) Combine (a) and (b). (d) A list spanning $\operatorname{im} T$ is a basis of it iff it is independent, iff T is injective by (a).

(a) If T is injective, then $T(\mathcal{B})$ is linearly independent by Proposition 2.6.14, since $\mathcal{B}$ is. Conversely, suppose that $T(\mathcal{B})$ is linearly independent, and let $\mathbf{v} \in \ker T$. Since $\mathcal{B}$ spans V, $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ for some scalars a_i, and by linearity

$$\mathbf{0} = T(\mathbf{v}) = a_1 T(\mathbf{v}_1) + \cdots + a_n T(\mathbf{v}_n).$$

As $T(\mathcal{B})$ is linearly independent, $a_1 = \cdots = a_n = 0$ (Theorem 1.5.5), so $\mathbf{v} = \mathbf{0}$. Hence $\ker T = \{\mathbf{0}\}$, and T is injective by Theorem 2.6.11.

(b) Since $\mathcal{B}$ spans V, Proposition 2.1.16(a) gives

$$\operatorname{im} T = \operatorname{span} T(\mathcal{B}).$$

So $\operatorname{im} T = W$, which is surjectivity, holds if and only if $\operatorname{span} T(\mathcal{B}) = W$.

(c) By Definition 2.6.3 and Proposition 1.6.2, (c) is (a) and (b) together.

(d) The first claim was shown in (b). The list $T(\mathcal{B})$ lies in $\operatorname{im} T$ and spans it, so it is a basis of $\operatorname{im} T$ if and only if it is linearly independent (Proposition 1.6.2), which by (a) happens if and only if T is injective. □

So whether T is injective, and whether it is surjective, can be decided on one basis of V: a list of n vectors instead of all of V. This is Technique 2.2.7 in a new form. There, the property to be checked was an equation between linear maps; here it is a property of the list of values.

Remark 2.6.23 (Which half of "basis" is used where). The proof of (a) used both halves of "basis": independence of $\mathcal{B}$ in the "only if" direction (through Proposition 2.6.14), and spanning in the "if" direction, to expand $\mathbf{v}$. Part (b) used only that $\mathcal{B}$ spans V, so it holds for every spanning list: if $\mathcal{L}$ spans V, then T is surjective if and only if $T(\mathcal{L})$ spans W. The "if" direction of (a) fails for lists that do not span. The map $P(x, y) = x$ sends the independent list $\mathbf{e}_1$ to the independent list 1, yet P is not injective. Compare Remark 2.2.2, where the two halves played similar roles in constructing maps.

For a matrix, the values on the standard basis are the columns.

Corollary 2.6.24 (The columns decide). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$.
(a) T_A is injective if and only if $\mathbf{a}_1, \dots, \mathbf{a}_n$ is linearly independent.
(b) T_A is surjective if and only if $\mathbf{a}_1, \dots, \mathbf{a}_n$ spans F^m.
(c) T_A is bijective if and only if $\mathbf{a}_1, \dots, \mathbf{a}_n$ is a basis of F^m.

Proof. ROUTINE

Uses. Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Proposition 1.6.7 (Standard bases), Proposition 2.3.13 (Linearity of the product)
Hint 1. Apply Proposition 2.6.22 to T_A and a basis of F^n whose values you know.
Hint 2. The standard basis: $T_A(\mathbf{e}_j) = A\mathbf{e}_j = \mathbf{a}_j$.
Sketch. $\mathbf{e}_1, \dots, \mathbf{e}_n$ is a basis of F^n and $T_A(\mathbf{e}_j) = \mathbf{a}_j$, so $T_A(\mathcal{E})$ is the list of columns. Apply Proposition 2.6.22(a)–(c).

The standard basis $\mathcal{E} = \mathbf{e}_1, \dots, \mathbf{e}_n$ is a basis of F^n (Proposition 1.6.7), and $T_A(\mathbf{e}_j) = A\mathbf{e}_j = \mathbf{a}_j$ by Proposition 2.3.13(c). So $T_A(\mathcal{E})$ is the list of columns of A, and (a)–(c) are parts (a)–(c) of Proposition 2.6.22. □

Part (a) can also be read off from the definition of the product. Since $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$, a nonzero vector of $\ker T_A$ is the same thing as a nontrivial linear relation among the columns, so $\ker T_A = \{\mathbf{0}\}$ says exactly that the columns are linearly independent (Theorem 1.5.5). The three parts of the corollary are the equivalences (I) ⇔ (III) in the Fundamental Theorems of Section 2.7. Combined with Proposition 2.6.20, they decide injectivity and surjectivity of any map between spaces with bases from the columns of one of its matrices.

Example 2.6.25 (Evaluating at three points). Let $E : F[x]_{\le 2} \to F^3$, $E(p) = (p(0), p(1), p(2))$, where $2 = 1 + 1$. Each entry of E is an evaluation map, so E is linear (Example 2.1.9). On the basis $1, x, x^2$,

$$E(1) = (1, 1, 1), \quad E(x) = (0, 1, 2), \quad E\big(x^2\big) = (0, 1, 4).$$

Suppose $a(1,1,1) + b(0,1,2) + c(0,1,4) = \mathbf{0}$. The first entry gives $a = 0$, then $b + c = 0$ and $2b + 4c = 0$, so $2c = (2b + 4c) - 2(b + c) = 0$.

If $2 \neq 0$ in F, then $c = 0$ and $b = 0$, so the three values are linearly independent and E is injective by Proposition 2.6.22(a). They also span F^3, because

$$\mathbf{e}_3 = 2^{-1}\big(E\big(x^2\big) - E(x)\big), \quad \mathbf{e}_2 = E(x) - 2\mathbf{e}_3, \quad \mathbf{e}_1 = E(1) - \mathbf{e}_2 - \mathbf{e}_3,$$

so E is bijective by Proposition 2.6.22(c): for every $(y_0, y_1, y_2) \in F^3$ there is exactly one $p \in F[x]_{\le 2}$ with $p(0) = y_0$, $p(1) = y_1$, and $p(2) = y_2$. This is a case of polynomial interpolation (Theorem 3.4.24).

Over $\mathbb{F}_2$, $2 = 4 = 0$, so $E(x) = E\big(x^2\big) = (0, 1, 0)$, and the values are linearly dependent. So E is not injective; indeed $E\big(x^2 + x\big) = (0, 1 + 1, 0) = \mathbf{0}$. Nor is it surjective: every value of E has equal first and third entries, because over $\mathbb{F}_2$ the points 0 and 2 coincide, so $(1, 0, 0) \notin \operatorname{im} E$.

Exercise 2.6.26. ROUTINE For $c \in F$, let $T_c : F[x]_{\le 2} \to F[x]_{\le 2}$, $T_c(p) = D(xp) - cp$, where D is differentiation.

(a) Compute $T_c(1)$, $T_c(x)$, and $T_c\big(x^2\big)$.

(b) Over $\mathbb{Q}$, find all c for which T_c is injective, and all c for which it is surjective.

(c) Answer the same questions over $\mathbb{F}_2$, over $\mathbb{F}_3$, and over $\mathbb{F}_5$.

Solution on page 1162

2.6.6 Computing kernels and images

Here are the kernels and images of the main examples of Section 2.1.

Example 2.6.27 (Differentiation on $F[x]$). Let $D : F[x] \to F[x]$ be differentiation. For $f = \sum_{k=0}^{n} a_k x^k$, Example 2.1.9 gives $D(f) = \sum_{k=1}^{n} (ka_k) x^{k-1}$, where $ka_k = (k * 1)a_k$, so the coefficient of x^{k-1} in $D(f)$ is $(k * 1)a_k$.

Kernel. $D(f) = \mathbf{0}$ if and only if $(k * 1)a_k = 0$ for every $k \geq 1$. If F has characteristic 0 (Definition 1.1.28), then $k * 1 \neq 0$ for $k \geq 1$, so this forces $a_k = 0$ for all $k \geq 1$ (Proposition 1.1.17). Hence $\ker D = \operatorname{span}(1)$, the constant polynomials. Over $\mathbb{F}_p$, $k * 1 = 0$ exactly when p divides k, so $D(f) = \mathbf{0}$ if and only if $a_k = 0$ for every k not divisible by p. Hence $\ker D$ consists of the polynomials $b_0 + b_1 x^p + b_2 x^{2p} + \cdots$, those in which only powers of x^p occur.

Image. In characteristic 0, $DJ = \mathrm{id}$ for the antiderivative J of Example 2.5.24, so D is surjective by Proposition 2.5.5. Over $\mathbb{F}_p$ it is not. The coefficient of x^{k-1} in $D(f)$ is $(k * 1)a_k$, which is 0 whenever p divides k, so every value of D has coefficient 0 at $x^{p-1}, x^{2p-1}, x^{3p-1}, \ldots$. Conversely, if $q = \sum_j c_j x^j$ has $c_j = 0$ whenever p divides $j + 1$, then

$$q = D\left(\sum_{p \nmid j+1} ((j+1) * 1)^{-1} c_j x^{j+1} \right).$$

So $\operatorname{im} D$ consists of the polynomials whose coefficients at $x^{p-1}, x^{2p-1}, \ldots$ vanish. In particular $x^p \in \ker D$ and $x^{p-1} \notin \operatorname{im} D$, as in Exercise 2.5.42.

Example 2.6.28 (The trace). For $n \geq 1$, the trace $\operatorname{tr} : F^{n \times n} \to F$ is linear (Proposition 2.4.27). Write $U_{ij} = \mathbf{e}_i \mathbf{e}_j^{\mathsf{T}}$ for the matrix units, which form a basis of $F^{n \times n}$ (Lemma 2.4.24). Then $\operatorname{tr} U_{ij}$ is 1 if $i = j$ and 0 otherwise.

Image. The values $\operatorname{tr} U_{ij}$ include 1, which spans F, so tr is surjective by Proposition 2.6.22(b). Explicitly, $c = \operatorname{tr}(cU_{11})$.

Kernel. $\ker \operatorname{tr}$ is the subspace of **trace-zero** matrices, $\{A \in F^{n \times n} \mid a_{11} + \cdots + a_{nn} = 0\}$. For $n \geq 2$ it contains $U_{12} \neq 0$, so tr is not injective. To find a spanning list, let $\operatorname{tr} A = 0$. Then $a_{nn} = -\big(a_{11} + \cdots + a_{n-1,n-1}\big)$, and expanding A in matrix units gives

$$A = \sum_{i \neq j} a_{ij} U_{ij} + \sum_{i=1}^{n-1} a_{ii} (U_{ii} - U_{nn}).$$

So $\ker \operatorname{tr}$ is spanned by the $n^2 - n$ matrices U_{ij} with $i \neq j$ and the $n - 1$ matrices $U_{ii} - U_{nn}$, all of which have trace 0. By Proposition 2.4.28, every commutator $AB - BA$ lies in $\ker \operatorname{tr}$, which is why $AB - BA = I_n$ is impossible when $n * 1 \neq 0$ (Proposition 2.4.30).

Example 2.6.29 (Evaluation at 0, and multiplication by x). Let $\mathrm{ev}_0 : F[x] \to F$, $\mathrm{ev}_0(p) = p(0)$, and let $M : F[x] \to F[x]$, $M(p) = xp$. For $p = a_0 + a_1 x + \cdots + a_n x^n$, $\mathrm{ev}_0(p) = a_0$. So ev_0 is surjective, since $c = \mathrm{ev}_0(c)$ for each constant c, and its kernel is the set of polynomials with constant term 0.

That kernel is exactly $\operatorname{im} M$. Every xq has constant term 0, and conversely $a_1 x + a_2 x^2 + \cdots + a_n x^n = M\big(a_1 + a_2 x + \cdots + a_n x^{n-1}\big)$. So in

$$F[x] \xrightarrow{M} F[x] \xrightarrow{\mathrm{ev}_0} F,$$

M is injective (Example 2.6.4(e)), ev_0 is surjective, and the vectors that the second map sends to 0 are precisely the values of the first.

For $t \in F$, the kernel of $\mathrm{ev}_t(p) = p(t)$ consists of the polynomials with root t. It contains every value of $M - t\,\mathrm{id}$: if $q = \sum_k b_k x^k$, then $(xq)(t) = \sum_k b_k t^{k+1} = tq(t)$, so $(xq - tq)(t) = tq(t) - tq(t) = 0$ by Proposition 1.2.18(c). In fact $\ker \mathrm{ev}_t = \operatorname{im}(M - t\,\mathrm{id})$: this is the factor theorem, a consequence of division with remainder (Theorem 7.1.14).

2.6.7 Preimages and the equation $T(\mathbf{v}) = \mathbf{w}$

The kernel describes the solutions of $T(\mathbf{v}) = \mathbf{0}$, and the image records which equations $T(\mathbf{v}) = \mathbf{w}$ have solutions at all. Together they describe the solutions of every such equation.

Notation 2.6.30 (Preimages). For $Y \subseteq W$, the **preimage** of Y under T is

$$T^{-1}(Y) = \{\mathbf{v} \in V \mid T(\mathbf{v}) \in Y\}.$$

For $\mathbf{w} \in W$, we write $T^{-1}(\mathbf{w})$ for $T^{-1}(\{\mathbf{w}\})$, the set of solutions of $T(\mathbf{v}) = \mathbf{w}$, also called the **fiber** of T over $\mathbf{w}$.

The notation does not presuppose that T is invertible: $T^{-1}(\mathbf{w})$ is a set, possibly empty. If T happens to be invertible, it is the one-element set $\{T^{-1}(\mathbf{w})\}$, and we will never need to tell the two apart. In this notation $\ker T = T^{-1}(\mathbf{0})$, and T is injective if and only if every fiber has at most one element, surjective if and only if every fiber is nonempty.

Corollary 2.6.31 (Fibers are translates of the kernel). Let $\mathbf{w} \in W$.

(a) If $\mathbf{w} \notin \operatorname{im} T$, then $T^{-1}(\mathbf{w}) = \emptyset$.

(b) If $\mathbf{w} \in \operatorname{im} T$ and $\mathbf{v}_0 \in T^{-1}(\mathbf{w})$, then $T^{-1}(\mathbf{w}) = \mathbf{v}_0 + \ker T$, and $\mathbf{z} \mapsto \mathbf{v}_0 + \mathbf{z}$ is a bijection from $\ker T$ onto $T^{-1}(\mathbf{w})$.

(c) Every $\mathbf{v} \in V$ lies in exactly one fiber, namely $T^{-1}(T(\mathbf{v}))$. So V is the disjoint union of the fibers $T^{-1}(\mathbf{w})$ with $\mathbf{w} \in \operatorname{im} T$, each of which is a translate of $\ker T$.

Proof. ROUTINE

Uses. Definition 2.6.6 (Image), Definition 2.6.5 (Kernel), Notation 3.1.17 (Translates), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Particular plus homogeneous (Technique 3.1.18), Double containment (Technique 1.3.14)
Hint 1. (a) is the definition of the image. For (b), prove the two containments: apply T to $\mathbf{v}_0 + \mathbf{z}$ with $\mathbf{z} \in \ker T$, and for a solution $\mathbf{v}$, look at $\mathbf{v} - \mathbf{v}_0$.
Hint 2. For (c), $\mathbf{v} \in T^{-1}(\mathbf{w})$ means $T(\mathbf{v}) = \mathbf{w}$. How many $\mathbf{w}$ satisfy this for a fixed $\mathbf{v}$?
Sketch. (a) No $\mathbf{v}$ has $T(\mathbf{v}) = \mathbf{w}$. (b) $T(\mathbf{v}_0 + \mathbf{z}) = \mathbf{w} + \mathbf{0}$ for $\mathbf{z} \in \ker T$, and if $T(\mathbf{v}) = \mathbf{w}$, then $\mathbf{v} - \mathbf{v}_0 \in \ker T$; this is Proposition 3.1.19. The map $\mathbf{z} \mapsto \mathbf{v}_0 + \mathbf{z}$ is injective by cancellation. (c) $\mathbf{v}$ lies in $T^{-1}(\mathbf{w})$ exactly when $\mathbf{w} = T(\mathbf{v})$.

(a) If $\mathbf{w} \notin \operatorname{im} T$, then no $\mathbf{v} \in V$ satisfies $T(\mathbf{v}) = \mathbf{w}$ (Definition 2.6.6).

(b) This is Proposition 3.1.19; we repeat the argument. If $\mathbf{z} \in \ker T$, then $T(\mathbf{v}_0 + \mathbf{z}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$, so $\mathbf{v}_0 + \ker T \subseteq T^{-1}(\mathbf{w})$. If $T(\mathbf{v}) = \mathbf{w}$, then $T(\mathbf{v} - \mathbf{v}_0) = \mathbf{w} - \mathbf{w} = \mathbf{0}$, so $\mathbf{v} = \mathbf{v}_0 + (\mathbf{v} - \mathbf{v}_0) \in \mathbf{v}_0 + \ker T$. Hence $T^{-1}(\mathbf{w}) = \mathbf{v}_0 + \ker T$, which says that $\mathbf{z} \mapsto \mathbf{v}_0 + \mathbf{z}$ maps $\ker T$ onto the fiber. It is injective, because $\mathbf{v}_0 + \mathbf{z} = \mathbf{v}_0 + \mathbf{z}'$ implies $\mathbf{z} = \mathbf{z}'$.

(c) $\mathbf{v} \in T^{-1}(\mathbf{w})$ means $T(\mathbf{v}) = \mathbf{w}$, which holds for exactly one $\mathbf{w}$, namely $\mathbf{w} = T(\mathbf{v}) \in \operatorname{im} T$. So the fibers over the vectors of $\operatorname{im} T$ are disjoint and cover V. By (b), each of them is a translate of $\ker T$. □

So the domain is sliced into parallel copies of $\ker T$, one for each vector of $\operatorname{im} T$. The map T collapses each slice to a single vector, and different slices go to different vectors. In this picture, Theorem 2.6.11 says that T is injective exactly when the slices are single points. Figure 2.6.1 shows the slices for

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad T_A(\mathbf{x}) = (x_1 + 2x_2)\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Here $\ker T_A$ is the line $x_1 + 2x_2 = 0$, the column space is $\operatorname{im} T_A = \operatorname{span}((1,1))$, and the fiber over $t(1,1)$ is the line $x_1 + 2x_2 = t$, parallel to the kernel.

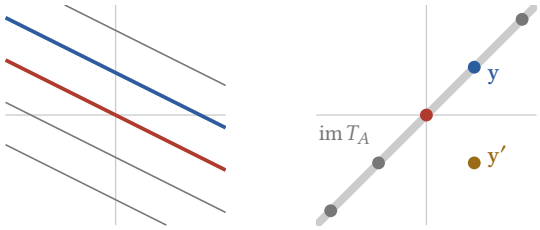


Figure 2.6.1. The fibers of T_A for $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ over $\mathbb{R}$. Left: $\ker T_A$ (red) and the fibers $x_1 + 2x_2 = t$ for $t = -4, -2, 2, 4$, parallel translates of the kernel. The blue one is the fiber over $\mathbf{y} = (2, 2)$, the solution set of $A\mathbf{x} = \mathbf{y}$. Right: the column space $\operatorname{im} T_A = \operatorname{span}((1, 1))$ (grey), with the vectors $t(1, 1)$ onto which the fibers collapse. The vector $\mathbf{y}' = (2, -2)$ is not in $\operatorname{im} T_A$, so its fiber is empty.

The row and column pictures in Section 3.4.4 show the same phenomenon interactively for a single fiber: there $\mathbf{y}$ can be dragged along the column space, where the solutions form a line parallel to the kernel, and off it, where there are none.

Remark 2.6.32 (What the slicing leads to). The slicing suggests a count: V consists of as many copies of $\ker T$ as there are vectors in $\operatorname{im} T$. Over a finite field this is literally true, $|V| = |\ker T| \cdot |\operatorname{im} T|$ (Exercise 2.6.44). For finite-dimensional spaces over any field, the count becomes the dimension formula $\dim V = \dim \ker T + \dim \operatorname{im} T$, which is rank–nullity (Theorem 4.2.11). Treating each slice as a single object leads to the quotient space $V/\ker T$, and the bijection between the slices and $\operatorname{im} T$ becomes the first isomorphism theorem (Theorem 4.6.34).

Exercise 2.6.33. ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix}$ as in Example 2.6.19(a). Find one solution of $A\mathbf{x} = (1, 3)$ by inspection, and then describe all of them.

Solution on page 1162

Exercises

Exercise 2.6.34. ROUTINE Let $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (x + y, y + z, x - z)$.

(a) Find $\ker T$, and describe $\operatorname{im} T$ by a single linear equation.

(b) Decide whether $(1, 2, 3)$ and $(2, 1, 1)$ lie in $\operatorname{im} T$. For the one that does, find all of its preimages.

Solution on page 1163

Exercise 2.6.35. ROUTINE Let $A = \begin{pmatrix} 1 & i & 1+i \\ i & -1 & i-1 \end{pmatrix} \in \mathbb{C}^{2\times 3}$.

(a) Show that every column of A is a multiple of $(1, i)$, and find $\operatorname{im} T_A$.

(b) Find $\ker T_A$ as the span of two vectors.

(c) Is T_A injective? Is it surjective? Find all solutions of $A\mathbf{x} = (1 + i, i - 1)$.

Solution on page 1163

Exercise 2.6.36. ROUTINE Let $S, T \in \mathrm{L}(V, W)$.

(a) Show that $\ker S \cap \ker T \subseteq \ker(S + T)$.

(b) Show that $\operatorname{im}(S + T) \subseteq \operatorname{im} S + \operatorname{im} T$, the sum of subspaces from Proposition 1.4.24.

(c) Show that both inclusions can be strict.

Solution on page 1163

Exercise 2.6.37. STANDARD Prove or disprove each statement.

(a) If $S, T \in \mathrm{L}(V, W)$ satisfy $\ker S = \ker T$ and $\operatorname{im} S = \operatorname{im} T$, then $S = T$.

(b) For every $T \in \mathrm{L}(V)$, $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$.

(c) If $P \in \mathrm{L}(V)$ satisfies $P^2 = P$, then $\ker P \cap \operatorname{im} P = \{\mathbf{0}\}$.

(d) If $T \in \mathrm{L}(V, W)$ is injective and U is a subspace of V, then the restriction $T|_U : U \to W$ is injective.

(e) If $T \in \mathrm{L}(V, W)$ is not injective, then T is not surjective.

Solution on page 1164

Exercise 2.6.38. STANDARD

(a) Find $T \in \mathrm{L}(F^2)$ with $\ker T = \operatorname{im} T$.

(b) Find $T \in \mathrm{L}(F^4)$ with $\ker T = \operatorname{im} T$.

(c) Show that every $T \in \mathrm{L}(V)$ with $\ker T = \operatorname{im} T$ satisfies $T^2 = 0$, and that $T^2 = 0$ does not imply $\ker T = \operatorname{im} T$.

Solution on page 1164

Exercise 2.6.39. STANDARD Let $S \in \mathrm{L}(U, V)$ and $T \in \mathrm{L}(V, W)$.

(a) Show that $\ker(TS) = \ker S$ if and only if $\ker T \cap \operatorname{im} S = \{\mathbf{0}\}$.

(b) Show that $\operatorname{im}(TS) = \operatorname{im} T$ if and only if $\operatorname{im} S + \ker T = V$.

(c) Deduce that for $T \in \mathrm{L}(V)$, $\ker T^2 = \ker T$ if and only if $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$, and $\operatorname{im} T^2 = \operatorname{im} T$ if and only if $\ker T + \operatorname{im} T = V$.

Solution on page 1164

Exercise 2.6.40. ROUTINE Find the error in the following argument, show that its claim is false, and state a correct version.

Claim. Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of V and $T \in \mathrm{L}(V, W)$. If $T(\mathbf{v}_i) \neq \mathbf{0}$ for every i, then T is injective.

Proof. Let $\mathbf{v} \in \ker T$ and write $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Then $\mathbf{0} = T(\mathbf{v}) = a_1T(\mathbf{v}_1) + \cdots + a_nT(\mathbf{v}_n)$. None of the vectors $T(\mathbf{v}_i)$ is $\mathbf{0}$, so every coefficient a_i must be 0. Hence $\mathbf{v} = \mathbf{0}$, and T is injective by Theorem 2.6.11.

Solution on page 1165

Exercise 2.6.41. STANDARD

(a) Let $D \in \mathrm{L}(\mathbb{F}_2[x]_{\leq 3})$ be differentiation. Find $\ker D$ and $\operatorname{im} D$, and show that they are equal.

(b) Find $\ker D$ and $\operatorname{im} D$ for differentiation $D \in \mathrm{L}(\mathbb{F}_3[x]_{\leq 3})$. Are they equal?

(c) Show that $D^2 = 0$ on all of $\mathbb{F}_2[x]$.

Solution on page 1165

Exercise 2.6.42. STANDARD Let $T \in \mathrm{L}(V, W)$ be injective, and let $S \in \mathrm{L}(U, W)$.

(a) Suppose $\operatorname{im} S \subseteq \operatorname{im} T$. Show that there is exactly one linear map $R : U \to V$ with $TR = S$.

(b) Show that the hypothesis $\operatorname{im} S \subseteq \operatorname{im} T$ in (a) is necessary.

Solution on page 1166

Exercise 2.6.43. HARD Let $T \in \mathrm{L}(V,W)$ be surjective, and let $S \in \mathrm{L}(V,U)$.

(a) Suppose $\ker T \subseteq \ker S$. Show that there is exactly one linear map $R : W \to U$ with $RT = S$.

(b) Show that the hypothesis $\ker T \subseteq \ker S$ in (a) is necessary.

Solution on page 1166

Exercise 2.6.44. STANDARD Let p be a prime, let V and W be vector spaces over $\mathbb{F}_p$ with V finite, and let $T \in \mathrm{L}(V,W)$.

(a) Show that $|V| = |\ker T| \cdot |\operatorname{im} T|$.

(b) Deduce that if $T : \mathbb{F}_p^n \to \mathbb{F}_p^m$ is linear and injective, then $n \leq m$, and if it is linear and surjective, then $n \geq m$.

Solution on page 1167

Exercise 2.6.45. HARD Let p be a prime. Show that exactly $\left(p^3-1\right)\left(p^3-p\right)$ linear maps $\mathbb{F}_p^3 \to \mathbb{F}_p^2$ are surjective. (Every such map is T_A for exactly one $A \in \mathbb{F}_p^{2\times 3}$, by Theorem 2.3.20 and Proposition 2.3.16. Use Corollary 2.6.24 and Theorem 1.4.25.)

Solution on page 1167

Exercise 2.6.46. STANDARD Let $\Phi : F^{n\times n} \to F^{n\times n}$, $\Phi(A) = A - A^{\mathsf{T}}$.

(a) Show that Φ is linear and that $\ker \Phi$ is the subspace of symmetric matrices.

(b) Suppose $1+1 \neq 0$ in F. Show that $\operatorname{im} \Phi$ is the subspace of skew-symmetric matrices.

(c) Over $\mathbb{F}_2$, show that $\operatorname{im} \Phi$ consists of the symmetric matrices with all diagonal entries 0. Deduce that $\operatorname{im} \Phi$ is strictly smaller than the set of skew-symmetric matrices when $n \geq 1$.

Solution on page 1168

Exercise 2.6.47. HARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$.

(a) Show that if $\mathbf{x} \in \ker T_{I_n - BA}$, then $A\mathbf{x} \in \ker T_{I_m - AB}$, and that if $\mathbf{y} \in \ker T_{I_m - AB}$, then $B\mathbf{y} \in \ker T_{I_n - BA}$.

(b) Show that $\mathbf{x} \mapsto A\mathbf{x}$ is a bijection from $\ker T_{I_n - BA}$ onto $\ker T_{I_m - AB}$, with inverse $\mathbf{y} \mapsto B\mathbf{y}$.

(c) Deduce that $T_{I_n - BA}$ is injective if and only if $T_{I_m - AB}$ is.

Solution on page 1168

Exercise 2.6.48. ROUTINE Let $\ell = \operatorname{span}(\mathbf{u})$ be a line in $\mathbb{R}^2$, with $\mathbf{u} = (\cos\theta, \sin\theta)$, and let $\ell^\perp = \operatorname{span}((-\sin\theta, \cos\theta))$. Let Pr_ℓ be the projection onto ℓ and $S_\ell = 2\operatorname{Pr}_\ell - \mathrm{id}$ the reflection across ℓ (Example 2.1.11). Find the kernel and the image of each map.

(a) $\operatorname{Pr}_\ell - \mathrm{id}$.

(b) $S_\ell - \mathrm{id}$.

(c) $S_\ell + \mathrm{id}$.

(d) $R_\varphi - \mathrm{id}$, for a rotation angle φ that is not a multiple of 2π.

Solution on page 1169

Exercise 2.6.49. HARD Let $T \in \mathrm{L}(V)$ satisfy $T^3 = T$, and suppose $1+1 \neq 0$ in F.

(a) Show that every $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_+ + \mathbf{v}_-$ with $\mathbf{v}_0 \in \ker T$, $\mathbf{v}_+ \in \ker(T - \mathrm{id})$, and $\mathbf{v}_- \in \ker(T + \mathrm{id})$.

(b) Show that the decomposition in (a) is unique.

(c) Show that (a) fails over $\mathbb{F}_2$ for $T = T_A$, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Solution on page 1169

Exercise 2.6.50. CHALLENGE

(a) Let $S, T \in \mathrm{L}(V, W)$ both be nonzero. Show that some $\mathbf{v} \in V$ has $S(\mathbf{v}) \neq \mathbf{0}$ and $T(\mathbf{v}) \neq \mathbf{0}$.

(b) Find three nonzero linear maps $T_1, T_2, T_3 : \mathbb{F}_2^2 \to \mathbb{F}_2$ such that every $\mathbf{v} \in \mathbb{F}_2^2$ lies in some $\ker T_i$.

(c) Let F be infinite, and let $T_1, \ldots, T_k \in \mathrm{L}(V, W)$ all be nonzero. Show that some $\mathbf{v} \in V$ has $T_i(\mathbf{v}) \neq \mathbf{0}$ for every i.

Solution on page 1170

Exercise 2.6.51. STANDARD Let $T \in \mathrm{L}(V, W)$.

(a) Show that $\mathrm{im}(T\restriction_X)$ is a subspace of W for every subspace X of V, and that $T^{-1}(U)$ is a subspace of V containing $\ker T$ for every subspace U of W.

(b) Show that $\mathrm{im}\left(T\restriction_{T^{-1}(U)}\right) = U \cap \mathrm{im}\, T$ for every subspace U of W.

(c) Show that $T^{-1}(\mathrm{im}(T\restriction_X)) = X + \ker T$ for every subspace X of V.

Solution on page 1171

Exercise 2.6.52. STANDARD Let $S, T \in \mathrm{L}(V)$ with $ST = TS$.

(a) Show that S maps $\ker T$ into $\ker T$ and $\mathrm{im}\, T$ into $\mathrm{im}\, T$.

(b) Deduce that for every $\lambda \in F$, T maps $\ker(T - \lambda\,\mathrm{id})$ into itself and $\mathrm{im}(T - \lambda\,\mathrm{id})$ into itself.

(c) Show by an example that (a) can fail if $ST \neq TS$.

Solution on page 1171

2.7 The Fundamental Theorems

For a matrix $A \in F^{m \times n}$, the last two sections produced a long list of properties. The map T_A can be injective or surjective, its kernel can be $\{\mathbf{0}\}$ and its image F^m, the columns of A can be linearly independent or span F^m, A can have a left or a right inverse, and the equation $A\mathbf{x} = \mathbf{y}$ can have at most one or at least one solution for each $\mathbf{y}$. Many of these turned out to be the same property in different words. The three Fundamental Theorems sort them into three groups: the conditions that say "at most one", those that say "at least one", and those that say "exactly one". Almost every equivalence among them is a result we already have, read the right way. The exception is that an injective T_A has a linear left inverse, and a surjective one a linear right inverse. These need a construction, which the lecture postpones until row reduction is available (Section 3.5). We give a direct construction here as well.

Each theorem is a list of equivalent statements, so we start with how such lists are proved. Throughout, $A \in F^{m \times n}$ has columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and $T_A : F^n \to F^m$ is the linear map $\mathbf{x} \mapsto A\mathbf{x}$ (Definition 2.3.15).

2.7.1 Proving that statements are equivalent

A theorem of the form "the following are equivalent" (TFAE) about statements $P_1, \ldots, P_k$ asserts that they are all true or all false. Proving every implication $P_i \Rightarrow P_j$ separately would take $k(k-1)$ arguments, which is 20 for the five conditions of a Fundamental Theorem. Implications can be chained, so far fewer are needed.

Technique 2.7.1 (Proving TFAE). To prove that $P_1, \ldots, P_k$ are equivalent, prove some implications $P_i \Rightarrow P_j$ such that every statement can be reached from every other by following implications. Then any P_i implies any P_j, by chaining. Two shapes cover most cases.

- A **cycle** $P_1 \Rightarrow P_2 \Rightarrow \cdots \Rightarrow P_k \Rightarrow P_1$ uses k implications, the fewest possible. The order is ours to choose, and a good order makes every step short. Usually one goes from the most informative statement (one that hands us an object, such as an inverse) toward the least informative, and closes the loop with the one step that needs an idea.
- A **hub with spokes** fixes one statement, say P_1, and proves $P_i \Leftrightarrow P_1$ for every other i. It uses more implications than a cycle, but it is the natural shape when earlier results already relate each P_i to P_1. Each spoke is then a citation, and the citations are a complete proof, because equivalence is transitive.

Mixed shapes are fine. Draw the statements as points and the proved implications as arrows, and check that you can travel from any point to any other along arrows. Two mistakes are common. Equivalences inside two separate groups of statements say nothing about one group versus the other (Exercise 2.7.21). And each arrow must be proved without the theorem it belongs to, and without any result that depends on that theorem. Going around a cycle is not circular reasoning, since every arrow is proved on its own; quoting the theorem to prove one of its own arrows would be.

The payoff comes after the proof. To show that one of the statements holds, check whichever is easiest; to use one, use whichever is strongest. We have already seen cycles in Theorem 1.4.3 and Theorem 2.1.3. The canonical example is the First Fundamental Theorem (Theorem 2.7.2), whose proof is a hub with three spokes plus two implications about inverses.

Used in. Theorem 1.4.3, Theorem 2.7.2, Theorem 2.7.4, Exercise 2.7.6, Theorem 2.7.8, Exercise 2.7.15, Corollary 2.7.16, Exercise 2.7.21, Exercise 2.7.23, Exercise 2.8.11, Corollary 3.4.20, Theorem 3.5.12, Theorem 3.5.13, Theorem 3.5.14, Corollary 3.5.25, Theorem 4.4.7, Corollary 4.5.21, Theorem 4.5.25, Lemma 4.6.14, Theorem 4.7.22, Theorem 4.7.37, Proposition 6.3.34, Theorem 7.4.16, Proposition 7.5.2, Theorem 7.6.15, Theorem 8.2.5, Theorem 8.2.12, Theorem 8.3.5, Proposition 8.3.15, Theorem 8.4.4, Theorem 8.4.11, Theorem 8.4.15, Theorem A.2.1, Lemma A.3.5, Exercise 2.7.19, Exercise 4.2.50, Exercise 6.4.57, Exercise 6.6.40.

2.7.2 The First and Second Fundamental Theorems

Intuition (Five ways to say that nothing is lost). Each condition of the First Fundamental Theorem says, in its own language, that T_A loses no information. In (I), different inputs give different outputs. In (II), each equation $A\mathbf{x} = \mathbf{y}$ pins down $\mathbf{x}$ whenever it has a solution at all. In (III), no nontrivial combination of the columns vanishes, so the coefficients of a combination can be recovered from its value. In (IV), no nonzero vector is sent to $\mathbf{0}$. In (V), a linear map recovers the input from the output. The first four are visibly one statement in four languages, and Section 2.6 proved as much. The fifth is of a different kind, since it asserts that a certain map exists.

Theorem 2.7.2 (First Fundamental Theorem). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is injective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has at most one solution $\mathbf{x} \in F^n$.
(III) The columns of A are linearly independent.
(IV) $\ker T_A = \{\mathbf{0}\}$.
(V) T_A has a left inverse: a linear map $S : F^m \to F^n$ with $S \circ T_A = \mathrm{id}_{F^n}$.

We use (I) as the hub. Parts 1 and 2 of the proof only reread earlier results. Part 3 needs a construction, and it is given in Section 2.7.4.

Proof, part 1 of 3: (I)–(IV) are equivalent. ROUTINE

> *Uses.* Definition 2.6.1 (Injective), Definition 2.3.15 (The linear map of a matrix), Corollary 2.6.24 (The columns decide), Proposition 2.3.13 (Linearity of the product), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* Each of (II), (III), and (IV) is equivalent to (I) by a definition or by a result of Section 2.6. Find the right one for each.
> *Hint 2.* In the language of the map T_A, what is a solution of $A\mathbf{x} = \mathbf{y}$? What are the values of T_A on the standard basis?
> *Hint 3.* (I) ⇔ (III) is Corollary 2.6.24(a), and (I) ⇔ (IV) is Theorem 2.6.11.
> *Sketch.* Hub at (I). A solution of $A\mathbf{x} = \mathbf{y}$ is a preimage of $\mathbf{y}$ under T_A, so (II) is the definition of injectivity. The columns are the values $T_A(\mathbf{e}_j)$, so (I) ⇔ (III) is the basis test of Corollary 2.6.24(a). Since T_A is linear, (I) ⇔ (IV) is Theorem 2.6.11.

We prove (I) ⇔ (II), (I) ⇔ (III), and (I) ⇔ (IV).

(I) ⇔ (II). By Definition 2.3.15, $\mathbf{x}$ is a solution of $A\mathbf{x} = \mathbf{y}$ exactly when $T_A(\mathbf{x}) = \mathbf{y}$, that is, when $\mathbf{x}$ is a preimage of $\mathbf{y}$ under T_A. So (II) says that every $\mathbf{y} \in F^m$ has at most one preimage under T_A, which is the definition of injectivity (Definition 2.6.1).

(I) ⇔ (III). This is Corollary 2.6.24(a).

(I) ⇔ (IV). T_A is linear (Proposition 2.3.13), so this is Theorem 2.6.11 applied to $T = T_A$.

Proof, part 2 of 3: (V) ⇒ (I). ROUTINE

> *Uses.* Proposition 2.5.4 (Left inverses give uniqueness)
> *Hint 1.* Suppose $S \circ T_A = \mathrm{id}_{F^n}$ and $A\mathbf{x} = A\mathbf{x}'$. Apply S to both sides.
> *Hint 2.* This is Proposition 2.5.4.
> *Sketch.* Applying S to $A\mathbf{x} = A\mathbf{x}'$ gives $\mathbf{x} = \mathbf{x}'$. This is Proposition 2.5.4.

Let S be a left inverse of T_A. By Proposition 2.5.4, T_A is injective. Explicitly, if $T_A(\mathbf{x}) = T_A(\mathbf{x}')$, then

$$\mathbf{x} = S(T_A(\mathbf{x})) = S(T_A(\mathbf{x}')) = \mathbf{x}'.$$

What remains is to show that an injective T_A has a left inverse. Since (I)–(IV) are already known to be equivalent, we may start from whichever of them is most convenient, and in Section 2.7.4 we start from (III). Figure 2.7.1 shows the whole web of implications.

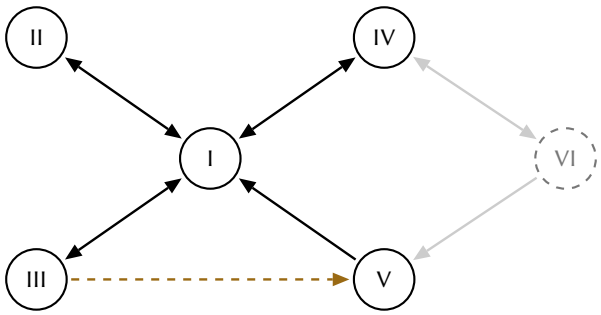


Figure 2.7.1. The implications proved for Theorem 2.7.2. The solid arrows are parts 1 and 2 of the proof, which reread earlier results. The dashed arrow (III) ⇒ (V) is the construction of part 3 (Section 2.7.4). Every node can be reached from every other, so (I)–(V) are equivalent. The faint node is a sixth condition, rank $A = n$, which Section 3.5.2 attaches by the faint arrows. For Theorem 2.7.4 the picture is the same, except that the construction starts from (II), and (VI) attaches at (II) rather than (IV).

Remark 2.7.3 (Reinterpreting earlier results). Part 1 of the proof cites one definition and two earlier results, and nothing else, yet it is a complete proof. Each spoke is a definition or a theorem proved earlier, the hypotheses of the theorems hold here, and equivalence is transitive. What part 1 adds is the translation that lets each earlier result be read as a statement about A. A solution of $A\mathbf{x} = \mathbf{y}$ is a preimage of $\mathbf{y}$ under T_A. The columns of A are the values of T_A on the standard basis. And T_A is linear, so every theorem about linear maps applies to it.

When you cite a result in this way, check two things: that its hypotheses really hold, and that its conclusion, after translation, is literally the implication you need. For instance, Proposition 2.6.22(a) decides injectivity from the values on a *basis* of the domain; Corollary 2.6.24 is its special case for the standard basis of F^n, whose values are the columns. The real work was done in Section 2.6. Collecting the conditions into one theorem is still worthwhile, because in a problem we rarely know in advance which of them is easiest to check, and the theorem lets us check that one and use any other.

> *Intuition (Five ways to say that everything is reached).* The Second Fundamental Theorem is the mirror image of the first. Each of its conditions says that T_A reaches every vector of F^m: as a map (I); as a statement about equations (II); as a statement about combinations of the columns (III); as a statement about the image (IV); and, in (V), by a linear map that produces, for each $\mathbf{y}$, an input that reaches it.

Theorem 2.7.4 (Second Fundamental Theorem). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is surjective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has at least one solution $\mathbf{x} \in F^n$.
(III) The columns of A span F^m.
(IV) $\operatorname{im} T_A = F^m$.
(V) T_A has a right inverse: a linear map $S : F^m \to F^n$ with $T_A \circ S = \mathrm{id}_{F^m}$.

Proof, part 1 of 3: (I)–(IV) are equivalent. ROUTINE

> *Uses.* Definition 2.6.2 (Surjective), Definition 2.6.6 (Image), Definition 2.3.15 (The linear map of a matrix), Corollary 2.6.24 (The columns decide)
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* Use (I) as the hub again. Two of the spokes are the definition of surjectivity, read in two ways.
> *Hint 2.* (I) ⇔ (III) is the other half of Corollary 2.6.24.
> *Sketch.* A solution of $A\mathbf{x} = \mathbf{y}$ is a preimage of $\mathbf{y}$, so (I) ⇔ (II) is the definition of surjectivity, and so is (I) ⇔ (IV), since $\operatorname{im} T_A$ is the set of values of T_A. (I) ⇔ (III) is Corollary 2.6.24(b).

(I) ⇔ (II). As in the proof of Theorem 2.7.2, the solutions of $A\mathbf{x} = \mathbf{y}$ are the preimages of $\mathbf{y}$ under T_A. So (II) says that every $\mathbf{y} \in F^m$ has at least one preimage, which is Definition 2.6.2.

(I) ⇔ (IV). By Definition 2.6.6, $\operatorname{im} T_A$ is the set of values of T_A, so $\operatorname{im} T_A = F^m$ says that every vector of F^m is a value, which is again Definition 2.6.2.

(I) ⇔ (III). This is Corollary 2.6.24(b).

Instead of the last spoke we could have linked (III) and (IV) directly, since $\operatorname{im} T_A$ is the span of the columns (Proposition 2.6.18). The resulting chain (II) ⇔ (I) ⇔ (IV) ⇔ (III) is just as good.

Proof, part 2 of 3: (V) ⇒ (I). ROUTINE

> *Uses.* Proposition 2.5.5 (Right inverses give existence)
> *Hint 1.* Suppose $T_A \circ S = \mathrm{id}_{F^m}$, and let $\mathbf{y} \in F^m$. Which vector could T_A send to $\mathbf{y}$?
> *Hint 2.* This is Proposition 2.5.5.
> *Sketch.* $T_A(S(\mathbf{y})) = \mathbf{y}$ for every $\mathbf{y}$. This is Proposition 2.5.5.

Let S be a right inverse of T_A. For every $\mathbf{y} \in F^m$, $T_A(S(\mathbf{y})) = \mathbf{y}$, so $\mathbf{y}$ is a value of T_A. This is Proposition 2.5.5.

Again the implication from (I)–(IV) to (V) needs a construction, and it is part 3 of the proof, in Section 2.7.4.

Example 2.7.5 (A tall matrix). Let

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \in F^{3\times 2}, \quad T_A(x, y) = (x, x + y, y),$$

over any field F. Its columns $\mathbf{a}_1 = (1, 1, 0)$ and $\mathbf{a}_2 = (0, 1, 1)$ are linearly independent: if $x\mathbf{a}_1 + y\mathbf{a}_2 = (x, x + y, y) = \mathbf{0}$, then $x = y = 0$. So condition (III) of Theorem 2.7.2 holds, and with it all of (I)–(V). For instance, by (II) each system $A\mathbf{x} = \mathbf{y}$ has at most one solution, and by (V) there is a left inverse. Here one is easy to find, since x and y can be read off the first and third entries of $(x, x + y, y)$:

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad CA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

They can also be read off differently, using $x = (x + y) - y$:

$$C' = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad C'A = I_2.$$

Indeed, the first row of C' sends $\mathbf{a}_1$ to $0 + 1 - 0 = 1$ and $\mathbf{a}_2$ to $0 + 1 - 1 = 0$.

On the other hand, T_A is not surjective. The vector $(0, 1, 0)$ would need $x = 0$, $y = 0$, and $x + y = 1$, which is impossible. So no condition of Theorem 2.7.4 holds: the columns do not span F^3, and there is no right inverse. (Corollary 2.5.11 also rules out a right inverse, since $C \neq C'$.) Uniqueness of solutions and existence of solutions really are independent properties.

Over $\mathbb{R}$, the span explorer below shows the columns of this A and a target $\mathbf{w}$. While $\mathbf{w}$ lies in the plane spanned by the columns, the explorer finds coefficients x, y with $x\mathbf{a}_1 + y\mathbf{a}_2 = \mathbf{w}$, and they are unique: this is (II) of Theorem 2.7.2. Drag $\mathbf{w}$ off the plane and there are none, as Theorem 2.7.4 fails. Adding a third vector off the plane makes the list independent (Proposition 1.5.14), and the explorer reports that it also spans $\mathbb{R}^3$; for three vectors in $\mathbb{R}^3$ that is no accident (Corollary 3.5.16). The 3×3 matrix with these columns then satisfies all three theorems.

$$(1, 2, 1) = \mathbf{a}_1 + \mathbf{a}_2,$$
$$\operatorname{span}(\mathbf{a}_1, \mathbf{a}_2) = \{(a, b, c) \in \mathbb{R}^3 \mid b = a + c\}.$$

The columns $\mathbf{a}_1 = (1, 1, 0)$ and $\mathbf{a}_2 = (0, 1, 1)$ of the matrix A of Example 2.7.5. They are independent, so a target in their span has exactly one pair of coefficients; the target $\mathbf{w} = (1, 2, 1) = \mathbf{a}_1 + \mathbf{a}_2$ gives $(x, y) = (1, 1)$. A target off the plane has none. (Interactive in the web version.)

Exercise 2.7.6. ROUTINE Let $\mathbf{a} \in F^m$, regarded as a matrix in $F^{m\times 1}$, so that $T_{\mathbf{a}} : F \to F^m$ is $x \mapsto x\mathbf{a}$. Show directly, without quoting Theorem 2.7.2, that the following are equivalent: (i) $T_{\mathbf{a}}$ is injective; (ii) $\mathbf{a} \neq \mathbf{0}$; (iii) $\mathbf{a}$ has a left inverse $\mathbf{c} \in F^{1\times m}$, that is, $\mathbf{ca} = 1$. Which step corresponds to the construction in part 3 of the proof of Theorem 2.7.2?

Solution on page 1171

2.7.3 The Third Fundamental Theorem

For the Third Fundamental Theorem, the implication from bijectivity to invertibility is much easier than the corresponding implications for one-sided inverses. As the lecture puts it, there is only one plausible choice for the inverse.

Intuition (Only one candidate). Let $T \in \mathrm{L}(V, W)$ be bijective. If S is a two-sided inverse of T, then $T(S(\mathbf{w})) = \mathbf{w}$ for every $\mathbf{w} \in W$, so $S(\mathbf{w})$ is a preimage of $\mathbf{w}$. Since T is bijective, $\mathbf{w}$ has exactly one preimage. So S has to be the inverse function of T, which sends each $\mathbf{w}$ to its unique preimage, and the only question

is whether that function is linear. For a left inverse of an injective map there is no such uniqueness: its values off the image are free, and many choices are not linear (Remark 2.5.8).

Proposition 2.7.7 (Bijective linear maps are invertible). Let $T \in \mathrm{L}(V, W)$ be bijective, and let $S : W \to V$ be its inverse function: $S(\mathbf{w})$ is the unique $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$. Then S is linear. Consequently T is invertible, and $T^{-1} = S$.

Proof. ROUTINE

Uses. Definition 2.6.3 (Bijective), Definition 2.1.1 (Linear map), Definition 2.5.12 (Invertible), Section A.1
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. To compute S at a vector, find the vector that T sends there. Fix $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ and scalars $a_1, \ldots, a_k$, and look for the vector that T sends to $a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$.
Hint 2. Put $\mathbf{v}_i = S(\mathbf{w}_i)$, so that $T(\mathbf{v}_i) = \mathbf{w}_i$. What is $T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k)$?
Sketch. With $\mathbf{v}_i = S(\mathbf{w}_i)$, linearity of T gives $T(\sum a_i\mathbf{v}_i) = \sum a_i\mathbf{w}_i$, so by the definition of S, $S(\sum a_i\mathbf{w}_i) = \sum a_i\mathbf{v}_i = \sum a_iS(\mathbf{w}_i)$. Since S is the inverse function, $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$.

Since T is bijective, every $\mathbf{w} \in W$ has exactly one preimage (Definition 2.6.3), so S is a well-defined function $W \to V$, and $T(S(\mathbf{w})) = \mathbf{w}$ for every $\mathbf{w}$.

Fix $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ and $a_1, \ldots, a_k \in F$. We will show that $S(a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k) = a_1S(\mathbf{w}_1) + \cdots + a_kS(\mathbf{w}_k)$. Put $\mathbf{v}_i = S(\mathbf{w}_i)$, so that $T(\mathbf{v}_i) = \mathbf{w}_i$. Since T is linear,

$$\begin{aligned} T(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) &= a_1T(\mathbf{v}_1) + \cdots + a_kT(\mathbf{v}_k) \\ &= a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k. \end{aligned}$$

So $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$ is the preimage of $a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$, and by the definition of S,

$$\begin{aligned} S(a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k) &= a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \\ &= a_1S(\mathbf{w}_1) + \cdots + a_kS(\mathbf{w}_k), \end{aligned}$$

as desired. Thus S is linear (Definition 2.1.1).

Finally, S is the inverse function of T (Section A.1): $S(T(\mathbf{v})) = \mathbf{v}$, because $\mathbf{v}$ is the preimage of $T(\mathbf{v})$, and $T(S(\mathbf{w})) = \mathbf{w}$. So S is a linear map that is both a left and a right inverse of T, and T is invertible with $T^{-1} = S$ (Definition 2.5.12). □

No basis appeared in the proof, so the proposition holds for every bijective linear map, between any two vector spaces. Together with Proposition 2.5.4 and Proposition 2.5.5 it says that a linear map is invertible if and only if it is bijective.

Intuition (The third theorem from the first two). "Exactly one" means "at most one and at least one". Each condition (I)–(IV) of the Third Fundamental Theorem below is, in the same way, the combination of condition (I)–(IV) of Theorem 2.7.2 with the same condition of Theorem 2.7.4: bijective means injective and surjective, a basis is an independent spanning list, and (IV) is written as a combination. So the equivalence of (I)–(IV) is inherited from the first two theorems, and only (V) needs an argument.

Theorem 2.7.8 (Third Fundamental Theorem). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is bijective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has exactly one solution $\mathbf{x} \in F^n$.
(III) The columns of A form a basis of F^m.
(IV) $\ker T_A = \{\mathbf{0}\}$ and $\operatorname{im} T_A = F^m$.
(V) T_A is invertible.

Proof, part 1 of 3: (I)–(IV) are equivalent. ROUTINE

Uses. Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Definition 2.6.3 (Bijective), Proposition 1.6.2 (Bases are independent spanning lists)

Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Write each of (I)–(IV) as "P and Q", where P is the corresponding condition of Theorem 2.7.2 and Q that of Theorem 2.7.4.
Hint 2. If $P_i \Leftrightarrow P_j$ and $Q_i \Leftrightarrow Q_j$, then (P_i and Q_i) $\Leftrightarrow$ (P_j and Q_j). Use only parts 1 of the proofs of the first two theorems.
Sketch. Bijective = injective and surjective; exactly one = at most one and at least one; basis = independent and spanning (Proposition 1.6.2); (IV) is the combination of the two (IV)s. By parts 1 of the proofs of Theorem 2.7.2 and Theorem 2.7.4, the first halves are equivalent to each other, and so are the second halves; hence so are the conjunctions.

Each of the conditions (I)–(IV) of the present theorem combines the condition with the same numeral in Theorem 2.7.2 with the one in Theorem 2.7.4:

- (I) says that T_A is injective and surjective (Definition 2.6.3);
- (II) says that $A\mathbf{x} = \mathbf{y}$ has at most one and at least one solution, for every $\mathbf{y}$;
- (III) says that the columns of A are linearly independent and span F^m (Proposition 1.6.2);
- (IV) says that $\ker T_A = \{\mathbf{0}\}$ and $\operatorname{im} T_A = F^m$.

By part 1 of the proof of Theorem 2.7.2, the first halves of these four conditions are equivalent to one another, and by part 1 of the proof of Theorem 2.7.4, so are the second halves. If $P \Leftrightarrow P'$ and $Q \Leftrightarrow Q'$, then (P and Q) $\Leftrightarrow$ (P' and Q'). So (I)–(IV) are equivalent.

Proof, part 2 of 3: (V) $\Rightarrow$ (I). ROUTINE

Uses. Definition 2.5.12 (Invertible), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence)
Hint 1. An inverse is both a left and a right inverse.
Hint 2. Apply Proposition 2.5.4 and Proposition 2.5.5.
Sketch. T_A^{-1} is a left inverse, so T_A is injective, and a right inverse, so T_A is surjective.

If T_A is invertible, then T_A^{-1} is a left inverse and a right inverse of T_A (Definition 2.5.12). By Proposition 2.5.4 and Proposition 2.5.5, T_A is injective and surjective, that is, bijective.

For the last step we give two proofs. The first is the lecture's; the second builds the inverse on a basis, and it is the model for the constructions in Section 2.7.4.

Proof (via the inverse function), part 3 of 3: (I) $\Rightarrow$ (V). ROUTINE

Uses. Proposition 2.7.7 (Bijective linear maps are invertible), Proposition 2.3.13 (Linearity of the product)
Hint. T_A is a bijective linear map. What do we know about the inverse function of such a map?
Sketch. The inverse function of the bijective linear map T_A is linear (Proposition 2.7.7), so it is a linear two-sided inverse.

T_A is linear (Proposition 2.3.13) and, by (I), bijective. By Proposition 2.7.7 it is invertible. □

Proof (by building the inverse on a basis), part 3 of 3: (III) $\Rightarrow$ (V). ROUTINE

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 1.6.7 (Standard bases), Proposition 2.3.13 (Linearity of the product)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
Hint 1. An inverse S must send $\mathbf{a}_j = T_A(\mathbf{e}_j)$ back to $\mathbf{e}_j$. Since $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is a basis of F^m, these values can be prescribed.
Hint 2. Let S be the linear map with $S(\mathbf{a}_j) = \mathbf{e}_j$. Check $S \circ T_A = \mathrm{id}$ on $\mathbf{e}_1, \ldots, \mathbf{e}_n$, and $T_A \circ S = \mathrm{id}$ on $\mathbf{a}_1, \ldots, \mathbf{a}_n$.
Sketch. By Theorem 2.2.1 there is a linear $S : F^m \to F^n$ with $S(\mathbf{a}_j) = \mathbf{e}_j$. Then $S(T_A(\mathbf{e}_j)) = \mathbf{e}_j$ and $T_A(S(\mathbf{a}_j)) = \mathbf{a}_j$, so $S \circ T_A$ and $T_A \circ S$ agree with the identity on a basis, and are the identity.

Suppose the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ form a basis of F^m. By Theorem 2.2.1 there is a linear map $S : F^m \to F^n$ with $S(\mathbf{a}_j) = \mathbf{e}_j$ for $j = 1, \ldots, n$. Recall that $T_A(\mathbf{e}_j) = A\mathbf{e}_j = \mathbf{a}_j$ (Proposition 2.3.13(c)). Then

$$(S \circ T_A)(\mathbf{e}_j) = S(\mathbf{a}_j) = \mathbf{e}_j,$$
$$(T_A \circ S)(\mathbf{a}_j) = T_A(\mathbf{e}_j) = \mathbf{a}_j$$

for every j. So the linear maps $S \circ T_A$ and id_{F^n} agree on the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of F^n (Proposition 1.6.7), and $T_A \circ S$ and id_{F^m} agree on the basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of F^m. By the uniqueness in Theorem 2.2.1, $S \circ T_A = \mathrm{id}_{F^n}$ and $T_A \circ S = \mathrm{id}_{F^m}$. So S is a linear two-sided inverse of T_A. □

The first proof works for any bijective linear map, with no bases at all. The second uses that the columns are a basis of the target, and it explains where the inverse comes from: it undoes T_A on the columns, and the columns are a basis, so nothing else needs to be decided. In Section 2.7.4 the columns will only be independent, or only spanning, and then something does have to be decided.

Example 2.7.9 (One matrix over three fields). Let $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. The system $A\mathbf{x} = \mathbf{y}$ reads

$$x_1 + x_2 = y_1, \quad x_1 - x_2 = y_2.$$

Adding and subtracting the equations shows that every solution satisfies $2x_1 = y_1 + y_2$ and $2x_2 = y_1 - y_2$, where $2 = 1 + 1$.

Over $\mathbb{Q}$, or any field with $2 \neq 0$, this gives $x_1 = 2^{-1}(y_1 + y_2)$ and $x_2 = 2^{-1}(y_1 - y_2)$, and conversely these values solve the system. So the system has exactly one solution for every $\mathbf{y}$, which is (II) of Theorem 2.7.8, and A is invertible. The solution formula is the inverse:

$$A^{-1} = 2^{-1}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Over $\mathbb{F}_3$, where $2^{-1} = 2$ and $-1 = 2$, this reads $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ and $A^{-1} = \begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}$, and indeed $AA^{-1} = \begin{pmatrix} 4 & 3 \\ 6 & 4 \end{pmatrix} = I_2$ in $\mathbb{F}_3$.

Over $\mathbb{F}_2$, $-1 = 1$, so both columns are $(1, 1)$, and $A(1, 1) = \mathbf{0}$. So (IV) of Theorem 2.7.2 fails, and with it every condition of Theorem 2.7.2 and Theorem 2.7.8. Every value $A\mathbf{x} = (x_1 + x_2)(1, 1)$ is a multiple of $(1, 1)$, so $(1, 0)$ is not a value, and every condition of Theorem 2.7.4 fails as well.

2.7.4 Constructing one-sided inverses

It remains to prove (I) $\Rightarrow$ (V) in the first two theorems: an injective T_A has a left inverse, and a surjective T_A has a right inverse.

> *Intuition (What has to be chosen).* A left inverse S of an injective T_A is forced on the image: it must send $A\mathbf{x}$ to $\mathbf{x}$. Off the image nothing forces it, but whatever we choose there has to fit together into a linear map on all of F^m. A right inverse of a surjective T_A must pick, for each $\mathbf{y} \in F^m$, one solution $S(\mathbf{y})$ of $A\mathbf{x} = \mathbf{y}$, and when T_A is not injective there are many solutions to pick from. Again the picks must fit together linearly. Choices made one vector at a time usually do not, as the next example shows. Choices made on a basis always do, by Theorem 2.2.1, and that is all the proofs below use (Technique 2.2.4).

The lecture does not prove these implications at this point. It returns to them in Section 3.5, after row reduction, where left and right inverses are computed from $\operatorname{rref}(A)$ as explicit matrices (Proposition 3.5.9, Proposition 3.5.8), and the completed theorems (Theorem 3.5.12, Theorem 3.5.13) close the loops through a sixth condition. That is the route to use for computing. Here is a second route, which needs only Theorem 2.2.1 (and, for left inverses, Proposition 1.5.14) and no row reduction.

Non-example 2.7.10 (Choices that are not linear).

(a) Let $T : F \to F^2$, $T(x) = (x, 0)$, the map of Example 2.5.3. Undo T on its image and send everything else to 0:

$$g(x, 0) = x, \quad g(x, y) = 0 \text{ if } y \neq 0.$$

Then $g(T(x)) = x$ for every x, but g is not linear, since $g(1, 1) = 0$ while $g(1, 0) + g(0, 1) = 1 + 0 = 1$.

(b) Let $A = (1 \;\; 1) \in \mathbb{R}^{1\times 2}$, so $T_A(x_1, x_2) = x_1 + x_2$. For each $y \in \mathbb{R}$, choose the solution

$$h(y) = (y, 0) \text{ if } y \geq 0, \quad h(y) = (0, y) \text{ if } y < 0.$$

Then $T_A(h(y)) = y$ for every y, but h is not linear, since $h(1) + h(-1) = (1, -1)$ while $h(0) = (0, 0)$.

For a right inverse, the basis to choose on is the standard basis of the target.

Proof of Theorem 2.7.4, part 3 of 3: (II) ⇒ (V). ROUTINE

> *Uses.* Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 1.6.7 (Standard bases), Definition 2.5.2 (Right inverse)
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
> *Hint 1.* A right inverse has to choose a solution of $A\mathbf{x} = \mathbf{y}$ for every $\mathbf{y}$. Make the choice only for $\mathbf{y} = \mathbf{e}_1, \dots, \mathbf{e}_m$.
> *Hint 2.* Choose solutions $\mathbf{x}_i$ of $A\mathbf{x} = \mathbf{e}_i$, let S be the linear map with $S(\mathbf{e}_i) = \mathbf{x}_i$, and check $T_A \circ S = \mathrm{id}$ on the standard basis.
> *Sketch.* By (II), choose $\mathbf{x}_i$ with $A\mathbf{x}_i = \mathbf{e}_i$ for $i = 1, \dots, m$. Let S be the linear map with $S(\mathbf{e}_i) = \mathbf{x}_i$ (Theorem 2.2.1). Then $T_A(S(\mathbf{e}_i)) = \mathbf{e}_i$, so $T_A \circ S$ agrees with id_{F^m} on a basis and equals it. In matrices, $AB = I_m$ for the matrix B with columns $\mathbf{x}_1, \dots, \mathbf{x}_m$.

Suppose (II). For each $i = 1, \dots, m$, the equation $A\mathbf{x} = \mathbf{e}_i$ has a solution; choose one and call it $\mathbf{x}_i$. Since $\mathbf{e}_1, \dots, \mathbf{e}_m$ is a basis of F^m (Proposition 1.6.7), Theorem 2.2.1 gives a linear map $S : F^m \to F^n$ with $S(\mathbf{e}_i) = \mathbf{x}_i$ for every i. Then

$$(T_A \circ S)(\mathbf{e}_i) = A\mathbf{x}_i = \mathbf{e}_i = \mathrm{id}_{F^m}(\mathbf{e}_i) \quad (i = 1, \dots, m),$$

so the linear maps $T_A \circ S$ and id_{F^m} agree on a basis, and they are equal by the uniqueness in Theorem 2.2.1. So S is a right inverse of T_A (Definition 2.5.2). In matrix terms, $S = T_B$ for the matrix $B \in F^{n \times m}$ with columns $\mathbf{x}_1, \dots, \mathbf{x}_m$, and $AB = I_m$. □

For a left inverse the values on the columns are forced, $S(\mathbf{a}_j) = \mathbf{e}_j$. The columns are independent but need not span F^m, so before we can use Theorem 2.2.1 we must complete them to a basis.

Proof of Theorem 2.7.2, part 3 of 3: (III) ⇒ (V). STANDARD

> *Uses.* Proposition 1.5.14 (Appending a vector), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 1.6.7 (Standard bases), Proposition 1.6.2 (Bases are independent spanning lists), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.3.13 (Linearity of the product), Definition 2.5.1 (Left inverse)
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
> *Hint 1.* A left inverse must send $\mathbf{a}_j = T_A(\mathbf{e}_j)$ to $\mathbf{e}_j$. To prescribe these values with Theorem 2.2.1, the $\mathbf{a}_j$ have to be part of a basis of F^m.
> *Hint 2.* Go through $\mathbf{e}_1, \dots, \mathbf{e}_m$ in turn, and append each one that is not in the span of the list built so far. Why does the final list span F^m, and why is it still independent?
> *Hint 3.* Define S on the resulting basis by $S(\mathbf{a}_j) = \mathbf{e}_j$ and $S = \mathbf{0}$ on the appended vectors. Check $S \circ T_A = \mathrm{id}$ on the standard basis of F^n.
> *Sketch.* Append to $\mathbf{a}_1, \dots, \mathbf{a}_n$ each $\mathbf{e}_i$ that is not in the span of the list so far. By Proposition 1.5.14 the list stays independent, and every $\mathbf{e}_i$ ends up in its span, so it becomes a basis $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_t}$ of F^m. Let $S(\mathbf{a}_j) = \mathbf{e}_j$ and $S(\mathbf{e}_{i_s}) = \mathbf{0}$. Then $S(T_A(\mathbf{e}_j)) = \mathbf{e}_j$, so $S \circ T_A = \mathrm{id}_{F^n}$.

Suppose the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ are linearly independent.

Completing the columns to a basis. Let $\mathcal{L}_0$ be the list $\mathbf{a}_1, \dots, \mathbf{a}_n$. For $i = 1, \dots, m$ in turn, let $\mathcal{L}_i = \mathcal{L}_{i-1}$ if $\mathbf{e}_i \in \mathrm{span}(\mathcal{L}_{i-1})$, and let $\mathcal{L}_i$ be $\mathcal{L}_{i-1}$ with $\mathbf{e}_i$ appended otherwise. By Proposition 1.5.14 and induction on i, every $\mathcal{L}_i$ is linearly independent. For each i, the vector $\mathbf{e}_i$ lies in $\mathrm{span}(\mathcal{L}_i)$, either because it was already in $\mathrm{span}(\mathcal{L}_{i-1})$ or because it was appended, and $\mathrm{span}(\mathcal{L}_i) \subseteq \mathrm{span}(\mathcal{L}_m)$, since $\mathcal{L}_i$ is an initial part of $\mathcal{L}_m$. So the subspace $\mathrm{span}(\mathcal{L}_m)$ contains $\mathbf{e}_1, \dots, \mathbf{e}_m$, hence all of their linear combinations (Proposition 1.4.14), which is all of F^m (Proposition 1.6.7). Thus

$$\mathcal{L}_m = \mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_t}$$

is linearly independent and spans F^m, so it is a basis of F^m (Proposition 1.6.2).

Defining S. By Theorem 2.2.1 there is a linear map $S : F^m \to F^n$ with

$$S(\mathbf{a}_j) = \mathbf{e}_j \quad (j = 1, \dots, n),$$
$$S(\mathbf{e}_{i_s}) = \mathbf{0} \quad (s = 1, \dots, t).$$

Since $T_A(\mathbf{e}_j) = \mathbf{a}_j$ (Proposition 2.3.13(c)), we get $(S \circ T_A)(\mathbf{e}_j) = S(\mathbf{a}_j) = \mathbf{e}_j$ for every j. So $S \circ T_A$ and id_{F^n} agree on the standard basis of F^n, and they are equal by the uniqueness in Theorem 2.2.1. That is, S is a left inverse of T_A (Definition 2.5.1). □

This completes the proofs of all three theorems.

Remark 2.7.11 (What was chosen, and what was not). In both constructions, the choices are exactly the freedom that one-sided inverses have. A right inverse is determined by its values $S(\mathbf{e}_i)$, which may be any solutions of $A\mathbf{x} = \mathbf{e}_i$; when T_A is not injective, each of these equations has many solutions. A left inverse is forced on the columns, and its values on the appended vectors $\mathbf{e}_{i_s}$ were set to $\mathbf{0}$ only for definiteness; any values work. So when T_A is injective but not surjective, it has many left inverses, in line with Corollary 2.5.11.

The completion step is the special case of Theorem 4.1.16 in which the spanning list is $\mathbf{e}_1, \dots, \mathbf{e}_m$. In F^m it needs no theory of dimension, because each of the m standard basis vectors is tested once. The same proof works for $T \in \mathrm{L}(V, W)$ whenever V and W have finite bases: if T is injective and $\mathcal{B}$ is a basis of V, complete the independent list $T(\mathcal{B})$ to a basis of W, testing the vectors of a basis of W in place of $\mathbf{e}_1, \dots, \mathbf{e}_m$ (compare Corollary 2.7.16 below). For arbitrary vector spaces one needs Zorn's lemma (Remark 2.5.25).

Example 2.7.12 (All left inverses of the tall matrix). Return to $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$ from Example 2.7.5, with columns $\mathbf{a}_1 = (1, 1, 0)$ and $\mathbf{a}_2 = (0, 1, 1)$. Their span consists of the vectors $(x, x + y, y)$, whose middle entry is the sum of the other two. The vector $\mathbf{e}_1 = (1, 0, 0)$ fails this test, so it is appended. Now every vector is a combination of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}_1$, since

$$(p, q, r) = \alpha\mathbf{a}_1 + \beta\mathbf{a}_2 + \gamma\mathbf{e}_1 = (\alpha + \gamma, \alpha + \beta, \beta)$$

holds for $\beta = r$, $\alpha = q - r$, and $\gamma = p - q + r$. So $\mathbf{e}_2$ and $\mathbf{e}_3$ are not appended, and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}_1$ is a basis of F^3.

A linear map $S : F^3 \to F^2$ is a left inverse of T_A exactly when $S(\mathbf{a}_1) = \mathbf{e}_1$ and $S(\mathbf{a}_2) = \mathbf{e}_2$ (check $S \circ T_A = \mathrm{id}$ on $\mathbf{e}_1, \mathbf{e}_2$). Its value $S(\mathbf{e}_1) = (c, d)$ is free. With the coordinates above,

$$S(p, q, r) = (q - r)\mathbf{e}_1 + r\mathbf{e}_2 + (p - q + r)(c, d).$$

Collecting the coefficients of p, q, and r in each entry, the left inverses of A are exactly the matrices

$$S_{c,d} = \begin{pmatrix} c & 1 - c & c - 1 \\ d & -d & 1 + d \end{pmatrix} \quad (c, d \in F).$$

The construction in the proof, with $S(\mathbf{e}_1) = \mathbf{0}$, gives $S_{0,0} = C'$, and $S_{1,0} = C$; these are the two left inverses found by inspection in Example 2.7.5. Over $\mathbb{F}_2$ there are exactly four left inverses, and over $\mathbb{Q}$ infinitely many.

With the constructions in hand, the theorems can be used in the direction that the propositions of Section 2.5 could not reach: a matrix without a one-sided inverse fails the other conditions too.

> **Exercise 2.7.13.** ROUTINE Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$. Show that the following are equivalent: T_A is injective; T_A is surjective; $ad - bc \neq 0$. (Use Proposition 2.5.28.)

Solution on page 1172

2.7.5 Using the theorems

Once a list of conditions is known to be equivalent, a question posed in one language can be answered in another. The following is a typical instance.

Example 2.7.14 (Products of matrices with independent columns). Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$ both have linearly independent columns. We claim that AB does too.

Stated with columns, the claim takes some unwinding, since the columns of AB are the vectors $A\mathbf{b}_k$, each a combination of the columns of A with coefficients from a column of B (Proposition 2.4.8). In the language of kernels it takes one line. By Theorem 2.7.2, (III) ⇒ (IV), $\ker T_A = \{\mathbf{0}\}$ and $\ker T_B = \{\mathbf{0}\}$. If $(AB)\mathbf{x} = \mathbf{0}$, then $A(B\mathbf{x}) = \mathbf{0}$, so $B\mathbf{x} \in \ker T_A$, so $B\mathbf{x} = \mathbf{0}$, so $\mathbf{x} = \mathbf{0}$. Thus $\ker T_{AB} = \{\mathbf{0}\}$, and by (IV) ⇒ (III) for AB, the columns of AB are independent.

Condition (V) gives another one-line proof. By Proposition 2.5.20(c), (V) says that the matrix has a left inverse. If $CA = I_n$ and $DB = I_p$, then $(DC)(AB) = D(CA)B = DB = I_p$, so AB has a left inverse (Proposition 2.5.26(a)).

> **Exercise 2.7.15.** ROUTINE Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$. Show that if the columns of A span F^m and the columns of B span F^n, then the columns of AB span F^m.

Solution on page 1172

2.7.6 Linear maps between spaces with bases

The theorems are stated for maps $F^n \to F^m$, but every linear map between spaces with finite bases is represented by a matrix (Technique 2.3.27), and injectivity, surjectivity, and one-sided inverses all survive the translation.

> **Corollary 2.7.16** (The Fundamental Theorems for linear maps). Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V and $\mathcal{C}$ a basis of W of length m, let $T \in \mathrm{L}(V, W)$, and let $A = [T]_{\mathcal{C}\mathcal{B}}$, whose columns are the coordinate vectors $[T(\mathbf{v}_j)]_{\mathcal{C}}$.
>
> (a) The following are equivalent: T is injective; $\ker T = \{\mathbf{0}\}$; the list $T(\mathcal{B})$ is linearly independent; the columns of A are linearly independent; T has a left inverse.
>
> (b) The following are equivalent: T is surjective; $\operatorname{im} T = W$; the list $T(\mathcal{B})$ spans W; the columns of A span F^m; T has a right inverse.
>
> (c) The following are equivalent: T is bijective; the list $T(\mathcal{B})$ is a basis of W; the columns of A form a basis of F^m; T is invertible.

Proof. STANDARD

Uses. Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Proposition 2.6.20 (Kernel and image in coordinates), Proposition 2.5.20 (Inverses and matrices), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.7.8 (Third Fundamental Theorem), Proposition 2.7.7 (Bijective linear maps are invertible), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence)
Techniques. Proving TFAE (Technique 2.7.1), Translate between maps and matrices (Technique 2.3.27)
Hint 1. Use "T is injective" (in (b): surjective; in (c): bijective) as a hub. Two spokes are results of Section 2.6 about T itself.
Hint 2. The other spokes cross over to T_A. Which results transfer injectivity, surjectivity, and one-sided inverses between T and its matrix?
Hint 3. Proposition 2.6.20(c) and Proposition 2.5.20(b), (c). Then apply the Fundamental Theorems to A.
Sketch. Hub at "T is injective". $\ker T = \{\mathbf{0}\}$ by Theorem 2.6.11; $T(\mathcal{B})$ independent by Proposition 2.6.22(a). T is injective iff T_A is (Proposition 2.6.20(c)), iff the columns of A are independent, iff T_A has a left inverse (Theorem 2.7.2), iff A does, iff T does (Proposition 2.5.20). Part (b) is the same with Theorem 2.7.4, and (c) uses Theorem 2.7.8 and Proposition 2.7.7.

(a) We use "T is injective" as the hub. By Theorem 2.6.11 it is equivalent to $\ker T = \{\mathbf{0}\}$, and by Proposition 2.6.22(a) to the linear independence of $T(\mathcal{B})$. By Proposition 2.6.20(c), T is injective if and only if T_A is, and by Theorem 2.7.2, (I) ⇔ (III), this holds if and only if the columns of A are linearly independent. Finally, by Proposition 2.5.20(b) and (c), T has a left inverse if and only if A does, if and only if T_A does, and by Theorem 2.7.2, (I) ⇔ (V), this holds if and only if T_A is injective, that is, if and only if T is.

(b) The same argument with "surjective" in place of "injective". T is surjective if and only if $\operatorname{im} T = W$ (Definition 2.6.2), if and only if $T(\mathcal{B})$ spans W (Proposition 2.6.22(b)), if and only if T_A is surjective

(Proposition 2.6.20(c)), and Theorem 2.7.4 and Proposition 2.5.20(b), (c) handle the columns of A and right inverses.

(c) T is bijective if and only if $T(\mathcal{B})$ is a basis of W (Proposition 2.6.22(c)). By Proposition 2.6.20(c), T is bijective if and only if T_A is, and by Theorem 2.7.8, (I) $\Leftrightarrow$ (III), if and only if the columns of A form a basis of F^m. Finally, a bijective linear map is invertible (Proposition 2.7.7), and an invertible one is injective and surjective (Proposition 2.5.4, Proposition 2.5.5). □

So injectivity, surjectivity, and invertibility of T can be read off any one matrix of T, whatever bases were used to compute it.

Example 2.7.17 (Differentiation plus the identity). Let $T : F[x]_{\le 2} \to F[x]_{\le 2}$, $T(p) = p + D(p)$, where D is differentiation (Example 2.1.9). On the basis $\mathcal{B} = 1, x, x^2$,

$$T(1) = 1, \quad T(x) = 1 + x, \quad T\big(x^2\big) = 2x + x^2,$$

so

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

The system $[T]_{\mathcal{B}\mathcal{B}}\mathbf{c} = \mathbf{y}$ reads $c_1 + c_2 = y_1$, $c_2 + 2c_3 = y_2$, $c_3 = y_3$, and working from the bottom up it has exactly one solution,

$$c_3 = y_3, \quad c_2 = y_2 - 2y_3, \quad c_1 = y_1 - y_2 + 2y_3.$$

By Theorem 2.7.8, (II) $\Rightarrow$ (III), the columns of the matrix form a basis of F^3, so by Corollary 2.7.16(c), T is invertible and $1, 1 + x, 2x + x^2$ is a basis of $F[x]_{\le 2}$, over every field. The solution formula gives the inverse: for $q = y_1 + y_2 x + y_3 x^2$,

$$\begin{aligned} T^{-1}(q) &= (y_1 - y_2 + 2y_3) + (y_2 - 2y_3)x + y_3 x^2 \\ &= q - D(q) + D^2(q). \end{aligned}$$

So $T^{-1} = D^2 - D + \mathrm{id}$. This is the geometric series of Technique 2.5.16, cut off because $D^3 = 0$ on $F[x]_{\le 2}$: indeed $T = D + \mathrm{id}$, and $(D + \mathrm{id})\big(D^2 - D + \mathrm{id}\big) = D^3 + \mathrm{id} = \mathrm{id}$.

2.7.7 The theorems side by side

Table 2.7.1 lines up the three theorems condition by condition. In each row, the entry for the third theorem is the combination of the other two.

	First	Second	Third
(I)	injective	surjective	bijective
(II)	at most one	at least one	exactly one
(III)	independent	span F^m	basis of F^m
(IV)	$\ker T_A = \{\mathbf{0}\}$	$\operatorname{im} T_A = F^m$	both
(V)	left inverse	right inverse	inverse
(VI)	$\operatorname{rank} A = n$	$\operatorname{rank} A = m$	$\operatorname{rank} A = m = n$
size	$n \leq m$	$n \geq m$	$n = m$

Table 2.7.1. The three Fundamental Theorems (Theorem 2.7.2, Theorem 2.7.4, Theorem 2.7.8) for $A \in F^{m\times n}$. Rows (I), (IV), and (V) are about the map T_A; row (II) gives the number of solutions of $A\mathbf{x} = \mathbf{y}$, for every $\mathbf{y} \in F^m$; row (III) is about the columns of A. The last two rows are added in Section 3.5.2.

The lecture's statements of the theorems contain two more pieces. Each theorem has a sixth condition, about a single number computed from A by row reduction, and each ends with a constraint on the shape of A: "moreover, $n \leq m$" for the first theorem, $n \geq m$ for the second, and $n = m$ for the third. The sixth condition needs row reduction. The constraints on the shape can already be proved with the tools of this chapter (Exercise 2.7.32), but the text proves both pieces together in Section 3.5.2. There the rank of A is defined as the number of pivots of $\operatorname{rref}(A)$ (Definition 3.5.1); the sixth conditions are $\operatorname{rank} A = n$, $\operatorname{rank} A = m$, and $\operatorname{rank} A = m = n$ (Theorem 3.5.12, Theorem 3.5.13, Theorem 3.5.14); and the constraints on the shape follow from $\operatorname{rank} A \leq \min(m, n)$. In particular, invertible matrices are square (Corollary 3.5.15).

Intuition (Why the shape is constrained). An injective T_A sends the n independent vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to the n independent columns of A in F^m, and F^m should not have room for more than m independent vectors. A surjective T_A produces all of F^m from combinations of n columns, and m "directions" should need at least m vectors. Making "room" and "directions" precise is what dimension does (Section 4.1). With what we already have, the constraints are within reach by an induction on m (Exercise 2.7.32), and over $\mathbb{F}_p$ by counting (Exercise 2.6.44).

The dictionary in Table 3.1.1 arranges the same facts by the three ways of reading a linear system: as equations, as the map T_A, and as a question about the columns.

Remark 2.7.18 (On the name). The name follows the lecture, and the course notes abbreviate the three theorems FTLA1, FTLA2, and FTLA3. Elsewhere, "fundamental theorem" refers to other results. In LADR, the *fundamental theorem of linear maps* is rank–nullity, $\dim V = \dim \ker T + \dim \operatorname{im} T$ (Theorem 4.2.11). Strang's *fundamental theorem of linear algebra* describes the four subspaces attached to a matrix and their dimensions (Definition 4.3.12). The statements of this section are closest to what many books call the *invertible matrix theorem*, a list of conditions equivalent to invertibility of a square matrix (Corollary 3.5.25). The Fundamental Theorems split it into an injective half and a surjective half, and do not assume that A is square.

Exercises

Exercise 2.7.19. ROUTINE For each matrix, decide which of the three Fundamental Theorems have their conditions satisfied, and justify your answer by checking one convenient condition of each theorem.

(a) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$ over $\mathbb{Q}$.

(b) $\begin{pmatrix} 1 & i \\ i & -1 \\ 0 & 1 \end{pmatrix}$ over $\mathbb{C}$.

(c) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ over $\mathbb{F}_2$.

(d) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ over $\mathbb{F}_3$.

Solution on page 1172

Exercise 2.7.20. ROUTINE

(a) Let $T : F[x]_{\leq 2} \to F^3$, $T(p) = (p(0), (Dp)(0), p(1))$, where D is differentiation. Use Corollary 2.7.16 to show that T is invertible over every field, and find T^{-1}.

(b) Let $U : F[x]_{\leq 2} \to F^3$, $U(p) = (p(0), (Dp)(0), (Dp)(1))$. For which fields F is U invertible?

Solution on page 1173

Exercise 2.7.21. ROUTINE Find the error in the following argument, and show that its claim is false.

Claim. For $A \in F^{m\times n}$, the following are equivalent: (i) T_A is injective; (ii) the columns of A are linearly independent; (iii) T_A is surjective; (iv) the columns of A span F^m.

Proof. (i) $\Leftrightarrow$ (ii) is Corollary 2.6.24(a), and (iii) $\Leftrightarrow$ (iv) is Corollary 2.6.24(b). Hence the four statements are equivalent.

Solution on page 1173

Exercise 2.7.22. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$. Prove or disprove each statement.

(a) If $A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$, then $A\mathbf{x} = \mathbf{y}$ has a solution for every $\mathbf{y} \in F^m$.

(b) If A has two different left inverses, then T_A is not surjective.

(c) If T_{AB} is bijective, then T_A and T_B are bijective.

(d) If T_{AB} is bijective, then the columns of A span F^m and the columns of B are linearly independent.

Solution on page 1174

Exercise 2.7.23. STANDARD Let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$. Show that the following are equivalent: (i) P is injective; (ii) P is surjective; (iii) P is invertible; (iv) $P = \mathrm{id}_V$. Here V need not have a finite basis.

Solution on page 1174

Exercise 2.7.24. STANDARD Find the error in the following argument, and repair it.

Claim. If $T_A : F^n \to F^m$ is injective, then T_A has a left inverse.

Proof. Regard T_A as a map $F^n \to \operatorname{im} T_A$. As such it is linear, injective, and surjective, so by Proposition 2.7.7 it has a linear inverse $R : \operatorname{im} T_A \to F^n$, with $R(T_A(\mathbf{x})) = \mathbf{x}$ for every $\mathbf{x} \in F^n$. So R is a left inverse of T_A.

Solution on page 1174

Exercise 2.7.25. STANDARD Let $A \in F^{m\times n}$. The columns of A^{T} are the rows of A, read as vectors in F^n. Show that:

(a) the columns of A are linearly independent if and only if the rows of A span F^n;

(b) the columns of A span F^m if and only if the rows of A are linearly independent.

Solution on page 1175

Exercise 2.7.26. STANDARD Let p be a prime.

(a) Show that exactly $(p^3-1)(p^3-p)$ linear maps $\mathbb{F}_p^2 \to \mathbb{F}_p^3$ are injective.

(b) By Exercise 2.6.45, the same number of linear maps $\mathbb{F}_p^3 \to \mathbb{F}_p^2$ are surjective. Explain the coincidence with Exercise 2.7.25.

Solution on page 1175

Exercise 2.7.27. STANDARD Let $A \in F^{m\times n}$. Prove or disprove: if the columns of A are linearly independent, then so are the columns of $A^{\mathsf{T}}A \in F^{n\times n}$.

(a) For $F = \mathbb{R}$. (b) For $F = \mathbb{C}$. (c) For $F = \mathbb{F}_2$.

Solution on page 1176

Exercise 2.7.28. STANDARD For $c \in F$, let

$$A_c = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & c \\ c & 0 & 1 \end{pmatrix}.$$

(a) Show that A_c is invertible if and only if $1 + c^3 \neq 0$, and that otherwise T_{A_c} is neither injective nor surjective.

(b) List the c for which A_c is not invertible over $\mathbb{R}$, over $\mathbb{C}$, and over $\mathbb{F}_7$.

Solution on page 1176

Exercise 2.7.29. HARD Let $A \in F^{m\times n}$.

(a) Show that A has exactly one left inverse if and only if A is invertible.

(b) Show that A has exactly one right inverse if and only if A is invertible.

Solution on page 1177

Exercise 2.7.30. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $AB = I_m$.

(a) Show that if T_B is surjective, then $BA = I_n$.

(b) Show that if T_A is injective, then $BA = I_n$.

Solution on page 1177

Exercise 2.7.31. STANDARD Let $n \geq 2$, let $J \in F^{n\times n}$ be the matrix with every entry 1, and let $A = J - I_n$, the matrix with 0 on the diagonal and 1 elsewhere. Write n also for $n * 1 \in F$.

(a) Show that $J^2 = nJ$.

(b) Show that if $n - 1 \neq 0$ in F, then A is invertible with $A^{-1} = (n-1)^{-1}J - I_n$.

(c) Show that if $n - 1 = 0$ in F, then T_A is neither injective nor surjective.

(d) For which primes p is the 5×5 matrix A invertible over $\mathbb{F}_p$?

Solution on page 1178

Exercise 2.7.32. HARD This exercise proves the constraints on the shape of A without rank.

(a) Show by induction on $m \geq 1$ that every list of $m+1$ vectors in F^m is linearly dependent.

(b) Deduce that if $T_A : F^n \to F^m$ is injective, then $n \leq m$; that if T_A is surjective, then $n \geq m$; and that if A is invertible, then $m = n$.

Solution on page 1178

Exercise 2.7.33. ROUTINE Let $A \in F^{m\times n}$, and let $G \in F^{m\times m}$ and $H \in F^{n\times n}$ be invertible. Show that T_{GAH} is injective if and only if T_A is, and surjective if and only if T_A is. Deduce that GAH is invertible if and only if A is.

Solution on page 1179

Exercise 2.7.34. STANDARD Let $A \in F^{m\times n}$ and $p \geq 1$, and let $L_A : F^{n\times p} \to F^{m\times p}$, $L_A(X) = AX$.

(a) Show that L_A is linear.

(b) Show that L_A is injective if and only if T_A is injective.

(c) Show that L_A is surjective if and only if T_A is surjective.

(d) Deduce that L_A is invertible if and only if A is invertible.

Solution on page 1179

Exercise 2.7.35. CHALLENGE For a matrix A with integer entries and a prime p, let $\overline{A} \in \mathbb{F}_p^{m\times n}$ be the matrix of remainders of the entries of A modulo p.

(a) Let $A \in \mathbb{Z}^{m\times n}$. Show that if the columns of $\overline{A}$ are linearly independent over $\mathbb{F}_p$, then the columns of A are linearly independent over $\mathbb{Q}$.

(b) Show that every $n \times n$ matrix with odd integers on the diagonal and even integers off the diagonal has linearly independent columns over $\mathbb{Q}$.

(c) Show that the converse of (a) is false.

Solution on page 1180

2.8 Proof patterns

A linear map can be handled at four levels: as a map, through its values on a basis, through a matrix, or through the entries of that matrix. Once bases are fixed, all four carry the same information (Section 2.3.6), so most proofs in this chapter begin by choosing a level, and the choice decides how long the proof is. This section sorts the chapter's proofs by the question they answer, compares the levels on examples from the chapter, and says when a tool is the wrong one. The exercises at the end practice one move at a time, and the problem set (Problems for Chapter 2) mixes them.

2.8.1 Showing that a map is linear, or building one

There are three moves, and they answer different questions.

- *Check the definition.* Expand $T(a\mathbf{u} + \mathbf{v})$ and regroup, as in (iii) of Theorem 2.1.3. This verifies a map that is given by a formula, and it is the default for formulas in coordinates or coefficients (Example 2.1.7, Example 2.1.9).
- *Assemble the map* from sums, scalar multiples, and composites of maps already known to be linear (Technique 2.1.26). This transfers linearity without computing, but it only covers the operations of $\mathrm{L}(V, W)$. A pointwise product of two linear maps, such as $x \mapsto x \cdot x$, is not covered, and neither is an inverse function, whose linearity needed its own argument (Proposition 2.7.7).

- *Define the map on a basis* and extend it (Technique 2.2.4). This creates a map with prescribed behaviour, and linearity is free. It does not show that a given formula is linear: to identify a given map with the one built on the basis, we have to evaluate the given map at a general vector, which is the first move again. The exception is a formula that already has the form (2.2.1), as for the maps of $F[x]$ that move each coefficient to a fixed place (Proposition 2.2.20).

Constructions go wrong at the choices. A map defined by choosing a value for each vector separately is usually not linear (Non-example 2.7.10), and values prescribed on a spanning list must respect its linear relations (Proposition 2.2.11, Exercise 2.2.24, Exercise 2.2.26). Choices made on a basis are always safe. To show that a map is *not* linear, the tests $T(\mathbf{0})$ and $T(2\mathbf{v})$ catch most formulas, but not over $\mathbb{F}_2$, where $2 = 0$ and every function that fixes $\mathbf{0}$ is homogeneous. There, test additivity (Exercise 2.1.15(d), Exercise 2.8.5).

Example 2.8.1 (One map, three routes). The reflection S_ℓ across a line ℓ through $\mathbf{0}$ appeared three times in the chapter. In Example 2.1.11 it came with a coordinate formula, obtained from the double angle formulas, and the formula has the shape of Example 2.1.7, so S_ℓ is linear by the first move. In Example 2.1.27(a) we noticed that $S_\ell = 2\,\mathrm{Pr}_\ell - \mathrm{id}_{\mathbb{R}^2}$, which proves linearity with no trigonometry, once Pr_ℓ is known to be linear. In Example 2.2.3, for the line $y = x$, we chose $\mathbf{b}_1 = (1, 1)$ on the line and $\mathbf{b}_2 = (1, -1)$ across it, and *defined* S by $S(\mathbf{b}_1) = \mathbf{b}_1$ and $S(\mathbf{b}_2) = -\mathbf{b}_2$. Linearity came for free, the formula $S(x, y) = (y, x)$ came out of (2.2.1), and identities such as $S^2 = \mathrm{id}$ became one-line checks on the basis (Example 2.2.8).

The formula verifies a map that is already given, assembly transfers linearity from maps we trust, and the basis creates a map from a description of what it should do. The last route has a cost: the basis must exist and must be found. Over $\mathbb{F}_2$ the vectors $(1, 1)$ and $(1, -1)$ coincide, and the construction of Example 2.2.3 breaks down, although $(x, y) \mapsto (y, x)$ is still linear there.

Practice: Exercise 2.8.5 and Exercise 2.8.7.

2.8.2 Showing that two maps or two matrices are equal

For linear maps, the default is to compare them on a basis, or on any spanning list (Technique 2.2.7, Corollary 2.2.6); on $F[x]$ the monomials play the role of the basis (Proposition 2.2.20). Matrices can be compared entry by entry, column by column ($A = B$ as soon as $A\mathbf{e}_j = B\mathbf{e}_j$ for every j, by Lemma 2.4.24(b)), or through their maps ($A = B$ as soon as $T_A = T_B$, by Proposition 2.3.17). Two maps are also equal when their matrices in one pair of bases are (Theorem 2.3.20(a)). Often the shortest route compares no values at all and derives the identity from identities already known, with the algebra of Proposition 2.1.24 and Proposition 2.4.14.

Before checking on a basis, check two things. Both sides must be linear in the vector being tested. An identity between sums and composites of linear maps is, but a condition such as "$T(\mathbf{v}) \in \mathrm{span}(\mathbf{v})$" is not, and it can hold on a basis without holding everywhere (Exercise 2.2.25). And the list must be a basis over the field at hand: $(1, 1), (1, -1)$ is a basis of F^2 only when $1 + 1 \neq 0$ (Exercise 2.3.31).

Example 2.8.2 ($S^2 = \mathrm{id}$, three ways). Let ℓ be the line $y = 2x$ in $\mathbb{R}^2$ and $S = S_\ell$, whose standard matrix is $\frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$ (Example 2.3.19).

Entries. $\frac{1}{25}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}^2 = \frac{1}{25}\begin{pmatrix} 9+16 & -12+12 \\ -12+12 & 16+9 \end{pmatrix} = I_2$. This settles one line; for a general line the same computation needs $\cos^2\theta + \sin^2\theta = 1$ again.

On an adapted basis. With $\mathbf{u}_1 = (1, 2)$ on ℓ and $\mathbf{u}_2 = (-2, 1)$ across it, $S(\mathbf{u}_1) = \mathbf{u}_1$ and $S(\mathbf{u}_2) = -\mathbf{u}_2$ (Example 2.3.23). So S^2 fixes both basis vectors, and $S^2 = \mathrm{id}$ by Corollary 2.2.6. This explains the identity: reflecting twice flips the perpendicular direction twice.

By algebra. $S = 2P - \mathrm{id}$ with $P = \mathrm{Pr}_\ell$ and $P^2 = P$ (Exercise 2.1.42), so by Proposition 2.1.24,

$$S^2 = 4P^2 - 4P + \mathrm{id} = \mathrm{id}.$$

This uses no basis and no coordinates, and it proves more: $(2P - \mathrm{id})^2 = \mathrm{id}$ for every operator P with $P^2 = P$, on every vector space over every field.

Each route is the wrong one somewhere. The entries do not show why the identity holds, the adapted basis must be found first, and the algebra needs the right identity to start from. When no convenient identity is at hand, as for $DM - MD = \mathrm{id}$ on $F[x]$, the computation has to be done, and the monomials are the natural place to do it (Exercise 2.4.51); the product rule gives another proof, but it is a fact about products of polynomials, which is more than the identity needs.

Practice: Exercise 2.8.6 and Exercise 2.8.8.

2.8.3 Proving a matrix identity

The chapter proved its matrix identities in four ways, and many identities admit more than one.

- *Translate to maps* (Technique 2.3.27). If both sides are built from products, sums, and scalar multiples, read them as composites and use a fact about maps. Associativity and distributivity came this way (first proof of Proposition 2.4.14), and all the work went into the dictionary, Theorem 2.4.2. The translation does not see entries, which depend on the bases while maps do not. So it does not reach the transpose, which needs more structure to mean anything for a map (Section 13.1).
- *Compute entry by entry* (Technique 2.4.15). This is the tool for statements about entries: transposes (Proposition 2.4.23), traces (Proposition 2.4.28), triangular shapes (Proposition 2.4.36). It is mechanical, and it uses only the ring laws of the entries, which is why it also proves associativity for block matrices, whose entries do not commute (second proof of Proposition 2.4.14).
- *Use a view of the product* (Proposition 2.4.8). Statements about the columns or rows of a product take one line in the right view: right multiplication recombines columns (Exercise 2.4.11), and left multiplication by a diagonal matrix scales rows (Example 2.4.10).
- *Split into blocks* (Technique 2.4.35). This exploits zero blocks and identity blocks, and it sets up induction on the size.

When the identity is linear in one of the matrices, Technique 2.2.7 on the matrix units is a fifth option, as in the second proof of Proposition 2.4.28.

Example 2.8.3 (Strictly upper triangular matrices are nilpotent, three ways). Let $N \in F^{n\times n}$ be strictly upper triangular, $N_{ij} = 0$ whenever $i \geq j$. We show $N^n = 0$.

Maps. Let $V_k = \mathrm{span}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$, with $V_0 = \{\mathbf{0}\}$. Column j of N is zero from row j down, so $N\mathbf{e}_j \in V_{j-1}$. Hence N maps V_k into V_{k-1} (Proposition 2.1.16), N^k maps V_k into V_0, and N^n kills $F^n = V_n$. This is Exercise 2.2.36 for T_N.

Entries. We claim $\left(N^k\right)_{ij} = 0$ whenever $j - i < k$. For $k = 1$ this is the hypothesis. If it holds for k, then a term of

$$\left(N^{k+1}\right)_{ij} = \sum_{l=1}^{n} \left(N^k\right)_{il} N_{lj}$$

can be nonzero only if $l - i \geq k$ and $j - l \geq 1$, so only if $j - i \geq k + 1$. For $k = n$, every pair has $j - i \leq n - 1 < n$.

Blocks. We induct on n; for $n = 1$, $N = 0$. For $n \geq 2$, write $N = \begin{pmatrix} 0 & \mathbf{r}^{\mathsf{T}} \\ \mathbf{0} & N' \end{pmatrix}$ with $\mathbf{r} \in F^{n-1}$ and N' strictly upper triangular of size $n - 1$. Block multiplication and induction on k give $N^k = \begin{pmatrix} 0 & \mathbf{r}^{\mathsf{T}} N'^{k-1} \\ \mathbf{0} & N'^k \end{pmatrix}$, and by the induction hypothesis $N'^{n-1} = 0$, so $N^n = 0$.

The map proof explains the result: each application of N pushes vectors one step down the chain $V_n \supseteq \cdots \supseteq V_0$, so N^k maps V_j into V_{j-k}, which is exactly the band of zeros $\left(N^k\right)_{ij} = 0$ for $j - i < k$ that the entrywise proof computes. In the form of Exercise 2.2.36 it needs no particular basis: any basis in which each vector is sent into the span of the earlier ones will do, while the entrywise proof needs the matrix already in triangular shape. The entrywise proof, in turn, is the most mechanical, and needs no subspaces at all. For $n \geq 2$ the shift $N = \sum_{i<n} \mathbf{e}_i \mathbf{e}_{i+1}^{\mathsf{T}}$ has $\left(N^{n-1}\right)_{1n} = 1$, so the bound n cannot be lowered. The block proof is the longest here, but it is the template for every later induction on the size of a triangular matrix (Proposition 2.4.36).

Practice: Exercise 2.8.9 and Exercise 2.8.12.

2.8.4 Proving injectivity, surjectivity, or invertibility

The chapter collected certificates for each property.

- *Injective:* a nonzero vector has a nonzero value (the kernel test, Technique 2.6.12); the values on a basis are independent (Proposition 2.6.22(a)); a left inverse (Proposition 2.5.4); an injective composite TS, which makes S injective (Proposition 2.6.15(c)).
- *Surjective:* a solution of $T(\mathbf{v}) = \mathbf{w}$ for a general $\mathbf{w}$; the values on a spanning list span (Proposition 2.6.22(b), Remark 2.6.23); a right inverse (Proposition 2.5.5); a surjective composite TS, which makes T surjective.
- *Invertible:* a candidate inverse, checked on both sides (Technique 2.5.16); injective and surjective, after which the inverse is linear for free (Proposition 2.7.7); the values on a basis form a basis (Proposition 2.6.22(c)); an invertible matrix in some pair of bases (Proposition 2.5.20, Corollary 2.7.16); for a 2×2 matrix, $ad - bc \neq 0$ (Proposition 2.5.28).

The kernel test is usually the shortest, but it proves injectivity and nothing else. Nothing in this chapter turns injectivity into surjectivity. For square matrices that is a theorem of the next chapter (Theorem 3.5.23), for operators on finite-dimensional spaces it is Corollary 4.2.18, and on $F[x]$ it is false (Example 2.6.4(e)).

Example 2.8.4 (One operator, four certificates). Let $T = \mathrm{id} + D$ on $F[x]$, over any field, where D is differentiation.

The kernel test. For $p \neq \mathbf{0}$ the degrees of p and $D(p)$ differ, so $T(p)$ has the degree of p and is nonzero (Example 2.6.13(a)). So T is injective, and that is all this argument gives.

Values on a basis. On $F[x]_{\leq 2}$, T sends $1, x, x^2$ to $1, 1+x, 2x+x^2$. Their span contains 1, then $x = (1+x) - 1$, then $x^2 = \left(2x + x^2\right) - 2x$, so it is $F[x]_{\leq 2}$. Together with the kernel test, Proposition 2.6.22 shows that T is bijective on $F[x]_{\leq 2}$, and the same works on each $F[x]_{\leq n}$. Since every polynomial lies in some $F[x]_{\leq n}$, which T maps onto itself, T is surjective on $F[x]$ as well. This reaches $F[x]$ only through its finite pieces, and it gives no formula for the inverse.

A guessed inverse. The series $1/(1+t) = 1 - t + t^2 - \cdots$ suggests $S = \mathrm{id} - D + D^2 - \cdots$. On a polynomial of degree at most d the series stops, since $D^{d+1}(p) = \mathbf{0}$ (Example 2.1.19(a)). So set $S(p) = \sum_{k=0}^{d} (-1)^k D^k(p)$ for any $d \geq \deg p$. The extra terms are $\mathbf{0}$, so the choice of d does not matter, and S agrees on each $F[x]_{\leq d}$ with the linear map $\sum_{k \leq d} (-1)^k D^k$, so S is linear. For $\deg p \leq d$ the sum telescopes:

$$\begin{aligned} T(S(p)) &= \sum_{k=0}^{d} (-1)^k D^k(p) + \sum_{k=0}^{d} (-1)^k D^{k+1}(p) \\ &= p + (-1)^d D^{d+1}(p) = p, \end{aligned}$$

and $S(T(p)) = p$ in the same way, since $T(p)$ also has degree at most d. By Technique 2.5.16, T is invertible on all of $F[x]$, with $T^{-1} = \mathrm{id} - D + D^2 - \cdots$.

A matrix. In the basis $1, x, x^2$ of $F[x]_{\leq 2}$, the matrix of T is upper triangular with every diagonal entry equal to 1, and back substitution solves every system (Example 2.7.17).

The kernel test cannot finish on $F[x]$, and the basis argument needs pieces $F[x]_{\leq n}$ that T maps into themselves. The guessed inverse works on all of $F[x]$ at once and gives a formula. The heuristic behind it is not a proof. The same guess for $\mathrm{id} + M$, with M multiplication by x, gives $\mathrm{id} - M + M^2 - \cdots$, which never stops; $\mathrm{id} + M$ does not map any $F[x]_{\leq n}$ into itself either, and indeed it is injective but not surjective: every value $(1 + x)q$ vanishes at -1, and the constant 1 does not (Example 2.6.29). The series becomes a proof when it terminates.

Practice: Exercise 2.8.10 and Exercise 2.8.15.

2.8.5 Proving that conditions are equivalent

The chapter's equivalence theorems use each shape of Technique 2.7.1. Theorem 2.1.3 is a cycle: (i) ⇒ (iii) ⇒ (ii) are choices of special values, and (ii) ⇒ (i) is an induction on the length of a combination. Theorem 2.7.2 and Theorem 2.7.4 are hubs: the three spokes and the arrow (V) ⇒ (I) are earlier results read in the right language (Remark 2.7.3), and one arrow, the construction of a one-sided inverse, needs an idea. Theorem 2.7.8 inherits the equivalence of (I)–(IV) from the first two, since each of these conditions is the conjunction of the corresponding conditions there, and only (V) needs an argument of its own. Corollary 2.7.16 carries all three from F^n to spaces with bases by composing with coordinate maps.

Two decisions shape such a proof. The first is where to spend the construction. Arrows out of a condition that hands us an object, such as a left inverse, are short (Proposition 2.5.4). Arrows into it are the hard ones, so they should start from whichever equivalent condition makes the construction easiest. For a left inverse that is (III) of Theorem 2.7.2, since independent columns can be completed to a basis. For a right inverse it is (II) of Theorem 2.7.4, since the solutions of $A\mathbf{x} = \mathbf{e}_i$ are exactly the values the construction needs. The second decision is to check connectivity. Two separately proved groups of equivalent statements say nothing about each other (Exercise 2.7.21), and no arrow may use the theorem it belongs to. Practice: Exercise 2.8.11.

2.8.6 Reading the chapter's proofs

Definitions that are theorems in disguise. The products $A\mathbf{x}$ and AB were not chosen. Each was derived from a statement we wanted to be true, Theorem 2.3.9 and Theorem 2.4.2, and shown to be the only choice that makes it true (Proposition 2.3.10, Proposition 2.4.3). So the entry formula $\sum_k a_{ik}b_{kj}$ is a consequence, not a convention to memorize, and every algebraic law of the product is a law of composition seen through coordinates. Associativity is the clearest case. The composite of three functions does not depend on how they are grouped, for trivial reasons, so $(AB)C = A(BC)$ as soon as the dictionary is in place (first proof of Proposition 2.4.14). The failures come from composition too: $AB \neq BA$ and zero divisors are already visible for maps of the plane (Example 2.1.19, Exercise 2.1.40). The entrywise proof of associativity looks more elementary, but its reach is different: it uses only the ring laws of the entries, so it covers block matrices, where no single linear map is in sight (Section 2.4.8).

Uniqueness first, then existence. The proof of Theorem 2.2.1 assumed that a linear map with the given values exists and computed what it had to be. That proved uniqueness and produced the only candidate, which was then checked. The same move found the inverse of a bijective linear map, which can only be the inverse function (Proposition 2.7.7), and it compared a left and a right inverse through the single expression $S_l T S_r$ (Theorem 2.5.10).

Well-definedness. A definition that refers to a choice must not depend on it. When values are prescribed on a spanning list, a vector may have several representations, and the first proof of Proposition 2.2.11 checks

that any two give the same value: two representations differ by a linear relation, and the hypothesis says that the prescribed values respect it. Skipping this check is the error in Exercise 2.2.24.

Which hypothesis was used where. Remark 2.2.2 and Remark 2.6.23 record which half of "basis" each argument used, and the record tells us how far a result extends: spanning alone gives uniqueness of a map and the surjectivity test, and independence alone gives existence on the span. In the same way, Proposition 2.4.30 uses $n * 1 \neq 0$ at exactly one step, and over $\mathbb{F}_p$ with p dividing n that step fails (Remark 2.4.31).

Forced and free. A left inverse of T is forced on the image of T and free off it. That is why an injective map that is not surjective has many left inverses, and why two different left inverses rule out a right inverse (Corollary 2.5.11, Remark 2.7.11). Counting the free choices counts the one-sided inverses (Problem 2.3).

2.8.7 Common errors

Most of these appear in the chapter as find-the-error exercises.

- *Commuting factors.* In general $AB \neq BA$ (Example 2.4.18), so $(A+B)^2 \neq A^2 + 2AB + B^2$ (Exercise 2.4.17) and $(AB)^2 \neq A^2B^2$ (Exercise 2.4.44(c)). Inside a trace, a product may be rotated but not rearranged (Exercise 2.4.32).
- *Cancelling.* $AB = AC$ gives $B = C$ when A is left-invertible (Corollary 2.5.6), but not in general (Example 2.4.19), and $AB = 0$ does not force $A = 0$ or $B = 0$ (Exercise 2.4.20). The idempotent version, $P(P - \mathrm{id}) = 0$ hence $P \in \{0, \mathrm{id}\}$, is Exercise 2.1.41, and cancelling on the side where no inverse exists is Exercise 2.8.15.
- *Inverting or transposing a product.* $(AB)^{-1} = B^{-1}A^{-1}$ and $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$, in reversed order (Exercise 2.8.14, Exercise 2.4.48).
- *One side for two.* A one-sided inverse need not be an inverse: $AB = I_m$ does not give $BA = I_n$ (Exercise 2.5.34), and on $\mathbb{Q}[x]$, $DJ = \mathrm{id}$ while $JD \neq \mathrm{id}$ (Example 2.5.24).
- *Injective for surjective.* For operators on $F[x]$ neither implies the other: multiplication by x is injective but not surjective, and its left inverse L of Exercise 2.5.41 is surjective but not injective (Example 2.6.4(e), Exercise 2.6.37(e)). For square matrices the implication is a theorem that this chapter does not have (Theorem 3.5.23).
- *Swapped bases.* $[T]_{\mathcal{C}\mathcal{B}}$ takes $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates, so $[\mathrm{id}]_{\mathcal{E}\mathcal{B}}$ converts $\mathcal{B}$-coordinates into standard ones and not the other way (Exercise 2.3.30). In a product of matrices of maps, the inner subscripts must match (Exercise 2.8.13, Problem 2.2).
- *Tests that are not tests.* A list that is a basis over one field need not be one over another (Exercise 2.3.31); a condition that is not linear in the vector cannot be checked on a basis (Exercise 2.2.25); and nonzero values on a basis do not make a map injective (Exercise 2.6.40).
- *Values on a dependent list.* They must respect every relation of the list (Exercise 2.2.24, Exercise 2.2.26).
- *Integer multiples for all scalars.* Additivity gives $T(n\mathbf{v}) = nT(\mathbf{v})$ for integers n, not for all scalars (Exercise 2.1.35).
- *Partial equivalences and partial inverses.* Equivalences proved inside two groups of conditions do not connect the groups (Exercise 2.7.21), and an inverse defined on the image of T only is not a left inverse of T (Exercise 2.7.24).
- *The kernel test without linearity.* On $\mathbb{R}$, $x \mapsto x^2$ sends only 0 to 0 and is not injective; for maps that are homogeneous but not additive the test fails as well (Problem 2.1).

Exercises

Exercise 2.8.5. ROUTINE Decide which of the following maps are linear, and give the shortest argument you can.

(a) $F^{n\times n} \to F^{n\times n}$, $X \mapsto AXB - X^{\mathsf{T}}$, where $A, B \in F^{n\times n}$ are fixed.

(b) $F[x] \to F[x]$, $p \mapsto p(x+1) - 2p + p(x-1)$.

(c) $F^{n\times n} \to F$, $X \mapsto \operatorname{tr}(X^2)$, where $n \geq 1$. (The answer depends on F.)

Solution on page 1180

Exercise 2.8.6. STANDARD Let $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} \in F^{n\times n}$ be the matrix units.

(a) Show that $\sum_{i=1}^n \sum_{j=1}^n U_{ij} A U_{ji} = (\operatorname{tr} A) I_n$ for every $A \in F^{n\times n}$, twice: by checking on the matrix units, and entry by entry.

(b) Suppose A commutes with every matrix in $F^{n\times n}$. Show that $(n * 1)A = (\operatorname{tr} A) I_n$. Deduce that if $n * 1 \neq 0$ in F, the matrices that commute with every matrix are exactly the scalar multiples of I_n.

Solution on page 1181

Exercise 2.8.7. STANDARD In $\mathbb{Q}^4$, let $K = \operatorname{span}((1,1,0,0),(0,0,1,1))$ and $L = \operatorname{span}((1,0,1,0),(0,1,0,1))$.

(a) Construct $T \in \mathrm{L}(\mathbb{Q}^4)$ with $\ker T = K$ and $\operatorname{im} T = L$, and give its standard matrix.

(b) Show that no $T \in \mathrm{L}(\mathbb{Q}^4)$ with $\ker T = K$ and $\operatorname{im} T = L$ satisfies $T^2 = T$.

Solution on page 1182

Exercise 2.8.8. STANDARD Let p be a prime and D differentiation on $\mathbb{F}_p[x]$.

(a) Show that $D^p = 0$.

(b) Let $n \geq 1$, and let N be the matrix of D on $\mathbb{F}_p[x]_{\leq n}$ with respect to $1, x, \ldots, x^n$. Show that $N^p = 0$ without multiplying matrices. Show also that over $\mathbb{Q}$ the corresponding matrix satisfies $N^n \neq 0$.

(c) Show that $\mathrm{id} + D$ is invertible on $\mathbb{F}_p[x]$, with inverse $(\mathrm{id} + D)^{p-1}$. You may use the binomial theorem for commuting operators, which has the same proof as for numbers, and that p divides $\binom{p}{k}$ for $0 < k < p$.

Solution on page 1182

Exercise 2.8.9. STANDARD Let $p, q, r \geq 1$, let $A \in F^{p\times q}$, $B \in F^{p\times r}$, and $C \in F^{q\times r}$, and let

$$M = \begin{pmatrix} I_p & A & B \\ 0 & I_q & C \\ 0 & 0 & I_r \end{pmatrix}.$$

(a) Write $M = I + N$, compute N^2 and N^3 by block multiplication, and deduce M^{-1}.

(b) For numbers,

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix},$$

and the corner entry $ac - b$ could also be written $ca - b$. Which block version is correct, and why does the other one fail?

Solution on page 1183

Exercise 2.8.10. STANDARD Let $c \in F$, and let M and D be multiplication by x and differentiation on $F[x]$.

(a) Show that $M - c\,\mathrm{id}$ is injective but not surjective.

(b) Suppose $c \neq 0$. Show that $D - c\,\mathrm{id}$ is invertible, and find its inverse.

(c) Explain why the kernel test alone could not have decided (b).

Solution on page 1184

Exercise 2.8.11. STANDARD Let $n \geq 1$, $A \in F^{n\times n}$, and $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$. Show that the following are equivalent, and arrange the proof so that as many arrows as possible are citations of earlier results.

(i) $A = \lambda I_n$ for some $\lambda \in F$.
(ii) $AX = XA$ for every $X \in F^{n\times n}$.
(iii) $AU_{ij} = U_{ij}A$ for all i, j.
(iv) $A\mathbf{x} \in \operatorname{span}(\mathbf{x})$ for every $\mathbf{x} \in F^n$.
(v) $[T_A]_{\mathcal{B}\mathcal{B}} = A$ for every basis $\mathcal{B}$ of F^n of length n.

Solution on page 1185

Exercise 2.8.12. STANDARD Let $n \geq 1$.

(a) Show that there are no $A, B \in \mathbb{R}^{n\times n}$ with $AB - BA = AA^{\mathsf{T}} + I_n$.

(b) Show that the statement of (a) is false over $\mathbb{C}$ and over $\mathbb{F}_2$.

(c) Show that $\operatorname{tr}\left(AA^{\mathsf{T}}\right) = 0$ forces $A = 0$ for $A \in \mathbb{R}^{n\times n}$, but find nonzero $A \in \mathbb{C}^{2\times 2}$ and $A \in \mathbb{F}_2^{2\times 2}$ with $\operatorname{tr}\left(AA^{\mathsf{T}}\right) = 0$.

Solution on page 1185

Exercise 2.8.13. ROUTINE Let $T \in \mathrm{L}\left(\mathbb{Q}^2\right)$, $T(x, y) = (y, x)$, and let $\mathcal{B}$ be the basis $\mathbf{b}_1 = (1, 0)$, $\mathbf{b}_2 = (1, 1)$. Find the error in the following computation, and compute $[T]_{\mathcal{B}\mathcal{B}}$ correctly.

Computation. Let $P = [\mathrm{id}]_{\mathcal{E}\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, whose columns are $\mathbf{b}_1$ and $\mathbf{b}_2$. Changing from the standard basis to $\mathcal{B}$ means applying P, so

$$[T]_{\mathcal{B}\mathcal{B}} = P[T]_{\mathcal{E}\mathcal{E}}P^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

Solution on page 1186

Exercise 2.8.14. ROUTINE Find the error in the following argument, show that its claim is false, and determine exactly when its conclusion holds.

Claim. If $A, B \in F^{n\times n}$ are invertible, then $(AB)^{-1} = A^{-1}B^{-1}$.

Proof. By associativity, $(AB)\left(A^{-1}B^{-1}\right) = AA^{-1}BB^{-1} = I_n$, and similarly $\left(A^{-1}B^{-1}\right)(AB) = I_n$. So $(AB)^{-1} = A^{-1}B^{-1}$ by Technique 2.5.16.

Solution on page 1186

Exercise 2.8.15. STANDARD Find the error in the following argument, show that its claim is false, and add a hypothesis under which the argument becomes correct.

Claim. If $T \in \mathrm{L}(V)$ has a left inverse, then T is invertible.

Proof. Let $ST = \mathrm{id}_V$. Then $TST = T = \mathrm{id}_V T$, so $(TS - \mathrm{id}_V)T = 0 = 0T$. Cancelling T on the right gives $TS - \mathrm{id}_V = 0$. So $TS = \mathrm{id}_V$, and S is a two-sided inverse of T.

Solution on page 1187

Problems for Chapter 2

These problems combine the sections of Chapter 2 with each other, and several combine them with Chapters 1 and 3. The first three are review prompts: state a main result, prove it without looking, and then answer a question that the text did not ask. The rest are ordered roughly by difficulty. Problems of kind "competition" are harder than anything the course expects. Some problems multiply polynomials, which Section 7.1 develops; each of them says what may be used. Section 2.8 compares the techniques these problems call for.

Review: state and prove

Problem 2.1. STANDARD

(a) State Theorem 2.6.11 and Proposition 2.6.14.

(b) Prove both.

(c) The proof of the theorem uses only part of linearity. Show that its statement holds for every additive map $T : V \to W$, with $\ker T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$ as before. Then show that it fails for homogeneous maps: find a function $T : \mathbb{F}_2^2 \to \mathbb{F}_2$ with $T(a\mathbf{v}) = aT(\mathbf{v})$ for all $a \in \mathbb{F}_2$ and $\mathbf{v} \in \mathbb{F}_2^2$, such that $T(\mathbf{v}) = 0$ only for $\mathbf{v} = \mathbf{0}$, but T is not injective.

Solution on page 1187

Problem 2.2. STANDARD

(a) State Theorem 2.4.2(a).

(b) Prove it from the definitions, and deduce from it, without computing any entry, that $(AB)C = A(BC)$ for all matrices of compatible sizes.

(c) Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V of the same length n. Show that $[T]_{\mathcal{C}\mathcal{B}}[T]_{\mathcal{C}\mathcal{B}} = \left[T^2\right]_{\mathcal{C}\mathcal{B}}$ holds for every $T \in \mathrm{L}(V)$ if and only if $\mathcal{B} = \mathcal{C}$.

Solution on page 1188

Problem 2.3. STANDARD

(a) State Theorem 2.7.4.

(b) Prove (II) $\Rightarrow$ (V). Explain why the construction chooses solutions of $A\mathbf{x} = \mathbf{y}$ only for $\mathbf{y} = \mathbf{e}_1, \ldots, \mathbf{e}_m$, and not for every $\mathbf{y}$ separately.

(c) Let p be a prime, and let $A \in \mathbb{F}_p^{m \times n}$ with T_A surjective. Show that A has exactly $p^{m(n-m)}$ right inverses. Check the count against Exercise 2.5.36(c).

Solution on page 1189

Synthesis problems

Problem 2.4. STANDARD For $A, B \in F^{n \times n}$ let $L_{A,B} : F^{n \times n} \to F^{n \times n}$ be $L_{A,B}(X) = AXB$, and let $\tau(X) = X^{\mathsf{T}}$. Let $\mathcal{U}$ be the basis of $F^{n \times n}$ formed by the matrix units $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Lemma 2.4.24), in any fixed order. For an operator Φ on $F^{n \times n}$, write $\operatorname{tr}\Phi$ for the trace of $[\Phi]_{\mathcal{U}\mathcal{U}}$.

(a) Show that $L_{A,B}$ is linear, that $L_{A,B}L_{C,D} = L_{AC,DB}$, and that $\tau L_{A,B}\tau = L_{B^\intercal,A^\intercal}$.

(b) Show that $\operatorname{tr} L_{A,B} = (\operatorname{tr} A)(\operatorname{tr} B)$. Compute $\operatorname{tr}\tau$, and $\operatorname{tr} C_A$ for $C_A(X) = AX - XA$.

(c) Show that every linear map $\Phi : F^{n\times n} \to F^{n\times n}$ is a sum of maps of the form $L_{A,B}$, and write τ in this form.

Solution on page 1189

Problem 2.5. STANDARD For $c \in F$, let $S_c \in \mathrm{L}(F[x])$ be the substitution $S_c(f) = f(x+c)$ (Exercise 2.1.32), and let D be differentiation. You may use the binomial theorem $(x+c)^m = \sum_{k=0}^m \binom{m}{k} c^k x^{m-k}$ in $F[x]$, that $S_c S_d = S_{c+d}$, and that p divides $\binom{p}{k}$ for a prime p and $0 < k < p$.

(a) Let $F = \mathbb{Q}$ and $n \geq 0$. Show that on $\mathbb{Q}[x]_{\leq n}$,

$$S_c = \sum_{k=0}^{n} \frac{c^k}{k!} D^k.$$

In particular $f(x+1) = \sum_k D^k(f)/k!$, Taylor's formula.

(b) Deduce that S_1 is invertible on $\mathbb{Q}[x]_{\leq n}$, and find its inverse as a combination of powers of D.

(c) Let p be a prime and $F = \mathbb{F}_p$. Show that $S_1 = \sum_{k=0}^n (k!)^{-1} D^k$ on $\mathbb{F}_p[x]_{\leq n}$ when $n < p$, but that for $n \geq p$ there are no scalars $a_0, \ldots, a_r$ with $S_1 = \sum_{k=0}^r a_k D^k$ on $\mathbb{F}_p[x]_{\leq n}$.

Solution on page 1190

Problem 2.6. STANDARD Let M and D be multiplication by x and differentiation on $F[x]$. You may use that multiplication of polynomials is associative and commutative and distributes over addition, that $\deg(fg) = \deg f + \deg g$ for nonzero f, g, and the product rule $D(fg) = D(f)g + fD(g)$ (Exercise 2.1.49(a)).

(a) Show that $T \in \mathrm{L}(F[x])$ satisfies $TM = MT$ if and only if there is $q \in F[x]$ with $T(f) = qf$ for every f, and that then $q = T(1)$.

(b) For such T, show that T is injective if and only if $q \neq \mathbf{0}$, and surjective if and only if q is a nonzero constant.

(c) Find all $T \in \mathrm{L}(F[x])$ that commute with both M and D, when F has characteristic 0 and when $F = \mathbb{F}_p$.

Solution on page 1191

Problem 2.7. STANDARD Let $n \geq 0$, let $c_0, \ldots, c_n \in F$ be distinct, and let $L_0, \ldots, L_n \in F[x]_{\leq n}$ be the Lagrange polynomials (Definition 3.4.26), so that $L_i(c_k)$ is 1 for $k = i$ and 0 otherwise. Write $V = F[x]_{\leq n}$ and $\mathrm{ev}_c : V \to F$, $f \mapsto f(c)$.

(a) Show that $L_0, \ldots, L_n$ is a basis of V, and that every $\varphi \in \mathrm{L}(V, F)$ satisfies $\varphi = \sum_{i=0}^n \varphi(L_i)\,\mathrm{ev}_{c_i}$. Deduce that $\mathrm{ev}_{c_0}, \ldots, \mathrm{ev}_{c_n}$ is a basis of $\mathrm{L}(V, F)$.

(b) Let $n = 2$ and $(c_0, c_1, c_2) = (0, 1, 2)$, and suppose $2 \neq 0$ in F. Find the scalars w_0, w_1, w_2 with $f'(0) = w_0 f(0) + w_1 f(1) + w_2 f(2)$ for every $f \in F[x]_{\leq 2}$, where $f' = D(f)$. Write the formula over $\mathbb{F}_3$.

(c) Show that over $\mathbb{Q}$ no w_0, w_1, w_2 make the formula of (b) hold for every $f \in \mathbb{Q}[x]_{\leq 3}$.

Solution on page 1191

Problem 2.8. STANDARD A matrix P is **idempotent** if $P^2 = P$. By Exercise 2.5.46(a), every $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$ satisfies $A^2 - (a+d)A + (ad - bc)I_2 = 0$.

(a) Show that $A \in F^{2\times 2}$ is idempotent if and only if $A = 0$, $A = I_2$, or ($a + d = 1$ and $ad - bc = 0$).

(b) Let p be a prime. Show that exactly $p^2 + p + 2$ matrices in $\mathbb{F}_p^{2\times 2}$ are idempotent.

(c) Suppose $1 + 1 \neq 0$ in F. Show that $P \mapsto 2P - I_n$ is a bijection from the idempotent matrices in $F^{n\times n}$ onto the matrices S with $S^2 = I_n$. Compare (b) with Exercise 2.5.46(c).

(d) Over $\mathbb{F}_2$, count the $S \in \mathbb{F}_2^{2\times 2}$ with $S^2 = I_2$, and show that every $A \in \mathbb{F}_2^{2\times 2}$ is invertible, idempotent, or satisfies $A^2 = 0$. Show that this fails over $\mathbb{F}_3$.

Solution on page 1192

Problem 2.9. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$. A square matrix N is **nilpotent** if $N^k = 0$ for some $k \geq 1$.

(a) Show that $(BA)^{k+1} = B(AB)^k A$ for every $k \geq 0$.

(b) Deduce that AB is nilpotent if and only if BA is, and that $(AB)^k = 0$ implies $(BA)^{k+1} = 0$.

(c) Find $A, B \in F^{2\times 2}$ with $AB = 0$ and $BA \neq 0$. So the exponents in (b) can differ.

(d) Show that $\operatorname{tr}\left((AB)^k\right) = \operatorname{tr}\left((BA)^k\right)$ for every $k \geq 1$.

Solution on page 1193

Problem 2.10. STANDARD Call $N \in \mathrm{L}(V)$ **nilpotent** if $N^k = 0$ for some $k \geq 1$. You may use the binomial theorem for commuting operators, which has the same proof as for numbers.

(a) Show that the sum of two commuting nilpotent operators is nilpotent.

(b) Show that the sum of two nilpotent matrices in $F^{2\times 2}$ can be invertible. Deduce that the nilpotent matrices in $F^{2\times 2}$ do not form a subspace.

(c) Let $G \in \mathrm{L}(V)$ be invertible and $N \in \mathrm{L}(V)$ nilpotent, with $GN = NG$. Show that $G + N$ is invertible, and find its inverse.

(d) Show that (c) can fail if $GN \neq NG$.

Solution on page 1194

Problem 2.11. STANDARD Let $n \geq 2$, and let $X \in F^{n\times n}$ be invertible with columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Let $Y \in F^{n\times n}$ have columns $\mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{0}$, the columns of X moved one place to the left with a zero column at the end.

(a) Find a matrix N, not depending on X, with $Y = XN$.

(b) Show that $X^{-1}Y$ and YX^{-1} are nilpotent, and that $\left(X^{-1}Y\right)^{n-1} \neq 0$ and $\left(YX^{-1}\right)^{n-1} \neq 0$.

Solution on page 1194

Problem 2.12. STANDARD Let $n \geq 1$, let W be a subspace of $F^{n\times n}$, and let $U_{ij} = \mathbf{e}_i \mathbf{e}_j^{\mathsf{T}}$.

(a) Suppose that $AX \in W$ and $XA \in W$ for all $A \in F^{n\times n}$ and $X \in W$. Show that $W = \{0\}$ or $W = F^{n\times n}$.

(b) Let $n \geq 2$, and let W_1 be the set of matrices whose columns $2, \ldots, n$ are zero. Show that $AX \in W_1$ for all $A \in F^{n\times n}$ and $X \in W_1$, but that $XA \notin W_1$ for some $X \in W_1$ and A.

(c) Let $\psi : F^{n\times n} \to F^{m\times m}$ be linear and nonzero, with $\psi(XY) = \psi(X)\psi(Y)$ for all X, Y. Show that ψ is injective, and give an example with $m = 2n$.

Solution on page 1195

Problem 2.13. STANDARD For $Z \in \mathbb{C}^{m\times n}$ write $Z = X + iY$ with $X, Y \in \mathbb{R}^{m\times n}$ (the entrywise real and imaginary parts), and let

$$\rho(Z) = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \in \mathbb{R}^{2m\times 2n}.$$

(a) Show that $\rho(ZW) = \rho(Z)\rho(W)$ for $Z \in \mathbb{C}^{m\times n}$ and $W \in \mathbb{C}^{n\times p}$, that ρ is $\mathbb{R}$-linear, and that $\rho(I_n) = I_{2n}$.

(b) Regard $\mathbb{C}^n$ as a vector space over $\mathbb{R}$ (Proposition 1.2.19), with the list $\mathcal{B}_n = \mathbf{e}_1, \ldots, \mathbf{e}_n, i\mathbf{e}_1, \ldots, i\mathbf{e}_n$. Show that $\mathcal{B}_n$ is a basis, that $T_Z : \mathbb{C}^n \to \mathbb{C}^m$ is $\mathbb{R}$-linear, and that $[T_Z]_{\mathcal{B}_m\mathcal{B}_n} = \rho(Z)$.

(c) Show that $Z \in \mathbb{C}^{n\times n}$ is invertible if and only if $\rho(Z)$ is, and that then $\rho\big(Z^{-1}\big) = \rho(Z)^{-1}$.

(d) Show that $\operatorname{tr}\rho(Z) = 2\operatorname{Re}(\operatorname{tr} Z)$ for $Z \in \mathbb{C}^{n\times n}$.

Solution on page 1195

Problem 2.14. STANDARD Prove or disprove each statement.

(a) If $P, Q \in F^{n\times n}$ satisfy $P^2 = P$, $Q^2 = Q$, and $PQ = QP$, then $P + Q - PQ$ is idempotent.

(b) If $T \in \mathrm{L}(V)$ and $\ker T = \ker T^2$, then $\operatorname{im} T = \operatorname{im} T^2$.

(c) If $A \in F^{n\times n}$ and $\operatorname{tr}\big(A^k\big) = 0$ for every $k \geq 1$, then $A^k = 0$ for some k.

(d) If $P, Q \in F^{n\times n}$ are idempotent, with $n \geq 2$, then $P + Q$ is idempotent.

(e) If $T \in \mathrm{L}(V, W)$ and $S \in \mathrm{L}(W, V)$ satisfy $ST = \mathrm{id}_V$ and TS is injective, then T is invertible.

Solution on page 1196

Problem 2.15. HARD Let p be a prime, $\mathbf{u} = (1, 1, 1) \in \mathbb{F}_p^3$, and $\ell = \operatorname{span}(\mathbf{u})$.

(a) Show that exactly p^4 linear maps $T : \mathbb{F}_p^3 \to \mathbb{F}_p^2$ satisfy $T(\mathbf{u}) = \mathbf{0}$.

(b) Show that exactly $\big(p^2 - 1\big)\big(p^2 - p\big)$ of them have $\ker T = \ell$, and that each of these is surjective.

(c) Show that every surjective linear map $\mathbb{F}_p^3 \to \mathbb{F}_p^2$ has a line $\operatorname{span}(\mathbf{v})$, $\mathbf{v} \neq \mathbf{0}$, as its kernel. Use (b) and the number of lines in $\mathbb{F}_p^3$ (Problem 1.13) to count the surjective linear maps $\mathbb{F}_p^3 \to \mathbb{F}_p^2$, and compare with Exercise 2.6.45.

Solution on page 1197

Problem 2.16. HARD

(a) Let $P \in F^{n\times n}$ satisfy $P^2 = P$, and let $k = \operatorname{rank} P$. Show that $\operatorname{tr} P = k * 1$.

(b) Let $A \in F^{n\times n}$ satisfy $A^2 = 0$. Show that $\operatorname{tr} A = 0$.

(c) Let $P_1, \ldots, P_r \in \mathbb{Q}^{n\times n}$ be idempotent with $P_1 + \cdots + P_r = I_n$. Show that $\operatorname{rank} P_1 + \cdots + \operatorname{rank} P_r = n$. Show that this fails over $\mathbb{F}_2$.

Solution on page 1198

Problem 2.17. HARD Let $A, B \in F^{n\times n}$ satisfy $AB = 0$.

(a) Show that $\operatorname{tr}\big((A + B)^k\big) = \operatorname{tr}\big(A^k\big) + \operatorname{tr}\big(B^k\big)$ for every $k \geq 1$.

(b) Show that $(A + B)^k = A^k + B^k$ can fail: find $A, B \in F^{2\times 2}$ with $AB = 0$ and $(A + B)^2 \neq A^2 + B^2$.

Solution on page 1198

Problem 2.18. HARD A town has n residents and k clubs. Every club has an odd number of members, and every two different clubs have an even number of members in common.

(a) Show that $k \leq n$.

(b) Show that $k = n$ is possible for every $n \geq 1$.

(c) Suppose instead that every club has an *even* number of members, and every two different clubs still share an even number. Show that there can be more than n clubs: find 7 nonempty clubs in a town of 6 residents.

Solution on page 1199

Problem 2.19. HARD Let $A, B \in \mathbb{R}^{n \times n}$ be different matrices with $A^3 = B^3$ and $A^2B = B^2A$.

(a) Can $A^2 + B^2$ be invertible?

(b) Show that $A^2 + B^2$ need not be 0.

Solution on page 1200

Problem 2.20. CHALLENGE Let $A, B \in F^{n \times n}$ satisfy $ABAB = 0$.

(a) Show that if $n = 2$, then $BABA = 0$.

(b) Show that for every $n \geq 3$ there are $A, B \in F^{n \times n}$ with $ABAB = 0$ but $BABA \neq 0$.

Solution on page 1200

Chapter 3

Linear systems and row reduction

Gauss–Jordan elimination as the engine of the theory, rank, and the Fundamental Theorems completed.

3.1 Linear systems

The Fundamental Theorems (Section 2.7) are statements about the equation $A\mathbf{x} = \mathbf{y}$: whether it has a solution for every $\mathbf{y}$, and whether solutions are unique. They do not tell us how to settle either question for a particular matrix, let alone how to find the solutions. This chapter develops the tool that does, Gauss–Jordan elimination, and then uses it to finish the proofs of the theorems. This section sets up the problem. A linear system can be read as a list of equations, as one matrix equation, or as a question about linear combinations of columns, and its solution set always has the same shape: it is empty, or it is one solution plus the kernel of T_A.

3.1.1 Linear systems

Definition 3.1.1 (Linear system). Let F be a field. A **linear equation** in the unknowns $x_1, \dots, x_n$ over F is an equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where the **coefficients** $a_1, \dots, a_n$ and the **right-hand side** b are fixed elements of F. A **linear system** (or **system of linear equations**) of m equations in n unknowns is a list of m linear equations in the same unknowns,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m, \end{aligned}$$

with all a_{ij} and y_i in F. A **solution** of the system is a vector $\mathbf{x} = (x_1, \dots, x_n) \in F^n$ that satisfies all m equations, and the **solution set** is the set of all solutions. The system is **consistent** if it has at least one solution and **inconsistent** otherwise. Two systems in the same unknowns are **equivalent** if they have the same solution set.

The word "linear" refers to the unknowns: each term is a scalar times a single unknown, and nothing else involves the unknowns. The coefficients can be any scalars. An equation such as $2x_1 + 3 = x_2 - x_1$ also counts, since adding $-3 - x_2 + x_1$ to both sides turns it into $3x_1 - x_2 = -3$ without changing its solutions.

Example 3.1.2 (A system with a line of solutions). Over $\mathbb{R}$, consider the system of three equations in three unknowns

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1, \\ 2x_1 + 5x_2 + x_3 &= 5, \\ 3x_1 + 7x_2 &= 6. \end{aligned}$$

The vector $(2, 0, 1)$ is a solution, since $2 + 0 - 1 = 1$, $4 + 0 + 1 = 5$, and $6 + 0 = 6$. So is $(-5, 3, 0)$. The vector $(1, 0, 0)$ satisfies the first equation but not the second, so it is not a solution. We will find the whole solution set in Example 3.1.13: it is a line in $\mathbb{R}^3$. We call it the *running system* and return to it several times.

Non-example 3.1.3 (Equations that are not linear). Over $\mathbb{R}$, the equations $x_1 x_2 = 1$, $x_1^2 + x_2^2 = 1$, and $x_1 + \sin x_2 = 0$ are not linear: the unknowns are multiplied together, squared, or fed into a function. Their solution sets in $\mathbb{R}^2$ are a hyperbola, a circle, and a sine curve, and none of them contains a straight line. By Corollary 3.1.23, the solution set of a linear system contains the whole line through any two of its points, so none of these curves is the solution set of a linear system. On the other hand, $\sqrt{2}x_1 + \pi x_2 = e$ is a linear equation; its coefficients are just unusual real numbers.

3.1.2 Three faces of one problem

A linear system is determined by its coefficients a_{ij} and its right-hand sides y_i; what we call the unknowns plays no role. Collecting the coefficients into a matrix and the right-hand sides into a vector turns the m equations into a single equation, and that equation can be read in two further ways.

Definition 3.1.4 (Coefficient matrix; matrix and vector forms). The **coefficient matrix** of the system in Definition 3.1.1 is the matrix $A \in F^{m \times n}$ whose (i, j) entry is a_{ij}, the coefficient of x_j in the ith equation. Write $\mathbf{y} = (y_1, \ldots, y_m) \in F^m$ for the vector of right-hand sides, and $\mathbf{a}_1, \ldots, \mathbf{a}_n \in F^m$ for the columns of A. The **matrix form** of the system is the equation

$$A\mathbf{x} = \mathbf{y}$$

in the unknown $\mathbf{x} \in F^n$, and its **vector form** is the equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{y}$$

in the unknown scalars $x_1, \ldots, x_n$. We refer to the system itself as "the system $A\mathbf{x} = \mathbf{y}$".

For the running system (Example 3.1.2),

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 1 \\ 3 & 7 & 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 5 \\ 6 \end{pmatrix}, \quad x_1\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + x_2\begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 6 \end{pmatrix}.$$

The missing x_3 in the third equation is a coefficient 0.

Proposition 3.1.5 (Three faces of a linear system). Let $A \in F^{m \times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, let $\mathbf{y} \in F^m$, and let $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$. The following are equivalent.

(i) $\mathbf{x}$ is a solution of the linear system with coefficient matrix A and right-hand sides $\mathbf{y}$.
(ii) $A\mathbf{x} = \mathbf{y}$.
(iii) $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{y}$.

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 1.2.4 (The space F^n)
Hint 1. Two of the three statements are the same equation by a definition. Which two?
Hint 2. For the remaining equivalence, compare the ith entries of the two sides of (iii).

Sketch. By the definition of $A\mathbf{x}$ as a combination of the columns of A, (ii) and (iii) are the same equation. The ith entry of $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ is $a_{i1}x_1 + \cdots + a_{in}x_n$, the left side of the ith equation, so (iii) holds exactly when all m equations do.

By Definition 2.3.8, $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$, so (ii) and (iii) are the same equation.

For (i) $\Leftrightarrow$ (iii), recall that vectors in F^m are added and scaled entry by entry, and that two vectors in F^m are equal exactly when their entries are (Definition 1.2.4). The ith entry of $\mathbf{a}_j$ is a_{ij}, so the ith entry of $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ is

$$x_1a_{i1} + x_2a_{i2} + \cdots + x_na_{in} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n.$$

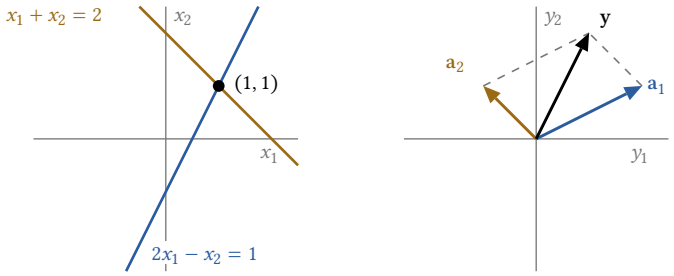
Hence (iii) holds if and only if, for each $i = 1, \ldots, m$, the left side of the ith equation equals y_i. That is, (iii) holds if and only if $\mathbf{x}$ is a solution. □

Example 3.1.6 (Two pictures of one system). The system $2x_1 - x_2 = 1$, $x_1 + x_2 = 2$ over $\mathbb{R}$ has matrix form and vector form

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad x_1\begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Read as equations, it asks for the points of $\mathbb{R}^2$ that lie on two lines; the lines cross at $(1, 1)$ (Figure 3.1.1, left). Read as a vector equation, it asks how to reach $(1, 2)$ by scaling $\mathbf{a}_1 = (2, 1)$ and $\mathbf{a}_2 = (-1, 1)$ and adding; the answer is one of each, since $(2, 1) + (-1, 1) = (1, 2)$ (Figure 3.1.1, right).

Realize that the two pictures live in different spaces. The **row picture** is drawn in F^n, the space where the unknown $\mathbf{x}$ lives, and has one line (in $\mathbb{R}^3$, one plane) per equation. The **column picture** is drawn in F^m, the space where the right-hand side $\mathbf{y}$ lives, and has one vector per unknown.



Row picture Column picture

Figure 3.1.1. The system $2x_1 - x_2 = 1$, $x_1 + x_2 = 2$. Left: the solution is the point where the two lines meet. Right: the same solution, $(x_1, x_2) = (1, 1)$, gives the coefficients that combine the columns into $\mathbf{y}$.

Exercise 3.1.7. ROUTINE Write the system

$$x_1 - x_3 = 2, \quad x_2 + 2x_3 = -1, \quad x_1 + x_2 = 1$$

over $\mathbb{R}$ in matrix form and in vector form. Is $(2, -1, 0)$ a solution? Is $(3, -3, 1)$?

Solution on page 1201

What each face is good for

Intuition (Three ways to read $A\mathbf{x} = \mathbf{y}$). The *equations* are what we manipulate. Adding a multiple of one equation to another, swapping two equations, or scaling one by a nonzero scalar does not change the solution set, and elimination is built from exactly these moves (Section 3.1.3). Geometrically, the solution set of a system is the intersection of the solution sets of its equations, which is the row picture.

The *matrix equation* brings in the linear map $T_A : F^n \to F^m, \mathbf{x} \mapsto A\mathbf{x}$ (Definition 2.3.15). Solving the system means finding every $\mathbf{x}$ with $T_A(\mathbf{x}) = \mathbf{y}$, that is, the preimage of $\mathbf{y}$ under T_A. Consistency asks whether $\mathbf{y} \in \operatorname{im} T_A$, and the solution set of $A\mathbf{x} = \mathbf{0}$ is $\ker T_A$. Everything we know about kernels and images applies, and so does linearity: $A(\mathbf{x} + \mathbf{x}') = A\mathbf{x} + A\mathbf{x}'$ tells us how solutions of systems with the same coefficient matrix combine. This is the view in which the structure of the solution set (Proposition 3.1.19) is proved.

The *vector equation* asks whether $\mathbf{y}$ is a linear combination of the columns of A, and if so in how many ways. Read backwards, it says that every question about a span or about linear independence of a list in F^m is a question about a linear system. So a procedure that solves linear systems also answers, by computation, the questions of Chapter 1.

The first two readings are related by the following proposition, which is little more than a translation of definitions.

Proposition 3.1.8 (Systems and the map T_A). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and let $\mathbf{y} \in F^m$.

(a) The system $A\mathbf{x} = \mathbf{y}$ is consistent if and only if $\mathbf{y} \in \operatorname{im} T_A$, if and only if $\mathbf{y} \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

(b) The solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ is $\ker T_A$.

Proof. ROUTINE

Uses. Proposition 3.1.5 (Three faces of a linear system), Definition 2.3.15 (The linear map of a matrix), Definition 2.6.6 (Image), Definition 2.6.5 (Kernel), Definition 1.3.10 (Span)
Hint 1. Unwind the definitions: what does $\mathbf{y} \in \operatorname{im} T_A$ say? What does $\mathbf{y} \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ say?
Hint 2. Each of them says that some $\mathbf{x} \in F^n$ satisfies one of the forms in Proposition 3.1.5.
Sketch. $\mathbf{y} \in \operatorname{im} T_A$ means $A\mathbf{x} = \mathbf{y}$ for some $\mathbf{x}$, and $\mathbf{y} \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ means $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{y}$ for some scalars x_j. By Proposition 3.1.5, each says that the system has a solution. For (b), $\mathbf{x}$ solves $A\mathbf{x} = \mathbf{0}$ exactly when $T_A(\mathbf{x}) = \mathbf{0}$.

(a) By Definition 2.6.6 and Definition 2.3.15, $\mathbf{y} \in \operatorname{im} T_A$ means that $A\mathbf{x} = \mathbf{y}$ for some $\mathbf{x} \in F^n$. By Definition 1.3.10, $\mathbf{y} \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ means that $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{y}$ for some $x_1, \ldots, x_n \in F$. By Proposition 3.1.5, each of these says that the system has a solution $\mathbf{x} = (x_1, \ldots, x_n)$.

(b) By Proposition 3.1.5, $\mathbf{x} \in F^n$ solves $A\mathbf{x} = \mathbf{0}$ if and only if $T_A(\mathbf{x}) = A\mathbf{x} = \mathbf{0}$, that is, if and only if $\mathbf{x} \in \ker T_A$ (Definition 2.6.5). □

Combining Proposition 3.1.8 with the Fundamental Theorems gives a dictionary between the three faces (Table 3.1.1). Each row says the same thing about A in three ways. Most of the questions we ask about a matrix in the next two chapters are entries of this table, and Gauss–Jordan elimination will answer each of them by inspecting a single matrix.

The system $A\mathbf{x} = \mathbf{y}$	The map T_A	The columns of A
is consistent	$\mathbf{y} \in \operatorname{im} T_A$	$\mathbf{y}$ is in their span
with $\mathbf{y} = \mathbf{0}$ has a solution $\mathbf{x} \neq \mathbf{0}$	$\ker T_A \neq \{\mathbf{0}\}$	they are linearly dependent
is consistent for every $\mathbf{y} \in F^m$	T_A is surjective	they span F^m
has at most one solution for every $\mathbf{y}$	T_A is injective	they are linearly independent
has exactly one solution for every $\mathbf{y}$	T_A is bijective	they form a basis of F^m

Table 3.1.1. Three readings of the same property. The first two rows come from Proposition 3.1.8 and the definition of linear independence (Definition 1.5.1). The last three are parts of Theorem 2.7.4, Theorem 2.7.2, and Theorem 2.7.8.

The third reading turns questions about vectors into linear systems. This move is used so often that it deserves a name.

Technique 3.1.9 (Undetermined coefficients). To decide whether an object with prescribed properties exists, and to find all such objects, write the unknown object in terms of unknown scalars, the **undetermined coefficients**, and translate each requirement into an equation in those scalars. When the requirements are linear in the scalars, the result is a linear system, and the question becomes whether that system is consistent and what its solution set is.

Reach for it whenever you are asked whether a vector is a linear combination of given vectors, whether a list is linearly independent, or whether a polynomial or a linear map with given properties exists. Membership of $\mathbf{w}$ in $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ is the system $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{w}$ in the unknowns c_i. Linear independence of $\mathbf{v}_1, \dots, \mathbf{v}_k$ asks whether the homogeneous system $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ has only the trivial solution. The canonical example is Example 3.1.10.

Used in. Exercise 3.1.11, Exercise 3.1.29, Exercise 3.1.37, Exercise 3.1.38, Exercise 3.1.39, Exercise 3.1.40, Exercise 3.1.41, Theorem 3.4.24, Exercise 3.6.12, Problem 3.9, Problem 3.11, Problem 3.13, Exercise 3.4.43, Exercise 6.3.49, Exercise 6.6.48.

Example 3.1.10 (Membership and independence over $\mathbb{R}$ and over $\mathbb{F}_2$). Let $\mathbf{v}_1 = (1, 0, 1)$, $\mathbf{v}_2 = (1, 1, 0)$, and $\mathbf{v}_3 = (0, 1, 1)$. Is $\mathbf{w} = (1, 1, 1)$ a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$? Writing $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{w}$ and comparing entries gives the system

$$c_1 + c_2 = 1, \quad c_2 + c_3 = 1, \quad c_1 + c_3 = 1.$$

Adding the three equations shows that every solution satisfies $2(c_1 + c_2 + c_3) = 3$.

Over $\mathbb{R}$ we may divide by 2, so every solution has $c_1 + c_2 + c_3 = 3/2$. Subtracting each of the three equations from this one gives $c_3 = 1/2$, $c_1 = 1/2$, and $c_2 = 1/2$. Conversely, $(1/2, 1/2, 1/2)$ satisfies all three equations. So $\mathbf{w} = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3)$, and this is the only way to write $\mathbf{w}$ as a combination of the $\mathbf{v}_i$. The same computation with right-hand side $\mathbf{0}$ shows that $c_1 + c_2 = c_2 + c_3 = c_1 + c_3 = 0$ has only the trivial solution, so $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly independent over $\mathbb{R}$.

Over $\mathbb{F}_2$ we have $2 = 0$ and $3 = 1$, so every solution would satisfy $0 = 1$. The system is inconsistent, and $\mathbf{w} \notin \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. The homogeneous system now has the nontrivial solution $(1, 1, 1)$, because $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = (2, 2, 2) = \mathbf{0}$, so the list is linearly dependent over $\mathbb{F}_2$.

The only step that distinguished the two fields was a division by 2. Elimination divides by nonzero scalars all the time, and which scalars are nonzero depends on the field.

Exercise 3.1.11. ROUTINE Set up a linear system whose solutions are exactly the triples $(c_1, c_2, c_3) \in \mathbb{R}^3$ with

$$x^2 = c_1(x + 1)^2 + c_2(x - 1)^2 + c_3$$

in $\mathbb{R}[x]$, and solve it.

Solution on page 1201

3.1.3 Augmented matrices and elimination

Definition 3.1.12 (Augmented matrix). The **augmented matrix** of the system $A\mathbf{x} = \mathbf{y}$, where $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$, is the $m \times (n + 1)$ matrix $(A \mid \mathbf{y})$ obtained by adjoining $\mathbf{y}$ to A as a last column. A vertical bar separates the coefficients from the right-hand sides.

The augmented matrix of the running system (Example 3.1.2) is

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 5 & 1 & 5 \\ 3 & 7 & 0 & 6 \end{array}\right).$$

Row i of the augmented matrix records equation i, so operations on equations become operations on rows, and we can skip writing the unknowns.

Example 3.1.13 (Elimination on the running system). We solve the running system by eliminating unknowns. Subtract 2 times the first equation from the second, and 3 times the first from the third:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 3 & 3 \\ 0 & 1 & 3 & 3 \end{array}\right).$$

Now subtract the second equation from the third, and 2 times the second from the first:

$$\left(\begin{array}{ccc|c} 1 & 0 & -7 & -5 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

The last row is the equation $0 = 0$, which every vector satisfies. The other two rows say $x_1 - 7x_3 = -5$ and $x_2 + 3x_3 = 3$. Whatever value t we give x_3, these determine $x_1 = -5 + 7t$ and $x_2 = 3 - 3t$. So the solution set is

$$\{(-5 + 7t, 3 - 3t, t) \mid t \in \mathbb{R}\} = \{(-5, 3, 0) + t(7, -3, 1) \mid t \in \mathbb{R}\},$$

the line through $(-5, 3, 0)$ in the direction of $(7, -3, 1)$. The solutions $(-5, 3, 0)$ and $(2, 0, 1)$ from Example 3.1.2 are the points with $t = 0$ and $t = 1$.

If the third right-hand side were 7 instead of 6, the same operations would end with the row $(0, 0, 0 \mid 1)$, which is the equation $0 = 1$. No vector satisfies it, so that system is inconsistent.

These moves never change the solution set. Each new equation is a combination of the old ones, so every solution of the old system also solves the new one. And each move can be undone by a move of the same kind: subtracting 2 times the first equation from the second is undone by adding 2 times the first equation back, which is possible because the first equation is still there. (In each step of the example, the equations we subtracted were themselves left unchanged in that step.) So the old equations are combinations of the new ones as well, and every solution of the new system solves the old one. The two systems are equivalent. Section 3.2 makes this precise by writing each move as left multiplication by an invertible matrix, and Exercise 3.1.35 shows what goes wrong when a move cannot be undone.

> *Intuition (Why we need an algorithm).* Elimination by hand worked here, but each step was a choice, and we stopped when the answer looked readable. To prove theorems with elimination we need a precise procedure: one that always terminates, whose steps provably never change the solution set, and whose output has a standard shape from which the solution set, and every row of Table 3.1.1, can be read off directly. Gauss–Jordan elimination (Section 3.3) is that procedure. Section 3.2 sets up its bookkeeping, Section 3.4 reads solutions off its output, and Section 3.5 uses it to finish the Fundamental Theorems. It stays the computational engine through Chapter 4, where it computes bases, dimensions, and ranks.

3.1.4 Homogeneous systems and the shape of the solution set

Definition 3.1.14 (Homogeneous system). A linear system $A\mathbf{x} = \mathbf{y}$ is **homogeneous** if $\mathbf{y} = \mathbf{0}$. The **associated homogeneous system** of $A\mathbf{x} = \mathbf{y}$ is $A\mathbf{x} = \mathbf{0}$. Every homogeneous system has the **trivial solution** $\mathbf{x} = \mathbf{0}$; any other solution is **nontrivial**.

Example 3.1.15 (The running system, made homogeneous). The associated homogeneous system of the running system is $x_1 + 2x_2 - x_3 = 0, 2x_1 + 5x_2 + x_3 = 0, 3x_1 + 7x_2 = 0$. The row operations of Example 3.1.13

never mix the last column into the others, and a last column of zeros stays zero, so they end with $x_1 - 7x_3 = 0$ and $x_2 + 3x_3 = 0$. The solution set is

$$\{t(7, -3, 1) \mid t \in \mathbb{R}\} = \operatorname{span}((7, -3, 1)),$$

a line through the origin, parallel to the solution line of the running system.

Corollary 3.1.16 (Homogeneous solution sets are subspaces). The solution set of a homogeneous system $A\mathbf{x} = \mathbf{0}$, where $A \in F^{m\times n}$, is the subspace $\ker T_A$ of F^n.

Proof. ROUTINE

Uses. Proposition 3.1.8 (Systems and the map T_A), Proposition 2.6.8 (Kernels and images are subspaces)
Hint 1. Which subset of F^n is the solution set, by Proposition 3.1.8?
Hint 2. Kernels of linear maps are subspaces.
Sketch. By Proposition 3.1.8 the solution set is $\ker T_A$, and the kernel of a linear map is a subspace of its domain.

By part (b) of Proposition 3.1.8, the solution set is $\ker T_A$. Since $T_A : F^n \to F^m$ is linear, $\ker T_A$ is a subspace of F^n by Proposition 2.6.8. □

Concretely: sums and scalar multiples of solutions of a homogeneous system are again solutions. A system with $\mathbf{y} \neq \mathbf{0}$ has no such luck. Its solution set never contains $\mathbf{0}$, since $A\mathbf{0} = \mathbf{0} \neq \mathbf{y}$, so it is never a subspace. The running system shows what it is instead: the line of solutions in Example 3.1.13 is the line of Example 3.1.15, moved over so that it passes through one solution. To say this in general we need a name for a moved-over subset.

Notation 3.1.17 (Translates). Let V be a vector space, $\mathbf{v} \in V$, and $U \subseteq V$. The **translate** of U by $\mathbf{v}$ is

$$\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\}.$$

Intuition. Suppose $\mathbf{x}$ and $\mathbf{x}'$ both solve $A\mathbf{x} = \mathbf{y}$. Then $A(\mathbf{x}' - \mathbf{x}) = \mathbf{y} - \mathbf{y} = \mathbf{0}$, so any two solutions differ by a solution of the homogeneous system. Conversely, adding a homogeneous solution $\mathbf{z}$ to a solution $\mathbf{x}$ gives another solution, since $A(\mathbf{x} + \mathbf{z}) = \mathbf{y} + \mathbf{0}$. So once we know one solution, the others are exactly the ones we reach from it by adding elements of $\ker T_A$. Only linearity is involved, so the same holds for any linear map in place of T_A.

Technique 3.1.18 (Particular plus homogeneous). To find every solution of a linear equation $T(\mathbf{x}) = \mathbf{y}$, find *one* solution $\mathbf{x}_0$, called a **particular solution**, and *all* solutions of the homogeneous equation $T(\mathbf{z}) = \mathbf{0}$. Every solution is $\mathbf{x}_0 + \mathbf{z}$ for exactly one such $\mathbf{z}$. This splits the problem in two: existence of a solution depends on $\mathbf{y}$, while the shape of the solution set, and in particular uniqueness, depends on $\ker T$ alone.

Reach for it whenever you need all solutions of a linear problem, such as a linear system, a linear recurrence, or a linear differential equation, and whenever you need to show that a solution is unique: a solution, if one exists, is unique exactly when the homogeneous equation has only the trivial solution. The canonical example is Proposition 3.1.19.
Used in. Corollary 2.6.31, Exercise 2.6.33, Proposition 3.1.19, Exercise 3.1.21, Corollary 3.1.22, Corollary 3.1.23, Exercise 3.1.31, Exercise 3.1.37, Exercise 3.1.38, Exercise 3.1.46, Exercise 3.1.47, Theorem 3.4.9, Exercise 3.6.12, Problem 3.1, Problem 3.5, Problem 3.9, Problem 3.16, Theorem 4.7.8, Corollary 7.1.30, Exercise 2.6.34, Exercise 3.1.29, Exercise 3.1.30.

Proposition 3.1.19 (Solution sets are translates of the kernel). Let $T : V \to W$ be a linear map, let $\mathbf{y} \in W$, and let $S = \{\mathbf{x} \in V \mid T(\mathbf{x}) = \mathbf{y}\}$. If $\mathbf{x}_0 \in S$, then

$$S = \mathbf{x}_0 + \ker T,$$

and the map $\ker T \to S$, $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$, is a bijection. In particular, the solution set of a linear system $A\mathbf{x} = \mathbf{y}$ is either empty or equal to $\mathbf{x}_0 + \ker T_A$, where $\mathbf{x}_0$ is any one solution.

Proof. ROUTINE

> *Uses.* Definition 2.1.1 (Linear map), Definition 2.6.5 (Kernel), Notation 3.1.17 (Translates), Proposition 1.2.34 (Negatives and the scalar -1), Proposition 3.1.5 (Three faces of a linear system)
> *Techniques.* Double containment (Technique 1.3.14), Particular plus homogeneous (Technique 3.1.18)
> *Hint 1.* Prove $\mathbf{x}_0 + \ker T \subseteq S$ and $S \subseteq \mathbf{x}_0 + \ker T$ separately.
> *Hint 2.* For the first containment, apply T to $\mathbf{x}_0 + \mathbf{z}$, where $\mathbf{z} \in \ker T$.
> *Hint 3.* For the second, given $\mathbf{x} \in S$, the only possible choice of $\mathbf{z}$ is $\mathbf{x} - \mathbf{x}_0$. Show that it lies in $\ker T$.
> *Sketch.* If $T(\mathbf{z}) = \mathbf{0}$, then $T(\mathbf{x}_0 + \mathbf{z}) = \mathbf{y} + \mathbf{0} = \mathbf{y}$. If $T(\mathbf{x}) = \mathbf{y}$, then $T(\mathbf{x} - \mathbf{x}_0) = \mathbf{y} - \mathbf{y} = \mathbf{0}$, and $\mathbf{x} = \mathbf{x}_0 + (\mathbf{x} - \mathbf{x}_0)$. So $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ maps $\ker T$ onto S, and it is injective by cancellation. For a system, take $T = T_A$ and use Proposition 3.1.5.

Suppose $\mathbf{x}_0 \in S$, so that $T(\mathbf{x}_0) = \mathbf{y}$. We prove the two containments.

$\mathbf{x}_0 + \ker T \subseteq S$. Let $\mathbf{z} \in \ker T$. Since T is linear,

$$T(\mathbf{x}_0 + \mathbf{z}) = T(\mathbf{x}_0) + T(\mathbf{z}) = \mathbf{y} + \mathbf{0} = \mathbf{y},$$

so $\mathbf{x}_0 + \mathbf{z} \in S$.

$S \subseteq \mathbf{x}_0 + \ker T$. Let $\mathbf{x} \in S$, and put $\mathbf{z} = \mathbf{x} - \mathbf{x}_0 = \mathbf{x} + (-1)\mathbf{x}_0$ (Proposition 1.2.34). Since T is linear,

$$T(\mathbf{z}) = T(\mathbf{x}) + (-1)T(\mathbf{x}_0) = \mathbf{y} - \mathbf{y} = \mathbf{0},$$

so $\mathbf{z} \in \ker T$, and $\mathbf{x} = \mathbf{x}_0 + \mathbf{z} \in \mathbf{x}_0 + \ker T$.

Hence $S = \mathbf{x}_0 + \ker T$, which says exactly that $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ maps $\ker T$ onto S. This map is also injective: if $\mathbf{x}_0 + \mathbf{z} = \mathbf{x}_0 + \mathbf{z}'$, adding $-\mathbf{x}_0$ to both sides gives $\mathbf{z} = \mathbf{z}'$. So it is a bijection.

For the last statement, let $A \in F^{m \times n}$ and $\mathbf{y} \in F^m$. By Proposition 3.1.5, the solution set of the system $A\mathbf{x} = \mathbf{y}$ is $\{\mathbf{x} \in F^n \mid T_A(\mathbf{x}) = \mathbf{y}\}$. If it is nonempty, we apply what we proved to $T = T_A$ and any solution $\mathbf{x}_0$. □

So the solution set of a consistent system is $\ker T_A$ slid over to pass through any one solution, and which solution we slide it to does not matter. For the running system, Example 3.1.13 and Example 3.1.15 give

$$\left\{\mathbf{x} \in \mathbb{R}^3 \;\middle|\; A\mathbf{x} = \mathbf{y}\right\} = (-5, 3, 0) + \operatorname{span}((7, -3, 1)) = (2, 0, 1) + \operatorname{span}((7, -3, 1)).$$

The proof used nothing about V and W beyond the linearity of T, so the proposition applies to linear equations far from F^n.

Example 3.1.20 (A linear equation in $\mathbb{R}[x]$). Let $D : \mathbb{R}[x] \to \mathbb{R}[x]$ be differentiation, $D(f) = f'$, which is linear (Section 2.1). Which polynomials f satisfy $f' = 2x + 1$? One solution is $x^2 + x$. The kernel of D consists of the constant polynomials: if $f = c_0 + c_1 x + \cdots + c_d x^d$ and $f' = c_1 + 2c_2 x + \cdots + dc_d x^{d-1} = 0$, then $kc_k = 0$, hence $c_k = 0$, for every $k \geq 1$. By Proposition 3.1.19, the solutions are exactly the polynomials $x^2 + x + c$ with $c \in \mathbb{R}$: the "$+C$" of an antiderivative is the kernel of D.

Over $\mathbb{F}_p$ the step from $kc_k = 0$ to $c_k = 0$ fails when p divides k. For instance $D(x^p) = px^{p-1} = 0$, so the kernel of D on $\mathbb{F}_p[x]$ is larger, and so is every nonempty solution set.

ASIDE: *Linear differential equations*

Proposition 3.1.19 is the reason behind the recipe "general solution = particular solution + complementary solution" for linear differential equations. For instance, $f \mapsto f'' + f$ is a linear map on the space of twice-differentiable functions $\mathbb{R} \to \mathbb{R}$, its kernel consists of the functions $a\cos x + b\sin x$, and $f(x) = x$ is one solution of $f'' + f = x$. So every solution is $x + a\cos x + b\sin x$. Identifying that kernel takes some analysis; the shape of the answer is linear algebra.

Exercise 3.1.21. ROUTINE Let $A \in \mathbb{R}^{2\times 3}$, and suppose that $(1, 2, 0)$ and $(0, 1, 1)$ both solve $A\mathbf{x} = (3, 1)$. Find a nonzero vector in $\ker T_A$, a solution of $A\mathbf{x} = (6, 2)$, and infinitely many solutions of $A\mathbf{x} = (3, 1)$.

Solution on page 1201

Consequences

Proposition 3.1.19 separates the two questions every system raises. Whether $A\mathbf{x} = \mathbf{y}$ has a solution depends on $\mathbf{y}$. How many solutions it has, once it has one, depends only on $\ker T_A$.

Corollary 3.1.22 (Uniqueness depends only on A). Let $A \in F^{m\times n}$.

(a) A consistent system $A\mathbf{x} = \mathbf{y}$ has exactly one solution if and only if $\ker T_A = \{\mathbf{0}\}$.

(b) If $A\mathbf{x} = \mathbf{y}$ has exactly one solution for some $\mathbf{y} \in F^m$, then $A\mathbf{x} = \mathbf{y}'$ has at most one solution for every $\mathbf{y}' \in F^m$.

Proof. ROUTINE

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. Proposition 3.1.19 gives a bijection between the solution set and $\ker T_A$. When does a set in bijection with $\ker T_A$ have exactly one element?
Hint 2. $\ker T_A$ always contains $\mathbf{0}$. For (b), apply (a) twice, and remember that $A\mathbf{x} = \mathbf{y}'$ might be inconsistent.
Sketch. The solution set of a consistent system is in bijection with $\ker T_A$, which contains $\mathbf{0}$, so it is a single vector exactly when $\ker T_A = \{\mathbf{0}\}$. For (b), (a) gives $\ker T_A = \{\mathbf{0}\}$, and then (a) again says that each consistent $A\mathbf{x} = \mathbf{y}'$ has one solution, while an inconsistent one has none.

(a) Let S be the solution set and $\mathbf{x}_0 \in S$. By Proposition 3.1.19, $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ is a bijection from $\ker T_A$ onto S, so S has exactly one element if and only if $\ker T_A$ does. Since $\mathbf{0} \in \ker T_A$ (Proposition 2.1.4), $\ker T_A$ has exactly one element if and only if $\ker T_A = \{\mathbf{0}\}$.

(b) By (a), applied to the system with exactly one solution, $\ker T_A = \{\mathbf{0}\}$. Let $\mathbf{y}' \in F^m$. If $A\mathbf{x} = \mathbf{y}'$ is inconsistent, it has no solutions. Otherwise it has exactly one solution, by (a) again. □

Part (a) recovers the implication "if $\ker T_A = \{\mathbf{0}\}$, then $A\mathbf{x} = \mathbf{y}$ has at most one solution for every $\mathbf{y}$" of Theorem 2.7.2, now seen from the shape of the solution set. Part (b) cannot be improved to "exactly one solution for every $\mathbf{y}'$": see Exercise 3.1.34.

Over $\mathbb{R}$, two solutions of a system determine a line, and the whole line consists of solutions. In general:

Corollary 3.1.23 (Lines through solutions). If $\mathbf{u}$ and $\mathbf{v}$ are solutions of the system $A\mathbf{x} = \mathbf{y}$ over F, then so is $\mathbf{u} + t(\mathbf{v} - \mathbf{u})$ for every $t \in F$.

Proof. ROUTINE

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 3.1.16 (Homogeneous solution sets are subspaces)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. Use Proposition 3.1.19 with $\mathbf{x}_0 = \mathbf{u}$. Where does $\mathbf{v} - \mathbf{u}$ lie?
Hint 2. $\ker T_A$ is a subspace, so it is closed under scalar multiplication.
Sketch. $\mathbf{v} - \mathbf{u} \in \ker T_A$ since both are solutions; hence $t(\mathbf{v} - \mathbf{u}) \in \ker T_A$, and $\mathbf{u} + t(\mathbf{v} - \mathbf{u}) \in \mathbf{u} + \ker T_A$, which is the solution set.

Let S be the solution set. By Proposition 3.1.19 with $\mathbf{x}_0 = \mathbf{u}$, we have $S = \mathbf{u} + \ker T_A$. Since $\mathbf{v} \in S$, $\mathbf{v} = \mathbf{u} + \mathbf{z}$ for some $\mathbf{z} \in \ker T_A$, namely $\mathbf{z} = \mathbf{v} - \mathbf{u}$. As $\ker T_A$ is a subspace (Corollary 3.1.16), $t\mathbf{z} \in \ker T_A$ for each $t \in F$, so $\mathbf{u} + t(\mathbf{v} - \mathbf{u}) = \mathbf{u} + t\mathbf{z} \in S$. □

Corollary 3.1.24 (No, one, or infinitely many solutions). Let F be an infinite field, such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. Every linear system over F has no solutions, exactly one solution, or infinitely many solutions.

Proof. ROUTINE

Uses. Corollary 3.1.23 (Lines through solutions), Proposition 1.2.35 (Zero products)
Hint 1. Suppose the system has two different solutions $\mathbf{u}$ and $\mathbf{v}$. Produce infinitely many more.

Hint 2. Use Corollary 3.1.23, and check that different values of t give different points: if $\mathbf{u}+t(\mathbf{v}-\mathbf{u})=\mathbf{u}+t'(\mathbf{v}-\mathbf{u})$, what can you say about $t-t'$?
Sketch. If $\mathbf{u}\neq\mathbf{v}$ are solutions, the points $\mathbf{u}+t(\mathbf{v}-\mathbf{u})$ are solutions, and they are distinct for distinct t because $(t-t')(\mathbf{v}-\mathbf{u})=\mathbf{0}$ with $\mathbf{v}-\mathbf{u}\neq\mathbf{0}$ forces $t=t'$. An infinite field gives infinitely many values of t.

Suppose a system has at least two solutions; we show that it has infinitely many. Let $\mathbf{u}\neq\mathbf{v}$ be solutions. By Corollary 3.1.23, $\mathbf{u}+t(\mathbf{v}-\mathbf{u})$ is a solution for every $t\in F$. These solutions are pairwise distinct: if $\mathbf{u}+t(\mathbf{v}-\mathbf{u})=\mathbf{u}+t'(\mathbf{v}-\mathbf{u})$, then $(t-t')(\mathbf{v}-\mathbf{u})=\mathbf{0}$, and since $\mathbf{v}-\mathbf{u}\neq\mathbf{0}$, Proposition 1.2.35 gives $t-t'=0$. As F is infinite, the system has infinitely many solutions. □

For a finite field the third option is impossible, and the right replacement is a count.

Corollary 3.1.25 (Counting solutions over a finite field). Let F be a finite field and $A\in F^{m\times n}$. Every consistent system $A\mathbf{x}=\mathbf{y}$ has exactly as many solutions as the homogeneous system $A\mathbf{x}=\mathbf{0}$.

Proof. ROUTINE

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Proposition 3.1.8 (Systems and the map T_A)
Hint 1. What is the solution set of the homogeneous system, as a subset of F^n?
Hint 2. Proposition 3.1.19 gives a bijection from it onto the solution set of $A\mathbf{x}=\mathbf{y}$.
Sketch. The homogeneous solution set is $\ker T_A$, and $\mathbf{z}\mapsto\mathbf{x}_0+\mathbf{z}$ is a bijection from $\ker T_A$ onto the solution set of $A\mathbf{x}=\mathbf{y}$. Finite sets in bijection have the same size.

By Proposition 3.1.8, the solution set of $A\mathbf{x}=\mathbf{0}$ is $\ker T_A$. If $A\mathbf{x}=\mathbf{y}$ is consistent with a solution $\mathbf{x}_0$, then by Proposition 3.1.19 the map $\mathbf{z}\mapsto\mathbf{x}_0+\mathbf{z}$ is a bijection from $\ker T_A$ onto its solution set. Both sets are finite, as subsets of the finite set F^n, so they have the same number of elements. □

Example 3.1.26 (Exactly two solutions). Over $\mathbb{F}_2$, the equation $x_1+x_2=1$ has exactly two solutions, $(1,0)$ and $(0,1)$, and its homogeneous version $x_1+x_2=0$ has exactly two, $(0,0)$ and $(1,1)$. So Corollary 3.1.24 really needs F to be infinite. The "line" through the two solutions, $\mathbf{u}+t(\mathbf{v}-\mathbf{u})$ with $t\in\mathbb{F}_2$, consists of $\mathbf{u}$ and $\mathbf{v}$ alone, because $\mathbb{F}_2$ has only two values of t. Once we have dimension, Corollary 3.1.25 becomes a formula: over a field with q elements, a consistent system has exactly q^k solutions, where $k=\dim\ker T_A$ (Section 4.1).

3.1.5 Pictures in $\mathbb{R}^2$ and $\mathbb{R}^3$

Over $\mathbb{R}$ we can draw what the results of this section say.

The row picture

An equation $a_1x_1+a_2x_2=b$ with $(a_1,a_2)\neq(0,0)$ describes a line in $\mathbb{R}^2$. An equation with $a_1=a_2=0$ reads $0=b$; its solution set is all of $\mathbb{R}^2$ if $b=0$ and empty otherwise. The solution set of a system is the intersection of the solution sets of its equations, so two equations in two unknowns, each with a nonzero coefficient, describe the intersection of two lines. There are three possibilities (Figure 3.1.2): the lines cross in one point, they are parallel and distinct, or they coincide.

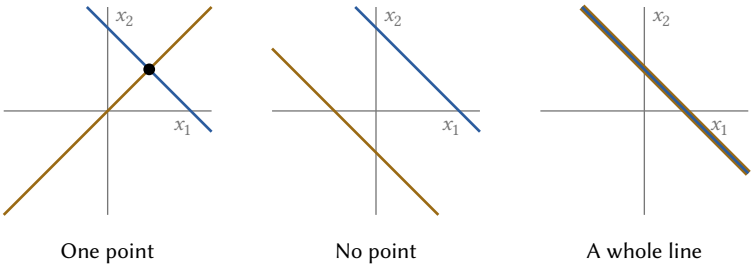


Figure 3.1.2. Two equations in two unknowns over $\mathbb{R}$. Left: $x_1 + x_2 = 2$, $x_1 - x_2 = 0$, with the single solution $(1, 1)$. Middle: $x_1 + x_2 = 2$, $x_1 + x_2 = -1$, which is inconsistent. Right: $x_1 + x_2 = 1$, $2x_1 + 2x_2 = 2$, two equations with the same solution set.

An equation in x_1, x_2, x_3 with some nonzero coefficient describes a plane in $\mathbb{R}^3$, and a system describes an intersection of planes. Two planes meet in a line, are parallel and distinct, or coincide. Three planes have more options. They can meet in a single point; pass through a common line, as the planes of the running system do (Example 3.1.13); coincide; or have no point in common. The last case occurs when two of the planes are parallel and distinct, but also in a less obvious way. Changing the right-hand side of the running system from $(1, 5, 6)$ to $(1, 5, 7)$ moves the third plane off the common line of the first two, but keeps it parallel to that line. The result is three planes, no two of them parallel, that meet in pairs along three parallel lines and have no point in common (Figure 3.1.3). Elimination detected this as the equation $0 = 1$.

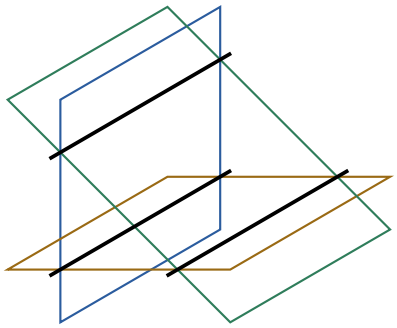


Figure 3.1.3. Three planes, no two of them parallel, with no common point. Each pair meets in a line (black), and the three lines are parallel. The planes of the running system with right-hand side $(1, 5, 7)$ are arranged like this, with the three lines parallel to $(7, -3, 1)$.

The column picture

The columns $(1, 2, 3)$, $(2, 5, 7)$, $(-1, 1, 0)$ of the running system all satisfy $y_3 = y_1 + y_2$, and this condition is preserved by sums and scalar multiples, so every combination of the columns satisfies it too. Hence $A\mathbf{x} = \mathbf{y}$ can only be consistent if $y_3 = y_1 + y_2$. The right-hand side $(1, 5, 6)$ passes this test and $(1, 5, 7)$ fails it, which explains the inconsistency without any elimination. Conversely, every $\mathbf{y}$ with $y_3 = y_1 + y_2$ is reached, since

$$(5y_1 - 2y_2)\begin{pmatrix}1\\2\\3\end{pmatrix} + (y_2 - 2y_1)\begin{pmatrix}2\\5\\7\end{pmatrix} = \begin{pmatrix}y_1\\y_2\\y_1 + y_2\end{pmatrix}.$$

So $\operatorname{im} T_A$ is the plane $y_3 = y_1 + y_2$ in $\mathbb{R}^3$. The row picture lives in the space of unknowns and the column picture in the space of right-hand sides, and the two describe the same system from opposite ends.

Many right-hand sides

Proposition 3.1.19 adds the last piece. For a fixed A, the nonempty solution sets of $A\mathbf{x} = \mathbf{y}$, as $\mathbf{y}$ varies, are all translates of the single subspace $\ker T_A$, so they are parallel to each other (Figure 3.1.4).

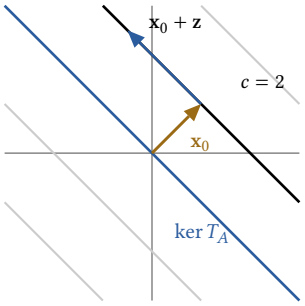


Figure 3.1.4. The solution sets of $x_1 + x_2 = c$ in $\mathbb{R}^2$ for several values of c. Each is a translate of $\ker T_A$, the solution set for $c = 0$ (blue). For $c = 2$, the particular solution $\mathbf{x}_0 = (1, 1)$ plus any $\mathbf{z} \in \ker T_A$ is again a solution.

Remark 3.1.27 (What a solution set can look like). The subspaces of $\mathbb{R}^2$ are $\{\mathbf{0}\}$, the lines through $\mathbf{0}$, and $\mathbb{R}^2$; the subspaces of $\mathbb{R}^3$ are $\{\mathbf{0}\}$, the lines and planes through $\mathbf{0}$, and $\mathbb{R}^3$. (This follows from the theory of dimension in Section 4.1.) So by Proposition 3.1.19, the solution set of a real linear system in two or three unknowns is empty, a point, a line, a plane, or the whole space. It is never a circle or a pair of points.

In any vector space, a translate $\mathbf{v} + U$ of a subspace U is called an *affine subspace* (Definition 4.7.1). In this language, Proposition 3.1.19 says that the solution set of a consistent system is an affine subspace, and that it is a translate of $\ker T_A$. Section 4.7 studies affine subspaces in their own right.

Exercises

Exercise 3.1.28. ROUTINE Which of the following are linear systems in the unknowns x_1, x_2, x_3 over $\mathbb{R}$? For each one that is, write its coefficient matrix and its augmented matrix.

(a) $x_1 = 2x_2 - x_3 + 1$ and $x_3 = 4x_1$.

(b) $x_1 x_2 = 0$ and $x_3 = 1$.

(c) $(\sin 1)x_1 + ex_2 = \pi x_3$.

(d) $2x_1 + 3 = x_2 - x_1$ and $x_2 + x_3 = \sqrt{2}$.

Solution on page 1202

Exercise 3.1.29. ROUTINE Let $\mathbf{a}_1 = (1, 2, 3)$, $\mathbf{a}_2 = (2, 1, 0)$, and $\mathbf{a}_3 = (3, 3, 3)$ in $\mathbb{R}^3$.

(a) Write $(1, 5, 9)$ as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ in every possible way.

(b) Show that $(1, 0, 0)$ is not a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

Solution on page 1202

Exercise 3.1.30. ROUTINE List all solutions of the system

$$x_1 + x_2 + x_3 = 1, \quad x_1 + 2x_2 = 0$$

over $\mathbb{F}_3$, and check that their number agrees with Corollary 3.1.25. How many solutions does the same system have over $\mathbb{R}$?

Solution on page 1203

Exercise 3.1.31. STANDARD Let p be a prime, $n \geq 1$, and $c \in \mathbb{F}_p$.

(a) Show that the equation $x_1 + x_2 + \cdots + x_n = c$ has exactly p^{n-1} solutions in $\mathbb{F}_p^n$.

(b) Deduce that a set with $n \geq 1$ elements has exactly 2^{n-1} subsets with an even number of elements.

Solution on page 1203

Solution on page 1203

Exercise 3.1.32. STANDARD Consider the system $x_1 + kx_2 = 1$, $kx_1 + x_2 = 1$.

(a) Over $\mathbb{R}$, for which k does it have no solution, exactly one solution, or infinitely many? Find the solution when it is unique.

(b) Over $\mathbb{F}_3$, find the number of solutions for each $k \in \mathbb{F}_3$.

Solution on page 1204

Exercise 3.1.33. ROUTINE Let $A \in F^{m \times n}$. Prove or disprove each statement.

(a) If $A\mathbf{x} = \mathbf{y}_1$ and $A\mathbf{x} = \mathbf{y}_2$ are consistent, then so are $A\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2$ and $A\mathbf{x} = c\mathbf{y}_1$ for every $c \in F$.

(b) If $A\mathbf{x} = \mathbf{y}_1$ and $A\mathbf{x} = \mathbf{y}_2$ are inconsistent, then so is $A\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2$.

Solution on page 1204

Exercise 3.1.34. ROUTINE Find the error in the following argument, and show by an example that its conclusion is false.

Claim. If $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then $A\mathbf{x} = \mathbf{y}$ has exactly one solution for every $\mathbf{y} \in F^m$.

Proof. Let $\mathbf{y} \in F^m$. By Proposition 3.1.19, the solution set of $A\mathbf{x} = \mathbf{y}$ is $\mathbf{x}_0 + \ker T_A = \mathbf{x}_0 + \{\mathbf{0}\} = \{\mathbf{x}_0\}$, a single vector.

Solution on page 1204

Exercise 3.1.35. STANDARD A student solves $x_1 + x_2 = 2$, $x_1 - x_2 = 0$ over $\mathbb{R}$ as follows. "Replace the first equation by the first minus the second, and the second by the second minus the first. This gives $2x_2 = 2$ and $-2x_2 = -2$. So $x_2 = 1$ and x_1 is arbitrary, and the solution set is $\{(t, 1) \mid t \in \mathbb{R}\}$." Find the error, and say exactly which solutions were gained or lost.

Solution on page 1204

Exercise 3.1.36. STANDARD Call the planes $a_1x_1 + a_2x_2 + a_3x_3 = b$ and $a_1'x_1 + a_2'x_2 + a_3'x_3 = b'$ in $\mathbb{R}^3$ (with nonzero coefficient vectors) **parallel** if $(a_1', a_2', a_3') = c(a_1, a_2, a_3)$ for some $c \in \mathbb{R}$. Find the error in the following argument, and give a counterexample to its conclusion.

Claim. Three planes in $\mathbb{R}^3$, no two of them parallel, always have a common point.

Proof. The first two planes are not parallel, so they meet in a line L. The third plane is parallel to neither of the first two, so it is not parallel to L. Hence it meets L in a single point, which lies on all three planes.

Solution on page 1205

Exercise 3.1.37. STANDARD

(a) Find all polynomials $p \in \mathbb{R}[x]_{\leq 2}$ with $p(-1) = 2$, $p(0) = 1$, and $p(1) = 4$.

(b) Find all $q \in \mathbb{R}[x]_{\leq 3}$ with the same three values, and explain the form of the answer using Proposition 3.1.19.

Solution on page 1205

Exercise 3.1.38. STANDARD Four one-way roads form a loop through the intersections A, B, C, D, carrying x_1, x_2, x_3, x_4 cars per hour as shown. In addition, 300 cars per hour enter at A, 200 leave at B, 100 enter at C, and 200 leave at D. At each intersection, the number of cars arriving per hour equals the number leaving.

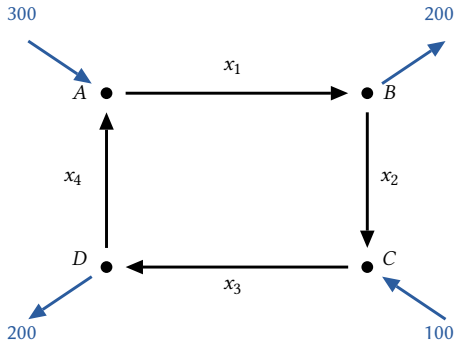


Figure 3.1.5. The road network. Black arrows are the four one-way roads; blue arrows are cars entering or leaving the network.

(a) Write down the linear system for x_1, x_2, x_3, x_4 and find all its solutions.

(b) Explain why any one of the four equations follows from the other three.

(c) Since the roads are one-way, $x_i \geq 0$ for all i. What is the smallest possible flow on each road? Can all four minima occur at once?

Solution on page 1206

Exercise 3.1.39. STANDARD A chemical equation is **balanced** when each element occurs equally often on both sides. Unknown coefficients make balancing a homogeneous linear system.

(a) Find all balancings $a\,\mathrm{Cu} + b\mathrm{HNO_3} \to c\mathrm{Cu(NO_3)_2} + d\,\mathrm{NO} + e\mathrm{H_2O}$, and the one with the smallest positive integer coefficients.

(b) Show that every balancing $a\mathrm{H_2} + b\mathrm{O_2} \to c\mathrm{H_2O} + d\mathrm{H_2O_2}$ is a linear combination of $(2, 1, 2, 0)$ and $(1, 1, 0, 1)$. What does this say chemically?

Solution on page 1206

Exercise 3.1.40. STANDARD A 2×2 grid of lights is controlled by pressing lights: pressing a light toggles it and the lights directly next to it horizontally and vertically (not diagonally). Number the lights 1 (top left), 2 (top right), 3 (bottom left), 4 (bottom right).

(a) Explain why a set of presses can be recorded as $\mathbf{x} \in \mathbb{F}_2^4$ and a pattern of lit lights as $\mathbf{y} \in \mathbb{F}_2^4$, and why pressing according to $\mathbf{x}$ switches off the pattern $\mathbf{y}$ exactly when $A\mathbf{x} = \mathbf{y}$ for a matrix $A \in \mathbb{F}_2^{4\times 4}$. Find A.

(b) Show that $A^2 = I_4$ over $\mathbb{F}_2$. Deduce that every pattern can be switched off, by exactly one set of presses.

(c) Which lights should be pressed to switch off a single lit corner?

Solution on page 1207

Exercise 3.1.41. STANDARD

(a) Find a homogeneous linear system in x_1, x_2, x_3 over $\mathbb{R}$ whose solution set is $\operatorname{span}((1,2,0),(0,1,1))$.

(b) Find a linear system in x_1, x_2, x_3 over $\mathbb{R}$ whose solution set is the line through $(1,0,2)$ and $(2,1,1)$.

Solution on page 1208

Exercise 3.1.42. STANDARD For a field F, let $S_F = \left\{(x_1, x_2) \in F^2 \;\middle|\; x_1^2 + x_2^2 = x_1 + x_2\right\}$. Prove or disprove: S_F is the solution set of some linear system in x_1, x_2 over F, when

(a) $F = \mathbb{R}$, (b) $F = \mathbb{F}_2$, (c) $F = \mathbb{F}_3$.

Solution on page 1208

Exercise 3.1.43. STANDARD Let $A \in F^{m\times n}$, and suppose that for each $i = 1, \ldots, m$ the system $A\mathbf{x} = \mathbf{e}_i$ has a solution $\mathbf{b}_i \in F^n$.

(a) Show that $A\mathbf{x} = \mathbf{y}$ is consistent for every $\mathbf{y} \in F^m$, and give an explicit solution in terms of the $\mathbf{b}_i$.

(b) Let $B \in F^{n\times m}$ be the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_m$. Show that $AB = I_m$, and conclude that T_A has a right inverse.

Solution on page 1209

Exercise 3.1.44. STANDARD

(a) Suppose that some linear system over a field F has exactly two solutions. Show that F has exactly two elements.

(b) More generally, suppose F is finite and a linear system over F has more than one solution. Show that it has at least $|F|$ solutions.

Solution on page 1209

Exercise 3.1.45. STANDARD Let V be a vector space.

(a) Let U, U' be subspaces of V and $\mathbf{v}, \mathbf{v}' \in V$. Show that if $\mathbf{v} + U = \mathbf{v}' + U'$, then $U = U'$ and $\mathbf{v}' - \mathbf{v} \in U$.

(b) Deduce that if two systems $A\mathbf{x} = \mathbf{y}$ and $B\mathbf{x} = \mathbf{z}$ in the same n unknowns have the same nonempty solution set, then $\ker T_A = \ker T_B$.

(c) Give an example of (b) with $A \neq B$, both in $\mathbb{R}^{2\times 2}$.

Solution on page 1209

Exercise 3.1.46. STANDARD Let $n \geq 3$. Find all real solutions of the cyclic system

$$x_1 + x_2 = 2, \quad x_2 + x_3 = 2, \quad \ldots, \quad x_{n-1} + x_n = 2, \quad x_n + x_1 = 2.$$

The answer depends on n. What are the solutions over $\mathbb{F}_2$?

Solution on page 1210

Exercise 3.1.47. HARD A metal plate is modelled by a finite set $G \subseteq \mathbb{Z}^2$ of interior grid points. The **boundary** ∂G is the set of points of $\mathbb{Z}^2 \setminus G$ that are adjacent (at distance 1, horizontally or vertically) to some point of G. The temperatures b_q at the boundary points $q \in \partial G$ are fixed real numbers, and in equilibrium the temperature u_p at each interior point is the average of the temperatures at its four neighbors. This is a linear system with one unknown u_p for each $p \in G$:

$$4u_p - \sum_{q\sim p, q\in G} u_q = \sum_{q\sim p, q\in \partial G} b_q \quad (p \in G),$$

where $q \sim p$ means that q and p are adjacent. Show that this system has at most one solution.

Solution on page 1210

Exercise 3.1.48. HARD BEYOND 21-242 Call a subset $S \subseteq F^n$ **line-closed** if $\mathbf{u} + t(\mathbf{v} - \mathbf{u}) \in S$ for all $\mathbf{u}, \mathbf{v} \in S$ and $t \in F$. By Corollary 3.1.23, solution sets of linear systems are line-closed.

(a) Suppose $1 + 1 \neq 0$ in F. Show that every nonempty line-closed set S is a translate $\mathbf{v} + U$ of a subspace U of F^n.

(b) Show that $\{(0,0),(1,0),(0,1)\} \subseteq \mathbb{F}_2^2$ is line-closed but is not a translate of any subspace.

Solution on page 1211

Exercise 3.1.49. CHALLENGE Thirteen weights with integer masses have the following property: whichever one of them is removed, the remaining twelve can be split into two groups of six with equal total mass. Prove that all thirteen masses are equal.

Solution on page 1211

3.2 Elementary matrices

In Example 3.1.13 we solved the running system by adding multiples of one row of its augmented matrix to another, and the paragraph after it argued that such steps never change the solution set, because each can be undone. This section turns that argument into algebra. Every row operation is left multiplication by an invertible matrix, an *elementary matrix*, so a sequence of row operations is left multiplication by a single invertible matrix G. Then "row operations do not change the solutions" becomes the statement $A\mathbf{x} = \mathbf{y} \Leftrightarrow GA\mathbf{x} = G\mathbf{y}$, and a run of Gauss–Jordan elimination (Section 3.3) is recorded in one equation, $GA = \operatorname{rref}(A)$, which we can reason about with matrix algebra.

3.2.1 Row operations as matrix multiplication

Definition 3.2.1 (Elementary row operations). Let A be a matrix with m rows, and let $1 \leq i, j \leq m$. An **elementary row operation** on A is one of the following.

1. **Swap** rows i and j, where $i \neq j$.
2. **Scale** row i by a scalar $c \neq 0$, that is, multiply every entry of row i by c.
3. **Add** c **times row** i **to row** j, where $i \neq j$ and $c \in F$. That is, replace row j by row j plus c times row i, and leave every other row, row i included, unchanged.

These are the operations of **type 1**, **type 2**, and **type 3**. An operation of type 3 is also called a **row replacement**.

These are the moves on equations from Section 3.1.3, performed on rows, with two restrictions. Both restrictions are there to make each operation reversible. Scaling a row by 0 turns it into a zero row, and nothing done afterwards can recover the old row. Adding c times row i to row i itself is scaling row i by $1 + c$, which erases the row when $c = -1$. When $i \neq j$, a replacement leaves row i in place, so subtracting c times row i from row j undoes it. (The error in Exercise 3.1.35 was to use a row that had already been replaced.)

Intuition (Why row operations should be matrix products). A row operation treats all columns of A alike: the new lth column depends only on the old lth column, and it is computed from it by the same rule for every l. Applied to a single column $\mathbf{x} = (x_1, \ldots, x_m) \in F^m$, the three operations are the maps that

- exchange the entries x_i and x_j;
- replace x_i by cx_i;
- replace x_j by $x_j + cx_i$, that is, $\mathbf{x} \mapsto \mathbf{x} + cx_i\mathbf{e}_j$.

Each is linear, since every entry of the output is a fixed linear combination of the entries of $\mathbf{x}$. Every linear map $F^m \to F^m$ is multiplication by the $m \times m$ matrix whose columns are the images of $\mathbf{e}_1, \ldots, \mathbf{e}_m$ (Proposition 2.3.17). Since $\mathbf{e}_1, \ldots, \mathbf{e}_m$ are the columns of the identity matrix I_m, those images are the columns of the matrix we get by performing the operation on I_m. So each row operation should be left multiplication by "the operation performed on I_m".

Definition 3.2.2 (Elementary matrices). Let $m \geq 1$, let $1 \leq i \leq m$, and let $c \in F$. We define three kinds of $m \times m$ matrices.

- For $1 \leq j \leq m$ with $j \neq i$, E_{ij} is the result of swapping rows i and j of I_m.
- For $c \neq 0$, E_{ic} is the result of scaling row i of I_m by c.
- For $1 \leq j \leq m$ with $j \neq i$, E_{icj} is the result of adding c times row i of I_m to row j.

A matrix of one of these forms is an **elementary matrix**, of type 1, 2, or 3 respectively. Equivalently, an $m \times m$ matrix is elementary if it is obtained from I_m by a single elementary row operation.

Each elementary matrix differs from I_m in very few entries. E_{ic} is diagonal, with c in position (i, i) and 1 elsewhere on the diagonal. E_{icj} is I_m with one more nonzero entry, c in position (j, i). That entry sits in row j, the row that changes, and in column i, the number of the row that is added. E_{ij} agrees with I_m outside rows i and j; its row i has its 1 in column j, and its row j has its 1 in column i.

Example 3.2.3 (Elementary matrices of size 3). With $m = 3$,

$$E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_{2,5} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{3,2,1} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_{1,-2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

The first swaps rows 1 and 3, the second scales row 2 by 5, the third adds 2 times row 3 to row 1, and the fourth adds -2 times row 1 to row 3.

Remark 3.2.4 (Reading the subscripts). The size m is not part of the notation. It is always clear from context, since E multiplies a matrix with m rows on the left (or, in Section 3.2.5, a matrix with m columns on the right). The number of subscripts tells the types apart: two for types 1 and 2, three for type 3. When the scalar is a specific number, we separate the subscripts by commas: $E_{2,5}$ scales row 2 by 5, and $E_{1,-2,3}$ adds -2 times row 1 to row 3. A swap carries no scalar, so its subscripts are never separated: E_{25} swaps rows 2 and 5. (We also use commas when an index is a longer expression, as in $E_{k,s,k+1}$.) A matrix can have several names. For instance $E_{ij} = E_{ji}$, and $I_m = E_{1,1}$ is itself elementary (when $m \geq 2$, it is also $E_{1,0,2}$). Some books write E_{ij} for the matrix with a single 1 in position (i, j) and zeros elsewhere; we never do.

Non-example 3.2.5 (Two operations at once). The matrix $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$ is not elementary. It differs from I_2 in all four entries, and the only elementary matrices that differ from I_2 in more than one entry are swaps, whose entries are all 0 or 1. It is the result of two operations on I_2: swapping the rows gives E_{12}, and then scaling row 1 by 2 gives $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$. By Proposition 3.2.7 below, this makes it the product $E_{1,2}E_{12}$, in which the swap acts first.

The intuition above says that we only need to know what an operation does to a single column. The following lemma makes this precise. We write $\rho[A]$ for the result of applying a map ρ to each column of A, to distinguish it from ρ applied to a vector.

Lemma 3.2.6 (Operations that act column by column). Let $\rho : F^m \to F^m$ be linear. For $A \in F^{m\times n}$ with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, let $\rho[A]$ be the $m \times n$ matrix with columns $\rho(\mathbf{a}_1), \dots, \rho(\mathbf{a}_n)$. Then

$$\rho[A] = \rho[I_m]A.$$

Proof. ROUTINE

> *Uses.* Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product), Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.8 (Standard bases), Definition 2.3.15 (The linear map of a matrix)
> *Techniques.* Check it on a basis (Technique 2.2.7)
> *Hint 1.* Let $R = \rho[I_m]$. What is the kth column of R, and what is $R\mathbf{e}_k$?
> *Hint 2.* The linear maps $\mathbf{x} \mapsto R\mathbf{x}$ and ρ agree on the standard basis. Then compare RA and $\rho[A]$ column by column.
> *Sketch.* $R = \rho[I_m]$ has columns $\rho(\mathbf{e}_1), \dots, \rho(\mathbf{e}_m)$, so $R\mathbf{e}_k = \rho(\mathbf{e}_k)$. Two linear maps that agree on a basis are equal, so $R\mathbf{x} = \rho(\mathbf{x})$ for all $\mathbf{x}$. The lth column of RA is $R\mathbf{a}_l = \rho(\mathbf{a}_l)$.

Let $R = \rho[I_m]$. The columns of I_m are $\mathbf{e}_1, \dots, \mathbf{e}_m$, so the kth column of R is $\rho(\mathbf{e}_k)$. By Definition 2.3.8, $R\mathbf{e}_k$ is the kth column of R. So the linear maps $T_R : \mathbf{x} \mapsto R\mathbf{x}$ and ρ agree on the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_m$, and by the uniqueness part of Theorem 2.2.1 they are equal: $R\mathbf{x} = \rho(\mathbf{x})$ for every $\mathbf{x} \in F^m$. By Definition 2.4.1, the lth column of RA is $R\mathbf{a}_l = \rho(\mathbf{a}_l)$, which is the lth column of $\rho[A]$. Hence $RA = \rho[A]$. □

Proposition 3.2.7 (Row operations are left multiplications). Let $A \in F^{m\times n}$, let $1 \le i, j \le m$ with $i \ne j$, and let the elementary matrices below be $m \times m$.

(a) $E_{ij}A$ is A with rows i and j swapped.

(b) If $c \ne 0$, then $E_{ic}A$ is A with row i scaled by c.

(c) $E_{icj}A$ is A with c times row i added to row j.

That is, performing an elementary row operation on A gives EA, where E is the result of performing the same operation on I_m.

The two proofs below differ in kind. The first explains why the matrix is "the operation performed on I_m", and it works for any operation that acts on each column by the same linear map. The second computes entries, and on the way it shows what left multiplication by an arbitrary matrix does to rows.

Proof (do it to the identity). ROUTINE

> *Uses.* Lemma 3.2.6 (Operations that act column by column), Definition 3.2.2 (Elementary matrices), Definition 1.2.4 (The space F^n)
> *Hint 1.* Show that performing the operation on A gives $\rho[A]$ for a linear map $\rho : F^m \to F^m$.
> *Hint 2.* Write $\rho(\mathbf{x})$ as $\mathbf{x}$ plus a multiple of $\mathbf{e}_i$ or $\mathbf{e}_j$, and apply Lemma 3.2.6.
> *Sketch.* Each operation changes every column by one map $\rho : F^m \to F^m$, for example $\rho(\mathbf{x}) = \mathbf{x} + cx_i\mathbf{e}_j$ for a replacement. Each ρ is linear because $\mathbf{x} \mapsto x_k$ is. By Lemma 3.2.6 the result is $\rho[I_m]A$, and $\rho[I_m]$ is the elementary matrix by definition.

Fix one of the three operations, and let $\rho : F^m \to F^m$ be the result of performing it on a single column $\mathbf{x} = (x_1, \dots, x_m)$. Explicitly,

$$\begin{aligned} \text{swap:} \quad & \rho(\mathbf{x}) = \mathbf{x} + (x_j - x_i)\mathbf{e}_i + (x_i - x_j)\mathbf{e}_j, \\ \text{scaling:} \quad & \rho(\mathbf{x}) = \mathbf{x} + (c-1)x_i\mathbf{e}_i, \\ \text{replacement:} \quad & \rho(\mathbf{x}) = \mathbf{x} + cx_i\mathbf{e}_j. \end{aligned}$$

(In the first line, the ith entry of $\rho(\mathbf{x})$ is $x_i + (x_j - x_i) = x_j$ and the jth entry is x_i.) Since vectors in F^m are added and scaled entry by entry (Definition 1.2.4), each map $\mathbf{x} \mapsto x_k$ is linear, and so each ρ is linear. A row operation changes every column of A by the same rule, so performing it on A gives $\rho[A]$, and performing

it on I_m gives $\rho[I_m]$, which is the elementary matrix E of Definition 3.2.2. By Lemma 3.2.6, the result of the operation on A is $\rho[A] = \rho[I_m]A = EA$. □

Proof (entry by entry). ROUTINE

> *Uses.* Definition 2.4.1 (Matrix product), Definition 2.3.8 (Matrix–vector product)
> *Hint 1.* Show that the (k,l) entry of GA is $g_{k1}a_{1l} + \cdots + g_{km}a_{ml}$ for any $G \in F^{m\times m}$.
> *Hint 2.* If row k of E equals row k of I_m, then row k of EA is row k of A. Check the remaining rows of E one type at a time.
> *Sketch.* Row k of GA is g_{k1} times row 1 of A, plus $\cdots$, plus g_{km} times row m. Rows of E that agree with I_m copy the corresponding rows of A. Row j of E_{icj} has 1 in column j and c in column i, so row j of $E_{icj}A$ is row j plus c times row i. Swaps and scalings are similar.

Let $G \in F^{m\times m}$ have entries g_{kp} and columns $\mathbf{g}_1, \ldots, \mathbf{g}_m$. By Definition 2.4.1 and Definition 2.3.8, the lth column of GA is $G\mathbf{a}_l = a_{1l}\mathbf{g}_1 + \cdots + a_{ml}\mathbf{g}_m$, whose kth entry is

$$(GA)_{kl} = g_{k1}a_{1l} + g_{k2}a_{2l} + \cdots + g_{km}a_{ml}.$$

In words: row k of GA is g_{k1} times row 1 of A, plus g_{k2} times row 2, and so on. In particular, if row k of G equals row k of I_m, with a single 1 in column k, then row k of GA is row k of A. So for $G = E$ elementary, we only have to look at the rows where E differs from I_m.

(a) Row i of E_{ij} has its 1 in column j, so row i of $E_{ij}A$ is row j of A. Likewise, row j of $E_{ij}A$ is row i of A.

(b) Row i of E_{ic} has c in column i and zeros elsewhere, so row i of $E_{ic}A$ is c times row i of A.

(c) Row j of E_{icj} has 1 in column j, c in column i, and zeros elsewhere (here $i \neq j$ matters). By the formula above, $\big(E_{icj}A\big)_{jl} = a_{jl} + ca_{il}$, so row j of $E_{icj}A$ is row j of A plus c times row i of A. □

The second proof shows more than we needed. For any $G \in F^{m\times m}$, writing $\mathbf{r}_1, \ldots, \mathbf{r}_m$ for the rows of A,

$$\text{row } k \text{ of } GA = g_{k1}\mathbf{r}_1 + g_{k2}\mathbf{r}_2 + \cdots + g_{km}\mathbf{r}_m,$$

the linear combination of the rows of A whose coefficients are the entries of row k of G. Left multiplication recombines rows. The elementary matrices are the simplest recombinations: each changes one row, using at most one other row.

Example 3.2.8 (The running system, recorded). The elimination of Example 3.1.13 performed four row operations on the augmented matrix $(A \mid \mathbf{y})$ of the running system: add -2 times row 1 to row 2; add -3 times row 1 to row 3; add -1 times row 2 to row 3; add -2 times row 2 to row 1. By Proposition 3.2.7 these are left multiplications by $E_{1,-2,2}$, $E_{1,-3,3}$, $E_{2,-1,3}$, $E_{2,-2,1}$, in that order. By associativity of matrix multiplication, the final matrix is

$$E_{2,-2,1}\big(E_{2,-1,3}\big(E_{1,-3,3}\big(E_{1,-2,2}(A \mid \mathbf{y})\big)\big)\big) = G(A \mid \mathbf{y}),$$

where

$$G = E_{2,-2,1}E_{2,-1,3}E_{1,-3,3}E_{1,-2,2}.$$

Realize that the first operation is the rightmost factor of G. To compute G, perform the same four operations on I_3 (each step is left multiplication by the next factor, by Proposition 3.2.7 again):

$$I_3 \to \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 5 & -2 & 0 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} = G.$$

Multiplying out confirms that G does the whole elimination at once:

$$\left(\begin{array}{ccc} 5 & -2 & 0 \\ -2 & 1 & 0 \\ -1 & -1 & 1 \end{array}\right)\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 5 & 1 & 5 \\ 3 & 7 & 0 & 6 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & -7 & -5 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

In Section 3.3 the same bookkeeping, applied to a complete run of Gauss–Jordan elimination, produces the equation $GA = \operatorname{rref}(A)$ (Theorem 3.3.10).

Exercise 3.2.9. ROUTINE Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$. Find elementary matrices E and E' with

$$EA = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 3 & 2 \end{pmatrix}, \quad E'A = \begin{pmatrix} 5 & 6 \\ 3 & 4 \\ 1 & 2 \end{pmatrix},$$

and write each of them out as a 3×3 array.

Solution on page 1212

3.2.2 Inverses

Each elementary row operation is undone by an operation of the same type: swap the same two rows again, scale row i by c^{-1}, or add $-c$ times row i to row j. In the language of matrices:

Proposition 3.2.10 (Inverses of elementary matrices). Every elementary matrix is invertible, and its inverse is an elementary matrix of the same type:

$$E_{ij}^{-1} = E_{ij}, \quad E_{ic}^{-1} = E_{ic^{-1}}, \quad E_{icj}^{-1} = E_{i(-c)j}.$$

Proof. ROUTINE

Uses. Proposition 3.2.7 (Row operations are left multiplications), Definition 3.2.2 (Elementary matrices), Definition 2.5.12 (Invertible)
Hint 1. Guess the inverse E' from the operation that undoes E. To verify $EE' = I_m$, apply Proposition 3.2.7 with $A = E'$: the product EE' is E' with the operation of E performed on it.
Hint 2. For type 3, you need row i of $E_{i(-c)j}$. Why is it row i of I_m?
Sketch. By Proposition 3.2.7, EE' is E' with the operation of E performed on it. Swapping rows i, j of E_{ij} gives I_m; scaling row i of $E_{ic^{-1}}$ by c gives I_m; adding c times row i of $E_{i(-c)j}$, which is row i of I_m because $i \neq j$, to its row j gives I_m. The same holds with the factors exchanged.

We check that each proposed inverse E' satisfies $EE' = I_m$ and $E'E = I_m$ (Definition 2.5.12). By Proposition 3.2.7, EE' is E' with the operation of E performed on it.

Type 1. $E_{ij}E_{ij}$ is E_{ij} with rows i and j swapped. Since E_{ij} is I_m with those rows swapped, swapping them back gives I_m.

Type 2. Let $c \neq 0$. Row i of $E_{ic^{-1}}$ has c^{-1} in column i and zeros elsewhere, and all other rows are rows of I_m. Scaling row i by c gives $cc^{-1} = 1$ in column i, so $E_{ic}E_{ic^{-1}} = I_m$. Exchanging the roles of c and c^{-1} gives $E_{ic^{-1}}E_{ic} = I_m$.

Type 3. The matrix $E_{i(-c)j}$ differs from I_m only in row j, so, as $i \neq j$, its row i is row i of I_m: a 1 in column i and zeros elsewhere. Its row j has 1 in column j and $-c$ in column i. Adding c times row i to row j turns the entry $-c$ into $-c + c = 0$ and changes nothing else, so $E_{icj}E_{i(-c)j} = I_m$. Exchanging the roles of c and $-c$ gives $E_{i(-c)j}E_{icj} = I_m$. □

The hypothesis $i \neq j$ was used exactly once, to know that the operation leaves row i alone. Without it, "adding -1 times row i to row i" would produce I_m with row i replaced by zeros, a matrix that sends $\mathbf{e}_i$ to $\mathbf{0}$ and so is not invertible.

Corollary 3.2.11 (Products of elementary matrices). Let $E_1, \dots, E_k \in F^{m\times m}$ be elementary. Then $G = E_k \cdots E_2 E_1$ is invertible, and

$$G^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

is again a product of elementary matrices. (For $k = 0$, the empty product is I_m.)

Proof. ROUTINE

Uses. Proposition 3.2.10 (Inverses of elementary matrices), Definition 2.5.12 (Invertible), induction on k
Hint 1. Induct on k. Write $G = E_k G'$, where G' is a product of $k-1$ elementary matrices.
Hint 2. Check that $H' E_k^{-1}$ is a two-sided inverse of $E_k G'$ when H' is a two-sided inverse of G'.
Sketch. If $G'H' = H'G' = I_m$, then $(E_k G')\left(H' E_k^{-1}\right) = E_k E_k^{-1} = I_m$ and $\left(H' E_k^{-1}\right)(E_k G') = H'G' = I_m$. Induction on k gives $G^{-1} = E_1^{-1} \cdots E_k^{-1}$, and each E_l^{-1} is elementary by Proposition 3.2.10.

We induct on k. For $k = 0$, $G = I_m$ is its own inverse. Let $k \geq 1$, and suppose $G' = E_{k-1} \cdots E_1$ is invertible with inverse $H' = E_1^{-1} \cdots E_{k-1}^{-1}$. Then $G = E_k G'$, and with $H = H' E_k^{-1}$, associativity gives

$$GH = E_k (G'H') E_k^{-1} = E_k E_k^{-1} = I_m, \quad HG = H'\left(E_k^{-1} E_k\right) G' = H'G' = I_m.$$

So G is invertible with $G^{-1} = H = E_1^{-1} \cdots E_k^{-1}$, as desired. By Proposition 3.2.10 each E_l^{-1} is elementary. □

Example 3.2.12 (Undoing the running elimination). For the matrix $G = E_{2,-2,1} E_{2,-1,3} E_{1,-3,3} E_{1,-2,2}$ of Example 3.2.8, Corollary 3.2.11 and Proposition 3.2.10 give

$$G^{-1} = E_{1,2,2} E_{1,3,3} E_{2,1,3} E_{2,2,1}.$$

To compute it, perform on I_3 first the operation of the rightmost factor (add 2 times row 2 to row 1), then the others from right to left. The result is

$$G^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 0 \\ 3 & 7 & 1 \end{pmatrix}.$$

Its first two columns are the first two columns $\mathbf{a}_1, \mathbf{a}_2$ of A. This is no accident: GA has first two columns $\mathbf{e}_1$ and $\mathbf{e}_2$, so $G\mathbf{a}_1 = \mathbf{e}_1$ and $G\mathbf{a}_2 = \mathbf{e}_2$, that is, $\mathbf{a}_1 = G^{-1}\mathbf{e}_1$ and $\mathbf{a}_2 = G^{-1}\mathbf{e}_2$.

Exercise 3.2.13. ROUTINE Over $\mathbb{F}_5$, write $\left(E_{12} E_{1,3} E_{2,4,1}\right)^{-1}$ as a product of 2×2 elementary matrices, and compute it as a matrix.

Solution on page 1212

3.2.3 Row operations preserve solution sets

We can now prove what Section 3.1.3 argued informally. Replacing the equations of a system by combinations of them never loses a solution, and when the replacement can be undone, it gains none either.

Proposition 3.2.14 (Invertible matrices preserve solution sets). Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$, and let $G \in F^{m\times m}$ be invertible. Then for every $\mathbf{x} \in F^n$,

$$A\mathbf{x} = \mathbf{y} \Leftrightarrow (GA)\mathbf{x} = G\mathbf{y}.$$

So the systems $A\mathbf{x} = \mathbf{y}$ and $(GA)\mathbf{x} = G\mathbf{y}$ are equivalent. Their augmented matrices are related by $G(A \mid \mathbf{y}) = (GA \mid G\mathbf{y})$.

Proof. ROUTINE

Uses. Theorem 2.4.2 (The matrix of a composite), Definition 2.5.12 (Invertible), Definition 2.4.1 (Matrix product), Definition 3.1.12 (Augmented matrix), Definition 3.1.1 (Linear system)
Hint 1. The direction $\Rightarrow$ holds for every G. Which property of G does $\Leftarrow$ need?

Hint 2. Multiply $G(A\mathbf{x}) = G\mathbf{y}$ on the left by G^{-1}, using $(GA)\mathbf{x} = G(A\mathbf{x})$.
Sketch. Since $(GA)\mathbf{x} = G(A\mathbf{x})$, applying G to $A\mathbf{x} = \mathbf{y}$ gives $(GA)\mathbf{x} = G\mathbf{y}$; applying G^{-1} to $G(A\mathbf{x}) = G\mathbf{y}$ gives back $A\mathbf{x} = \mathbf{y}$. The columns of $G(A \mid \mathbf{y})$ are $G\mathbf{a}_1, \ldots, G\mathbf{a}_n, G\mathbf{y}$.

By Theorem 2.4.2, $T_{GA} = T_G \circ T_A$, that is, $(GA)\mathbf{x} = G(A\mathbf{x})$ for every $\mathbf{x} \in F^n$; likewise $(HG)\mathbf{z} = H(G\mathbf{z})$ for all $H \in F^{m\times m}$ and $\mathbf{z} \in F^m$.

($\Rightarrow$) If $A\mathbf{x} = \mathbf{y}$, then $(GA)\mathbf{x} = G(A\mathbf{x}) = G\mathbf{y}$.

($\Leftarrow$) Suppose $(GA)\mathbf{x} = G\mathbf{y}$, that is, $G(A\mathbf{x}) = G\mathbf{y}$. Applying G^{-1} to both sides,

$$A\mathbf{x} = \big(G^{-1}G\big)(A\mathbf{x}) = G^{-1}(G(A\mathbf{x})) = G^{-1}(G\mathbf{y}) = \big(G^{-1}G\big)\mathbf{y} = \mathbf{y}.$$

So the two systems have the same solutions, which is what it means for them to be equivalent (Definition 3.1.1). Finally, by Definition 2.4.1, the columns of $G(A \mid \mathbf{y})$ are $G\mathbf{a}_1, \ldots, G\mathbf{a}_n, G\mathbf{y}$, where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the columns of A. The first n of these are the columns of GA, so $G(A \mid \mathbf{y}) = (GA \mid G\mathbf{y})$. □

Corollary 3.2.15 (Row operations preserve solution sets). If $(A' \mid \mathbf{y}')$ is obtained from $(A \mid \mathbf{y})$ by a finite sequence of elementary row operations, then the systems $A\mathbf{x} = \mathbf{y}$ and $A'\mathbf{x} = \mathbf{y}'$ are equivalent.

Proof. ROUTINE
Uses. Proposition 3.2.7 (Row operations are left multiplications), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.14 (Invertible matrices preserve solution sets)
Hint 1. Write the sequence of operations as left multiplication by one matrix G.
Hint 2. G is a product of elementary matrices. Now apply Proposition 3.2.14.
Sketch. If $E_1, \ldots, E_k$ perform the operations, then $(A' \mid \mathbf{y}') = G(A \mid \mathbf{y}) = (GA \mid G\mathbf{y})$ with $G = E_k \cdots E_1$ invertible, so $A' = GA$ and $\mathbf{y}' = G\mathbf{y}$, and Proposition 3.2.14 applies.

Let $E_1, \ldots, E_k$ be the elementary matrices of the operations, in the order performed. By Proposition 3.2.7, applied k times, and associativity,

$$(A' \mid \mathbf{y}') = E_k(\cdots(E_1(A \mid \mathbf{y}))\cdots) = G(A \mid \mathbf{y}), \quad G = E_k \cdots E_1.$$

By Proposition 3.2.14, $G(A \mid \mathbf{y}) = (GA \mid G\mathbf{y})$, so $A' = GA$ and $\mathbf{y}' = G\mathbf{y}$. Since G is invertible (Corollary 3.2.11), Proposition 3.2.14 says that $A\mathbf{x} = \mathbf{y}$ and $A'\mathbf{x} = \mathbf{y}'$ are equivalent. □

Remark 3.2.16 (Where invertibility is used). The direction $\Rightarrow$ of Proposition 3.2.14 holds for every matrix G, even a non-square one: replacing the equations of a system by combinations of them never loses a solution. Only $\Leftarrow$ needs G^{-1}, and without it solutions can be gained. The failed elimination in Exercise 3.1.35 replaced the equations of $x_1 + x_2 = 2$, $x_1 - x_2 = 0$ by "first minus second" and "second minus first" at the same time, which is left multiplication by

$$G = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

This G is not invertible: $G(1, 1) = \mathbf{0}$, whereas an invertible matrix sends only $\mathbf{0}$ to $\mathbf{0}$ (if $G\mathbf{z} = \mathbf{0}$, then $\mathbf{z} = G^{-1}G\mathbf{z} = \mathbf{0}$). Accordingly, the new system had solutions that the old one did not. Performed one after the other, the two replacements are $E_{2,-1,1}$ followed by $E_{1,-1,2}$, and their product $E_{1,-1,2}E_{2,-1,1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ is invertible. In fact a left inverse of G is all that $\Leftarrow$ needs; see Exercise 3.2.45.

Row operations change the columns of A, but they do not change the linear relations among them. This is the fact that will let us read off relations among the columns of A from a much simpler matrix.

Corollary 3.2.17 (Row operations preserve relations among columns). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and let $G \in F^{m\times m}$ be invertible, so that GA has columns $G\mathbf{a}_1, \ldots, G\mathbf{a}_n$. For all $x_1, \ldots, x_n \in F$,

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \Leftrightarrow x_1(G\mathbf{a}_1) + \cdots + x_n(G\mathbf{a}_n) = \mathbf{0}.$$

In particular:

(a) for distinct indices $l_1, \dots, l_r$, the list $\mathbf{a}_{l_1}, \dots, \mathbf{a}_{l_r}$ is linearly independent if and only if the list $G\mathbf{a}_{l_1}, \dots, G\mathbf{a}_{l_r}$ is;

(b) for scalars c_l, $\mathbf{a}_k = \sum_{l \neq k} c_l \mathbf{a}_l$ if and only if $G\mathbf{a}_k = \sum_{l \neq k} c_l G\mathbf{a}_l$.

Proof. ROUTINE

Uses. Proposition 3.2.14 (Invertible matrices preserve solution sets), Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Write both sides as matrix–vector products: $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}$.
Hint 2. Apply Proposition 3.2.14 with $\mathbf{y} = \mathbf{0}$. For (a) and (b), view a relation among some of the columns as a relation among all of them.
Sketch. By Definition 2.3.8 the two sides are $A\mathbf{x} = \mathbf{0}$ and $(GA)\mathbf{x} = \mathbf{0} = G\mathbf{0}$, which are equivalent by Proposition 3.2.14. A relation among $\mathbf{a}_{l_1}, \dots, \mathbf{a}_{l_r}$ is a relation among all columns with the other coefficients 0, so the two lists have the same relations, in particular the same nontrivial ones. For (b), use the relation with coefficient -1 on $\mathbf{a}_k$.

The columns of GA are $G\mathbf{a}_1, \dots, G\mathbf{a}_n$ by Definition 2.4.1. Let $\mathbf{x} = (x_1, \dots, x_n)$. By Definition 2.3.8, the left side of the displayed equivalence says $A\mathbf{x} = \mathbf{0}$, and the right side says $(GA)\mathbf{x} = \mathbf{0}$. Since $G\mathbf{0} = \mathbf{0}$, these are equivalent by Proposition 3.2.14 with $\mathbf{y} = \mathbf{0}$.

(a) Taking $x_l = 0$ for every l outside $\{l_1, \dots, l_r\}$, the equivalence says that $\sum_q x_{l_q}\mathbf{a}_{l_q} = \mathbf{0}$ if and only if $\sum_q x_{l_q} G\mathbf{a}_{l_q} = \mathbf{0}$. So the two lists satisfy the same linear relations. By Theorem 1.5.5, one of them is linearly independent exactly when it satisfies only the trivial relation, and then so does the other.

(b) $\mathbf{a}_k = \sum_{l \neq k} c_l \mathbf{a}_l$ says that $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$ with $x_k = -1$ and $x_l = c_l$ for $l \neq k$. By the equivalence, this holds if and only if the same relation holds among the $G\mathbf{a}_l$, which says $G\mathbf{a}_k = \sum_{l \neq k} c_l G\mathbf{a}_l$. □

Example 3.2.18 (Reading a relation off the reduced matrix). In Example 3.2.8, GA has columns $\mathbf{e}_1$, $\mathbf{e}_2$, and $(-7, 3, 0) = -7\mathbf{e}_1 + 3\mathbf{e}_2$. So the third column of GA is -7 times the first plus 3 times the second, and by part (b) of Corollary 3.2.17 the same holds in A:

$$\mathbf{a}_3 = -7\mathbf{a}_1 + 3\mathbf{a}_2, \quad \text{that is,} \quad \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = -7\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix}.$$

This is the relation $7\mathbf{a}_1 - 3\mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$ behind the homogeneous solution $(7, -3, 1)$ of Example 3.1.15. By part (a), $\mathbf{a}_1, \mathbf{a}_2$ is linearly independent, because $\mathbf{e}_1, \mathbf{e}_2$ is.

Non-example 3.2.19 (Row operations move the column space). The relations among the columns survive, but the columns themselves move, and so does their span. The columns of the running matrix A all lie in the plane $y_3 = y_1 + y_2$, which is $\operatorname{im} T_A$ (Section 3.1.5). The columns of GA all have third entry 0, so $\operatorname{im} T_{GA}$ is the plane $y_3 = 0$. A smaller example: for $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, the row operation $E_{1,-1,2}$ gives $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and $\operatorname{span}((1,1)) \neq \operatorname{span}((1,0))$. So Proposition 3.2.14 does not say that $A\mathbf{x} = \mathbf{y}$ and $(GA)\mathbf{x} = \mathbf{y}$ are equivalent: the right-hand side must be replaced by $G\mathbf{y}$.

Exercise 3.2.20. ROUTINE Let

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 0 & 2 \\ 1 & 3 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Find row operations that turn A into B. Then, using Corollary 3.2.17 and without further computation, write the third column of A as a combination of the first two, and decide whether the first two columns of A are linearly independent. Check your answers directly.

Solution on page 1213

3.2.4 Row equivalence

Because row operations can be undone, the relation "B can be reached from A by row operations" is symmetric, and it partitions matrices of a given size into classes.

Definition 3.2.21 (Row equivalence). Matrices $A, B \in F^{m\times n}$ are **row equivalent**, written $A \sim B$, if B can be obtained from A by a finite sequence of elementary row operations. By Proposition 3.2.7, this means that $B = E_k \cdots E_1 A$ for some elementary matrices $E_1, \ldots, E_k \in F^{m\times m}$, where $k \geq 0$ (for $k = 0$, $B = A$).

Proposition 3.2.22 (Row equivalence is an equivalence relation). For all $A, B, C \in F^{m\times n}$:

- (reflexivity) $A \sim A$;
- (symmetry) if $A \sim B$, then $B \sim A$;
- (transitivity) if $A \sim B$ and $B \sim C$, then $A \sim C$.

Proof. ROUTINE

Uses. Definition 3.2.21 (Row equivalence), Proposition 3.2.10 (Inverses of elementary matrices)
Hint 1. For symmetry, undo the operations in the reverse order.
Hint 2. Write $B = E_k \cdots E_1 A$ and multiply on the left by E_k^{-1}, then by E_{k-1}^{-1}, and so on. Why is each E_l^{-1} allowed?
Sketch. The empty sequence gives $A \sim A$. If $B = E_k \cdots E_1 A$, then $A = E_1^{-1} \cdots E_k^{-1} B$, and each E_l^{-1} is elementary by Proposition 3.2.10. If $B = E_k \cdots E_1 A$ and $C = E'_r \cdots E'_1 B$, then $C = E'_r \cdots E'_1 E_k \cdots E_1 A$.

Reflexivity. Take $k = 0$.

Symmetry. Let $B = E_k \cdots E_1 A$. Multiplying on the left by E_k^{-1}, then by E_{k-1}^{-1}, and so on, gives $A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} B$. By Proposition 3.2.10 each E_l^{-1} is elementary, so $B \sim A$: the operations of $E_k^{-1}, \ldots, E_1^{-1}$, performed in that order, turn B into A.

Transitivity. If $B = E_k \cdots E_1 A$ and $C = E'_r \cdots E'_1 B$, then $C = E'_r \cdots E'_1 E_k \cdots E_1 A$, so $A \sim C$. □

In this language, Corollary 3.2.15 says that row equivalent augmented matrices describe equivalent systems. Taking $\mathbf{y} = \mathbf{0}$ gives an invariant that is easy to compute.

Corollary 3.2.23 (Row equivalent matrices have the same kernel). If $A \in F^{m\times n}$ and $G \in F^{m\times m}$ is invertible, then $\ker T_{GA} = \ker T_A$. In particular, if $A \sim B$, then $\ker T_A = \ker T_B$.

Proof. ROUTINE

Uses. Proposition 3.2.14 (Invertible matrices preserve solution sets), Proposition 3.1.8 (Systems and the map T_A), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. The kernel is the solution set of a homogeneous system.
Hint 2. Apply Proposition 3.2.14 with $\mathbf{y} = \mathbf{0}$.
Sketch. $\mathbf{x} \in \ker T_{GA}$ means $(GA)\mathbf{x} = \mathbf{0} = G\mathbf{0}$, which by Proposition 3.2.14 is equivalent to $A\mathbf{x} = \mathbf{0}$. If $A \sim B$, then $B = GA$ with G a product of elementary matrices, which is invertible.

By Proposition 3.1.8, $\ker T_A$ and $\ker T_{GA}$ are the solution sets of $A\mathbf{x} = \mathbf{0}$ and $(GA)\mathbf{x} = \mathbf{0}$. Since $G\mathbf{0} = \mathbf{0}$, these systems are equivalent by Proposition 3.2.14, so the kernels are equal. If $A \sim B$, then $B = GA$ with $G = E_k \cdots E_1$, which is invertible by Corollary 3.2.11. □

Example 3.2.24 (Telling matrices apart by their kernels). The matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ are not row equivalent: the first sends (x_1, x_2) to $(x_1, 0)$ and the second to $(x_2, 0)$, so their kernels are $\operatorname{span}(\mathbf{e}_2)$ and $\operatorname{span}(\mathbf{e}_1)$. On the other hand, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix}$: swap the rows, then scale row 2 by 3.

Similarly, a matrix $A \in F^{m\times m}$ with $A \sim I_m$ is invertible: by symmetry (Proposition 3.2.22) $I_m \sim A$, so $A = GI_m = G$ for a product G of elementary matrices. The converse, that every invertible matrix is row equivalent to I_m, is proved in Section 3.3 (Theorem 3.3.23); it gives Theorem 3.5.19 in Section 3.5.

Remark 3.2.25. The converse of Corollary 3.2.23 is also true: two $m \times n$ matrices with the same kernel are row equivalent. The proof needs the reduced row echelon form of Section 3.3. Each matrix is row equivalent

to its reduced row echelon form, and a matrix in reduced row echelon form turns out to be determined by its kernel together with its size. So the kernel is a *complete* invariant of row equivalence.

Exercise 3.2.26. ROUTINE Which of the following matrices are row equivalent to $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$?

$$\text{(i)} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \quad \text{(ii)} \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}, \quad \text{(iii)} \begin{pmatrix} 0 & 0 \\ 3 & 6 \end{pmatrix}.$$

Solution on page 1213

3.2.5 Column operations

The **elementary column operations** on a matrix with n columns are defined like the row operations: swap columns i and j ($i \neq j$), scale column i by $c \neq 0$, and add c times column j to column i ($i \neq j$). Right multiplication recombines columns: by Definition 2.4.1 and Definition 2.3.8, column l of AH is

$$A\mathbf{h}_l = h_{1l}\mathbf{a}_1 + h_{2l}\mathbf{a}_2 + \cdots + h_{nl}\mathbf{a}_n,$$

the combination of the columns of A with coefficients from column l of H. So we expect column operations to be right multiplications. They are, by the same elementary matrices, with the roles of the indices exchanged.

Rows and columns are exchanged by the transpose, so we can obtain this from Proposition 3.2.7 instead of proving it from scratch. The move is worth naming.

Technique 3.2.27 (Transpose to trade rows for columns). The transpose turns the rows of a matrix into the columns of A^{T}, and it reverses products: $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$ (Proposition 2.4.23). So a statement about the rows of A, or about left multiplication, becomes a statement about the columns of A^{T}, or about right multiplication, once everything is transposed. Prove whichever version is easier, and transpose to get the other.

Reach for it whenever you have a result about rows and need its mirror image about columns, or the other way around. The canonical example is Proposition 3.2.28, where column operations come from row operations. Later instances are the equality of row rank and column rank (Section 4.3) and $\det A^{\mathsf{T}} = \det A$, which makes every column property of the determinant a row property too (Section 5.4).
Used in. Exercise 2.5.45, Proposition 3.2.28, Exercise 3.2.46, Theorem 3.5.27, Exercise 3.6.8, Problem 3.2, Theorem 4.3.8, Corollary 4.3.10, Proposition 4.3.23, Exercise 4.3.29, Problem 4.2, Theorem 5.4.14, Corollary 5.4.15, Theorem 5.5.3, Problem 7.7.

Proposition 3.2.28 (Column operations are right multiplications). Let $A \in F^{m\times n}$, let $1 \leq i, j \leq n$ with $i \neq j$, and let the elementary matrices below be $n \times n$.

(a) AE_{ij} is A with columns i and j swapped.

(b) If $c \neq 0$, then AE_{ic} is A with column i scaled by c.

(c) AE_{icj} is A with c times column j added to column i.

Proof (transpose). ROUTINE

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 2.4.23 (Transpose of a product), Definition 2.4.21 (Transpose)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. Compute $(AE)^{\mathsf{T}}$ with Proposition 2.4.23. What is E^{T} for each type?
Hint 2. E_{icj} has its extra entry c in position (j, i). Which elementary matrix has c in position (i, j)?
Sketch. $(AE)^{\mathsf{T}} = E^{\mathsf{T}}A^{\mathsf{T}}$. Swaps and scalings are symmetric, and $E_{icj}^{\mathsf{T}} = E_{jci}$. So $E^{\mathsf{T}}A^{\mathsf{T}}$ is A^{T} with a row operation performed, by Proposition 3.2.7; the rows of A^{T} are the columns of A, and E_{jci} adds c times row j to row i. Transposing back gives the column statement.

By Proposition 2.4.23, $(AE)^{\mathsf{T}} = E^{\mathsf{T}}A^{\mathsf{T}}$, so $AE = \left(E^{\mathsf{T}}A^{\mathsf{T}}\right)^{\mathsf{T}}$. Row k of A^{T} is column k of A, written as a row (Definition 2.4.21).

(c) E_{icj} is I_n with an extra entry c in position (j,i), so E_{icj}^{T} is I_n with c in position (i,j), which is E_{jci}. By Proposition 3.2.7, $E_{jci}A^{\mathsf{T}}$ is A^{T} with c times row j added to row i. Transposing turns rows into columns, so $AE_{icj} = \left(E_{jci}A^{\mathsf{T}}\right)^{\mathsf{T}}$ is A with c times column j added to column i.

(a), (b) E_{ij} and E_{ic} are symmetric: $E_{ij}^{\mathsf{T}} = E_{ij}$ and $E_{ic}^{\mathsf{T}} = E_{ic}$. By Proposition 3.2.7, $E_{ij}A^{\mathsf{T}}$ and $E_{ic}A^{\mathsf{T}}$ are A^{T} with rows i and j swapped, respectively with row i scaled by c, and transposing gives the claims. □

Proof (directly). ROUTINE

> *Uses.* Definition 2.4.1 (Matrix product), Definition 2.3.8 (Matrix–vector product)
> *Hint 1.* By Definition 2.4.1, column l of AE is A times column l of E.
> *Hint 2.* Find the columns of E_{icj}. Which of them differs from a standard basis vector?
> *Sketch.* Column l of AE is A times column l of E. Column i of E_{icj} is $\mathbf{e}_i + c\mathbf{e}_j$ and the others are $\mathbf{e}_l$, so column i of AE_{icj} is $\mathbf{a}_i + c\mathbf{a}_j$ and the others are unchanged. Columns i and j of E_{ij} are $\mathbf{e}_j$ and $\mathbf{e}_i$, and column i of E_{ic} is $c\mathbf{e}_i$.

By Definition 2.4.1, column l of AE is $A\mathbf{u}_l$, where $\mathbf{u}_l$ is column l of E, and $A\mathbf{e}_l = \mathbf{a}_l$ by Definition 2.3.8.

(c) Column l of E_{icj} is $\mathbf{e}_l$ for $l \neq i$, and column i is $\mathbf{e}_i + c\mathbf{e}_j$, since the extra entry c is in row j of column i. So column l of AE_{icj} is $\mathbf{a}_l$ for $l \neq i$, and column i is $A\left(\mathbf{e}_i + c\mathbf{e}_j\right) = \mathbf{a}_i + c\mathbf{a}_j$.

(a), (b) Columns i and j of E_{ij} are $\mathbf{e}_j$ and $\mathbf{e}_i$, and column i of E_{ic} is $c\mathbf{e}_i$; all other columns are those of I_n. So AE_{ij} has columns i and j equal to $\mathbf{a}_j$ and $\mathbf{a}_i$, and AE_{ic} has column i equal to $c\mathbf{a}_i$, with the other columns unchanged. □

The direct proof is shorter here. The transposed proof shows that the column statement carries no new information: it is the row statement seen in a mirror.

Remark 3.2.29 (One matrix, two operations). The matrix E_{icj}, which has c in position (j,i), is at once the result of adding c times row i to row j of the identity and the result of adding c times column j to column i of the identity. So the subscript icj reads left to right for rows ("from row i into row j") and right to left for columns ("from column j into column i"). For example, with $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$,

$$AE_{1,-2,2} = \begin{pmatrix} 1-4 & 2 & 3 \\ 4-10 & 5 & 6 \end{pmatrix} = \begin{pmatrix} -3 & 2 & 3 \\ -6 & 5 & 6 \end{pmatrix},$$

where $E_{1,-2,2}$ is 3×3 and adds -2 times column 2 to column 1. The 2×2 matrix with the same name acts on the left and adds -2 times row 1 to row 2: $E_{1,-2,2}A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$. Also note the order of factors: if column operations with matrices $E_1, \ldots, E_k$ are performed on A in that order, the result is $AE_1E_2\cdots E_k$, so the first operation is now the *leftmost* factor.

Row operations do not change the kernel (Corollary 3.2.23) but can change the image (Non-example 3.2.19). Column operations behave the other way around.

> **Proposition 3.2.30** (Column operations preserve the image). Let $A \in F^{m\times n}$, and let $H \in F^{n\times n}$ be invertible. Then $\operatorname{im} T_{AH} = \operatorname{im} T_A$. In particular, column operations do not change the span of the columns of a matrix.

Proof. ROUTINE

> *Uses.* Definition 2.6.6 (Image), Theorem 2.4.2 (The matrix of a composite), Definition 2.5.12 (Invertible), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.1.8 (Systems and the map T_A)
> *Techniques.* Double containment (Technique 1.3.14)
> *Hint 1.* Prove the two containments separately.
> *Hint 2.* $AH\mathbf{x} = A(H\mathbf{x})$, and $A\mathbf{x} = (AH)\left(H^{-1}\mathbf{x}\right)$.
> *Sketch.* Every $(AH)\mathbf{x} = A(H\mathbf{x})$ lies in $\operatorname{im} T_A$. Conversely $A\mathbf{x} = (AH)\left(H^{-1}\mathbf{x}\right) \in \operatorname{im} T_{AH}$. Column operations multiply on the right by a product of elementary matrices, which is invertible.

We prove both containments, using $(AH)\mathbf{x} = A(H\mathbf{x})$ (Theorem 2.4.2) throughout.

$\operatorname{im} T_{AH} \subseteq \operatorname{im} T_A$. For $\mathbf{x} \in F^n$, $(AH)\mathbf{x} = A(H\mathbf{x}) \in \operatorname{im} T_A$.

$\operatorname{im} T_A \subseteq \operatorname{im} T_{AH}$. For $\mathbf{x} \in F^n$, $A\mathbf{x} = A\big(\big(HH^{-1}\big)\mathbf{x}\big) = (AH)\big(H^{-1}\mathbf{x}\big) \in \operatorname{im} T_{AH}$.

A sequence of column operations multiplies A on the right by a product of elementary matrices (Proposition 3.2.28), which is invertible by Corollary 3.2.11, and $\operatorname{im} T_A$ is the span of the columns of A (Proposition 3.1.8). □

Non-example 3.2.31 (Column operations change the kernel). For $A = \begin{pmatrix} 1 & 1 \end{pmatrix}$, $\ker T_A = \operatorname{span}((1,-1))$. Adding -1 times column 2 to column 1 gives $AE_{1,-1,2} = \begin{pmatrix} 0 & 1 \end{pmatrix}$, whose kernel is $\operatorname{span}((1,0))$. In general $\ker T_{AH} = \left\{H^{-1}\mathbf{x} \;\middle|\; \mathbf{x} \in \ker T_A\right\}$: the kernel is moved by H^{-1}.

This is why we solve systems with row operations. The solution set of a consistent system $A\mathbf{x} = \mathbf{y}$ is a translate of $\ker T_A$ (Proposition 3.1.19), and row operations keep the kernel and the solution set. Column operations keep the image, so they suit questions about the span of the columns, and both kinds together reduce every matrix to a very simple form, the rank normal form of Theorem 3.5.27.

> **Exercise 3.2.32.** ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Find a 3×3 elementary matrix E with $AE = \begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 6 \end{pmatrix}$. What does the same matrix E do to a 3×3 matrix B when it multiplies on the left?

Solution on page 1213

3.2.6 Elementary matrices in the plane

Over $\mathbb{R}$, a 2×2 matrix E is a linear map T_E of the plane, and we can watch what it does. Realize that a vector $\mathbf{x} \in \mathbb{R}^2$ is itself a 2×1 matrix, so by Proposition 3.2.7, $E\mathbf{x}$ is $\mathbf{x}$ with the row operation of E performed on its two entries. The elementary 2×2 matrices are

$$E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_{1c} = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}, \quad E_{2c} = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}, \quad E_{1c2} = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}, \quad E_{2c1} = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix},$$

with $c \neq 0$ for the two scalings, and they act as follows.

Example 3.2.33 (The elementary maps of the plane).

- *Swap.* $T_{E_{12}}(x_1, x_2) = (x_2, x_1)$ is the reflection across the line $x_1 = x_2$. It fixes that line pointwise and reverses orientation: in Figure 3.2.1 the letter F is mirrored.
- *Scaling.* $T_{E_{1c}}(x_1, x_2) = (cx_1, x_2)$. For $c > 0$ it stretches the plane horizontally by the factor c (a compression if $c < 1$) and fixes the x_2-axis pointwise. For $c < 0$ it stretches by $|c|$ and also reflects across the x_2-axis. E_{2c} does the same vertically.
- *Replacement.* $T_{E_{2c1}}(x_1, x_2) = (x_1 + cx_2, x_2)$ is a horizontal **shear**. It slides each horizontal line $x_2 = t$ along itself by ct, so the x_1-axis is fixed pointwise, and the higher a point is, the farther it moves. The unit square becomes a parallelogram with the same base and the same height. E_{1c2} is a vertical shear.

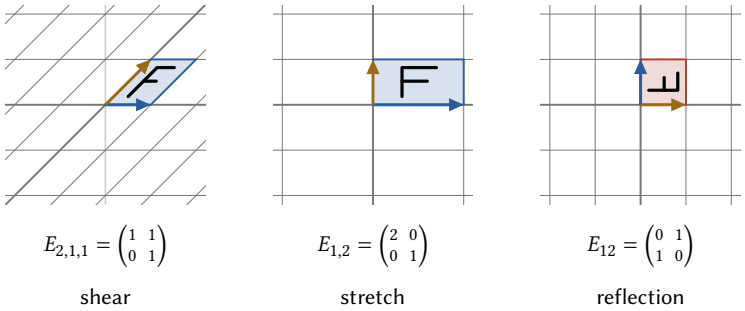
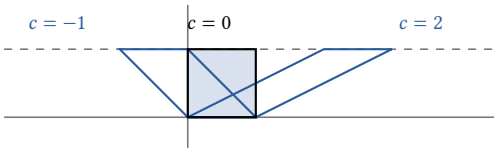


Figure 3.2.1. The three types of 2×2 elementary matrices acting on the plane. Each panel shows the images of the grid, of the unit square, and of $\mathbf{e}_1$ (blue) and $\mathbf{e}_2$ (ochre). The reflection reverses orientation, so its image of the letter F is mirrored and its square is drawn in red.

The inverses of Proposition 3.2.10 are visible here: the shear by c is undone by the shear by $-c$, the stretch by c by the stretch by c^{-1}, and the reflection by itself.



The plane transformer, starting at the shear $E_{2,1,1}$. Move $A\mathbf{e}_2$ along the horizontal grid line $x_2 = 1$ (focus it and use the left and right arrow keys): every position gives a shear $E_{2c1} = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$. Move $A\mathbf{e}_1$ along the x_1-axis back and forth through 0 to see the scalings E_{1c}, and how the square flattens at $c = 0$, which is why $c = 0$ is excluded. For the swap, type the entries $0, 1, 1, 0$. (Interactive in the web version.)

Example 3.2.34 (A rotation built from shears). The rotation of the plane by 90° counterclockwise sends $\mathbf{e}_1 \mapsto \mathbf{e}_2$ and $\mathbf{e}_2 \mapsto -\mathbf{e}_1$, so its matrix is $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. It is not elementary, but it is a product of three shears. With $S = E_{2,-1,1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $L = E_{1,1,2} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$,

$$SLS = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = R.$$

Figure 3.2.2 shows the three steps: a horizontal shear, a vertical shear, and the first shear again. The same R is also a product of two reflections, $R = E_{1,-1}E_{12}$: first reflect across the line $x_1 = x_2$, then across the x_2-axis. So a matrix can be written as a product of elementary matrices in several ways, and the number and types of the factors are not determined by the matrix.

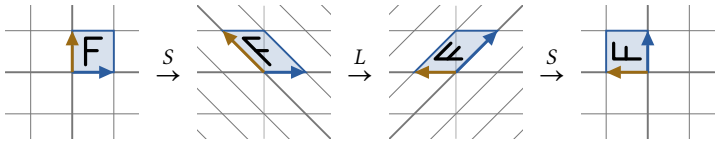


Figure 3.2.2. The rotation by 90° as three shears, $R = SLS$. From left to right: the unit square and the letter F, then their images under S, under LS, and under $SLS = R$.

Intuition (Elementary maps as building blocks). Section 3.5 proves that every invertible matrix is a product of elementary matrices (Theorem 3.5.19). For the plane this says that every invertible linear map of $\mathbb{R}^2$ is a composition of shears, stretches along the axes (possibly with a reflection), and the reflection across

$x_1 = x_2$. The figures also show how each building block changes area. A shear keeps the area of the unit square. The stretch E_{1c} multiplies it by $|c|$. The swap keeps it but reverses orientation. Recorded with signs, these are the numbers 1, c, and -1, which will be the determinants of the three types, and the determinant of any invertible matrix is then pinned down by its factorization into elementary matrices.

Remark 3.2.35 (Where elementary matrices reappear). Elementary matrices are bookkeeping for elimination, and they come back several times.

- In Section 3.3, a run of Gauss–Jordan elimination on A is recorded by a product $G = E_k \cdots E_1$ with $GA = \operatorname{rref}(A)$ (Theorem 3.3.10). By Proposition 3.2.14, the system $A\mathbf{x} = \mathbf{y}$ can then be solved in the form $\operatorname{rref}(A)\mathbf{x} = G\mathbf{y}$ (Section 3.4).
- In Section 3.5, every invertible matrix turns out to be a product of elementary matrices (Theorem 3.5.19). This yields a proof technique, reducing to elementary matrices (Technique 5.4.6): to prove a statement about all invertible matrices that is preserved under products, prove it for each of the three types.
- In Section 5.4, that technique computes $\det(AE) = \det A \cdot \det E$ type by type, and from it $\det(AB) = \det A \cdot \det B$ (Theorem 5.4.10) and $\det A^{\mathsf{T}} = \det A$ (Theorem 5.4.14).

Exercises

Exercise 3.2.36. ROUTINE Which of the following matrices are elementary? For each one that is, name the row operation it performs and write down its inverse. All matrices are over $\mathbb{R}$, except (h), which is over $\mathbb{F}_3$.

$$\text{(a)}\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\quad \text{(b)}\ \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix},\quad \text{(c)}\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad \text{(d)}\ \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix},\quad \text{(e)}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{(f)}\ \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad \text{(g)}\ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},\quad \text{(h)}\ \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},\quad \text{(i)}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Solution on page 1214

Exercise 3.2.37. ROUTINE For each matrix A below, find elementary matrices $E_1, \ldots, E_k$ with $E_k \cdots E_1 A = I_2$, compute $G = E_k \cdots E_1$, and check that $AG = I_2$ as well.

(a) $A = \begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}$ over $\mathbb{F}_5$.

(b) $A = \begin{pmatrix} 1 & i \\ i & 2 \end{pmatrix}$ over $\mathbb{C}$.

Solution on page 1214

Exercise 3.2.38. ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Compute $E_{1,2,3}A$ and $AE_{1,2,3}$ and describe each in words. Then find the elementary matrix E for which AE is A with 2 times column 1 added to column 3.

Solution on page 1215

Exercise 3.2.39. ROUTINE Let $i \neq j$ and $c, d \in F$.

(a) Show that $E_{icj}E_{idj} = E_{i(c+d)j}$ and, for $c, d \neq 0$, that $E_{ic}E_{id} = E_{i(cd)}$.

(b) Deduce that $E_{icj}^k = E_{i(kc)j}$ for every positive integer k, and that $E_{icj}^p = I_m$ when $F = \mathbb{F}_p$.

Solution on page 1215

Exercise 3.2.40. STANDARD Let $m = 3$ and $c, d \in F$.

(a) Compute $E_{2d3}E_{1c2}$ and $E_{1c2}E_{2d3}$, and explain the difference in terms of row operations.

(b) Show that $E_{2d3}E_{1c2} = E_{1(cd)3}E_{1c2}E_{2d3}$.

(c) Show that E_{1c2} commutes with E_{1d3} and with E_{3d2}.

Solution on page 1216

Exercise 3.2.41. STANDARD Let $i \neq j$.

(a) Show that $E_{ij} = E_{i,-1}E_{i,1,j}E_{j,-1,i}E_{i,1,j}$. So a swap can be performed with three replacements and one scaling.

(b) Show that over $\mathbb{F}_2$, E_{ij} is a product of three elementary matrices of type 3.

(c) Deduce that every product of elementary matrices is a product of elementary matrices of types 2 and 3 only.

Solution on page 1216

Exercise 3.2.42. STANDARD For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, let $D(A) = ad - bc$.

(a) Show that $D(EA) = D(A)$ for every 2×2 elementary matrix E of type 3, that $D(E_{12}A) = -D(A)$, and that $D(E_{is}A) = sD(A)$ for $i = 1, 2$ and $s \neq 0$.

(b) Suppose $1 + 1 \neq 0$ in F. Show that E_{12} is not a product of elementary matrices of type 3. Why does this not contradict Exercise 3.2.41(b)?

(c) Is $E_{1,2} \in \mathbb{R}^{2\times 2}$ a product of elementary matrices of type 3?

Solution on page 1217

Exercise 3.2.43. ROUTINE Find the error in the following argument, and give a counterexample to its conclusion.

Claim. If $A \in F^{m\times n}$ and $G \in F^{m\times m}$ is invertible, then $\operatorname{im} T_{GA} = \operatorname{im} T_A$.

Proof. Let $\mathbf{y} \in F^m$. By Proposition 3.2.14, the system $(GA)\mathbf{x} = \mathbf{y}$ has a solution if and only if $A\mathbf{x} = \mathbf{y}$ has one. So $\mathbf{y} \in \operatorname{im} T_{GA}$ if and only if $\mathbf{y} \in \operatorname{im} T_A$.

Solution on page 1217

Exercise 3.2.44. STANDARD Let $A, B \in F^{m\times n}$ with $A \sim B$. Prove or disprove each statement.

(a) $A^{\mathsf{T}} \sim B^{\mathsf{T}}$.

(b) $AC \sim BC$ for every $C \in F^{n\times p}$.

(c) $CA \sim CB$ for every $C \in F^{q\times m}$.

(d) If also $C \sim D$ in $F^{m\times n}$, then $A + C \sim B + D$.

(e) If $m = n$ and A is invertible, then B is invertible.

Solution on page 1217

Exercise 3.2.45. STANDARD Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$, and let $G \in F^{p\times m}$, where p need not equal m.

(a) Suppose G has a left inverse, that is, $HG = I_m$ for some $H \in F^{m\times p}$. Show that $A\mathbf{x} = \mathbf{y}$ and $(GA)\mathbf{x} = G\mathbf{y}$ have the same solutions.

(b) Give an example with $m = 2$ and $p = 3$, and interpret it as an operation on the equations of a system.

(c) Conversely, suppose that for every $A \in F^{m\times 1}$ and every $\mathbf{y} \in F^m$ the systems $A\mathbf{x} = \mathbf{y}$ and $(GA)\mathbf{x} = G\mathbf{y}$ have the same solutions. Show that T_G is injective.

Solution on page 1218

Exercise 3.2.46. STANDARD The **row space** of $A \in F^{m\times n}$ is the span of its rows, viewed as vectors in F^n; equivalently, it is $\operatorname{im} T_{A^{\mathsf{T}}}$. Show that row equivalent matrices have the same row space.

Solution on page 1218

Solution on page 1219

Exercise 3.2.47. STANDARD

(a) Show that two $1 \times n$ matrices are row equivalent if and only if one is a nonzero scalar multiple of the other.

(b) Find all $B \in \mathbb{R}^{2\times 2}$ with $B \sim \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$.

Solution on page 1219

Exercise 3.2.48. STANDARD Let p be a prime and $m \geq 1$. How many different matrices in $\mathbb{F}_p^{m\times m}$ are elementary? List them for $p = 2$ and $m = 2$.

Solution on page 1219

Exercise 3.2.49. STANDARD Let $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ be the rotation of $\mathbb{R}^2$ by θ, and for $t \in \mathbb{R}$ let $S(t) = E_{2t1} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ and $L(t) = E_{1t2} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$.

(a) Show that if θ is not an odd multiple of π, then $R_\theta = S(-\tan(\theta/2))\, L(\sin\theta)\, S(-\tan(\theta/2))$.

(b) Write $R_\pi = -I_2$ as a product of five shears.

Solution on page 1220

Exercise 3.2.50. STANDARD

(a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$ with $ad - bc \neq 0$. Find elementary matrices $E_1, \ldots, E_k$ with $E_k \cdots E_1 A = I_2$, and conclude that A is a product of elementary matrices.

(b) Carry this out for $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, and describe T_A as a composition of shears, stretches, and reflections.

Solution on page 1220

Exercise 3.2.51. STANDARD Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} \in \mathbb{R}^{2\times 3}$.

(a) Find elementary column operations that turn A into $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, and compute the product H of the corresponding elementary matrices, so that $AH = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(b) Let $\mathbf{h}_3$ be the third column of H. Show that $\ker T_A = \operatorname{span}(\mathbf{h}_3)$.

Solution on page 1221

Exercise 3.2.52. STANDARD Let $m \geq 2$ and $M \in F^{m\times m}$. Show that M commutes with every elementary matrix of type 3 if and only if $M = \lambda I_m$ for some $\lambda \in F$.

Solution on page 1221

Exercise 3.2.53. HARD Let F be any field and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$ with $ad - bc = 1$. Show that A is a product of at most four elementary matrices of type 3, that is, of shears.

Solution on page 1222

Exercise 3.2.54. HARD Let a, b be integers, not both 0, and let $g = \gcd(a, b)$. Call a 2×2 matrix an **integer elementary matrix** if it is E_{12}, $E_{1,-1}$, $E_{2,-1}$, or E_{icj} with $c \in \mathbb{Z}$.

(a) Show that integer elementary matrices have integer entries, and that their inverses are integer elementary matrices.

(b) Show that there are integer elementary matrices $E_1, \ldots, E_k$ with $E_k \cdots E_1 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$.

(c) Deduce Bézout's identity: there are integers u, v with $ua + vb = g$.

(d) Carry out (b) and (c) for $(a, b) = (42, 30)$.

Solution on page 1222

Exercise 3.2.55. HARD Start with the integer matrix $M = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$. A **move** adds an integer multiple of one row to the other row, or an integer multiple of one column to the other column.

(a) Show that a move changes neither $ad - bc$ nor the greatest common divisor of the four entries of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(b) Which of $\begin{pmatrix} 1 & 0 \\ 0 & -8 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, and $\begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$ can be reached from M by a sequence of moves?

Solution on page 1223

Exercise 3.2.56. CHALLENGE BEYOND 21-242 Let $n \geq 2$ and let $A \in F^{n\times n}$ be invertible. Using Theorem 3.5.19, show that

$$A = PE_{nd}$$

for some product P of elementary matrices of type 3 and some $d \neq 0$. (In Chapter 5, d turns out to be $\det A$.)

Solution on page 1224

3.3 Gauss–Jordan elimination

In Example 3.1.13 we solved the running system by eliminating unknowns, choosing each move by eye and stopping when the answer looked readable. To prove theorems with elimination we need a procedure with no judgment calls in it. This section states one, Gauss–Jordan elimination, proves that it terminates, and studies its output, the reduced row echelon form $\operatorname{rref}(A)$. With the bookkeeping of Section 3.2, a whole run of the algorithm is recorded by one invertible matrix G with $GA = \operatorname{rref}(A)$. The main theorem of the section says that $\operatorname{rref}(A)$ depends only on A: every way of row reducing A to a matrix of this shape ends at the same matrix. As a first application, we decide whether a square matrix is invertible, and compute its inverse, by row reduction.

3.3.1 The algorithm

Intuition (How elimination is organized). Read the rows of A as equations and the columns as unknowns. Elimination uses one equation to remove one unknown from all the other equations, and the natural order is left to right. Find the first unknown x_j that occurs in some equation, move such an equation to the top, divide it by its coefficient of x_j, and subtract multiples of it from every other equation, so that x_j occurs in no other equation. The top equation is now used up; we call its row *pivoted*. Then do the same with the equations that are left: find the leftmost unknown that still occurs in one of them, and so on.

Gauss–Jordan elimination clears the column of each new pivot above it as well as below it, so at the end each pivot is alone in its column. (Clearing only below gives Gaussian elimination, Section 3.3.6.) A column whose nonzero entries all lie in pivoted rows is skipped: it is already a combination of earlier pivot columns, and in Section 3.4 its unknown becomes a free variable.

In the algorithm below, M is the current matrix and k is the number of pivoted rows. The pivoted rows are always rows $1, \ldots, k$ of M, and the other rows are **unpivoted**.

Algorithm 3.3.1 (Gauss–Jordan elimination)

Input A matrix $A \in F^{m \times n}$.
Output The matrix $\operatorname{rref}(A) \in F^{m \times n}$, and its pivot columns $p_1 < p_2 < \cdots < p_r$.

Start with $M = A$ and $k = 0$.

1. **Scan.** Find the leftmost column j of M that has a nonzero entry in an unpivoted row, that is, in one of the rows $k+1, \ldots, m$. If there is no such column, halt and output M.
2. **Swap.** If the entry of M in position $(k+1, j)$ is 0, swap row $k+1$ with the topmost row below it whose entry in column j is nonzero.
3. **Scale.** Scale row $k+1$ by the inverse of its entry in column j, so that this entry becomes 1. (If it is 1 already, there is nothing to do.) Declare this entry a **pivot**, set $p_{k+1} = j$, and increase k by 1. Row k is now pivoted.
4. **Clear.** For each row $i \neq k$ whose entry c in column j is nonzero, add $-c$ times row k to row i.
5. **Repeat.** Go back to step 1.

One execution of step 1 through step 4 that does not halt in step 1 is a **pass**. Each pass creates exactly one pivot, so after r passes the pivots sit in positions $(1, p_1), (2, p_2), \ldots, (r, p_r)$. (Later passes never move a pivoted row: swaps only involve unpivoted rows.) We call $p_1, \ldots, p_r$ the **pivot columns** and rows $1, \ldots, r$ the **pivot rows**. The **leading entry** of a nonzero row is its first nonzero entry, reading from the left.

Each step can be carried out. In step 2, column j has a nonzero entry in some unpivoted row, so if row $k+1$ has a 0 there, a row below it does not. In step 3 the entry is nonzero, so it has an inverse in F. In step 4 every operation adds a multiple of row k, which the step never changes, so the operations can be performed one at a time in any order with the same result. (Compare Exercise 3.1.35, where a row was used after it had been replaced.)

Remark 3.3.2 (The one free choice). Every step is determined by the current matrix except, in principle, step 2. The lecture's version of that step says only "swap rows to make the topmost unpivoted entry of this column nonzero", leaving open which row to bring up. We have fixed the choice (the topmost suitable row, and no swap when none is needed) so that the output is a function of A. Other choices have their uses. By hand one might bring up a row whose entry is 1, to avoid fractions; in floating-point arithmetic one brings up the entry of largest absolute value to limit rounding errors (*partial pivoting*). Whether a different choice can change the output is a real question. Corollary 3.3.18 answers it: it cannot.

Example 3.3.3 (A complete run). We run the algorithm on

$$A = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 2 & 4 & 2 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix} \in \mathbb{Q}^{3\times 4}.$$

Pass 1. No row is pivoted, and column 1 has nonzero entries, so $j = 1$. The entry in position $(1, 1)$ is 0, and the topmost row below it with a nonzero entry in column 1 is row 2, so we swap rows 1 and 2. Then we scale row 1 by $1/2$ and declare the pivot $(1, 1)$, so $p_1 = 1$. In column 1, row 2 already has a 0 and row 3 has a 1, so we add -1 times row 1 to row 3:

$$A \xrightarrow{E_{12}} \begin{pmatrix} 2 & 4 & 2 & 0 \\ 0 & 0 & 2 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \xrightarrow{E_{1,1/2}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

$$\xrightarrow{E_{1,-1,3}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{pmatrix}.$$

Pass 2. Rows 2 and 3 are unpivoted. Column 2 has entries $0, 0$ in these rows (the 2 in row 1 does not count, since row 1 is pivoted), so the scan moves on and finds $j = 3$. The entry in position $(2, 3)$ is $2 \neq 0$, so there is no swap. We scale row 2 by $1/2$ and declare the pivot $(2, 3)$, so $p_2 = 3$. Then we clear column 3 above and below the pivot:

$$\xrightarrow{E_{2,1/2}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix} \xrightarrow{E_{2,-1,1}} \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix}$$

$$\xrightarrow{E_{2,-2,3}} \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

End. Only row 3 is unpivoted, and it is zero, so the next scan finds no column and the algorithm halts after two passes. The output is

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

with pivot columns $p_1 = 1$ and $p_2 = 3$. The run used six row operations, with elementary matrices (in the notation of Definition 3.2.2) E_{12}, $E_{1,1/2}$, $E_{1,-1,3}$, $E_{2,1/2}$, $E_{2,-1,1}$, $E_{2,-2,3}$.

The row reducer below performs the same run, one row operation at a time. Next to the current matrix M it shows the product G of the elementary matrices used so far, and $GA = M$ at every step (Theorem 3.3.10).

$$\begin{pmatrix} 0 & 0 & 2 & 4 \\ 2 & 4 & 2 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Gauss–Jordan elimination on the matrix of Example 3.3.3. Step forward and back to see each elementary matrix; pivots are boxed. Use "Edit" to row reduce your own matrix, over $\mathbb{Q}$ or over $\mathbb{F}_p$. (Interactive in the web version.)

The running system is also a good test. In Example 3.1.13 the four operations on its coefficient matrix were exactly the two passes of the algorithm (in pass 2 we cleared row 3 before row 1, which does not matter). So Example 3.2.8 already computed

$$\operatorname{rref}\begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 1 \\ 3 & 7 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -7 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix},$$

together with the product G of the four elementary matrices.

Exercise 3.3.4. ROUTINE Run Gauss–Jordan elimination on

$$B = \begin{pmatrix} 0 & 2 & 4 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{pmatrix} \in \mathbb{Q}^{3\times 3},$$

listing the elementary matrix of every row operation. What are the pivot columns?

Solution on page 1224

3.3.2 Termination and the shape of the output

An algorithm is a construction, and a construction can be a proof: running Gauss–Jordan elimination on A will prove that A is row equivalent to a matrix of a very special shape. For that we have to show two things about every run, not just the one in Example 3.3.3. It stops, and what it stops with has the shape we want. The method is worth naming.

Technique 3.3.5 (Algorithms as proofs: invariants and termination). To prove that something exists, it is often enough to give a procedure that builds it. Two things then have to be proved about the procedure.

Termination. Find a quantity that moves in one direction at every step and cannot do so forever, such as a nonnegative integer that decreases, or an integer that increases and is bounded above. (Such a quantity is sometimes called a *monovariant*.)

Correctness. Find an **invariant**: a statement about the current state that is true at the start and that each step keeps true. When the procedure halts, the invariant together with the reason for halting gives the property of the output. Formally this is induction on the number of steps.

The art is in choosing the invariant. The property we want at the end is rarely preserved step by step on its own, and usually has to be strengthened, by adding facts about the intermediate states, until it is. Reach for this technique whenever a result says that some reduction is always possible ("every matrix can be brought to the form ...", "the process ends with ..."). The canonical example is Gauss–Jordan elimination: termination is Proposition 3.3.6, and the invariant is Lemma 3.3.7. Later examples are Gram–Schmidt orthogonalization (Section 6.4) and the Euclidean algorithm for polynomials (Section 7.1); Exercise 3.2.54 is a small one.

Used in. Proposition 3.3.6, Lemma 3.3.7, Proposition 3.3.8, Theorem 3.3.10, Theorem 3.3.12, Exercise 3.3.53, Exercise 3.3.56, Exercise 3.6.11, Exercise 5.2.43, Algorithm 5.4.22, Theorem 6.4.15, Theorem 7.1.45, Proposition 8.3.25, Exercise 3.3.47, Exercise 5.4.48.

Proposition 3.3.6 (Termination). Let $A \in F^{m\times n}$. Gauss–Jordan elimination on A halts after completing at most m passes, and so performs at most $m(m+1)$ elementary row operations.

Proof. ROUTINE

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Find a quantity that changes in the same direction in every pass.

Hint 2. Track k, the number of pivoted rows. What does each pass do to it, and how large can it get?
Sketch. Each pass increases k by 1 in step 3, no step decreases it, and $k \leq m$, since once $k = m$ there is no unpivoted row and step 1 halts. A pass performs at most one swap, one scaling, and $m - 1$ replacements.

Every pass increases the number k of pivoted rows by 1, in step 3, and no step decreases k. A pass can only begin if some row is unpivoted, that is, if $k < m$; when $k = m$, step 1 finds no column and halts. So at most m passes are completed before the algorithm halts. Each pass performs at most one swap, at most one scaling, and at most $m - 1$ replacements (one for each row other than row k), so at most $m + 1$ row operations. In total there are at most $m(m + 1)$. □

Now for the shape of the output. In Example 3.3.3 the pivots formed a staircase descending to the right, each pivot was alone in its column, and the leftover row was zero. The first two facts are what each pass sets up for its own pivot; the real issue is that later passes must not undo the work of earlier ones.

Intuition (Choosing the invariant). The obvious invariant is that the current matrix is GA for a product G of elementary matrices, but it says nothing about the shape. For the shape we want three facts about the pivots found so far: each pivot column is a standard basis vector, the pivots descend to the right, and each pivot row vanishes to the left of its pivot. The first one maintains itself. A later pass swaps and scales unpivoted rows and adds multiples of an unpivoted row to other rows, and every unpivoted row has a 0 in column p_s, because that column is $\mathbf{e}_s$. The other two depend on where the next pivot is found. They are safe if the unpivoted rows vanish in every column up to the last pivot, so we add this to the invariant. It is exactly what choosing the *leftmost* column in step 1, followed by clearing, keeps true.

Lemma 3.3.7 (The loop invariant). Let $A \in F^{m \times n}$, and suppose Gauss–Jordan elimination on A has completed k passes, with pivot columns $p_1, \ldots, p_k$ and current matrix M. Put $p_0 = 0$. Then:

(a) $M = GA$ for some product G of elementary matrices;

(b) $p_1 < p_2 < \cdots < p_k$;

(c) for $1 \leq s \leq k$, column p_s of M is $\mathbf{e}_s$;

(d) for $1 \leq s \leq k$, row s of M has zeros in columns $1, \ldots, p_s - 1$;

(e) rows $k + 1, \ldots, m$ of M have zeros in columns $1, \ldots, p_k$.

Proof. STANDARD

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Proposition 3.2.7 (Row operations are left multiplications), induction on k
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Induct on k. In pass $k + 1$, let j be the column chosen by step 1. What does the choice of j say about rows $k + 1, \ldots, m$ in columns $1, \ldots, j - 1$?
Hint 2. Let $\mathbf{u}$ be row $k + 1$ after step 3. Every operation of the pass either moves or scales unpivoted rows, or adds a multiple of $\mathbf{u}$ to another row. Where is $\mathbf{u}$ zero?
Hint 3. For (b), use (e) for the old k: the unpivoted rows vanish in columns $1, \ldots, p_k$, so the scan cannot stop before column $p_k + 1$.
Sketch. By the choice of j, rows $k + 1, \ldots, m$ vanish in columns $1, \ldots, j - 1$, and so does the new pivot row $\mathbf{u}$; by (e), $j > p_k$. The pass only swaps and scales unpivoted rows and adds multiples of $\mathbf{u}$, so columns p_s with $s \leq k$ and the zeros in (d) are untouched, while step 4 turns column j into $\mathbf{e}_{k+1}$. Rows $k + 2, \ldots, m$ are old unpivoted rows plus multiples of $\mathbf{u}$, so they vanish in columns $1, \ldots, j$.

We induct on k. Before the first pass, $k = 0$ and $M = A = I_m A$, and (b)–(e) say nothing.

Suppose that (a)–(e) hold after k passes, with current matrix M, and that pass $k + 1$ takes place, with column j chosen in step 1. Let M' be the matrix after the pass. By the choice of j as the *leftmost* suitable column,

$$\text{rows } k + 1, \ldots, m \text{ of } M \text{ have zeros in columns } 1, \ldots, j - 1. \tag{3.3.1}$$

The pass does three things. step 2 may exchange two rows among rows $k+1, \dots, m$. step 3 scales row $k+1$; call the resulting row $\mathbf{u}$. step 4 adds multiples of $\mathbf{u}$ to other rows and leaves row $k+1$ equal to $\mathbf{u}$. By (3.3.1), $\mathbf{u}$ has zeros in columns $1, \dots, j-1$, and by step 3 it has entry 1 in column j.

(a) Each operation of the pass is left multiplication by an elementary matrix (Proposition 3.2.7). So $M' = EM = (EG)A$, where E is the product of the elementary matrices of the pass, and EG is again a product of elementary matrices.

(b) If $k \geq 1$, then by (e), rows $k+1, \dots, m$ of M vanish in columns $1, \dots, p_k$, so the column j found by the scan satisfies $j > p_k$. Hence $p_1 < \cdots < p_k < p_{k+1} = j$.

(c) Let $s \leq k$. Column p_s of M is $\mathbf{e}_s$, so its entries in rows $k+1, \dots, m$ are 0. Steps step 2 and step 3 only exchange or scale entries in these rows, and step 4 adds multiples of $\mathbf{u}$, which is a multiple of one of these rows and so also has entry 0 in column p_s. So column p_s of M' is still $\mathbf{e}_s$. For the new pivot: after step 3 the entry of row $k+1$ in column j is 1, and each operation of step 4 changes a single row $i \neq k+1$, replacing its entry c in column j by $c - c \cdot 1 = 0$. So column j of M' is $\mathbf{e}_{k+1}$.

(d) Let $s \leq k$. Row s is not touched by steps step 2 and step 3, and in step 4 it changes by a multiple of $\mathbf{u}$, which vanishes in columns $1, \dots, j-1$, a range containing $1, \dots, p_s - 1$. So row s of M' still vanishes in columns $1, \dots, p_s - 1$. Row $k+1$ of M' is $\mathbf{u}$, which vanishes in columns $1, \dots, j-1 = p_{k+1} - 1$.

(e) Let $i \geq k+2$. Row i of M' is one of the rows $k+1, \dots, m$ of M (possibly moved by the swap) plus a multiple of $\mathbf{u}$. By (3.3.1) both vanish in columns $1, \dots, j-1$, so row i of M' does too, and its entry in column j is 0 by (c). So rows $k+2, \dots, m$ of M' vanish in columns $1, \dots, p_{k+1}$.

This completes the induction. □

Proposition 3.3.8 (Properties of the output). Let R be the output of Gauss–Jordan elimination on $A \in F^{m \times n}$, with pivot columns $p_1 < \cdots < p_r$. Then $r \leq \min(m, n)$, and:

(a) rows $r+1, \dots, m$ of R are zero;

(b) for $1 \leq s \leq r$, the leading entry of row s is the pivot, a 1 in column p_s;

(c) for $1 \leq s \leq r$, column p_s of R is $\mathbf{e}_s$.

In particular, R satisfies the following conditions.

(R1) Every zero row lies below every nonzero row.

(R2) The leading entry of every nonzero row is 1.

(R3) The leading entry of each nonzero row lies strictly to the right of the leading entry of every row above it.

(R4) In a column that contains the leading entry of some row, every other entry is 0.

Proof. ROUTINE

Uses. Lemma 3.3.7 (The loop invariant), Proposition 3.3.6 (Termination)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Apply Lemma 3.3.7 after the last pass, and ask why the algorithm halted.
Hint 2. The scan found no column with a nonzero entry in rows $r+1, \dots, m$. What does that say about those rows?
Sketch. There are at most m passes and the p_s are distinct columns, so $r \leq \min(m, n)$. The algorithm halted because rows $r+1, \dots, m$ have no nonzero entry, which is (a). Parts (b) and (c) are (c) and (d) of Lemma 3.3.7. Then (R1)–(R4) follow, using $p_1 < \cdots < p_r$ for (R3).

The algorithm completes r passes and then halts. By Proposition 3.3.6, $r \leq m$, and by Lemma 3.3.7(b), $p_1 < \cdots < p_r$ are r distinct columns, so $r \leq n$.

(a) The algorithm halted because step 1 found no column with a nonzero entry in one of the rows $r+1, \dots, m$. So these rows are zero.

(b), (c) By Lemma 3.3.7(d), row s of R vanishes in columns $1, \dots, p_s - 1$, and by Lemma 3.3.7(c), column p_s of R is $\mathbf{e}_s$, so the entry of row s in column p_s is 1.

By (a) and (b), the nonzero rows of R are exactly rows $1, \dots, r$, which gives (R1). Their leading entries are the pivots, which are 1 and lie in the columns $p_1 < \cdots < p_r$; this gives (R2) and (R3). The columns containing leading entries are $p_1, \dots, p_r$, so (R4) is (c). □

3.3.3 Reduced row echelon form

Definition 3.3.9 (Reduced row echelon form). Let $A \in F^{m\times n}$. The **reduced row echelon form** of A, written $\operatorname{rref}(A)$ or A_r, is the output of Gauss–Jordan elimination (Algorithm 3.3.1) on A. The **pivots** and **pivot columns** of A are those found by this run; by Proposition 3.3.8(b), the pivots are the leading entries of the nonzero rows of $\operatorname{rref}(A)$. A matrix R is **in reduced row echelon form** (or **in RREF**) if $\operatorname{rref}(R) = R$.

This is the lecture's definition: $\operatorname{rref}(A)$ is *whatever the algorithm outputs*. It has the advantage of being concrete, and the drawback that it is not yet clear which matrices are in RREF, or whether a different way of row reducing A could end somewhere else. We settle both questions in this section and the next. First, the bookkeeping of Section 3.2 turns the run into one matrix.

Theorem 3.3.10 ($GA = \operatorname{rref}(A)$). For every $A \in F^{m\times n}$ there is a matrix $G \in F^{m\times m}$ that is a product of elementary matrices, and hence invertible, with

$$GA = \operatorname{rref}(A).$$

Explicitly, $G = E_t \cdots E_2 E_1$, where $E_1, \dots, E_t$ are the elementary matrices of the row operations of the run, in the order performed.

Proof. ROUTINE

Uses. Lemma 3.3.7 (The loop invariant), Proposition 3.2.7 (Row operations are left multiplications), Corollary 3.2.11 (Products of elementary matrices)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Which part of Lemma 3.3.7 is about GA?
Hint 2. Apply it after the last pass, and recall that products of elementary matrices are invertible.
Sketch. By Lemma 3.3.7(a) after the last pass, $\operatorname{rref}(A) = GA$ with $G = E_t \cdots E_1$ the product of the elementary matrices of the run (Proposition 3.2.7). Such a product is invertible by Corollary 3.2.11.

By Proposition 3.2.7, performing the operations of the run one after another gives $E_1 A$, then $E_2 E_1 A$, and so on; this is the content of Lemma 3.3.7(a). After the last operation the matrix is $\operatorname{rref}(A) = E_t \cdots E_1 A = GA$. By Corollary 3.2.11, G is invertible. □

In other words, running the algorithm on A is the same as multiplying A on the left by one well-chosen invertible matrix. Two consequences are immediate. First, $A \sim \operatorname{rref}(A)$ (Definition 3.2.21). Second, by Corollary 3.2.23, $\ker T_A = \ker T_{\operatorname{rref}(A)}$, and by Proposition 3.2.14, $A\mathbf{x} = \mathbf{y}$ if and only if $\operatorname{rref}(A)\mathbf{x} = G\mathbf{y}$. Section 3.4 reads the solutions off this second system.

The matrix G is found by performing the operations of the run on I_m. It is convenient to do this alongside the run itself, by adjoining I_m to A.

Proposition 3.3.11 (Carrying extra columns along). Let $A \in F^{m\times n}$ and $B \in F^{m\times p}$, and let G be as in Theorem 3.3.10.

(a) $G(A \mid B) = (\operatorname{rref}(A) \mid GB)$.

(b) Performing on the $m \times (n + p)$ matrix $(A \mid B)$ the run of Gauss–Jordan elimination with step 1 restricted to the first n columns ends at $(\operatorname{rref}(A) \mid GB)$.

In particular, the restricted run on $(A \mid I_m)$ ends at $(\operatorname{rref}(A) \mid G)$.

Proof. ROUTINE

Uses. Definition 2.4.1 (Matrix product), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.2.7 (Row operations are left multiplications)
Hint 1. Compute $G(A \mid B)$ one column at a time.
Hint 2. For (b): the decisions in steps step 1 through step 4 depend only on entries in the scanned columns. How do row operations act on the individual columns of a matrix?
Sketch. Column l of GM is G times column l of M, which gives (a). A row operation changes each column by itself, so during the restricted run the first n columns evolve exactly as in the run on A; since every decision depends only on them, the restricted run performs the same operations $E_1, \dots, E_t$ and ends at $G(A \mid B)$.

(a) By Definition 2.4.1, column l of GM is G times column l of M, for any matrix M with m rows. The first n columns of $(A \mid B)$ are those of A and the last p are those of B, so the first n columns of $G(A \mid B)$ form $GA = \operatorname{rref}(A)$ and the last p form GB.

(b) For the same reason, if E is elementary, the first n columns of $E(A \mid B)$ form EA. So as long as the restricted run on $(A \mid B)$ and the run on A perform the same operations, the first n columns of the current matrix of the former are the current matrix of the latter. Every decision in steps step 1 through step 4 (which column, which rows, which scalars, whether to halt) depends only on the entries in the scanned columns. So, by induction on the number of operations, the two runs perform the same operations $E_1, \dots, E_t$, and the restricted run ends at $E_t \cdots E_1(A \mid B) = G(A \mid B)$, which is $(\operatorname{rref}(A) \mid GB)$ by (a). □

The row reducer does exactly this when it shows G: it carries I_m along. For Example 3.3.3,

$$(A \mid I_3) = \left(\begin{array}{cccc|ccc} 0 & 0 & 2 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cccc|ccc} 1 & 2 & 0 & -2 & -1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1/2 & 1 \end{array}\right),$$

so

$$G = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ -1 & -1/2 & 1 \end{pmatrix}, \quad GA = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \operatorname{rref}(A).$$

The last row of G is worth a look. Since the last row of GA is zero, -1 times row 1 of A, minus $1/2$ times row 2, plus row 3, is the zero row. So a zero row of $\operatorname{rref}(A)$ records a linear relation among the rows of A (see Exercise 3.3.57).

Which matrices are in RREF

We now characterize the matrices in RREF by their shape, without reference to the algorithm. One direction is Proposition 3.3.8. For the other, we run the algorithm on a matrix of the right shape and watch it do nothing.

Theorem 3.3.12 (Matrices in RREF). A matrix $R \in F^{m \times n}$ is in reduced row echelon form if and only if it satisfies conditions (R1)–(R4) of Proposition 3.3.8. In that case the pivots of R are its leading entries.

Proof. STANDARD

Uses. Proposition 3.3.8 (Properties of the output), Algorithm 3.3.1 (Gauss–Jordan elimination), Definition 3.3.9 (Reduced row echelon form), induction on the number of passes
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. One direction is Proposition 3.3.8, applied to $R = \operatorname{rref}(R)$.
Hint 2. For the other, let rows $1, \dots, r$ be the nonzero rows of R, with leading entries in columns $c_1 < \cdots < c_r$. Run the algorithm on R and show by induction that pass k changes nothing.

Hint 3. In pass k, which is the leftmost column with a nonzero entry in rows $k, \dots, m$? What is the entry in row k there, and what does (R4) say about the rest of that column?
Sketch. If $\operatorname{rref}(R) = R$, then R is an output, so (R1)–(R4) hold by Proposition 3.3.8. Conversely, if (R1)–(R4) hold, then in pass k the rows $k, \dots, m$ vanish to the left of c_k, the leading entry of row k, which is 1. So the scan picks c_k, there is no swap and no scaling, and by (R4) nothing needs clearing. After r passes only zero rows are unpivoted, and the algorithm halts with output R.

Suppose R is in RREF. Then $R = \operatorname{rref}(R)$ is an output of the algorithm, so it satisfies (R1)–(R4) by Proposition 3.3.8, and by Proposition 3.3.8(b) its pivots are its leading entries.

Conversely, suppose R satisfies (R1)–(R4). By (R1), its nonzero rows are rows $1, \dots, r$ for some r, and by (R3), their leading entries lie in columns $c_1 < c_2 < \cdots < c_r$. We show by induction on $k = 0, 1, \dots, r$ that after k passes of Gauss–Jordan elimination on R, the current matrix is still R and the pivots are the leading entries of rows $1, \dots, k$. For $k = 0$ there is nothing to show.

Suppose this holds for some $k < r$, and consider pass $k + 1$. Rows $k + 2, \dots, r$ have their leading entries in columns $c_{k+2}, \dots, c_r$, all to the right of c_{k+1}, and rows $r + 1, \dots, m$ are zero. Together with the leading entry of row $k + 1$, this shows that rows $k + 1, \dots, m$ vanish in columns $1, \dots, c_{k+1} - 1$, and that row $k + 1$ has a nonzero entry in column c_{k+1}. So step 1 picks $j = c_{k+1}$, and step 2 does nothing. By (R2) the entry in position $(k + 1, c_{k+1})$ is already 1, so step 3 changes nothing and declares this leading entry a pivot. By (R4) all other entries of column c_{k+1} are 0, so step 4 changes nothing. This completes the induction.

After r passes the unpivoted rows are rows $r + 1, \dots, m$, which are zero by (R1), so the algorithm halts and outputs R. Hence $\operatorname{rref}(R) = R$, and the pivots are the leading entries. □

Since $\operatorname{rref}(A)$ satisfies (R1)–(R4) by Proposition 3.3.8, Theorem 3.3.12 gives $\operatorname{rref}(\operatorname{rref}(A)) = \operatorname{rref}(A)$: running the algorithm a second time changes nothing. The shape described by (R1)–(R4) is a staircase. For instance, every 4×6 matrix in RREF with pivot columns $1, 3, 6$ looks like

$$\begin{pmatrix} 1 & * & 0 & * & * & 0 \\ 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where each $*$ is an arbitrary scalar. The pivot columns are $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, everything below the staircase is 0, and the non-pivot columns are free above it.

Example 3.3.13 (Matrices in RREF). The identity I_n and the zero matrix $0 \in F^{m \times n}$ are in RREF. (For I_n the leading entries are the diagonal entries; the zero matrix has no nonzero rows, so (R2)–(R4) say nothing.) So are

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

In the second matrix, the entry 5 sits in a non-pivot column, where (R4) imposes nothing.

Non-example 3.3.14 (Four ways to fail). Over $\mathbb{Q}$, each matrix below violates exactly one condition:

$$\text{(R1)}: \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{(R2)}: \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{(R3)}: \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{(R4)}: \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

In the first, a zero row lies above a nonzero row. In the second, a leading entry is 2. In the third, the leading entry of row 2 lies to the left of that of row 1. In the fourth, the leading entry of row 2 has a nonzero entry above it. The last three have reduced row echelon form I_2, and the first has $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Exercise 3.3.15. ROUTINE

(a) Which of the following matrices are in RREF? For each one that is not, name a condition among (R1)–(R4) that fails.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) Find all 2×2 matrices over a field F that are in RREF. How many are there over $\mathbb{F}_p$?

Solution on page 1225

3.3.4 Uniqueness

Two people row reducing the same matrix by hand will rarely perform the same operations. One may bring up a different row in step 2, or clear everything below the pivots first and the entries above them afterwards. If both end with a matrix satisfying (R1)–(R4), do they end with the same matrix? Equivalently, can one row equivalence class contain two different matrices in RREF? This matters beyond hand computation. Properties we read off $\operatorname{rref}(A)$, such as the number of pivots (the rank, Section 3.5) or which columns are pivot columns, should belong to A, and we will constantly want to replace A by a row equivalent matrix GA without changing them.

The answer is that the reduced row echelon form is unique. We first record how to read the columns of a matrix in RREF.

Lemma 3.3.16 (Reading the columns of a matrix in RREF). Let $R \in F^{m \times n}$ be in RREF, with columns $\mathbf{r}_1, \ldots, \mathbf{r}_n$, entries r_{ij}, and pivot columns $p_1 < \cdots < p_r$. Put $p_t = n + 1$ for $t > r$.

(a) For $0 \le s \le m$, every column $\mathbf{r}_j$ with $j < p_{s+1}$ has zeros in rows $s+1, \ldots, m$. That is, $\mathbf{r}_j = r_{1j}\mathbf{e}_1 + \cdots + r_{sj}\mathbf{e}_s$.

(b) $\mathbf{r}_{p_s} = \mathbf{e}_s$ for $1 \le s \le r$.

(c) Column j is a pivot column if and only if $\mathbf{r}_j$ is not a linear combination of $\mathbf{r}_1, \ldots, \mathbf{r}_{j-1}$.

(d) If column j is not a pivot column, and $p_1, \ldots, p_t$ are the pivot columns to the left of it, then

$$\mathbf{r}_j = r_{1j}\mathbf{r}_{p_1} + r_{2j}\mathbf{r}_{p_2} + \cdots + r_{tj}\mathbf{r}_{p_t}.$$

Proof. STANDARD

Uses. Theorem 3.3.12 (Matrices in RREF), Proposition 3.3.8 (Properties of the output), Definition 1.3.4 (Linear combination)
Hint 1. For (a), take a row $i \ge s+1$. Where is its leading entry, if it has one?
Hint 2. For (c), a pivot column $\mathbf{r}_{p_s} = \mathbf{e}_s$ has a 1 in row s. What is the entry in row s of every earlier column, and of every combination of them?
Hint 3. For (d), apply (a) with $s = t$ and substitute $\mathbf{e}_s = \mathbf{r}_{p_s}$.
Sketch. (a) A row $i > s$ is zero or has its leading entry in column $p_i \ge p_{s+1} > j$. (b) is (R2) with (R4). For (c), columns before p_s vanish in row s by (a), and so do their combinations, while $\mathbf{r}_{p_s} = \mathbf{e}_s$ does not. For a non-pivot column j with t pivot columns before it, $j < p_{t+1}$, so (a) and (b) give (d), which exhibits $\mathbf{r}_j$ as a combination of earlier columns.

By Theorem 3.3.12, R satisfies (R1)–(R4) and its pivots are its leading entries: the leading entry of row s is a 1 in column p_s for $s \le r$, and rows $r+1, \ldots, m$ are zero.

(a) Let $j < p_{s+1}$ and $i \ge s+1$. If $i \le r$, the leading entry of row i is in column $p_i \ge p_{s+1} > j$, so $r_{ij} = 0$. If $i > r$, row i is zero. So $r_{ij} = 0$ for all $i \ge s+1$.

(b) Column p_s contains the leading entry of row s, which is 1 by (R2), and its other entries are 0 by (R4).

(c), (d) Suppose first that $j = p_s$ is a pivot column. By (a), applied with $s-1$, every column $\mathbf{r}_l$ with $l < j$ has entry 0 in row s, and therefore so does every linear combination of $\mathbf{r}_1, \dots, \mathbf{r}_{j-1}$. But $\mathbf{r}_j = \mathbf{e}_s$ has entry 1 in row s. So $\mathbf{r}_j$ is not such a combination.

Now suppose j is not a pivot column, and let $p_1, \dots, p_t$ be the pivot columns to its left ($t = 0$ if there are none). Then $p_{t+1} > j$, so by (a) and (b),

$$\mathbf{r}_j = r_{1j}\mathbf{e}_1 + \cdots + r_{tj}\mathbf{e}_t = r_{1j}\mathbf{r}_{p_1} + \cdots + r_{tj}\mathbf{r}_{p_t}.$$

This is (d), and since $p_1, \dots, p_t < j$, it shows that $\mathbf{r}_j$ is a combination of earlier columns. (For $t = 0$, it says $\mathbf{r}_j = \mathbf{0}$, the empty combination.) Together with the first case, this proves (c). □

So a matrix in RREF records two kinds of information: *which* columns are new, in the sense of not being combinations of earlier columns (the pivot columns, which are $\mathbf{e}_1, \mathbf{e}_2, \dots$ in order), and *how* each other column is built from the pivot columns before it (its entries).

> *Intuition (Two strategies for uniqueness).* We want to show: if R and $R' = GR$ are both in RREF and G is invertible, then $R = R'$. There are two natural ways to think about this.
>
> *Track G.* Left multiplication by G sends column p_s of R, which is $\mathbf{e}_s$, to column p_s of R'. If we knew that R' had its pivots in the same columns, that column would be $\mathbf{e}_s$ too, so $G\mathbf{e}_s = \mathbf{e}_s$: the first r columns of G are $\mathbf{e}_1, \dots, \mathbf{e}_r$. Then G fixes every vector in $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_r)$, which contains every column of R, so $GR = R$. So everything comes down to showing that the pivot columns agree, and it is natural to do this one pivot at a time, from left to right, using invertibility of G to rule out a pivot that appears earlier in one matrix than in the other. This is an induction on the pivots.
>
> *Forget G.* Row operations preserve the linear relations among the columns (Corollary 3.2.17). By Lemma 3.3.16, these relations determine which columns are pivot columns, and how the other columns are expressed through them. So two RREF matrices with the same relations among their columns are equal.
>
> The first approach gives information about G that the second does not; the second is shorter and proves a stronger statement (Corollary 3.3.20).

Theorem 3.3.17 (Uniqueness of the reduced row echelon form). Let $R, R' \in F^{m\times n}$ be in RREF, and suppose $R' = GR$ for some invertible $G \in F^{m\times m}$. Then $R' = R$.

Proof (pivot by pivot). HARD

> *Uses.* Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.3.8 (Matrix–vector product), Definition 2.5.12 (Invertible), induction on s
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* Let $\mathbf{g}_1, \dots, \mathbf{g}_m$ be the columns of G. Column j of R' is $G\mathbf{r}_j = r_{1j}\mathbf{g}_1 + \cdots + r_{mj}\mathbf{g}_m$. Why are the $\mathbf{g}_i$ linearly independent?
> *Hint 2.* Let p_s and q_s be the pivot columns of R and R' (with value $n+1$ when there are fewer than s pivots). Prove by induction on s that $p_s = q_s$, and that $\mathbf{g}_s = \mathbf{e}_s$ when $p_s \le n$.
> *Hint 3.* If $p_s < q_s$, compute column p_s of R' in two ways: as $G\mathbf{e}_s = \mathbf{g}_s$, and with Lemma 3.3.16(a) for R'. If $q_s < p_s$, compute column q_s of R' in two ways and look at its sth entry.
> *Sketch.* The columns $\mathbf{g}_i$ of G are independent since $G\mathbf{c} = \mathbf{0}$ forces $\mathbf{c} = \mathbf{0}$. By induction on s: if $p_s < q_s$, then $\mathbf{g}_s = G\mathbf{e}_s$ is column p_s of R', which lies in the span of $\mathbf{e}_1, \dots, \mathbf{e}_{s-1}$, that is, of $\mathbf{g}_1, \dots, \mathbf{g}_{s-1}$: a contradiction. If $q_s < p_s$, column q_s of R lies in $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_{s-1})$, and G fixes that span, so column q_s of R' has sth entry 0, not $\mathbf{e}_s$. So $p_s = q_s$, and then $\mathbf{g}_s = G\mathbf{e}_s = \mathbf{e}_s$. Finally every column of R lies in $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_r)$, which G fixes.

Write $\mathbf{r}_j$, $\mathbf{r}'_j$, and $\mathbf{g}_i$ for the columns of R, R', and G, and $p_1 < \cdots < p_r$ and $q_1 < \cdots < q_{r'}$ for the pivot columns of R and R'. Put $p_s = n+1$ for $s > r$ and $q_s = n+1$ for $s > r'$. We use two facts.

- By Definition 2.3.8, column j of R' is

$$\mathbf{r}'_j = G\mathbf{r}_j = r_{1j}\mathbf{g}_1 + \cdots + r_{mj}\mathbf{g}_m.$$

In particular $G\mathbf{e}_i = \mathbf{g}_i$.

- The columns of G are linearly independent: if $c_1\mathbf{g}_1 + \cdots + c_m\mathbf{g}_m = \mathbf{0}$, then $G\mathbf{c} = \mathbf{0}$ for $\mathbf{c} = (c_1, \ldots, c_m)$, so $\mathbf{c} = G^{-1}G\mathbf{c} = \mathbf{0}$.

We prove by induction on $s = 1, 2, \ldots, m$ that

$$p_s = q_s, \quad \text{and} \quad \mathbf{g}_s = \mathbf{e}_s \text{ if } p_s \le n.$$

Fix s, and assume this holds for every $t < s$.

Suppose $p_s < q_s$. Then $p_s \le n$, so $s \le r$, and for $t < s$ we have $p_t < p_s \le n$, hence $q_t = p_t \le n$ and $\mathbf{g}_t = \mathbf{e}_t$ by the induction hypothesis. By Lemma 3.3.16(b), $\mathbf{r}'_{p_s} = G\mathbf{r}_{p_s} = G\mathbf{e}_s = \mathbf{g}_s$. On the other hand $p_s < q_s$, so by Lemma 3.3.16(a) for R', the column $\mathbf{r}'_{p_s}$ lies in $\operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{s-1})$, which is $\operatorname{span}(\mathbf{g}_1, \ldots, \mathbf{g}_{s-1})$. So $\mathbf{g}_s$ is a linear combination of $\mathbf{g}_1, \ldots, \mathbf{g}_{s-1}$, contradicting the independence of the columns of G.

Suppose $q_s < p_s$. Then $q_s \le n$, so $s \le r'$, and for $t < s$, $q_t < q_s \le n$; by the induction hypothesis $p_t = q_t \le n$ and $\mathbf{g}_t = \mathbf{e}_t$. By Lemma 3.3.16(a) for R, $\mathbf{r}_{q_s} = \sum_{t<s} r_{tq_s}\mathbf{e}_t$, so

$$\mathbf{r}'_{q_s} = G\mathbf{r}_{q_s} = \sum_{t<s} r_{tq_s}\mathbf{g}_t = \sum_{t<s} r_{tq_s}\mathbf{e}_t,$$

whose sth entry is 0. But $\mathbf{r}'_{q_s} = \mathbf{e}_s$ by Lemma 3.3.16(b) for R', a contradiction.

Hence $p_s = q_s$. If this number is at most n, then by Lemma 3.3.16(b) for both matrices, $\mathbf{e}_s = \mathbf{r}'_{p_s} = G\mathbf{r}_{p_s} = G\mathbf{e}_s = \mathbf{g}_s$. This completes the induction.

In particular $r = r'$, since r is the number of s with $p_s \le n$, and $\mathbf{g}_s = \mathbf{e}_s$ for $s \le r$. Finally, let j be any column. By Lemma 3.3.16(a) with $s = r$, $\mathbf{r}_j = \sum_{t\le r} r_{tj}\mathbf{e}_t$, so

$$\mathbf{r}'_j = G\mathbf{r}_j = \sum_{t\le r} r_{tj}\mathbf{g}_t = \sum_{t\le r} r_{tj}\mathbf{e}_t = \mathbf{r}_j.$$

So $R' = R$, as desired. □

Proof (via relations among columns). HARD

> *Uses.* Corollary 3.2.17 (Row operations preserve relations among columns), Lemma 3.3.16 (Reading the columns of a matrix in RREF)
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* By Corollary 3.2.17, R and $R' = GR$ satisfy the same linear relations among their columns. Use only this.
> *Hint 2.* By Lemma 3.3.16(c), being a pivot column is a statement about relations among the columns. Conclude that R and R' have the same pivot columns.
> *Hint 3.* A non-pivot column of R satisfies the relation of Lemma 3.3.16(d). Transfer that relation to R'.
> *Sketch.* The columns satisfy $\sum_l x_l\mathbf{r}_l = \mathbf{0} \Leftrightarrow \sum_l x_l\mathbf{r}'_l = \mathbf{0}$. "Column j is a combination of earlier columns" is such a relation, so by Lemma 3.3.16(c) the pivot columns agree, and pivot columns are $\mathbf{e}_s$ in both. A non-pivot column satisfies $\mathbf{r}_j = \sum_s r_{sj}\mathbf{r}_{p_s}$; the same relation in R' gives $\mathbf{r}'_j = \sum_s r_{sj}\mathbf{e}_s = \mathbf{r}_j$.

Let $\mathbf{r}_j$ and $\mathbf{r}'_j = G\mathbf{r}_j$ be the columns of R and R'. By Corollary 3.2.17, for all $x_1, \ldots, x_n \in F$,

$$x_1\mathbf{r}_1 + \cdots + x_n\mathbf{r}_n = \mathbf{0} \quad \Leftrightarrow \quad x_1\mathbf{r}'_1 + \cdots + x_n\mathbf{r}'_n = \mathbf{0}. \tag{3.3.2}$$

We show that (3.3.2) alone forces $R = R'$.

The pivot columns agree. For scalars $c_1, \ldots, c_{j-1}$, the equation

$$\mathbf{r}_j = c_1\mathbf{r}_1 + \cdots + c_{j-1}\mathbf{r}_{j-1}$$

is the relation in (3.3.2) with $x_l = c_l$ for $l < j$, $x_j = -1$, and $x_l = 0$ for $l > j$. So it holds if and only if the same equation holds with $\mathbf{r}'$ in place of $\mathbf{r}$. Hence $\mathbf{r}_j$ is a linear combination of the earlier columns of R if and only if $\mathbf{r}'_j$ is a linear combination of the earlier columns of R'. By Lemma 3.3.16(c), applied to R and to R', column j is a pivot column of R if and only if it is a pivot column of R'. Let $p_1 < \cdots < p_r$ be the common pivot columns.

The columns agree. By Lemma 3.3.16(b), $\mathbf{r}_{p_s} = \mathbf{e}_s = \mathbf{r}'_{p_s}$ for each s. Let j be a non-pivot column, with pivot columns $p_1, \dots, p_t$ to its left. By Lemma 3.3.16(d) for R,

$$\mathbf{r}_j = r_{1j}\mathbf{r}_{p_1} + \cdots + r_{tj}\mathbf{r}_{p_t}.$$

This is a relation among the columns of R, so by (3.3.2) the same relation holds in R':

$$\mathbf{r}'_j = r_{1j}\mathbf{r}'_{p_1} + \cdots + r_{tj}\mathbf{r}'_{p_t} = r_{1j}\mathbf{e}_1 + \cdots + r_{tj}\mathbf{e}_t = \mathbf{r}_j,$$

where the last equality is Lemma 3.3.16(d) and (b) for R again. So every column of R' equals the corresponding column of R. □

The two proofs teach different things. The first shows how rigid G is: whenever GR is again in RREF, the first r columns of G are $\mathbf{e}_1, \dots, \mathbf{e}_r$, although its other columns can be anything that keeps G invertible (Exercise 3.3.46). The second never looks at G. It used only that R and R' satisfy the same linear relations among their columns, that is, that $\ker T_R = \ker T_{R'}$, and we will exploit this in Corollary 3.3.20.

Corollary 3.3.18 (RREF and row equivalence). Let $A, B \in F^{m\times n}$.

(a) $\operatorname{rref}(GA) = \operatorname{rref}(A)$ for every invertible $G \in F^{m\times m}$.

(b) $A \sim B$ if and only if $\operatorname{rref}(A) = \operatorname{rref}(B)$.

(c) If R is in RREF and $A \sim R$, then $R = \operatorname{rref}(A)$. In particular, any sequence of elementary row operations that turns A into a matrix satisfying (R1)–(R4), performed in any order and with any choices, ends at $\operatorname{rref}(A)$.

Proof. STANDARD

Uses. Theorem 3.3.17 (Uniqueness of the reduced row echelon form), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 3.3.12 (Matrices in RREF), Proposition 3.3.8 (Properties of the output), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.22 (Row equivalence is an equivalence relation), Definition 3.2.21 (Row equivalence)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), write $\operatorname{rref}(A) = NA$ and $\operatorname{rref}(GA) = MGA$ with M, N invertible, and express $\operatorname{rref}(GA)$ as an invertible matrix times $\operatorname{rref}(A)$.
Hint 2. For (b), use (a) in one direction, and $A \sim \operatorname{rref}(A)$ with symmetry and transitivity of $\sim$ in the other.
Sketch. (a) $\operatorname{rref}(GA) = \left(MGN^{-1}\right)\operatorname{rref}(A)$, both sides are in RREF, and MGN^{-1} is invertible, so Theorem 3.3.17 applies. (b) If $B = E_k \cdots E_1 A$, apply (a). If $\operatorname{rref}(A) = \operatorname{rref}(B)$, then $A \sim \operatorname{rref}(A) = \operatorname{rref}(B) \sim B$. (c) is (b) with $B = R$, since $\operatorname{rref}(R) = R$; matrices satisfying (R1)–(R4) are in RREF by Theorem 3.3.12.

(a) By Theorem 3.3.10 there are invertible $N, M \in F^{m\times m}$ with $\operatorname{rref}(A) = NA$ and $\operatorname{rref}(GA) = M(GA)$. Then

$$\operatorname{rref}(GA) = MGA = \left(MGN^{-1}\right)(NA) = \left(MGN^{-1}\right)\operatorname{rref}(A),$$

and MGN^{-1} is invertible, with inverse $NG^{-1}M^{-1}$. Both $\operatorname{rref}(GA)$ and $\operatorname{rref}(A)$ are in RREF (by Proposition 3.3.8 and Theorem 3.3.12), so Theorem 3.3.17 gives $\operatorname{rref}(GA) = \operatorname{rref}(A)$.

(b) If $A \sim B$, then $B = GA$ with $G = E_k \cdots E_1$ a product of elementary matrices, which is invertible, so $\operatorname{rref}(B) = \operatorname{rref}(A)$ by (a). Conversely, suppose $\operatorname{rref}(A) = \operatorname{rref}(B)$. By Theorem 3.3.10, $A \sim \operatorname{rref}(A)$ and $B \sim \operatorname{rref}(B)$, so by symmetry and transitivity (Proposition 3.2.22), $A \sim \operatorname{rref}(A) = \operatorname{rref}(B) \sim B$.

(c) If R is in RREF, then $\operatorname{rref}(R) = R$, and $A \sim R$ gives $\operatorname{rref}(A) = \operatorname{rref}(R) = R$ by (b). A sequence of row operations that turns A into a matrix R satisfying (R1)–(R4) shows $A \sim R$, and R is in RREF by Theorem 3.3.12. □

Part (c) settles the question from Remark 3.3.2. A run of the algorithm that makes other choices in step 2 still ends at a matrix satisfying (R1)–(R4), since the proofs of Lemma 3.3.7 and Proposition 3.3.8 never used which row was brought up; so it ends at $\operatorname{rref}(A)$. The same goes for Gaussian elimination followed by clearing upward (Section 3.3.6), and for any improvised hand computation. It also shows that the pivot columns, and their number, are determined by A. Here is a description of them that does not mention the algorithm at all.

Corollary 3.3.19 (Pivot columns, intrinsically). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and let r_{ij} be the entries of $\operatorname{rref}(A)$.

(a) Column j is a pivot column of A if and only if $\mathbf{a}_j$ is not a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}$.

(b) If column j is not a pivot column and $p_1, \dots, p_t$ are the pivot columns to its left, then

$$\mathbf{a}_j = r_{1j}\mathbf{a}_{p_1} + \cdots + r_{tj}\mathbf{a}_{p_t}.$$

Proof. ROUTINE

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Corollary 3.2.17 (Row operations preserve relations among columns), Lemma 3.3.16 (Reading the columns of a matrix in RREF)
Hint 1. Write $\operatorname{rref}(A) = GA$. How are the columns of $\operatorname{rref}(A)$ related to those of A, and which relations do they share?
Hint 2. Combine Corollary 3.2.17(b) with Lemma 3.3.16(c) and (d).
Sketch. With $\operatorname{rref}(A) = GA$ and G invertible, the columns $\mathbf{r}_l = G\mathbf{a}_l$ satisfy the same relations as the $\mathbf{a}_l$. So $\mathbf{a}_j$ is a combination of earlier columns, with given coefficients, exactly when $\mathbf{r}_j$ is; now apply Lemma 3.3.16(c) and (d).

By Theorem 3.3.10, $\operatorname{rref}(A) = GA$ with G invertible, so the columns of $\operatorname{rref}(A)$ are $\mathbf{r}_l = G\mathbf{a}_l$. By Corollary 3.2.17(b), applied with $c_l = 0$ for $l > j$, we have $\mathbf{a}_j = \sum_{l<j} c_l\mathbf{a}_l$ if and only if $\mathbf{r}_j = \sum_{l<j} c_l\mathbf{r}_l$. So $\mathbf{a}_j$ is a combination of $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}$ if and only if $\mathbf{r}_j$ is a combination of $\mathbf{r}_1, \dots, \mathbf{r}_{j-1}$, and (a) follows from Lemma 3.3.16(c). Similarly, (b) follows from the relation of Lemma 3.3.16(d). □

For the matrix of Example 3.3.3, the pivot columns are 1 and 3, and Corollary 3.3.19(b) reads off $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_4 = -2\mathbf{a}_1 + 2\mathbf{a}_3$ from the entries of $\operatorname{rref}(A)$. Indeed $-2(0,2,1) + 2(2,2,3) = (4,0,4)$. Chapter 4 uses this description to extract bases from spanning lists.

Finally, we can prove the converse promised in Remark 3.2.25.

Corollary 3.3.20 (The kernel is a complete invariant). For $A, B \in F^{m\times n}$, the following are equivalent:

(i) $A \sim B$;
(ii) $\operatorname{rref}(A) = \operatorname{rref}(B)$;
(iii) $\ker T_A = \ker T_B$.

Proof. STANDARD

Uses. Corollary 3.3.18 (RREF and row equivalence), Corollary 3.2.23 (Row equivalent matrices have the same kernel), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Definition 2.3.8 (Matrix–vector product), Theorem 3.3.17 (Uniqueness of the reduced row echelon form)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. (i) ⇔ (ii) is Corollary 3.3.18(b), and (i) ⇒ (iii) is Corollary 3.2.23. Only (iii) ⇒ (ii) is new.
Hint 2. Reduce to RREF: $\ker T_{\operatorname{rref}(A)} = \ker T_A$. What does $\mathbf{x} \in \ker T_R$ say about the columns of R?
Hint 3. Reread the second proof of Theorem 3.3.17. Which hypothesis did it really use?
Sketch. If $\ker T_A = \ker T_B$, then $R = \operatorname{rref}(A)$ and $R' = \operatorname{rref}(B)$ have the same kernel by Corollary 3.2.23. By Definition 2.3.8, $\mathbf{x} \in \ker T_R$ says $x_1\mathbf{r}_1 + \cdots + x_n\mathbf{r}_n = \mathbf{0}$, so R and R' satisfy the same linear relations among their columns. The second proof of Theorem 3.3.17 used nothing else, so $R = R'$.

(i) ⇔ (ii) is Corollary 3.3.18(b), and (i) ⇒ (iii) is Corollary 3.2.23. We prove (iii) ⇒ (ii).

Suppose $\ker T_A = \ker T_B$, and let $R = \operatorname{rref}(A)$ and $R' = \operatorname{rref}(B)$. By Theorem 3.3.10 and Corollary 3.2.23, $\ker T_R = \ker T_A = \ker T_B = \ker T_{R'}$. By Definition 2.3.8, $R\mathbf{x} = x_1\mathbf{r}_1 + \cdots + x_n\mathbf{r}_n$, so $\ker T_R = \ker T_{R'}$ says exactly that for all $x_1, \dots, x_n \in F$,

$$x_1\mathbf{r}_1 + \cdots + x_n\mathbf{r}_n = \mathbf{0} \quad \Leftrightarrow \quad x_1\mathbf{r}'_1 + \cdots + x_n\mathbf{r}'_n = \mathbf{0},$$

where $\mathbf{r}_l$ and $\mathbf{r}'_l$ are the columns of R and R'. This is (3.3.2). The second proof of Theorem 3.3.17 deduced $R = R'$ from (3.3.2) and Lemma 3.3.16 alone, so $R = R'$. □

So the kernel of T_A, a subspace of F^n, determines the row equivalence class of A among $m \times n$ matrices, and $\operatorname{rref}(A)$ is a canonical representative of that class. For example, a 2×3 matrix A has $\ker T_A = \operatorname{span}((1,1,1))$ if and only if $\operatorname{rref}(A) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ (Exercise 3.3.49).

Exercise 3.3.21. ROUTINE Decide which two of the following real matrices are row equivalent:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \end{pmatrix}.$$

Solution on page 1225

3.3.5 Inverses by row reduction

For a square matrix A, row reduction decides whether A is invertible and, if it is, computes A^{-1}. The idea is one we will use again and again: invertibility does not change when A is multiplied on the left by an invertible matrix, and for a matrix in RREF it is visible at a glance.

Technique 3.3.22 (Reduce to RREF). Many properties of a matrix A do not change when A is replaced by GA with G invertible: for example, the kernel $\ker T_A$ (Corollary 3.2.23), the linear relations among the columns (Corollary 3.2.17), and, for square A, invertibility. To decide such a property, replace A by $\operatorname{rref}(A) = GA$ (Theorem 3.3.10), where the answer can be read off the pivots and the zero rows. Two things have to be checked: that the property really is unchanged by left multiplication by invertible matrices, and what the property says about a matrix in RREF.

Reach for it whenever a question about A is really about $\ker T_A$, about relations among the columns of A, or about solutions of $A\mathbf{x} = \mathbf{y}$ (where $\mathbf{y}$ must be replaced by $G\mathbf{y}$, Proposition 3.2.14). The canonical example is Theorem 3.3.23. It is the method of Section 3.4 and Section 3.5, and later of rank–nullity (Section 4.2) and of the equality of row rank and column rank (Section 4.3).

Used in. Corollary 3.3.20, Theorem 3.3.23, Proposition 3.4.1, Corollary 3.4.4, Theorem 3.4.9, Proposition 3.4.15, Proposition 3.4.16, Corollary 3.4.20, Exercise 3.4.45, Theorem 3.5.19, Theorem 3.5.27, Exercise 3.6.6, Exercise 3.6.7, Exercise 3.6.9, Exercise 3.6.10, Problem 3.3, Problem 3.4, Problem 3.5, Problem 3.6, Problem 3.8, Problem 3.10, Problem 3.11, Problem 3.12, Problem 3.14, Problem 3.17, Theorem 4.2.7, Theorem 4.2.11, Theorem 4.3.8, Problem 4.6, Exercise 3.3.49, Exercise 3.3.50, Exercise 3.5.42, Exercise 3.5.45.

Theorem 3.3.23 (Invertibility and RREF). Let $A \in F^{n \times n}$, and let G be a product of elementary matrices with $GA = \operatorname{rref}(A)$ (Theorem 3.3.10). Then A is invertible if and only if $\operatorname{rref}(A) = I_n$, and in that case $A^{-1} = G$.

Proof. STANDARD

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.3.8 (Properties of the output), Definition 2.5.12 (Invertible)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. If $\operatorname{rref}(A) = I_n$, then $GA = I_n$ with G invertible. Why does this make G a two-sided inverse of A?
Hint 2. Conversely, a square matrix in RREF with n pivots is I_n. If there are fewer pivots, what does the last row look like?
Hint 3. A matrix whose last row is zero has no right inverse. If A were invertible, would GA be?
Sketch. If $GA = I_n$, then $A = G^{-1}$, so $AG = I_n$ too and $A^{-1} = G$. If $R = \operatorname{rref}(A)$ has n pivots, its pivot columns are $1, \dots, n$ and $R = I_n$. Otherwise its last row is zero, so the last row of RX is zero for every X and R is not invertible; but if A were invertible, $R = GA$ would be.

Suppose $\operatorname{rref}(A) = I_n$, so $GA = I_n$. Since G is invertible (Corollary 3.2.11), multiplying on the left by G^{-1} gives $A = G^{-1}$. Hence $AG = G^{-1}G = I_n$ as well, so A is invertible with $A^{-1} = G$.

Conversely, suppose $R = \operatorname{rref}(A) \neq I_n$, and let $p_1 < \cdots < p_r$ be its pivot columns. If $r = n$, then $p_1, \dots, p_n$ are n increasing numbers among $1, \dots, n$, so $p_s = s$ for every s, and by Proposition 3.3.8(c) column s of R is $\mathbf{e}_s$; that is, $R = I_n$, which we have excluded. So $r < n$, and row n of R is zero by Proposition 3.3.8(a). Then for every

$X \in F^{n\times n}$, the (n,l) entry of RX is $\sum_k r_{nk}x_{kl} = 0$, so $RX \neq I_n$; thus R is not invertible. If A were invertible, then $R = GA$ would be invertible, with inverse $A^{-1}G^{-1}$. So A is not invertible. □

This gives an algorithm. To find the operations of the run and their product G at the same time, we carry I_n along as in Proposition 3.3.11.

Algorithm 3.3.24 (Inverting a matrix by row reduction)

Input A square matrix $A \in F^{n\times n}$.
Output A^{-1}, or the conclusion that A is not invertible.

1. Form the $n \times 2n$ matrix $(A \mid I_n)$.
2. Row reduce it with Gauss–Jordan elimination, scanning only the first n columns in step 1 of Algorithm 3.3.1. The result is $(\mathrm{rref}(A) \mid G)$.
3. If the left block is I_n, output $A^{-1} = G$. Otherwise A is not invertible.

The algorithm is correct by Proposition 3.3.11 and Theorem 3.3.23. Two practical remarks. If A is invertible, all n pivots lie in the left block, so restricting the scan makes no difference; it only matters when A is not invertible. And one may stop as soon as the left block acquires a zero row: at that moment the left block is $G'A$ with G' invertible, and a matrix with a zero row is not invertible (as in the proof above), so neither is A.

Intuition (Why carrying I_n along works). Column j of A^{-1} is the solution of $A\mathbf{x} = \mathbf{e}_j$, since $A\big(A^{-1}\mathbf{e}_j\big) = \mathbf{e}_j$. So inverting A means solving the n systems $A\mathbf{x} = \mathbf{e}_1, \ldots, A\mathbf{x} = \mathbf{e}_n$, which share the coefficient matrix A. Row reducing $(A \mid I_n)$ solves all of them at once: when the left block becomes I_n, the system $A\mathbf{x} = \mathbf{e}_j$ has become $I_n\mathbf{x} = G\mathbf{e}_j$ (Proposition 3.2.14), whose solution, column j of G, can be read off.

Example 3.3.25 (Inverting a 3×3 matrix). Let

$$A = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{pmatrix} \in \mathbb{Q}^{3\times 3}.$$

Swapping rows 1 and 2 of $(A \mid I_3)$ and then adding -2 times row 1 to row 3 finishes pass 1:

$$\left(\begin{array}{ccc|ccc} 0 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{E_{12}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{E_{1,-2,3}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -2 & 1 \end{array}\right).$$

Pass 2 scales row 2 by $1/2$ and clears column 2 with $E_{2,-1,1}$ and $E_{2,1,3}$; pass 3 scales row 3 by 2 and clears column 3 with $E_{3,1/2,1}$ and $E_{3,-1/2,2}$:

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -1/2 & -1/2 & 1 & 0 \\ 0 & 1 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & -2 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & -4 & 2 \end{array}\right).$$

The left block is I_3, so A is invertible and

$$A^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 2 & -1 \\ 1 & -4 & 2 \end{pmatrix}.$$

As a check, the first row of A times the columns of A^{-1} gives $(0+0+1, 0+4-4, 0-2+2) = (1,0,0)$, and similarly for the other rows. Fractions appeared along the way even though A and A^{-1} have integer entries; Exercise 3.3.61 explains when integer arithmetic suffices.

$$(A \mid I_3) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & -4 & 2 \end{array}\right)$$

Inverting the matrix of Example 3.3.25: $(A \mid I_3)$ is row reduced, with pivots sought only in the left block. Edit the left block to try other matrices; when it does not reduce to I_3, the summary says that A is not invertible. (Interactive in the web version.)

Example 3.3.26 (A matrix that is not invertible). Let $A \in \mathbb{Q}^{3\times 3}$ have rows $(1,2,3)$, $(4,5,6)$, $(7,8,9)$. The algorithm gives

$$(A \mid I_3) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -5/3 & 2/3 & 0 \\ 0 & 1 & 2 & 4/3 & -1/3 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right).$$

The left block is not I_3, so A is not invertible. The last row of G says more: since the last row of GA is zero, row 1 of A minus twice row 2 plus row 3 is zero, which is easy to confirm.

For square matrices, Theorem 3.3.23 says in particular that an invertible matrix is row equivalent to I_n. Section 3.5 turns this into the statement that every invertible matrix is a product of elementary matrices (Theorem 3.5.19), and fits it into the Fundamental Theorems. Also note that in the proof, $GA = I_n$ was enough to conclude $AG = I_n$, because G was known to be invertible. That a one-sided inverse of a square matrix is always two-sided is Theorem 3.5.23.

Exercise 3.3.27. ROUTINE Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Use Algorithm 3.3.24 to decide whether A is invertible over $\mathbb{Q}$, and over $\mathbb{F}_3$, and find A^{-1} when it exists.

Solution on page 1226

3.3.6 Row echelon form BEYOND 21-242

Echelon forms and triangular factorizations are standard in LADW and in numerical linear algebra, but not part of the course.

Clearing above the pivots is not needed to solve a system: once each pivot has zeros below it, the equations can be solved from the bottom up. This gives a cheaper algorithm and a weaker normal form.

Definition 3.3.28 (Row echelon form). A matrix $U \in F^{m\times n}$ is in **row echelon form** if it satisfies (R1) and (R3): every zero row lies below every nonzero row, and the leading entry of each nonzero row lies strictly to the right of the leading entry of every row above it. The leading entries of the nonzero rows are the **pivots** of U. A **row echelon form of** A is a matrix in row echelon form that is row equivalent to A.

Gaussian elimination is Algorithm 3.3.1 with two changes: step 3 of Algorithm 3.3.1 only declares the pivot, without scaling, and step 4 of Algorithm 3.3.1 clears only the rows *below* row k. If the pivot is a and row $i > k$ has entry c in the pivot column, the pass adds $-\ell_{ik}$ times row k to row i, where $\ell_{ik} = c/a$ is the **multiplier**. The proof of Lemma 3.3.7 goes through with (c) weakened to "column p_s of M has a nonzero entry in row s and zeros below it" (a pivot row is never changed after its own pass), and as in Proposition 3.3.8 the output is a row echelon form of A.

Example 3.3.29 (Row echelon forms are not unique). Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ over $\mathbb{Q}$. Gaussian elimination gives $\begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix}$. Swapping the rows first and then clearing gives $\begin{pmatrix} 3 & 4 \\ 0 & 2/3 \end{pmatrix}$, and $\operatorname{rref}(A) = I_2$ is also a row echelon form of A. All three have their pivots in columns 1 and 2.

The pivot positions, unlike the entries, are always the same.

Proposition 3.3.30 (Pivots of row echelon forms). If U is a row echelon form of A, then the columns containing the pivots of U are the pivot columns of A.

Proof. STANDARD

Uses. Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.12 (Matrices in RREF)
Hint 1. Turn U into a matrix in RREF by row operations that do not move any leading entry.
Hint 2. Scale each nonzero row by the inverse of its leading entry. Then clear above the pivots, one pivot at a time. Check that clearing column c_s never disturbs a leading entry or a column cleared earlier.
Sketch. Scaling makes the leading entries 1. Clearing above the pivots (say from the last one up) only adds multiples of a pivot row to rows above it, which changes those rows only to the right of their own leading entries, and never refills a column already cleared, since a pivot row is zero in the columns of later pivots once those are cleared. The result satisfies (R1)–(R4), has the same leading positions as U, and is row equivalent to A, so it is $\operatorname{rref}(A)$.

Let rows $1, \dots, r$ be the nonzero rows of U, with leading entries in columns $c_1 < \cdots < c_r$. First scale each row $s \le r$ by the inverse of its leading entry. Then, for $s = r, r-1, \dots, 1$ in this order, add suitable multiples of row s to rows $1, \dots, s-1$ so that column c_s has zeros above row s. Row s vanishes in the columns before c_s, so adding a multiple of it to a row $i < s$ changes row i only in columns $c_s, \dots, n$, all to the right of c_i. Moreover, when we process s, row s has already been cleared in the columns $c_{s+1}, \dots, c_r$ (it lies above those pivots), so the columns cleared earlier are not disturbed. The final matrix R therefore satisfies (R1)–(R4) with leading entries in columns $c_1, \dots, c_r$, and it is row equivalent to U, hence to A. By Corollary 3.3.18(c), $R = \operatorname{rref}(A)$, so $c_1, \dots, c_r$ are the pivot columns of A (Theorem 3.3.12). □

The second phase of this proof, scaling and clearing upward from the last pivot, is called **back substitution**. Gaussian elimination followed by back substitution is the usual way of computing $\operatorname{rref}(A)$ by hand, and by Corollary 3.3.18 it gives the same answer as Gauss–Jordan elimination.

3.3.7 LU and PLU factorizations BEYOND 21-242

Suppose we must solve $A\mathbf{x} = \mathbf{y}$ for one A and many right-hand sides $\mathbf{y}$, as happens constantly in applications. The work of elimination is almost all spent on A and does not depend on $\mathbf{y}$, so it is worth recording once and for all. Gaussian elimination without swaps only adds multiples of a row to rows *below* it, and the elementary matrices of such operations are lower triangular. This suggests factoring A into triangular pieces.

Definition 3.3.31 (Triangular matrices). A matrix $L \in F^{n \times n}$ is **lower triangular** if $l_{ij} = 0$ whenever $i < j$, and **unit lower triangular** if moreover $l_{ii} = 1$ for every i. **Upper triangular** and **unit upper triangular** matrices are defined in the same way with $i > j$. A **diagonal** matrix is one that is both.

Lemma 3.3.32 (Products and inverses of triangular matrices).

(a) If $L, L' \in F^{n\times n}$ are lower triangular, then so is LL', and $(LL')_{ii} = l_{ii}l'_{ii}$.

(b) A unit lower triangular matrix is invertible, and its inverse is unit lower triangular.

(c) The same statements hold for upper triangular matrices.

Proof. STANDARD

Uses. Definition 2.4.1 (Matrix product), Proposition 3.2.7 (Row operations are left multiplications), Corollary 3.2.11 (Products of elementary matrices), Proposition 2.4.23 (Transpose of a product)
Hint 1. For (a), in $(LL')_{ij} = \sum_k l_{ik}l'_{kj}$, which k can give a nonzero term?
Hint 2. For (b), reduce L to I_n by clearing below the diagonal one column at a time, from left to right, and check that every elementary matrix used is unit lower triangular.
Sketch. (a) A term $l_{ik}l'_{kj}$ can be nonzero only if $j \le k \le i$. (b) When column s is processed, row s is $\mathbf{e}_s^{\mathsf{T}}$, so adding multiples of it to the rows below only kills the entries below the diagonal in column s. This gives $GL = I_n$ with G a product of matrices E_{sci}, $s < i$, which are unit lower triangular; so $L^{-1} = G$ is unit lower triangular by (a). (c) Transpose.

(a) We have $(LL')_{ij} = \sum_k l_{ik}l'_{kj}$, and $l_{ik}l'_{kj}$ can be nonzero only if $k \le i$ and $k \ge j$. If $i < j$ there is no such k, so $(LL')_{ij} = 0$. If $i = j$, only $k = i$ contributes, giving $l_{ii}l'_{ii}$.

(b) Let L be unit lower triangular. For $s = 1, \dots, n-1$ in turn, and for each $i > s$, add $-c$ times row s to row i, where c is the current (i, s) entry. We claim that just before column s is processed, the current matrix is unit lower triangular with columns $1, \dots, s-1$ equal to $\mathbf{e}_1, \dots, \mathbf{e}_{s-1}$. This holds for $s = 1$. If it holds for s, then row s has zeros in columns $1, \dots, s-1$ (as those columns are $\mathbf{e}_1, \dots, \mathbf{e}_{s-1}$), a 1 in column s, and zeros after it; so row s is $\mathbf{e}_s^{\mathsf{T}}$. Adding a multiple of it to row $i > s$ changes only the (i, s) entry, which becomes 0. So afterwards column s is $\mathbf{e}_s$ and the claim holds for $s + 1$. At the end the matrix is I_n, so $GL = I_n$, where G is the product of the elementary matrices E_{sci} with $s < i$ that were used. Each of these is unit lower triangular (its only off-diagonal entry is in position (i, s)), so G is unit lower triangular by (a). Since G is invertible (Corollary 3.2.11), $L = G^{-1}$, and $L^{-1} = G$.

(c) A matrix is (unit) upper triangular if and only if its transpose is (unit) lower triangular. If U, U' are upper triangular, then $UU' = \left(U'^{\mathsf{T}}U^{\mathsf{T}}\right)^{\mathsf{T}}$ by Proposition 2.4.23, which is upper triangular by (a), with the same diagonal entries. If U is unit upper triangular, then U^{T} has a unit lower triangular inverse V by (b), and transposing $U^{\mathsf{T}}V = VU^{\mathsf{T}} = I_n$ shows that V^{T}, which is unit upper triangular, is the inverse of U. □

A **permutation matrix** is a product of swap matrices E_{ij} (equivalently, a matrix with exactly one 1 in each row and each column and zeros elsewhere; see Definition 5.2.29).

Theorem 3.3.33 (PLU and LU factorizations). Let $A \in F^{m\times n}$.

(a) There are a permutation matrix $P \in F^{m\times m}$, a unit lower triangular $L \in F^{m\times m}$, and a matrix $U \in F^{m\times n}$ in row echelon form such that $PA = LU$.

(b) If Gaussian elimination on A never needs a swap, then $A = LU$, where U is the output of Gaussian elimination and the entry of L in position (i, s), for $i > s$, is the multiplier ℓ_{is} of pass s (and 0 if there is no pass s).

Proof. HARD

Uses. Definition 2.4.33 (Block matrix), Definition 3.3.31 (Triangular matrices), Definition 3.3.28 (Row echelon form), Definition 5.2.29 (Permutation matrix), induction on m
Hint 1. Induct on the number of rows. Perform the first pass of Gaussian elimination and write the result in block form.
Hint 2. After moving a row with a nonzero entry a in the first nonzero column j to the top, write the first pass of Gaussian elimination as a factorization $P_1A = \begin{pmatrix} 1 & 0 \\ \ell & I \end{pmatrix}\begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & A' \end{pmatrix}$.
Hint 3. Apply the induction hypothesis $P'A' = L'U'$ and assemble P, L, U in block form. Where does P' have to act on ℓ?
Sketch. Bring a nonzero entry a of the first nonzero column j to the top with a swap P_1. With $\mathbf{b}$ the rest of column j and $\ell = a^{-1}\mathbf{b}$, $P_1A = \begin{pmatrix} 1 & 0 \\ \ell & I \end{pmatrix}\begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & A' \end{pmatrix}$, the first pass of elimination. If $P'A' = L'U'$, then $P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}P_1$, $L = \begin{pmatrix} 1 & 0 \\ P'\ell & L' \end{pmatrix}$,

$U = \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & U' \end{pmatrix}$ work, by block multiplication. Without swaps $P = I_m$, and the columns of L are the multipliers, pass by pass.

(a) We induct on m. If $m = 1$, take $P = L = (1)$ and $U = A$; a single row is in row echelon form. Let $m \geq 2$. If $A = 0$, take $P = L = I_m$ and $U = 0$. Otherwise, let j be the first nonzero column of A, choose a row i with $a_{ij} \neq 0$, and let $P_1 = E_{1i}$ (or $P_1 = I_m$ if $i = 1$). Then $B = P_1 A$ has first $j-1$ columns zero and a nonzero entry $a = b_{1j}$. Write

$$B = \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & \mathbf{b} & B' \end{pmatrix},$$

where the zero blocks have $j-1$ columns, $\mathbf{b} \in F^{m-1}$ is the rest of column j, $\mathbf{w} \in F^{1\times(n-j)}$, and $B' \in F^{(m-1)\times(n-j)}$. Let $\ell = a^{-1}\mathbf{b}$ and $A' = B' - \ell\mathbf{w}$. Multiplying blockwise (Definition 2.4.33),

$$B = \begin{pmatrix} 1 & 0 \\ \ell & I_{m-1} \end{pmatrix} \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & A' \end{pmatrix}, \tag{3.3.3}$$

since the lower blocks of the product are $\ell a = \mathbf{b}$ and $\ell\mathbf{w} + A' = B'$. (The right factor is the result of the first pass of Gaussian elimination on B.) If $j = n$, the blocks $\mathbf{w}$, B', A' are empty, and (3.3.3) is the required factorization with $P = P_1$. If $j < n$, the induction hypothesis gives a permutation matrix P', a unit lower triangular L', and a row echelon form U' with $P'A' = L'U'$. Put

$$P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} P_1, \quad L = \begin{pmatrix} 1 & 0 \\ P'\ell & L' \end{pmatrix}, \quad U = \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & U' \end{pmatrix}.$$

By (3.3.3) and blockwise multiplication,

$$PA = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} B = \begin{pmatrix} 1 & 0 \\ P'\ell & P' \end{pmatrix} \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & A' \end{pmatrix} = \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & P'\ell a & P'\ell\mathbf{w} + P'A' \end{pmatrix},$$

$$LU = \begin{pmatrix} 1 & 0 \\ P'\ell & L' \end{pmatrix} \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & 0 & U' \end{pmatrix} = \begin{pmatrix} 0 & a & \mathbf{w} \\ 0 & P'\ell a & P'\ell\mathbf{w} + L'U' \end{pmatrix},$$

and these agree because $P'A' = L'U'$. Here L is unit lower triangular because L' is. P is a permutation matrix, because if P' is a product of swaps E_{kl} of size $m-1$, then $\begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}$ is the product of the swaps $E_{(k+1)(l+1)}$ of size m. Finally, U is in row echelon form: its first row has leading entry a in column j, its other rows vanish in columns $1, \ldots, j$, and below the first row it has the shape of U', which is in row echelon form.

(b) If Gaussian elimination needs no swaps, then at every stage of the induction the $(1, j)$ entry is already nonzero and we may take $P_1 = I$; then $P = I_m$. In (3.3.3) the entries of $\ell = a^{-1}\mathbf{b}$ are the multipliers of the first pass, and A' is the lower right part of the matrix after the first pass, on which the later passes operate. So by induction column s of L holds the multipliers of pass s below the diagonal, and U is the output of Gaussian elimination. □

Example 3.3.34 (An LU factorization). Let

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 8 & 7 & 9 \end{pmatrix}.$$

Pass 1 uses the multipliers $\ell_{21} = 2$ and $\ell_{31} = 4$ and leaves rows $(0, 1, 1)$ and $(0, 3, 5)$; pass 2 uses $\ell_{32} = 3$ and leaves $(0, 0, 2)$. So

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

The multipliers drop into L with no further computation. By contrast, $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has no LU factorization: if $\left(\begin{smallmatrix} 1 & 0 \\ \ell & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} u & v \\ 0 & w \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, then $u = 0$ from the $(1,1)$ entry, and then the $(2,1)$ entry is $\ell u = 0 \neq 1$. A swap is unavoidable here, and $P = E_{12}$ gives $PA = I_2 = I_2 I_2$.

Once $PA = LU$ is known, $A\mathbf{x} = \mathbf{y}$ is equivalent to $L(U\mathbf{x}) = P\mathbf{y}$. One solves $L\mathbf{z} = P\mathbf{y}$ from the top down (z_1 first, then z_2, and so on: **forward substitution**) and then $U\mathbf{x} = \mathbf{z}$ from the bottom up (back substitution). For $A \in F^{n\times n}$ each of these costs about n^2 arithmetic operations, while the factorization itself costs a multiple of n^3:

Remark 3.3.35 (How much work elimination takes). Let $A \in F^{n\times n}$ be invertible, and suppose no swaps are needed. In pass k of Gauss–Jordan elimination the pivot is in position (k,k), and the pivot row vanishes in columns $1, \dots, k-1$ (Lemma 3.3.7(d)). So the scaling touches $n-k+1$ entries, and each of the $n-1$ replacements touches $n-k+1$ entries, with one multiplication and one addition each. The total number of multiplications is at most

$$\sum_{k=1}^{n} n(n-k+1) = n \cdot \frac{n(n+1)}{2} \approx \frac{n^3}{2},$$

and there are about as many additions: roughly n^3 arithmetic operations in all. Gaussian elimination only clears the $n-k$ rows below the pivot, which costs

$$\sum_{k=1}^{n} (n-k)(n-k+1) = \frac{n^3 - n}{3}$$

multiplications and divisions, and about as many additions: about $2n^3/3$ operations in all; back substitution adds only about n^2 per right-hand side. Either way, doubling the size of the matrix multiplies the work by about 8. A system with a few thousand unknowns is routine for a computer, whereas the formula for the solution that comes out of determinants in Section 5.5 involves sums with $n!$ terms.

Example 3.3.36 (A PLU factorization). For the matrix A of Example 3.3.25, the $(1,1)$ entry is 0, so the proof of Theorem 3.3.33 starts with $P_1 = E_{12}$. Then $\ell = (0, 2)$, and the remaining 2×2 block $A' = \left(\begin{smallmatrix} 2 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ needs no swap: its multiplier is $-1/2$, leaving $1/2$. So

$$E_{12}A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1/2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1/2 \end{pmatrix}.$$

3.3.8 Symmetric matrices: LDL^{T} BEYOND 21-242

A matrix A is **symmetric** if $A^{\mathsf{T}} = A$. Elimination does not see symmetry: the U of an LU factorization of a symmetric matrix is usually not symmetric. But the symmetry is hiding in the factorization, and it can be recovered.

Theorem 3.3.37 (LDL^{T} factorization). Let $A \in F^{n\times n}$ be symmetric, and suppose $A = LU$ with L unit lower triangular and U upper triangular with nonzero diagonal entries. Let D be the diagonal matrix with the same diagonal entries as U. Then $U = DL^{\mathsf{T}}$, so

$$A = LDL^{\mathsf{T}}.$$

Proof. HARD

Uses. Lemma 3.3.32 (Products and inverses of triangular matrices), Proposition 2.4.23 (Transpose of a product), Definition 3.3.31 (Triangular matrices)
Hint 1. Factor $U = DU_1$ with U_1 unit upper triangular, so $A = LDU_1$. Transpose this equation.
Hint 2. You get $LDU_1 = U_1^{\mathsf{T}} DL^{\mathsf{T}}$. Rearrange it so that one side is lower triangular and the other upper triangular.
Hint 3. A matrix that is both lower and upper triangular is diagonal. Compare diagonal entries.
Sketch. Write $U = DU_1$ with $U_1 = D^{-1}U$ unit upper triangular. Transposing $A = LDU_1$ gives $A = U_1^{\mathsf{T}} DL^{\mathsf{T}}$, and moving the outer factors across shows that $\left(U_1^{\mathsf{T}}\right)^{-1} LD$, which is lower triangular with diagonal D, equals $DL^{\mathsf{T}} U_1^{-1}$, which is upper triangular. So both equal D. Hence $\left(U_1^{\mathsf{T}}\right)^{-1} L = I_n$, so $U_1 = L^{\mathsf{T}}$ and $U = DL^{\mathsf{T}}$.

Since the diagonal entries $d_i = u_{ii}$ are nonzero, D is invertible, with D^{-1} diagonal with entries d_i^{-1}. Let $U_1 = D^{-1}U$; this scales row i of U by d_i^{-1}, so U_1 is unit upper triangular, and $U = DU_1$. Then $A = LDU_1$, and transposing (Proposition 2.4.23, with $D^{\mathsf{T}} = D$) gives

$$LDU_1 = A = A^{\mathsf{T}} = U_1^{\mathsf{T}} DL^{\mathsf{T}}.$$

Here U_1^{T} is unit lower triangular and L^{T} is unit upper triangular. Multiplying on the left by $\left(U_1^{\mathsf{T}}\right)^{-1}$ and on the right by U_1^{-1}, which exist by Lemma 3.3.32, gives

$$\left(U_1^{\mathsf{T}}\right)^{-1} LD = DL^{\mathsf{T}} U_1^{-1}.$$

By Lemma 3.3.32, $\left(U_1^{\mathsf{T}}\right)^{-1} L$ is unit lower triangular, so the left side is lower triangular with diagonal entries $d_1, \dots, d_n$; the right side is upper triangular. So both sides are diagonal, and comparing diagonals they equal D. Thus $\left(U_1^{\mathsf{T}}\right)^{-1} LD = D$, and multiplying on the right by D^{-1} gives $\left(U_1^{\mathsf{T}}\right)^{-1} L = I_n$, that is, $L = U_1^{\mathsf{T}}$. Hence $U_1 = L^{\mathsf{T}}$ and $U = DL^{\mathsf{T}}$. □

The hypothesis holds whenever A is invertible and Gaussian elimination on A needs no swaps. Then $A = LU$ by Theorem 3.3.33(b), and by Proposition 3.3.30 the pivots of U lie in the pivot columns of A, which are $1, \dots, n$ since $\operatorname{rref}(A) = I_n$ (Theorem 3.3.23); so the diagonal entries of U are its pivots, which are nonzero.

Example 3.3.38 (An LDL^{T} factorization). For the symmetric matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix},$$

Gaussian elimination uses the multipliers $\ell_{21} = 2$, $\ell_{31} = 3$, $\ell_{32} = 2$, and its output U has rows $(1, 2, 3)$, $(0, -1, -2)$, $(0, 0, 1)$. So $D = \operatorname{diag}(1, -1, 1)$ and

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Over $\mathbb{R}$, when every diagonal entry of D is positive, $D = D_0^2$ with D_0 diagonal with entries $\sqrt{d_i}$, and $A = (LD_0)(LD_0)^{\mathsf{T}}$ is a product of a lower triangular matrix and its transpose. This is the **Cholesky factorization** of a positive definite matrix (Section 8.3). The signs of the d_i are not an accident of the computation either: by Sylvester's law of inertia (Section 12.2), the numbers of positive and of negative d_i are determined by A.

Exercises

Exercise 3.3.39. ROUTINE Let

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & 4 & 3 & 1 & 0 \\ 1 & 2 & 2 & 2 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 5}.$$

(a) Find $\operatorname{rref}(A)$, its pivot columns, and a matrix G with $GA = \operatorname{rref}(A)$.

(b) Without further row reduction, write each non-pivot column of A as a linear combination of the pivot columns.

Solution on page 1226

Exercise 3.3.40. ROUTINE Find the reduced row echelon form of

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

over $\mathbb{Q}$ and over $\mathbb{F}_2$. Where do the two computations part ways, and why?

Solution on page 1227

Exercise 3.3.41. ROUTINE Over $\mathbb{C}$, let

$$A = \begin{pmatrix} 1 & i & 0 \\ 0 & 1 & i \\ i & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & i \\ 0 & 1 & 1 \\ i & 1 & 0 \end{pmatrix}.$$

Use Algorithm 3.3.24 to find A^{-1}, and to show that B is not invertible. For B, find the linear relation among its rows that the algorithm produces.

Solution on page 1227

Exercise 3.3.42. STANDARD For a scalar c, let

$$A_c = \begin{pmatrix} 1 & 2 & 1 \\ 2 & c & 2 \\ 1 & 2 & c \end{pmatrix}.$$

(a) Find $\operatorname{rref}(A_c)$ for every $c \in \mathbb{R}$. For which c is A_c invertible?

(b) Answer the same questions over $\mathbb{F}_3$, for $c \in \mathbb{F}_3$.

Solution on page 1228

Exercise 3.3.43. ROUTINE Let $A \in F^{m\times n}$ and $R = \operatorname{rref}(A)$.

(a) Show that $\operatorname{rref}(cA) = R$ for every nonzero $c \in F$.

(b) Show that $\operatorname{rref}((A \mid A)) = (R \mid R)$.

(c) Let $\begin{pmatrix} A \\ A \end{pmatrix} \in F^{2m\times n}$ be the matrix with A stacked on top of itself. Show that $\operatorname{rref}\begin{pmatrix} A \\ A \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix}$.

Solution on page 1229

Exercise 3.3.44. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{m\times p}$.

(a) Show that the first n columns of $\operatorname{rref}((A \mid B))$ form $\operatorname{rref}(A)$.

(b) Give an example in which the last p columns of $\operatorname{rref}((A \mid B))$ differ from the matrix GB of Proposition 3.3.11.

Solution on page 1229

Exercise 3.3.45. STANDARD Prove or disprove each statement. Throughout, A and B are matrices over F of sizes for which the expressions make sense.

(a) $\operatorname{rref}\left(A^{\mathsf{T}}\right) = \operatorname{rref}(A)^{\mathsf{T}}$ for every square A.

(b) $\operatorname{rref}(AB) = \operatorname{rref}(A)\operatorname{rref}(B)$.

(c) If $\operatorname{rref}(A) = \operatorname{rref}(B)$, then $\operatorname{im} T_A = \operatorname{im} T_B$.

(d) If A and B have the same size and $\operatorname{im} T_A = \operatorname{im} T_B$, then $\operatorname{rref}(A) = \operatorname{rref}(B)$.

(e) $\operatorname{rref}(\operatorname{rref}(A)B) = \operatorname{rref}(AB)$.

Solution on page 1230

Exercise 3.3.46. STANDARD

(a) Find the error in the following argument, and give a counterexample to its conclusion.

Claim. If $R \in F^{m\times n}$ is in RREF and $G \in F^{m\times m}$ is invertible with $GR = R$, then $G = I_m$.

Proof. Apply the first proof of Theorem 3.3.17 with $R' = R$. Its induction shows that $\mathbf{g}_s = \mathbf{e}_s$ for every s, so every column of G is the corresponding column of I_m.

(b) Suppose R has r pivots. Describe all invertible G with $GR = R$.

Solution on page 1230

Exercise 3.3.47. STANDARD Find the error in the following argument, and give a counterexample to its conclusion.

Claim. Change step 1 of Algorithm 3.3.1 of Algorithm 3.3.1 so that it picks *any* column with a nonzero entry in an unpivoted row, not necessarily the leftmost one. The output of the modified algorithm is still in RREF.

Proof. The proofs of Lemma 3.3.7 and Proposition 3.3.8 go through word for word.

Solution on page 1231

Exercise 3.3.48. STANDARD Let p be a prime.

(a) Show that exactly $2p^2 + 2p + 3$ matrices in $\mathbb{F}_p^{2\times 3}$ are in RREF.

(b) Deduce that $\mathbb{F}_p^{2\times 3}$ splits into exactly $2p^2 + 2p + 3$ row equivalence classes. How many is that for $p = 2$?

Solution on page 1231

Exercise 3.3.49. STANDARD Find all matrices $A \in \mathbb{R}^{2\times 3}$ with $\ker T_A = \operatorname{span}((1,1,1))$.

Solution on page 1232

Exercise 3.3.50. STANDARD For $\mathbf{u}, \mathbf{x} \in F^n$ write $\mathbf{u}\cdot\mathbf{x} = u_1x_1 + \cdots + u_nx_n$. The **row space** of $A \in F^{m\times n}$ is the span of its rows, viewed as vectors in F^n.

(a) Show that $\mathbf{x} \in \ker T_A$ if and only if $\mathbf{v}\cdot\mathbf{x} = 0$ for every $\mathbf{v}$ in the row space of A.

(b) Deduce that $A, B \in F^{m\times n}$ are row equivalent if and only if they have the same row space.

Solution on page 1232

Exercise 3.3.51. STANDARD Let $a \in F$, and let $B \in F^{n\times n}$ have 1 in every diagonal position, a in every position $(i, i+1)$ directly above the diagonal, and 0 elsewhere.

(a) Row reduce $(B \mid I_n)$ to show that B is invertible, and that B^{-1} has entry $(-a)^{j-i}$ in position (i, j) for $j \geq i$, and 0 for $j < i$.

(b) Find the inverse of the $n \times n$ matrix U with $u_{ij} = 1$ for $j \geq i$ and $u_{ij} = 0$ for $j < i$.

Solution on page 1232

Exercise 3.3.52. STANDARD Let $n \geq 1$, and let $A \in F^{n\times n}$ have every diagonal entry 0 and every other entry 1. Write $c = (n-1) * 1 = 1 + \cdots + 1$ ($n-1$ terms) in F. Show that A is invertible if and only if $c \neq 0$, and find A^{-1} in that case. For which n is A invertible over $\mathbb{F}_p$?

Solution on page 1233

Exercise 3.3.53. STANDARD A matrix $U \in F^{n\times n}$ is **upper triangular** if $u_{ij} = 0$ whenever $i > j$.

(a) Let U be upper triangular with every diagonal entry nonzero. Show that Gauss–Jordan elimination on U performs no swaps, and that every elementary matrix it uses is upper triangular. Deduce that U is invertible, and that U^{-1} is upper triangular with diagonal entries $u_{11}^{-1}, \ldots, u_{nn}^{-1}$.

(b) Let U be upper triangular with some diagonal entry 0. Find a nonzero $\mathbf{x}$ with $U\mathbf{x} = \mathbf{0}$, and conclude that U is not invertible.

Solution on page 1234

Exercise 3.3.54. STANDARD Let F be a subfield of a field K (for example $\mathbb{Q} \subseteq \mathbb{R}$ or $\mathbb{R} \subseteq \mathbb{C}$), and let $A \in F^{m\times n}$, so that A is also a matrix over K.

(a) Show that $\operatorname{rref}(A)$ is the same whether it is computed over F or over K.

(b) Deduce that if $A \in F^{n\times n}$ is invertible as a matrix over K, then A^{-1} has all its entries in F.

(c) Show that $A, B \in F^{m\times n}$ are row equivalent over K (using elementary matrices with entries in K) if and only if they are row equivalent over F.

Solution on page 1235

Exercise 3.3.55. STANDARD

(a) Describe a procedure that, given $A, B \in F^{m\times n}$, decides whether $A \sim B$ and, if so, finds an invertible G with $GA = B$.

(b) Carry it out for $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix}$ over $\mathbb{Q}$.

Solution on page 1235

Exercise 3.3.56. HARD Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. A matrix $A \in \mathbb{F}^{n\times n}$ is **strictly diagonally dominant** if in every row the diagonal entry outweighs the rest:

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad \text{for } i = 1, \ldots, n.$$

Show that Gauss–Jordan elimination on such an A never needs a swap, and conclude that A is invertible.

Solution on page 1236

Exercise 3.3.57. HARD Let $A \in F^{m\times n}$, and let $G \in F^{m\times m}$ be invertible with $GA = \operatorname{rref}(A)$, which has r pivots. For a row vector $\mathbf{z} = (z_1, \ldots, z_m)$, the product $\mathbf{z}A$ is z_1 times row 1 of A, plus $\cdots$, plus z_m times row

m. Call $\mathbf{z}$ a **relation among the rows** of A if $\mathbf{z}A = \mathbf{0}$. Show that the relations among the rows of A are exactly the linear combinations of the last $m - r$ rows of G.

Solution on page 1236

Exercise 3.3.58. HARD Let p be a prime. For $0 \le r \le n$, let $N(r,n)$ be the number of $r \times n$ matrices over $\mathbb{F}_p$ that are in RREF and have no zero row; put $N(0,n) = 1$, and $N(r,n) = 0$ for $r > n$.

(a) Show that for $1 \le r \le n$ and $n \ge 2$,

$$N(r,n) = p^r N(r,n-1) + N(r-1,n-1).$$

(b) For $p = 2$, compute $N(2,4)$, and the number of row equivalence classes in $\mathbb{F}_2^{2\times 4}$.

Solution on page 1237

Exercise 3.3.59. ROUTINE BEYOND 21-242 Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 7 & 2 \\ -2 & -1 & 5 \end{pmatrix}, \quad S = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 5 & 1 \\ -2 & 1 & 6 \end{pmatrix}.$$

(a) Find the LU factorization of A.

(b) Find the LDL^{T} factorization of the symmetric matrix S, and write $S = MM^{\mathsf{T}}$ with M lower triangular.

Solution on page 1238

Exercise 3.3.60. STANDARD BEYOND 21-242 Show that an invertible matrix $A \in F^{n\times n}$ has at most one factorization $A = LU$ with L unit lower triangular and U upper triangular.

Solution on page 1238

Exercise 3.3.61. CHALLENGE BEYOND 21-242 Call an $n \times n$ matrix an **integer elementary matrix** if it is a swap E_{ij}, a sign change $E_{i,-1}$, or a replacement E_{icj} with $c \in \mathbb{Z}$. Show that for $A \in \mathbb{Z}^{n\times n}$ the following are equivalent:

(i) A is invertible (over $\mathbb{Q}$), and A^{-1} has integer entries;

(ii) A is a product of integer elementary matrices.

In other words, the matrices that can be inverted without leaving $\mathbb{Z}$ are exactly those that can be row reduced to I_n without leaving $\mathbb{Z}$.

Solution on page 1239

3.4 Solving linear systems

By Theorem 3.3.10, Gauss–Jordan elimination turns A into $\operatorname{rref}(A) = GA$ with G invertible, and by Proposition 3.2.14 the system $A\mathbf{x} = \mathbf{y}$ has the same solutions as the system $\operatorname{rref}(A)\mathbf{x} = G\mathbf{y}$. In Example 3.1.13 we read the solutions of one reduced system off by eye. This section does it in general. A reduced system says two things. Its zero rows decide whether there is a solution at all, and its other rows split the unknowns into free ones, which can take any values, and pivot ones, which the free ones then determine. From this we get a criterion for uniqueness and a criterion for existence, both in terms of pivots, and a first substantial application: through $n+1$ points with distinct first coordinates there passes exactly one polynomial of degree at most n.

3.4.1 Reducing to RREF

As in Definition 3.3.9, we write $A_r = \operatorname{rref}(A)$.

Proposition 3.4.1 (Solving with the reduced matrix). Let $A \in F^{m\times n}$, and let $G \in F^{m\times m}$ be invertible with $GA = A_r$ (Theorem 3.3.10). Then for all $\mathbf{x} \in F^n$ and $\mathbf{y} \in F^m$,

$$A\mathbf{x} = \mathbf{y} \quad \Leftrightarrow \quad A_r\mathbf{x} = G\mathbf{y}.$$

Moreover, if G is the product of the elementary matrices of the run of Gauss–Jordan elimination on A (Theorem 3.3.10), then the run on the augmented matrix $(A \mid \mathbf{y})$, with step 1 of Algorithm 3.3.1 restricted to the first n columns, ends at $(A_r \mid G\mathbf{y})$.

Proof. ROUTINE

Uses. Proposition 3.2.14 (Invertible matrices preserve solution sets), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.11 (Carrying extra columns along)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. Which result of Section 3.2 says that multiplying a system on the left by an invertible matrix does not change its solutions?
Hint 2. For $\Leftarrow$, multiply $A_r\mathbf{x} = G\mathbf{y}$ on the left by G^{-1} and use $G^{-1}A_r = A$. For the augmented matrix, apply Proposition 3.3.11 with $B = \mathbf{y}$.
Sketch. If $A\mathbf{x} = \mathbf{y}$, then $A_r\mathbf{x} = GA\mathbf{x} = G\mathbf{y}$. If $A_r\mathbf{x} = G\mathbf{y}$, then $A\mathbf{x} = G^{-1}A_r\mathbf{x} = G^{-1}G\mathbf{y} = \mathbf{y}$. The statement about $(A \mid \mathbf{y})$ is Proposition 3.3.11(b) with $B = \mathbf{y}$.

($\Rightarrow$) If $A\mathbf{x} = \mathbf{y}$, then $A_r\mathbf{x} = (GA)\mathbf{x} = G(A\mathbf{x}) = G\mathbf{y}$.

($\Leftarrow$) Since $GA = A_r$, we have $A = G^{-1}A_r$. So if $A_r\mathbf{x} = G\mathbf{y}$, then

$$A\mathbf{x} = G^{-1}(A_r\mathbf{x}) = G^{-1}(G\mathbf{y}) = \mathbf{y}.$$

(This is Proposition 3.2.14 for our particular G.) The last statement is Proposition 3.3.11(b), applied with $B = \mathbf{y} \in F^{m\times 1}$; it concerns the particular G produced by the run. (When A has fewer than m pivots, other invertible matrices G' also satisfy $G'A = A_r$, and the equivalence holds for each of them.) □

So the analysis of an arbitrary system reduces to that of a system whose coefficient matrix is in RREF. We never need G separately: we row reduce $(A \mid \mathbf{y})$, and $\mathbf{y}$ rides along in the last column. To treat every right-hand side at once, we can let the last column hold the symbols $y_1, \ldots, y_m$, which amounts to carrying I_m along and producing G (Proposition 3.3.11).

Definition 3.4.2 (Pivot and free variables). Let $A \in F^{m\times n}$, and let $p_1 < p_2 < \cdots < p_k$ be its pivot columns (Definition 3.3.9). The other columns, $f_1 < f_2 < \cdots < f_\ell$, are the **free columns** of A; thus $k + \ell = n$. In the system $A\mathbf{x} = \mathbf{y}$, the unknowns $x_{p_1}, \ldots, x_{p_k}$ are the **pivot variables** and $x_{f_1}, \ldots, x_{f_\ell}$ are the **free variables**.

These notions belong to A alone. They do not involve $\mathbf{y}$, and they do not depend on how A was reduced (Corollary 3.3.18). In fact Corollary 3.3.19 describes them without any reduction: column j is free exactly when $\mathbf{a}_j$ is a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$. The k pivots lie in rows $1, \ldots, k$ and in distinct columns (Proposition 3.3.8), so $k \le \min(m, n)$. The number k is the *rank* of A, the subject of Section 3.5. For the matrix of Example 3.3.3,

$$A = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 2 & 4 & 2 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad A_r = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

the pivot columns are $p_1 = 1$ and $p_2 = 3$, and the free columns are $f_1 = 2$ and $f_2 = 4$. In $A\mathbf{x} = \mathbf{y}$ the pivot variables are x_1, x_3, and the free variables are x_2, x_4.

3.4.2 Reading off the solutions

Intuition (What a reduced system says). Let $R = A_r$, with entries r_{ij}, and $\mathbf{z} = G\mathbf{y}$. Row $i \le k$ of R has its pivot, a 1, in column p_i, and zeros in the other pivot columns, because the pivot columns of R are $\mathbf{e}_1, \dots, \mathbf{e}_k$ (Lemma 3.3.16(b)). So the ith equation of $R\mathbf{x} = \mathbf{z}$ contains exactly one pivot variable, with coefficient 1, plus some free variables:

$$x_{p_i} + \sum_{j=1}^{\ell} r_{if_j} x_{f_j} = z_i.$$

The rows below row k are zero, so their equations read $0 = z_i$. These contain no unknowns at all: each one is either automatically true or impossible. For instance, a 4×6 matrix in RREF with pivot columns $1, 3, 6$ gives a system of the shape

$$\begin{aligned} x_1 + r_{12}x_2 + r_{14}x_4 + r_{15}x_5 &= z_1, \\ x_3 + r_{24}x_4 + r_{25}x_5 &= z_2, \\ x_6 &= z_3, \\ 0 &= z_4. \end{aligned}$$

If $z_4 \neq 0$, there is nothing to be done. If $z_4 = 0$, we may give the free variables x_2, x_4, x_5 any values we like, and then each of the first three equations determines its own pivot variable.

Theorem 3.4.3 (Solving a system in reduced form). Let $R \in F^{m \times n}$ be in RREF, with entries r_{ij}, pivot columns $p_1 < \cdots < p_k$, and free columns $f_1 < \cdots < f_\ell$. Let $\mathbf{z} \in F^m$.

(a) If $z_i \neq 0$ for some $i > k$, then $R\mathbf{x} = \mathbf{z}$ has no solution.

(b) If $z_i = 0$ for every $i > k$, then $\mathbf{x} \in F^n$ solves $R\mathbf{x} = \mathbf{z}$ if and only if

$$x_{p_i} = z_i - \sum_{j=1}^{\ell} r_{if_j} x_{f_j} \quad \text{for } i = 1, \dots, k. \tag{3.4.1}$$

In particular, for every choice of $t_1, \dots, t_\ell \in F$ there is exactly one solution with $x_{f_1} = t_1, \dots, x_{f_\ell} = t_\ell$.

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Proposition 3.3.8 (Properties of the output), Theorem 3.3.12 (Matrices in RREF), Lemma 3.3.16 (Reading the columns of a matrix in RREF)
Hint 1. Write out the ith entry of $R\mathbf{x}$, and split the sum over the columns into a sum over pivot columns and a sum over free columns.
Hint 2. Rows $k+1, \dots, m$ of R are zero. For $i \le k$, use that column p_s of R is $\mathbf{e}_s$.
Hint 3. For the last claim: every index $1, \dots, n$ is exactly one of $p_1, \dots, p_k, f_1, \dots, f_\ell$. So the values t_j together with (3.4.1) specify each coordinate of $\mathbf{x}$ exactly once.
Sketch. Split the ith entry of $R\mathbf{x}$ into a sum over the pivot columns and a sum over the free columns. For $i > k$ row i of R is zero, so equation i reads $0 = z_i$, which is (a). For $i \le k$, r_{ip_s} is 1 if $s = i$ and 0 otherwise, so equation i involves the pivot variable x_{p_i} with coefficient 1 and otherwise only free variables, which is (3.4.1). Free values plus (3.4.1) define one vector, and every solution with those free values is that vector.

Since R is in RREF, $R = \operatorname{rref}(R)$ is an output of Gauss–Jordan elimination, and its pivots are its leading entries (Theorem 3.3.12). By Proposition 3.3.8, rows $k+1, \dots, m$ of R are zero, and by Lemma 3.3.16(b), column p_s of R is $\mathbf{e}_s$ for $s = 1, \dots, k$.

By Definition 2.3.8, the ith entry of $R\mathbf{x}$ is $\sum_{l=1}^{n} r_{il}x_l$. Each index $l \in \{1, \dots, n\}$ is exactly one of $p_1, \dots, p_k, f_1, \dots, f_\ell$, so

$$(R\mathbf{x})_i = \sum_{s=1}^{k} r_{ip_s} x_{p_s} + \sum_{j=1}^{\ell} r_{if_j} x_{f_j}. \tag{3.4.2}$$

(a) If $i > k$, row i of R is zero, so $(R\mathbf{x})_i = 0$ for every $\mathbf{x}$. If moreover $z_i \neq 0$, the ith equation fails for every $\mathbf{x} \in F^n$.

(b) For $i > k$, both sides of the ith equation are 0, so it holds for every $\mathbf{x}$. For $i \leq k$, the entry r_{ip_s} is the ith entry of $\mathbf{e}_s$, which is 1 if $s = i$ and 0 otherwise. So by (3.4.2),

$$(R\mathbf{x})_i = x_{p_i} + \sum_{j=1}^{\ell} r_{if_j} x_{f_j},$$

and the ith equation $(R\mathbf{x})_i = z_i$ holds if and only if (3.4.1) holds for this i. Hence $R\mathbf{x} = \mathbf{z}$ if and only if (3.4.1) holds for $i = 1, \ldots, k$.

For the last claim, let $t_1, \ldots, t_\ell \in F$. Define $\mathbf{x} \in F^n$ by $x_{f_j} = t_j$ for each j and

$$x_{p_i} = z_i - \sum_{j=1}^{\ell} r_{if_j} t_j$$

for each i. Since each index is exactly one pivot or free column, this specifies every coordinate exactly once, and $\mathbf{x}$ is a solution by what we just proved. Conversely, a solution with $x_{f_j} = t_j$ for all j has its pivot coordinates given by (3.4.1), so it is this $\mathbf{x}$. □

The theorem is the lecture's recipe, made precise: the free variables can be anything, and then the pivot variables are determined from the free variables and the corresponding z_i. Applied to $R = A_r$ and $\mathbf{z} = G\mathbf{y}$, it settles consistency at once.

Corollary 3.4.4 (Consistency). Let $A \in F^{m\times n}$ have k pivots, let G be invertible with $GA = A_r$, and let $\mathbf{y} \in F^m$. Then $A\mathbf{x} = \mathbf{y}$ is consistent if and only if

$$(G\mathbf{y})_i = 0 \quad \text{for every } i \text{ with } k < i \leq m.$$

Equivalently, $A\mathbf{x} = \mathbf{y}$ is inconsistent exactly when the reduced augmented matrix $(A_r \mid G\mathbf{y})$ has a row $(0\cdots0 \mid c)$ with $c \neq 0$. In particular,

$$\operatorname{im} T_A = \{\mathbf{y} \in F^m \mid (G\mathbf{y})_i = 0 \text{ for } k < i \leq m\}.$$

Proof. ROUTINE

Uses. Proposition 3.4.1 (Solving with the reduced matrix), Theorem 3.4.3 (Solving a system in reduced form), Proposition 3.3.8 (Properties of the output), Proposition 3.1.8 (Systems and the map T_A)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. By Proposition 3.4.1, $A\mathbf{x} = \mathbf{y}$ has the same solutions as $A_r\mathbf{x} = G\mathbf{y}$. Apply Theorem 3.4.3 to the latter.
Hint 2. Which rows of A_r are zero? The image of T_A is the set of $\mathbf{y}$ for which the system is consistent.
Sketch. The system has the same solutions as $A_r\mathbf{x} = \mathbf{z}$ with $\mathbf{z} = G\mathbf{y}$. If some $z_i \neq 0$ with $i > k$, there is no solution by Theorem 3.4.3(a); otherwise there is one, for instance with all free variables 0, by Theorem 3.4.3(b). Rows $1, \ldots, k$ of A_r are nonzero and the others are zero, which gives the reformulation. The image of T_A consists of the $\mathbf{y}$ for which the system is consistent.

Let $\mathbf{z} = G\mathbf{y}$. By Proposition 3.4.1, $A\mathbf{x} = \mathbf{y}$ and $A_r\mathbf{x} = \mathbf{z}$ have the same solutions. If $z_i \neq 0$ for some $i > k$, the latter has no solution by Theorem 3.4.3(a). If $z_i = 0$ for all $i > k$, it has a solution by Theorem 3.4.3(b), for instance the one whose free variables are all 0.

By Proposition 3.3.8, rows $1, \ldots, k$ of A_r are nonzero (they contain the pivots) and rows $k+1, \ldots, m$ are zero. So a row of $(A_r \mid \mathbf{z})$ of the form $(0\cdots0 \mid c)$ with $c \neq 0$ is exactly a row $i > k$ with $z_i = c \neq 0$.

Finally, by Proposition 3.1.8(a), $\mathbf{y} \in \operatorname{im} T_A$ if and only if $A\mathbf{x} = \mathbf{y}$ is consistent, which we have just characterized. □

The description of $\operatorname{im} T_A$ is worth a second look. The conditions $(G\mathbf{y})_i = 0$ for $i > k$ are $m-k$ homogeneous linear equations in $y_1, \dots, y_m$, whose coefficients are the last $m-k$ rows of G. So the image of T_A, which is defined as a set of outputs, is also the solution set of a homogeneous system, that is, a kernel. For the running system, the last row of G in Example 3.2.8 is $(-1,-1,1)$, and the single condition $-y_1 - y_2 + y_3 = 0$ is the plane $y_3 = y_1 + y_2$ that Section 3.1.5 found by inspecting the columns.

If we row reduce $(A \mid \mathbf{y})$ all the way, without restricting the scan, the test reads slightly differently, and it has a proof that never mentions G.

Corollary 3.4.5 (Consistency from the augmented matrix). Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$. Then $A\mathbf{x} = \mathbf{y}$ is consistent if and only if the last column of the augmented matrix $(A \mid \mathbf{y})$ is not one of its pivot columns, that is, if and only if $\operatorname{rref}((A \mid \mathbf{y}))$ has no row $(0\cdots0 \mid 1)$.

Proof. ROUTINE

Uses. Corollary 3.3.19 (Pivot columns, intrinsically), Proposition 3.1.8 (Systems and the map T_A), Definition 3.3.9 (Reduced row echelon form), Proposition 3.3.8 (Properties of the output)
Hint 1. Apply Corollary 3.3.19(a) to the $m \times (n+1)$ matrix $(A \mid \mathbf{y})$ and its last column.
Hint 2. When is $\mathbf{y}$ a linear combination of the columns of A?
Sketch. By Corollary 3.3.19(a), the last column $\mathbf{y}$ is a pivot column exactly when it is not a combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$, that is, by Proposition 3.1.8(a), exactly when the system is inconsistent. A pivot in the last column of a matrix in RREF is the leading 1 of a row that vanishes elsewhere, a row $(0\cdots0 \mid 1)$.

The columns of $(A \mid \mathbf{y})$ are $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{y}$. By Corollary 3.3.19(a), applied to this matrix and its column $n+1$, the last column is a pivot column if and only if $\mathbf{y}$ is not a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$. By Proposition 3.1.8(a), this happens if and only if $A\mathbf{x} = \mathbf{y}$ is inconsistent.

By Definition 3.3.9 and Proposition 3.3.8(b), the pivots of $(A \mid \mathbf{y})$ are the leading entries of the nonzero rows of $\operatorname{rref}((A \mid \mathbf{y}))$, and they are 1. So the last column is a pivot column if and only if some row has its leading entry, a 1, in the last column, that is, if and only if some row is $(0\cdots0 \mid 1)$. □

The two tests agree, as they must: by Exercise 3.3.44 the first n columns of $\operatorname{rref}((A \mid \mathbf{y}))$ form A_r, and the unrestricted run simply goes on to clear the last column once a row $(0\cdots0 \mid c)$ with $c \neq 0$ appears.

Example 3.4.6 (A complete solution). We solve $A\mathbf{x} = \mathbf{y}$ for the matrix A of Example 3.3.3 and every $\mathbf{y} \in \mathbb{Q}^3$. Carrying I_3 along in Section 3.3.3 produced

$$G = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ -1 & -1/2 & 1 \end{pmatrix},$$

so by Proposition 3.4.1 the system is equivalent to

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & -2 & (-y_1+y_2)/2 \\ 0 & 0 & 1 & 2 & y_1/2 \\ 0 & 0 & 0 & 0 & -y_1 - y_2/2 + y_3 \end{array}\right).$$

Here $k = 2$, so there is one consistency condition. By Corollary 3.4.4, the system is consistent if and only if $y_3 = y_1 + y_2/2$, and $\operatorname{im} T_A$ is the plane $2y_3 = 2y_1 + y_2$ in $\mathbb{Q}^3$. (Each column of A satisfies this, for example $2\cdot 4 = 2\cdot 4 + 0$ for $\mathbf{a}_4 = (4,0,4)$.)

Take $\mathbf{y} = (4,2,5)$. Then $G\mathbf{y} = (-1,2,0)$, the condition holds, and by (3.4.1) the solutions are exactly the vectors with

$$x_1 = -1 - 2x_2 + 2x_4, \quad x_3 = 2 - 2x_4, \quad x_2, x_4 \in \mathbb{Q} \text{ arbitrary}.$$

For $\mathbf{y} = (4,2,6)$, instead, $G\mathbf{y} = (-1,2,1)$, and the last equation of the reduced system reads $0 = 1$: there is no solution.

The row reducer below carries out the reduction of $(A \mid \mathbf{y})$ for $\mathbf{y} = (4, 2, 5)$, scanning only the columns of A. At the end it reports the pivot and free columns and whether the system is consistent. Edit the last column to $(4, 2, 6)$ to see the row $(0, 0, 0, 0 \mid 1)$ appear.

$$\left(\begin{array}{cccc|c} 0 & 0 & 2 & 4 & 4 \\ 2 & 4 & 2 & 0 & 2 \\ 1 & 2 & 3 & 4 & 5 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 0 & -2 & -1 \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Solving the system of Example 3.4.6. Pivots are boxed; the bar separates A from $\mathbf{y}$. (Interactive in the web version.)

Exercise 3.4.7. ROUTINE Let

$$A = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 2 & 4 & 3 \\ 0 & 1 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4}.$$

Find A_r and the pivot and free variables. For which $\mathbf{y} \in \mathbb{R}^3$ is $A\mathbf{x} = \mathbf{y}$ consistent? Find all solutions for $\mathbf{y} = (1, 2, 1)$.

Solution on page 1239

3.4.3 The shape of the solution set

Formula (3.4.1) expresses the pivot variables through the free ones. Collecting, for each free variable, the terms that involve it writes every solution as one fixed vector plus a combination of vectors that do not depend on $\mathbf{y}$, which is the form Proposition 3.1.19 predicts.

Definition 3.4.8 (Basic solutions). Let $A \in F^{m\times n}$, with $A_r = GA$ as above, entries r_{ij} of A_r, pivot columns $p_1 < \cdots < p_k$, and free columns $f_1 < \cdots < f_\ell$.

(a) For $j = 1, \ldots, \ell$, the jth **basic solution** of A is the vector $\mathbf{s}_j \in F^n$ with

$$(\mathbf{s}_j)_{f_j} = 1, \quad (\mathbf{s}_j)_{f_{j'}} = 0 \text{ for } j' \neq j, \quad (\mathbf{s}_j)_{p_i} = -r_{if_j} \text{ for } i = 1, \ldots, k.$$

(b) If $A\mathbf{x} = \mathbf{y}$ is consistent, its **particular solution** $\mathbf{x}_0 \in F^n$ has $(\mathbf{x}_0)_{p_i} = (G\mathbf{y})_i$ for $i = 1, \ldots, k$, and $(\mathbf{x}_0)_{f_j} = 0$ for $j = 1, \ldots, \ell$.

In words: $\mathbf{s}_j$ is the solution of $A\mathbf{x} = \mathbf{0}$ obtained by setting the free variable x_{f_j} to 1 and the other free variables to 0, and $\mathbf{x}_0$ is the solution of $A\mathbf{x} = \mathbf{y}$ obtained by setting every free variable to 0. (Both are solutions by Theorem 3.4.3(b) and Proposition 3.4.1.) The $\mathbf{s}_j$ depend only on A_r, which depends only on A. The vector $\mathbf{x}_0$ is the *only* solution whose free variables all vanish, by Theorem 3.4.3(b), so it does not depend on which G we used either. Realize also that $\mathbf{s}_j$ can be copied from column f_j of A_r: negate its first k entries and put them in the pivot positions $p_1, \ldots, p_k$, put a 1 in position f_j, and zeros elsewhere.

Theorem 3.4.9 (All solutions). Let $A \in F^{m\times n}$, and suppose $A\mathbf{x} = \mathbf{y}$ is consistent. Let $\mathbf{x}_0$ be its particular solution and $\mathbf{s}_1, \ldots, \mathbf{s}_\ell$ the basic solutions of A (Definition 3.4.8).

(a) The solutions of $A\mathbf{x} = \mathbf{y}$ are exactly the vectors

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{s}_1 + t_2\mathbf{s}_2 + \cdots + t_\ell\mathbf{s}_\ell \quad (t_1, \ldots, t_\ell \in F),$$

and each solution arises from exactly one choice of $t_1, \ldots, t_\ell$, namely $t_j = x_{f_j}$.

(b) $\ker T_A = \operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell)$, and each vector of $\ker T_A$ is a linear combination of $\mathbf{s}_1, \ldots, \mathbf{s}_\ell$ in exactly one way.

So the solution set of $A\mathbf{x} = \mathbf{y}$ is $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell) = \mathbf{x}_0 + \ker T_A$.

Proof. STANDARD

Uses. Theorem 3.4.3 (Solving a system in reduced form), Proposition 3.4.1 (Solving with the reduced matrix), Corollary 3.4.4 (Consistency), Proposition 3.1.8 (Systems and the map T_A), Definition 1.3.10 (Span)
Techniques. Particular plus homogeneous (Technique 3.1.18), Reduce to RREF (Technique 3.3.22)
Hint 1. Fix $t_1, \ldots, t_\ell$ and compute the coordinates of $\mathbf{w} = \mathbf{x}_0 + \sum_j t_j \mathbf{s}_j$: first those in the free positions, then those in the pivot positions.
Hint 2. The f_j-coordinate of $\mathbf{w}$ is t_j, and its p_i-coordinate is $z_i - \sum_j r_{if_j} t_j$, where $\mathbf{z} = G\mathbf{y}$. Compare with (3.4.1).
Hint 3. For (b), apply (a) with $\mathbf{y} = \mathbf{0}$. What is the particular solution then?
Sketch. With $\mathbf{z} = G\mathbf{y}$, the vector $\mathbf{w} = \mathbf{x}_0 + \sum_j t_j \mathbf{s}_j$ has f_j-coordinate t_j and p_i-coordinate $z_i - \sum_j r_{if_j} t_j$. By Theorem 3.4.3(b), $\mathbf{w}$ is a solution, and every solution $\mathbf{x}$ equals the $\mathbf{w}$ with $t_j = x_{f_j}$, because both have the same free coordinates and the pivot coordinates are determined by them. Reading off free coordinates gives uniqueness of the t_j. For $\mathbf{y} = \mathbf{0}$ the particular solution is $\mathbf{0}$, which gives (b).

Let $\mathbf{z} = G\mathbf{y}$. Since the system is consistent, $z_i = 0$ for $i > k$ (Corollary 3.4.4), and by Proposition 3.4.1 its solutions are those of $A_r\mathbf{x} = \mathbf{z}$.

(a) For $t_1, \ldots, t_\ell \in F$, let $\mathbf{w} = \mathbf{x}_0 + t_1\mathbf{s}_1 + \cdots + t_\ell\mathbf{s}_\ell$. In position f_j, the vector $\mathbf{x}_0$ has a 0, $\mathbf{s}_j$ has a 1, and $\mathbf{s}_{j'}$ has a 0 for $j' \neq j$, so

$$w_{f_j} = t_j.$$

In position p_i, $\mathbf{x}_0$ has z_i and $\mathbf{s}_j$ has $-r_{if_j}$, so

$$w_{p_i} = z_i - \sum_{j=1}^{\ell} r_{if_j} t_j = z_i - \sum_{j=1}^{\ell} r_{if_j} w_{f_j}.$$

This is (3.4.1) for $\mathbf{w}$, so $\mathbf{w}$ solves $A_r\mathbf{x} = \mathbf{z}$ by Theorem 3.4.3(b), and hence $A\mathbf{x} = \mathbf{y}$.

Conversely, let $\mathbf{x}$ be a solution, and let $\mathbf{w}$ be the vector above with $t_j = x_{f_j}$ for each j. Then $\mathbf{x}$ and $\mathbf{w}$ are solutions with the same free coordinates, so by the uniqueness in Theorem 3.4.3(b) they are equal. Finally, if $\mathbf{x} = \mathbf{x}_0 + \sum_j t_j\mathbf{s}_j$, then $x_{f_j} = t_j$ by the computation above, so $t_1, \ldots, t_\ell$ are determined by $\mathbf{x}$.

(b) By Proposition 3.1.8(b), $\ker T_A$ is the solution set of $A\mathbf{x} = \mathbf{0}$, which is consistent. Its particular solution is $\mathbf{0}$, since $G\mathbf{0} = \mathbf{0}$. By (a), its solutions are exactly the vectors $t_1\mathbf{s}_1 + \cdots + t_\ell\mathbf{s}_\ell$, each for exactly one choice of the t_j. By Definition 1.3.10, the set of these vectors is $\operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell)$.

The last sentence of the theorem combines (a) and (b). □

This is Proposition 3.1.19 made explicit. That proposition said that once we know one solution, the others are obtained by adding elements of $\ker T_A$; the theorem produces the one solution, $\mathbf{x}_0$, and describes $\ker T_A$ by ℓ vectors, one for each free variable. It is the canonical instance of Technique 3.1.18. There is also a quick way to see that $A\mathbf{s}_j = \mathbf{0}$ without the theorem. By Corollary 3.3.19(b), column f_j of A is $\sum_i r_{if_j}\mathbf{a}_{p_i}$ (the terms with $p_i > f_j$ are zero by Lemma 3.3.16(a)). Moving everything to one side gives a relation among the columns of A whose coefficient vector is exactly $\mathbf{s}_j$.

Example 3.4.10 (The complete solution, in vector form). In Example 3.4.6, with $\mathbf{y} = (4, 2, 5)$, we have $G\mathbf{y} = (-1, 2, 0)$ and

$$A_r = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The particular solution has -1 and 2 in the pivot positions $1, 3$, and 0 in the free positions $2, 4$. The basic solutions come from columns 2 and 4 of A_r:

$$\mathbf{x}_0 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{s}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

By Theorem 3.4.9, the solution set is

$$\{(-1,0,2,0) + t_1(-2,1,0,0) + t_2(2,0,-2,1) \mid t_1, t_2 \in \mathbb{Q}\},$$

a plane in $\mathbb{Q}^4$, and $\ker T_A = \operatorname{span}((-2,1,0,0),(2,0,-2,1))$. The relations among the columns behind $\mathbf{s}_1$ and $\mathbf{s}_2$ are $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_4 = -2\mathbf{a}_1 + 2\mathbf{a}_3$, which we already read off in Section 3.3.4.

Remark 3.4.11 (A basis of the kernel). Part (b) of Theorem 3.4.9 says that every vector of $\ker T_A$ is a linear combination of $\mathbf{s}_1, \ldots, \mathbf{s}_\ell$ in exactly one way. In the language of Section 1.5 and Section 1.6, the list $\mathbf{s}_1, \ldots, \mathbf{s}_\ell$ is linearly independent (Definition 1.5.1) and spans $\ker T_A$, so it is a basis of $\ker T_A$ (Definition 1.6.1). It has one vector for each free variable, so $\ell = n - k$ vectors. Once dimension is defined, this says that $\dim \ker T_A$ is the number of free columns of A, which is the starting point of rank–nullity (Section 4.2).

Remark 3.4.12 (Checking a description of the solutions). Solutions are often written down by some other method, or the parameters are chosen differently. A description $\mathbf{x}_{0'} + \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$ is easy to test in one direction: if $A\mathbf{x}_{0'} = \mathbf{y}$ and $A\mathbf{u}_i = \mathbf{0}$ for every i, then every vector in it is a solution. But this test cannot detect missing solutions: dropping one of the $\mathbf{u}_i$ still passes it. The description is complete exactly when $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r) = \ker T_A$, and by Theorem 3.4.9 this holds if and only if each basic solution $\mathbf{s}_j$ is a combination of the $\mathbf{u}_i$. (Section 4.2 gives a quicker test: $r = \ell$ linearly independent vectors of $\ker T_A$ always span it.) Exercise 3.4.34 is an example.

The whole procedure fits in a few lines.

Algorithm 3.4.13 (Solving a linear system)

Input A matrix $A \in F^{m \times n}$ and a vector $\mathbf{y} \in F^m$.
Output The solution set of $A\mathbf{x} = \mathbf{y}$, in the form $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell)$, or the conclusion that the system is inconsistent.

1. Row reduce $(A \mid \mathbf{y})$ by Gauss–Jordan elimination, scanning only the first n columns in step 1 of Algorithm 3.3.1. The result is $(A_r \mid \mathbf{z})$ with $\mathbf{z} = G\mathbf{y}$. Record the pivot columns $p_1, \ldots, p_k$ and the free columns $f_1, \ldots, f_\ell$.
2. If $z_i \neq 0$ for some $i > k$, stop: the system is inconsistent.
3. Otherwise, let $\mathbf{x}_0$ have z_i in position p_i for $i = 1, \ldots, k$, and 0 in every free position.
4. For each $j = 1, \ldots, \ell$, let $\mathbf{s}_j$ have 1 in position f_j, 0 in the other free positions, and $-r_{if_j}$ in position p_i, where r_{if_j} is the ith entry of column f_j of A_r.
5. Output $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell)$.

It is correct by Proposition 3.4.1, Corollary 3.4.4, and Theorem 3.4.9. The bijection $(t_1, \ldots, t_\ell) \mapsto \mathbf{x}_0 + \sum_j t_j \mathbf{s}_j$ of Theorem 3.4.9(a) also counts the solutions.

Corollary 3.4.14 (Counting solutions). Suppose $A\mathbf{x} = \mathbf{y}$ is consistent, where $A \in F^{m \times n}$ has ℓ free columns.

(a) The system has exactly one solution if and only if $\ell = 0$, that is, if and only if every column of A is a pivot column.

(b) If F is finite with q elements, the system has exactly q^ℓ solutions.

(c) If F is infinite and $\ell \geq 1$, the system has infinitely many solutions.

Proof. ROUTINE

> *Uses.* Theorem 3.4.9 (All solutions)
> *Hint 1.* Theorem 3.4.9(a) sets up a bijection between the solution set and the set of lists $(t_1, \dots, t_\ell)$ of scalars.
> *Hint 2.* For (a) with $\ell \geq 1$, compare $\mathbf{x}_0$ and $\mathbf{x}_0 + \mathbf{s}_1$.
> *Sketch.* By Theorem 3.4.9(a), $(t_1, \dots, t_\ell) \mapsto \mathbf{x}_0 + \sum_j t_j \mathbf{s}_j$ is a bijection from F^ℓ onto the solution set (for $\ell = 0$, the solution set is $\{\mathbf{x}_0\}$). Count: F^ℓ has q^ℓ elements when $|F| = q$, and infinitely many when F is infinite and $\ell \geq 1$. If $\ell \geq 1$, the solutions $\mathbf{x}_0$ and $\mathbf{x}_0 + \mathbf{s}_1$ differ in position f_1.

By Theorem 3.4.9(a), each solution is $\mathbf{x}_0 + \sum_j t_j \mathbf{s}_j$ for exactly one list $(t_1, \dots, t_\ell) \in F^\ell$, and every such vector is a solution. So the solution set is in bijection with F^ℓ. If $\ell = 0$, there are no parameters and the only solution is $\mathbf{x}_0$.

(a) If $\ell = 0$, the only solution is $\mathbf{x}_0$. If $\ell \geq 1$, then $\mathbf{x}_0$ and $\mathbf{x}_0 + \mathbf{s}_1$ are solutions, and they differ, since their f_1 -coordinates are 0 and 1.

(b) F^ℓ has q^ℓ elements (and $q^0 = 1$ covers $\ell = 0$).

(c) If $\ell \geq 1$ and F is infinite, the vectors $\mathbf{x}_0 + t\mathbf{s}_1$, $t \in F$, are infinitely many distinct solutions. □

Part (b) is the formula promised after Example 3.1.26, where the equation $x_1 + x_2 = 1$ over $\mathbb{F}_2$ has one free variable and 2^1 solutions. We did not need dimension: the number of free variables does the counting. Part (c) recovers Corollary 3.1.24, and adds that the number of solutions of a consistent system is decided by A alone.

3.4.4 Pivots in every column, pivots in every row

> *Intuition (Free variables and zero rows).* Every system $A\mathbf{x} = \mathbf{y}$ raises two questions: is there a solution, and is it unique? We have now located the source of each answer in A_r. Non-uniqueness comes from free variables, since each one can be varied independently. Inconsistency comes from zero rows, since each one produces an equation $0 = (G\mathbf{y})_i$, which fails for a suitable $\mathbf{y}$. A free variable is a column without a pivot, and a zero row is a row without a pivot. So we expect: a pivot in every column means uniqueness, and a pivot in every row means existence for every right-hand side. Uniqueness is a property of $\ker T_A$ (Corollary 3.1.22), so the first statement is about the homogeneous system.

Proposition 3.4.15 (A pivot in every column). Let $A \in F^{m \times n}$. The following are equivalent.
(I) A_r has a pivot in every column.
(II) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

The homogeneous system is consistent, so this is the case $\mathbf{y} = \mathbf{0}$ of Corollary 3.4.14(a). The lecture proves it directly, and the direct argument is worth having: it constructs a nonzero solution whenever one exists.

Proof (set one free variable to 1). ROUTINE

> *Uses.* Proposition 3.4.1 (Solving with the reduced matrix), Theorem 3.4.3 (Solving a system in reduced form), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.3.8 (Matrix–vector product)
> *Techniques.* Reduce to RREF (Technique 3.3.22), Contrapositive (Technique 1.5.4)
> *Hint 1.* Reduce: $A\mathbf{x} = \mathbf{0}$ if and only if $A_r\mathbf{x} = \mathbf{0}$. If every column of A_r is a pivot column, what are its columns?
> *Hint 2.* For (II) $\Rightarrow$ (I), prove the contrapositive. If some column f has no pivot, x_f is a free variable. Choose values of the free variables and solve for the pivot variables with Theorem 3.4.3.
> *Hint 3.* Set $x_f = 1$ and every other free variable to 0.
> *Sketch.* $A\mathbf{x} = \mathbf{0}$ iff $A_r\mathbf{x} = \mathbf{0}$, since $G\mathbf{0} = \mathbf{0}$. If every column is a pivot column, then $k = n$, $p_s = s$, and A_r has columns $\mathbf{e}_1, \dots, \mathbf{e}_n$, so $A_r\mathbf{x}$ lists $x_1, \dots, x_n$ followed by zeros, and vanishes only for $\mathbf{x} = \mathbf{0}$. If column f is free, Theorem 3.4.3(b) gives a solution with $x_f = 1$ and the other free variables 0; it is nonzero.

By Proposition 3.4.1 with $\mathbf{y} = \mathbf{0}$, and since $G\mathbf{0} = \mathbf{0}$, the equations $A\mathbf{x} = \mathbf{0}$ and $A_r\mathbf{x} = \mathbf{0}$ have the same solutions.

(I) ⇒ (II). Suppose every column of A_r is a pivot column. Then $k = n$, and $p_1 < \cdots < p_n$ are n increasing numbers in $\{1, \ldots, n\}$, so $p_s = s$ for every s. By Lemma 3.3.16(b), column s of A_r is $\mathbf{e}_s$, so by Definition 2.3.8

$$A_r\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n,$$

the vector of F^m whose first n entries are $x_1, \ldots, x_n$ and whose remaining $m - n$ entries are 0. It is $\mathbf{0}$ only if $\mathbf{x} = \mathbf{0}$.

(II) ⇒ (I). We prove the contrapositive. Suppose some column of A_r has no pivot, so it is a free column f_j. The right-hand side $\mathbf{0}$ certainly has $z_i = 0$ for $i > k$, so by Theorem 3.4.3(b) the system $A_r\mathbf{x} = \mathbf{0}$ has a solution $\mathbf{x}$ with $x_{f_j} = 1$ and every other free variable 0. This $\mathbf{x}$ is nonzero, as its f_j-coordinate is 1, and $A\mathbf{x} = \mathbf{0}$. So (II) fails. □

Proof (via relations among the columns). STANDARD

> *Uses.* Definition 2.3.8 (Matrix–vector product), Theorem 1.5.5 (Independence is decided at zero), Corollary 3.3.19 (Pivot columns, intrinsically), Lemma 1.5.10 (Linear dependence lemma)
> *Techniques.* Contrapositive (Technique 1.5.4)
> *Hint 1.* By Definition 2.3.8, a solution of $A\mathbf{x} = \mathbf{0}$ is a linear relation among the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Restate (II) in terms of linear independence.
> *Hint 2.* Corollary 3.3.19(a) restates (I): no column is a linear combination of the columns before it.
> *Hint 3.* Connect the two restatements with the linear dependence lemma (Lemma 1.5.10).
> *Sketch.* (II) says that $\mathbf{a}_1, \ldots, \mathbf{a}_n$ has only the trivial relation, that is, is linearly independent (Theorem 1.5.5). (I) says that no $\mathbf{a}_j$ lies in $\operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1})$ (Corollary 3.3.19(a)). If some $\mathbf{a}_j = \sum_{l<j} c_l\mathbf{a}_l$, then $\sum_{l<j} c_l\mathbf{a}_l - \mathbf{a}_j = \mathbf{0}$ is a nontrivial relation; if the columns are dependent, Lemma 1.5.10 provides such a j.

By Definition 2.3.8, $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$. So (II) says that the only linear relation among $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is the trivial one, which by Theorem 1.5.5 means that the list $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is linearly independent. By Corollary 3.3.19(a), (I) says that for no j is $\mathbf{a}_j$ a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$.

If (I) fails, then $\mathbf{a}_j = c_1\mathbf{a}_1 + \cdots + c_{j-1}\mathbf{a}_{j-1}$ for some j and some scalars c_l, and then

$$c_1\mathbf{a}_1 + \cdots + c_{j-1}\mathbf{a}_{j-1} + (-1)\mathbf{a}_j = \mathbf{0}$$

is a linear relation with a nonzero coefficient, so (II) fails. If (II) fails, the columns are linearly dependent, and the linear dependence lemma (Lemma 1.5.10) gives an index j with $\mathbf{a}_j \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1})$, so (I) fails. □

The first proof is constructive: when (I) fails, it hands us a nonzero solution, namely a basic solution. The second explains what a missing pivot means: a column that depends on the columns before it. It also shows that a pivot in every column is the same as linear independence of the columns. By Theorem 2.6.11, (II) says that T_A is injective, so a pivot in every column is a mechanical test for the equivalent conditions of Theorem 2.7.2. Section 3.5 adds it to that theorem.

Proposition 3.4.16 (A pivot in every row). Let $A \in F^{m\times n}$. The following are equivalent.

(I) A_r has a pivot in every row.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x} \in F^n$.

Proof. ROUTINE

> *Uses.* Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.8 (Properties of the output), Corollary 3.4.4 (Consistency), Theorem 3.4.3 (Solving a system in reduced form), Definition 2.5.12 (Invertible)
> *Techniques.* Reduce to RREF (Technique 3.3.22), Contrapositive (Technique 1.5.4)
> *Hint 1.* By Corollary 3.4.4, consistency is decided by the entries $(G\mathbf{y})_i$ with $i > k$. How many such i are there when every row has a pivot?
> *Hint 2.* For (II) ⇒ (I), prove the contrapositive. It suffices to find $\mathbf{z}$ such that $A_r\mathbf{x} = \mathbf{z}$ has no solution, and then take $\mathbf{y} = G^{-1}\mathbf{z}$.
> *Hint 3.* If some row has no pivot, the last row of A_r is zero. Try $\mathbf{z} = \mathbf{e}_m$.

Sketch. The pivots lie in rows $1, \dots, k$, so (I) says $k = m$. If $k = m$, the consistency conditions of Corollary 3.4.4 are vacuous; explicitly, $x_{p_i} = (G\mathbf{y})_i$ with all free variables 0 is a solution. If $k < m$, the last row of A_r is zero. With $\mathbf{y} = G^{-1}\mathbf{e}_m$, the system $A\mathbf{x} = \mathbf{y}$ is equivalent to $A_r\mathbf{x} = \mathbf{e}_m$, whose last equation reads $0 = 1$.

Let G be invertible with $GA = A_r$ (Theorem 3.3.10), and let k be the number of pivots. By Proposition 3.3.8 the pivots lie in rows $1, \dots, k$, one in each, so (I) says exactly that $k = m$.

(I) ⇒ (II). Suppose $k = m$, and let $\mathbf{y} \in F^m$. There is no index i with $k < i \le m$, so the condition in Corollary 3.4.4 holds trivially, and $A\mathbf{x} = \mathbf{y}$ is consistent. Explicitly, by Theorem 3.4.3(b) and Proposition 3.4.1, the vector with $x_{p_i} = (G\mathbf{y})_i$ for $i = 1, \dots, m$ and all free variables 0 is a solution.

(II) ⇒ (I). We prove the contrapositive. Suppose $k < m$, so that row m of A_r is zero (Proposition 3.3.8(a)). Let $\mathbf{y} = G^{-1}\mathbf{e}_m$. Then $G\mathbf{y} = \mathbf{e}_m$ has mth entry $1 \neq 0$, and $m > k$, so $A\mathbf{x} = \mathbf{y}$ is inconsistent by Corollary 3.4.4. (Directly: $A\mathbf{x} = \mathbf{y}$ if and only if $A_r\mathbf{x} = \mathbf{e}_m$, and the mth entry of $A_r\mathbf{x}$ is 0 for every $\mathbf{x}$.) So (II) fails. □

Remark 3.4.17 (Why G^{-1}). It is tempting to take $\mathbf{y} = \mathbf{e}_m$ itself in the last proof, but row operations move the image of T_A (Non-example 3.2.19), and $A\mathbf{x} = \mathbf{e}_m$ may well be consistent. For instance,

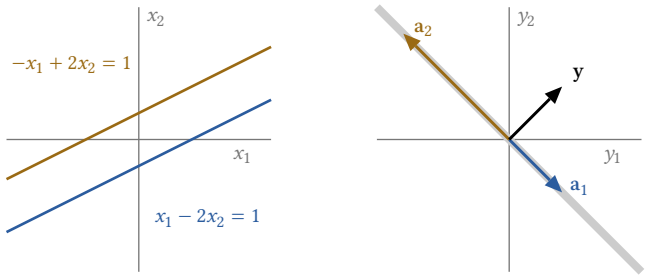
$$A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A_r = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = E_{12}A,$$

and $A\mathbf{x} = \mathbf{e}_2$ has the solution $x = 1$. With $G = E_{12}$, the bad right-hand side is $G^{-1}\mathbf{e}_2 = \mathbf{e}_1$: the equation $A\mathbf{x} = \mathbf{e}_1$ says $0 = 1$.

By Proposition 3.1.8(a), (II) says that T_A is surjective, equivalently that the columns of A span F^m, so a pivot in every row is a mechanical test for the conditions of Theorem 2.7.4. For a 2×2 matrix both failures can be watched at once. The component below starts with

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}, \quad A_r = \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix}.$$

The columns $\mathbf{a}_1 = (1, -1)$ and $\mathbf{a}_2 = (-2, 2) = -2\mathbf{a}_1$ lie on one line, and A_r has one pivot: its second column is free and its second row is zero. For $\mathbf{y} = (1, 1)$, off the line $\operatorname{span}(\mathbf{a}_1)$, the two lines of the row picture are parallel and there is no solution. Drag $\mathbf{y}$ onto the line spanned by the columns and the two lines of the row picture coincide: now there are infinitely many solutions, one for each value of the free variable x_2. Drag $\mathbf{a}_2$ off that line and both defects disappear at once, as they must for a square matrix (Corollary 3.4.20).



Row picture Column picture

The system $x_1 - 2x_2 = 1$, $-x_1 + 2x_2 = 1$. Left: the row picture, with one line per equation. Right: the column picture, where $\mathbf{y}$ must be reached as $x_1\mathbf{a}_1 + x_2\mathbf{a}_2$. Drag $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{y}$, or type new entries. (Interactive in the web version.)

Exercise 3.4.18. ROUTINE Each matrix below is A_r for some A. For each one, decide whether $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, and whether $A\mathbf{x} = \mathbf{y}$ is consistent for every $\mathbf{y}$.

$$R_1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \end{pmatrix}, \quad R_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad R_3 = \begin{pmatrix} 1 & 4 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad R_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution on page 1240

Since there are at most $\min(m, n)$ pivots, the shape of A alone can rule out one of the two properties.

Corollary 3.4.19 (More unknowns than equations). Let $A \in F^{m \times n}$.

(a) If $n > m$, then $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

(b) If $m > n$, then $A\mathbf{x} = \mathbf{y}$ is inconsistent for some $\mathbf{y} \in F^m$.

Proof. ROUTINE

Uses. Proposition 3.3.8 (Properties of the output), Proposition 3.4.15 (A pivot in every column), Proposition 3.4.16 (A pivot in every row)
Hint 1. How many pivots can A_r have?
Hint 2. If $n > m$, there are more columns than pivots.
Sketch. There are $k \leq \min(m, n)$ pivots. If $n > m$, then $k \leq m < n$, so some column has no pivot, and Proposition 3.4.15 applies. If $m > n$, then $k \leq n < m$, so some row has no pivot, and Proposition 3.4.16 applies.

By Proposition 3.3.8, A_r has $k \leq \min(m, n)$ pivots, in distinct rows and distinct columns. If $n > m$, then $k \leq m < n$, so at least one of the n columns has no pivot, and $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution by Proposition 3.4.15. If $m > n$, then $k \leq n < m$, so at least one of the m rows has no pivot, and some $A\mathbf{x} = \mathbf{y}$ is inconsistent by Proposition 3.4.16. □

Part (a) has a geometric reading over $\mathbb{R}$: two planes through the origin in $\mathbb{R}^3$ always share a line, since two homogeneous equations in three unknowns have a nontrivial solution. In terms of the Fundamental Theorems, (a) and (b) say that T_A can be injective only if $n \leq m$, and surjective only if $m \leq n$. Section 3.5 draws the consequences, among them that every invertible matrix is square.

For square matrices the two properties cannot be separated.

Corollary 3.4.20 (Square systems). Let $A \in F^{n \times n}$. The following are equivalent.

(i) A_r has a pivot in every column.
(ii) A_r has a pivot in every row.
(iii) $A_r = I_n$.
(iv) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
(v) $A\mathbf{x} = \mathbf{y}$ has a solution for every $\mathbf{y} \in F^n$.
(vi) $A\mathbf{x} = \mathbf{y}$ has exactly one solution for every $\mathbf{y} \in F^n$.
(vii) A is invertible.

Proof. STANDARD

Uses. Proposition 3.3.8 (Properties of the output), Proposition 3.4.15 (A pivot in every column), Proposition 3.4.16 (A pivot in every row), Corollary 3.1.22 (Uniqueness depends only on A), Proposition 3.1.8 (Systems and the map T_A), Theorem 3.3.23 (Invertibility and RREF)
Techniques. Proving TFAE (Technique 2.7.1), Reduce to RREF (Technique 3.3.22)
Hint 1. Each pivot occupies its own row and its own column. What do (i) and (ii) say about the number k of pivots of a square matrix?
Hint 2. (i) ⇔ (iv) and (ii) ⇔ (v) are the two propositions. For (vi), combine existence with Corollary 3.1.22.
Hint 3. (iii) ⇔ (vii) is Theorem 3.3.23.
Sketch. With k pivots in distinct rows and columns, (i) and (ii) both say $k = n$; and $k = n$ forces $p_s = s$ and $A_r = I_n$, while I_n has n pivots. The propositions give (i) ⇔ (iv) and (ii) ⇔ (v). (vi) implies (iv) (take $\mathbf{y} = \mathbf{0}$) and (v); conversely (iv) and (v) together give (vi) by Corollary 3.1.22. Finally (iii) ⇔ (vii) is Theorem 3.3.23.

Let k be the number of pivots of A. By Proposition 3.3.8 they lie in rows $1, \dots, k$ and in k distinct columns $p_1 < \cdots < p_k$. So (i) says $k = n$, and since A has n rows, (ii) also says $k = n$. Hence (i) $\Leftrightarrow$ (ii).

(i) $\Leftrightarrow$ (iii). If $k = n$, then $p_s = s$ for every s (as in the first proof of Proposition 3.4.15), and column s of A_r is $\mathbf{e}_s$ by Proposition 3.3.8(c), so $A_r = I_n$. Conversely, the pivots of I_n are its diagonal entries (Example 3.3.13), one in each column.

(i) $\Leftrightarrow$ (iv) and (ii) $\Leftrightarrow$ (v) are Proposition 3.4.15 and Proposition 3.4.16.

(iv), (v) $\Leftrightarrow$ (vi). If (vi) holds, then (v) holds, and the unique solution of $A\mathbf{x} = \mathbf{0}$ is the trivial one, which is (iv). Conversely, suppose (iv) and (v) hold. By (iv) and Proposition 3.1.8(b), $\ker T_A = \{\mathbf{0}\}$. Let $\mathbf{y} \in F^n$. By (v), $A\mathbf{x} = \mathbf{y}$ is consistent, and by Corollary 3.1.22(a) it has exactly one solution. Since (iv) $\Leftrightarrow$ (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (v), each of (iv) and (v) implies the other, so each is equivalent to (vi).

(iii) $\Leftrightarrow$ (vii) is Theorem 3.3.23. □

The equivalence of (iv) and (v) is the remarkable one. For a square system, knowing that the homogeneous system has only the trivial solution, a statement about uniqueness, gives a solution of $A\mathbf{x} = \mathbf{y}$ for every $\mathbf{y}$, a statement about existence, and conversely. For non-square matrices both implications fail: $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ has a pivot in every row but not in every column, and the 2×1 matrix with column $(1, 0)$ the reverse. The corollary also completes Exercise 3.1.47. That exercise showed that the equilibrium temperatures of a plate are unique, and its system has as many equations as unknowns, so they also exist, for every choice of boundary temperatures.

This is a proof strategy in its own right.

> **Technique 3.4.21** (Uniqueness gives existence). To show that a linear problem has a solution for *every* choice of data, it can be much easier to show that it has *at most one*. This works whenever the problem is a square linear system, with as many conditions as unknowns: by Corollary 3.4.20, if the homogeneous problem has only the zero solution, then every instance of the problem has exactly one solution. Uniqueness is usually the easier half because it is a statement about the homogeneous problem, where all data vanish, and it amounts to showing that some object is zero, often by a structural argument (a polynomial with too many roots, a maximum principle) rather than by computation. Existence then comes for free, without ever constructing a solution.
>
> The setup has two requirements: the conditions are linear in the unknowns, and the number of conditions equals the number of unknowns. With those in place, the one thing left to prove is uniqueness itself, namely that the homogeneous problem has only the zero solution. Counting conditions and unknowns alone proves nothing (Exercise 3.6.14). Reach for it when existence looks hard and uniqueness looks easy, typically for interpolation-type problems (find an object of a given size with prescribed values) and for discrete boundary value problems such as Exercise 3.1.47. The canonical example is polynomial interpolation (Theorem 3.4.24). Section 4.2 extends it from square matrices to linear maps $V \to W$ with $\dim V = \dim W$.
>
> *Used in.* Theorem 3.4.24, Exercise 3.4.42, Exercise 3.6.13, Exercise 3.6.14, Problem 3.13, Problem 3.15, Problem 3.16, Problem 3.19, Exercise 4.2.47, Theorem 7.1.14, Exercise 7.1.57, Exercise 3.5.39.

3.4.5 Polynomial interpolation

Two points in the plane with different first coordinates lie on exactly one line, the graph of a polynomial of degree at most 1. Three points with distinct first coordinates lie on exactly one graph of a polynomial of degree at most 2, as Exercise 3.1.37 found in one example. In general, through $n + 1$ points with distinct first coordinates there passes exactly one graph of a polynomial of degree at most n, and this holds over every field. The proof is an application of everything in this section.

Throughout, $c_0, c_1, \ldots, c_n \in F$ are distinct (the **nodes**) and $d_0, d_1, \ldots, d_n \in F$ are arbitrary (the **values**). We look for $p \in F[x]_{\leq n}$ with

$$p(c_0) = d_0, \quad p(c_1) = d_1, \quad \ldots, \quad p(c_n) = d_n. \tag{3.4.3}$$

By Technique 3.1.9, write $p = a_0 + a_1 x + \cdots + a_n x^n$ with unknown coefficients. By Definition 1.2.17, the condition $p(c_i) = d_i$ reads

$$a_0 + a_1 c_i + a_2 c_i^2 + \cdots + a_n c_i^n = d_i,$$

a linear equation in $a_0, \ldots, a_n$. So (3.4.3) is a linear system of $n+1$ equations in $n+1$ unknowns, with the following coefficient matrix.

Definition 3.4.22 (Vandermonde matrix). For $c_0, \ldots, c_n \in F$, the **Vandermonde matrix** $V(c_0, \ldots, c_n) \in F^{(n+1)\times(n+1)}$ is the matrix whose row $i+1$ is $\left(1, c_i, c_i^2, \ldots, c_i^n\right)$, for $i = 0, \ldots, n$:

$$V(c_0, \ldots, c_n) = \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 1 & c_1 & c_1^2 & \cdots & c_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{pmatrix}.$$

With $\mathbf{a} = (a_0, \ldots, a_n)$ and $\mathbf{d} = (d_0, \ldots, d_n)$ in F^{n+1}, the conditions (3.4.3) say exactly that $V\mathbf{a} = \mathbf{d}$, where $V = V(c_0, \ldots, c_n)$. The system is square, so by Technique 3.4.21 it is enough to show that $V\mathbf{a} = \mathbf{0}$ forces $\mathbf{a} = \mathbf{0}$, that is, that a polynomial of degree at most n that vanishes at $n+1$ distinct points is the zero polynomial.

Intuition (Splitting off a root). The familiar reason is that each root c of p contributes a factor $x - c$, and a polynomial of degree at most n has room for at most n such factors. We cannot multiply polynomials until Section 7.1, but we can split off the factor by hand, because of an identity in the field:

$$t^j - c^j = (t - c)\left(t^{j-1} + t^{j-2}c + \cdots + c^{j-1}\right).$$

Summing these over the terms of p shows that $p(t) - p(c)$ equals $t - c$ times the value at t of an explicit polynomial q of degree at most $n-1$. If p vanishes at $c_0, \ldots, c_n$, then q vanishes at $c_1, \ldots, c_n$, and induction on n takes over.

Lemma 3.4.23 (Too many roots). Let $n \geq 0$ and $p \in F[x]_{\leq n}$. If $p(c_0) = p(c_1) = \cdots = p(c_n) = 0$ for distinct $c_0, \ldots, c_n \in F$, then $p = \mathbf{0}$.

Proof. STANDARD

Uses. Definition 1.2.16 (Degree), Definition 1.2.17 (Evaluation), Proposition 1.1.17 (No zero divisors), induction on n

Hint 1. Induct on n. For the inductive step, look for a polynomial q of degree at most $n-1$ with $p(t) - p(c_0) = (t - c_0)q(t)$ for every $t \in F$.

Hint 2. Use the telescoping identity for $t^j - c^j$ from the intuition above, and collect the coefficient of t^i. It is $b_i = \sum_{j>i} a_j c^{j-1-i}$.

Hint 3. q vanishes at $c_1, \ldots, c_n$, so $q = \mathbf{0}$ by induction. Recover $a_n, a_{n-1}, \ldots, a_1$ from $b_{n-1} = a_n$ and $b_i = a_{i+1} + c_0 b_{i+1}$.

Sketch. Write $p = \sum_j a_j x^j$ and $c = c_0$. Telescoping gives $p(t) - p(c) = (t - c)q(t)$ for all t, where $q = \sum_{i<n} b_i x^i$ and $b_i = \sum_{j>i} a_j c^{j-1-i}$. For $i \geq 1$, $0 = (c_i - c)q(c_i)$ with $c_i - c \neq 0$, so q vanishes at the n distinct points $c_1, \ldots, c_n$, and $q = \mathbf{0}$ by induction. Since $b_{n-1} = a_n$ and $b_i = a_{i+1} + c b_{i+1}$, all of $a_1, \ldots, a_n$ vanish, and then $a_0 = p(c_0) = 0$.

We induct on n. If $n = 0$, then $p = a_0$ is constant, and $a_0 = p(c_0) = 0$.

Let $n \geq 1$, and assume the lemma for $n-1$. Write $p = a_0 + a_1 x + \cdots + a_n x^n$, and put $c = c_0$. For $j \geq 1$ and $t \in F$, the following sum telescopes:

$$(t - c)\sum_{i=0}^{j-1} t^i c^{j-1-i} = \sum_{i=0}^{j-1} t^{i+1} c^{j-1-i} - \sum_{i=0}^{j-1} t^i c^{j-i} = t^j - c^j.$$

Multiplying by a_j, summing over $j = 1, \dots, n$, and collecting the terms with the same power t^i (for a fixed i, the index j runs over $i+1, \dots, n$), we get

$$p(t) - p(c) = \sum_{j=1}^{n} a_j\left(t^j - c^j\right) = (t-c)\sum_{i=0}^{n-1} b_i t^i, \quad \text{where } b_i = \sum_{j=i+1}^{n} a_j c^{j-1-i}.$$

Let

$$q = b_0 + b_1 x + \cdots + b_{n-1}x^{n-1} \in F[x]_{\le n-1}.$$

We have shown that

$$p(t) - p(c_0) = (t - c_0)q(t) \quad \text{for every } t \in F. \tag{3.4.4}$$

For $i = 1, \dots, n$, since $p(c_i) = p(c_0) = 0$, (3.4.4) gives $(c_i - c_0)q(c_i) = 0$. Here $c_i - c_0 \neq 0$ because the nodes are distinct, so $q(c_i) = 0$ by Proposition 1.1.17. Thus q has degree at most $n-1$ and vanishes at the n distinct points $c_1, \dots, c_n$, and by the induction hypothesis $q = \mathbf{0}$, that is, $b_0 = b_1 = \cdots = b_{n-1} = 0$.

Now $b_{n-1} = a_n$, and for $0 \le i \le n-2$, splitting off the term $j = i+1$,

$$b_i = a_{i+1} + c\sum_{j=i+2}^{n} a_j c^{j-2-i} = a_{i+1} + c b_{i+1}.$$

Hence $a_n = b_{n-1} = 0$, and $a_{i+1} = b_i - c b_{i+1} = 0$ for $i = n-2, \dots, 0$. So $a_1 = \cdots = a_n = 0$, and finally $a_0 = p(c_0) = 0$. Thus $p = \mathbf{0}$. □

The recursion $b_i = a_{i+1} + c b_{i+1}$, started from $b_{n-1} = a_n$, is Horner's scheme (or synthetic division) for dividing p by $x - c$. Once polynomials can be multiplied, (3.4.4) becomes the factorization $p - p(c) = (x-c)q$, and the same induction shows that a nonzero polynomial has at most $\deg p$ roots (Proposition 7.1.22).

Theorem 3.4.24 (Polynomial interpolation). Let $c_0, \dots, c_n \in F$ be distinct, and let $d_0, \dots, d_n \in F$. There is exactly one polynomial $p \in F[x]_{\le n}$ with $p(c_i) = d_i$ for $i = 0, 1, \dots, n$. In particular, the Vandermonde matrix $V(c_0, \dots, c_n)$ is invertible.

Proof. STANDARD

Uses. Definition 3.4.22 (Vandermonde matrix), Definition 1.2.17 (Evaluation), Definition 1.2.16 (Degree), Definition 2.3.8 (Matrix–vector product), Lemma 3.4.23 (Too many roots), Corollary 3.4.20 (Square systems)
Techniques. Undetermined coefficients (Technique 3.1.9), Uniqueness gives existence (Technique 3.4.21)
Hint 1. Write $p = a_0 + a_1 x + \cdots + a_n x^n$, and translate the conditions $p(c_i) = d_i$ into a linear system for the a_j. What is its coefficient matrix?
Hint 2. The system is square. By Corollary 3.4.20 it is enough to show that the homogeneous system has only the trivial solution.
Hint 3. A solution of the homogeneous system gives a polynomial of degree at most n that vanishes at $n+1$ distinct points.
Sketch. For $\mathbf{a} \in F^{n+1}$ let $p_{\mathbf{a}} = \sum_j a_j x^j$. The entry of $V\mathbf{a}$ in the row of c_i is $p_{\mathbf{a}}(c_i)$, so the interpolating polynomials correspond to the solutions of $V\mathbf{a} = \mathbf{d}$. If $V\mathbf{a} = \mathbf{0}$, then $p_{\mathbf{a}}$ vanishes at $c_0, \dots, c_n$, so $p_{\mathbf{a}} = \mathbf{0}$ by Lemma 3.4.23 and $\mathbf{a} = \mathbf{0}$. Since V is square, Corollary 3.4.20 gives exactly one solution of $V\mathbf{a} = \mathbf{d}$ for each $\mathbf{d}$, and invertibility of V.

Let $V = V(c_0, \dots, c_n)$ and $\mathbf{d} = (d_0, \dots, d_n)$. For $\mathbf{a} = (a_0, \dots, a_n) \in F^{n+1}$, write $p_{\mathbf{a}} = a_0 + a_1 x + \cdots + a_n x^n$. By Definition 1.2.16, every $p \in F[x]_{\le n}$ equals $p_{\mathbf{a}}$ for exactly one $\mathbf{a}$, namely its list of coefficients $a_0, \dots, a_n$. By Definition 2.3.8 and Definition 1.2.17, the entry of $V\mathbf{a}$ in the row of c_i is

$$a_0 + a_1 c_i + \cdots + a_n c_i^n = p_{\mathbf{a}}(c_i).$$

So $p_{\mathbf{a}}$ satisfies $p_{\mathbf{a}}(c_i) = d_i$ for all i if and only if $V\mathbf{a} = \mathbf{d}$. Through $\mathbf{a} \mapsto p_{\mathbf{a}}$, the polynomials we want correspond exactly to the solutions of $V\mathbf{a} = \mathbf{d}$.

Suppose $V\mathbf{a} = \mathbf{0}$. Then $p_{\mathbf{a}}$ has degree at most n and vanishes at the $n+1$ distinct points $c_0, \ldots, c_n$, so $p_{\mathbf{a}} = \mathbf{0}$ by Lemma 3.4.23, that is, $\mathbf{a} = \mathbf{0}$. So the homogeneous system $V\mathbf{a} = \mathbf{0}$ has only the trivial solution. Since V is square, Corollary 3.4.20 ((iv) $\Rightarrow$ (vi) and (vii)) shows that $V\mathbf{a} = \mathbf{d}$ has exactly one solution, and that V is invertible. Hence there is exactly one polynomial $p \in F[x]_{\leq n}$ with $p(c_i) = d_i$ for all i. □

The proof never solved anything. Existence came from uniqueness: $n+1$ conditions on $n+1$ unknown coefficients form a square system, and the homogeneous version asks for a polynomial with too many roots. Both hypotheses are needed. If two nodes coincide, say $c_0 = c_1$, then V has two equal rows, and for $d_0 \neq d_1$ there is no solution at all. And the degree bound n is exactly right. With degree at most $n-1$ there are only n unknowns for $n+1$ conditions, so by Corollary 3.4.19(b) some values cannot be interpolated. With degree at most $n+1$ there are $n+2$ unknowns, so by Corollary 3.4.19(a) the homogeneous system has a nontrivial solution and uniqueness fails (Figure 3.4.1; see also Exercise 3.4.39).

Example 3.4.25 (A cubic through four points). Find the polynomial $p \in \mathbb{Q}[x]_{\leq 3}$ with

$$p(-1) = 2, \quad p(0) = 1, \quad p(1) = 0, \quad p(2) = 5.$$

With $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, the system $V\mathbf{a} = \mathbf{d}$ is

$$\left(\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 8 & 5 \end{array}\right).$$

The second equation gives $a_0 = 1$, and the others become $-a_1 + a_2 - a_3 = 1$, $a_1 + a_2 + a_3 = -1$, and $2a_1 + 4a_2 + 8a_3 = 4$. Adding the first two gives $a_2 = 0$. Then $a_1 + a_3 = -1$ and $a_1 + 4a_3 = 2$, so $a_3 = 1$ and $a_1 = -2$. Hence

$$p = 1 - 2x + x^3.$$

Check: $p(-1) = 1 + 2 - 1 = 2$, $p(0) = 1$, $p(1) = 1 - 2 + 1 = 0$, and $p(2) = 1 - 4 + 8 = 5$.

$$\left(\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 8 & 5 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

The system $V\mathbf{a} = \mathbf{d}$ of Example 3.4.25. The left block reduces to I_4, and the last column becomes the coefficient vector $(1, -2, 0, 1)$. Edit the last column to interpolate other values at the same nodes. (Interactive in the web version.)

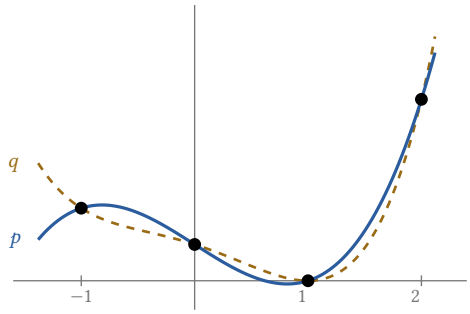


Figure 3.4.1. The cubic $p = 1 - 2x + x^3$ of Example 3.4.25 (solid) is the only polynomial of degree at most 3 through the four points. The quartic $q = 1 - x - x^2/2 + x^4/2$ (dashed) passes through them too: $q - p = (x^4 - 2x^3 - x^2 + 2x)/2$ vanishes at $-1, 0, 1, 2$. With degree at most 4 there are five unknown coefficients and only four conditions.

Lagrange's formula

The theorem gives the interpolating polynomial for each list of values separately. Linearity lets us do all of them at once: it is enough to interpolate the simplest values, a single 1 and zeros elsewhere.

Definition 3.4.26 (Lagrange polynomials). Let $c_0, \dots, c_n \in F$ be distinct. For $i = 0, \dots, n$, the ith **Lagrange polynomial** $L_i \in F[x]_{\le n}$ is the unique polynomial of degree at most n with

$$L_i(c_i) = 1 \quad \text{and} \quad L_i(c_k) = 0 \text{ for } k \neq i.$$

It exists and is unique by Theorem 3.4.24.

Proposition 3.4.27 (Lagrange's formula). Let $c_0, \dots, c_n \in F$ be distinct, and let $d_0, \dots, d_n \in F$. The unique $p \in F[x]_{\le n}$ with $p(c_i) = d_i$ for all i is

$$p = d_0 L_0 + d_1 L_1 + \cdots + d_n L_n.$$

Proof. ROUTINE

Uses. Definition 3.4.26 (Lagrange polynomials), Proposition 1.2.18 (Degree and evaluation versus the operations), Theorem 3.4.24 (Polynomial interpolation)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Evaluate the right side at c_k, using that evaluation is linear (Proposition 1.2.18(c)).
Hint 2. Conclude with the uniqueness in Theorem 3.4.24.
Sketch. The right side q has degree at most n, and $q(c_k) = \sum_i d_i L_i(c_k) = d_k$ by linearity of evaluation. By uniqueness in Theorem 3.4.24, $q = p$.

Let $q = d_0 L_0 + \cdots + d_n L_n$. By Proposition 1.2.18(a) and (b), sums and nonzero scalar multiples of polynomials of degree at most n have degree at most n (and $0 L_i = \mathbf{0}$ trivially does), so $q \in F[x]_{\le n}$. By Proposition 1.2.18(c), for each k,

$$q(c_k) = \sum_{i=0}^{n} d_i L_i(c_k) = d_k \cdot 1 + \sum_{i \neq k} d_i \cdot 0 = d_k.$$

So q satisfies the same conditions as p, and $q = p$ by the uniqueness in Theorem 3.4.24. □

Example 3.4.28 (Lagrange polynomials for the nodes $-1, 0, 1, 2$). For the nodes of Example 3.4.25, the Lagrange polynomials are

$$L_0 = -\frac{1}{3}x + \frac{1}{2}x^2 - \frac{1}{6}x^3, \quad L_1 = 1 - \frac{1}{2}x - x^2 + \frac{1}{2}x^3,$$

$$L_2 = x + \frac{1}{2}x^2 - \frac{1}{2}x^3, \qquad L_3 = -\frac{1}{6}x + \frac{1}{6}x^3.$$

Each claim is checked by evaluating at the four nodes; for instance $L_3(2) = -\frac{1}{3} + \frac{8}{6} = 1$ and $L_3(-1) = \frac{1}{6} - \frac{1}{6} = 0$. By Proposition 3.4.27, the cubic of Example 3.4.25 is

$$2L_0 + L_1 + 0L_2 + 5L_3 = 1 + \left(-\frac{2}{3} - \frac{1}{2} - \frac{5}{6}\right)x + (1-1)x^2 + \left(-\frac{1}{3} + \frac{1}{2} + \frac{5}{6}\right)x^3 = 1 - 2x + x^3,$$

as before. Once polynomials can be multiplied (Section 7.1), each L_i has the closed form

$$L_i = \prod_{j \neq i} \frac{x - c_j}{c_i - c_j},$$

because the product has degree n and takes the right values, so it is L_i by uniqueness. For instance

$$L_3 = \frac{(x+1)x(x-1)}{3 \cdot 2 \cdot 1} = \frac{x^3 - x}{6}.$$

Section 7.1 revisits interpolation from this side.

The Lagrange polynomials also describe V^{-1}. The coefficient vector $\mathbf{l}_i$ of L_i solves $V\mathbf{a} = \mathbf{e}_{i+1}$, since L_i takes the value 1 at c_i and 0 at the other nodes. So the columns of V^{-1} are the coefficient vectors of $L_0, \ldots, L_n$, and Lagrange's formula is the equation $V^{-1}\mathbf{d} = d_0\mathbf{l}_0 + \cdots + d_n\mathbf{l}_n$. For the nodes $-1, 0, 1, 2$, the first column of V^{-1} is $(0, -1/3, 1/2, -1/6)$, the coefficients of L_0.

Finally, V is a square matrix, so its invertibility can also be decided by a determinant. Section 5.5 computes

$$\det V(c_0, \ldots, c_n) = \prod_{i<j} (c_j - c_i)$$

(Theorem 5.5.21), which is nonzero exactly when the nodes are distinct; by Theorem 5.4.7 this gives a second proof that V is invertible.

Exercise 3.4.29. ROUTINE

(a) Over $\mathbb{Q}$, find the Lagrange polynomials for the nodes $0, 1, 3$, and use them to find the $p \in \mathbb{Q}[x]_{\leq 2}$ with $p(0) = 2$, $p(1) = -1$, and $p(3) = 5$.

(b) Find the $p \in \mathbb{F}_5[x]_{\leq 2}$ with the same three values, read in $\mathbb{F}_5$.

Solution on page 1240

Exercises

Exercise 3.4.30. ROUTINE Let

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 & 3 \\ 2 & 4 & 1 & 4 & 7 \\ -1 & -2 & 1 & 1 & -2 \end{pmatrix} \in \mathbb{Q}^{3\times 5}.$$

(a) Find all $\mathbf{y} \in \mathbb{Q}^3$ for which $A\mathbf{x} = \mathbf{y}$ is consistent.

(b) For $\mathbf{y} = (1, 3, 0)$, write the solution set in the form $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \ldots, \mathbf{s}_\ell)$ of Theorem 3.4.9.

Solution on page 1241

Exercise 3.4.31. ROUTINE Let A be the matrix of Exercise 3.3.40,

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

Solve $A\mathbf{x} = (1, 0, 1)$ over $\mathbb{Q}$ and over $\mathbb{F}_2$. Over $\mathbb{F}_2$, list all solutions, and compare their number with Corollary 3.4.14.

Solution on page 1241

Exercise 3.4.32. ROUTINE Over $\mathbb{C}$, consider

$$x_1 + ix_2 = 1, \quad ix_1 - x_2 + x_3 = 2 + i, \quad x_3 = 2.$$

(a) Find all solutions in $\mathbb{C}^3$.

(b) For which $\mathbf{y} \in \mathbb{C}^3$ is the system with these left sides and right side $\mathbf{y}$ consistent?

(c) Which of the solutions in (a) lie in $\mathbb{R}^3$?

Solution on page 1242

Exercise 3.4.33. STANDARD For a, b in a field F, consider the system

$$x_1 + x_2 + ax_3 = 1, \quad x_1 + ax_2 + x_3 = 1, \quad ax_1 + x_2 + x_3 = b.$$

(a) For $F = \mathbb{R}$, find all (a, b) for which the system has no solution, exactly one solution, and infinitely many solutions.

(b) Answer the same question for $F = \mathbb{F}_3$, and count the solutions in each case.

Solution on page 1242

Exercise 3.4.34. STANDARD Let A be the matrix of Example 3.4.6, and $\mathbf{y} = (4, 2, 5)$. Two students describe the solution set of $A\mathbf{x} = \mathbf{y}$.

Student 1. "The solution set is $(-1, 0, 2, 0) + \operatorname{span}((0, 1, -2, 1))$. Proof: $A(-1, 0, 2, 0) = (4, 2, 5)$ and $A(0, 1, -2, 1) = \mathbf{0}$, so every vector of this form is a solution, by Proposition 3.1.19."

Student 2. The solution set is

$$(1, -1, 2, 0) + \operatorname{span}((0, 1, -2, 1), (4, -1, -2, 1)).$$

(a) Find the error in Student 1's argument, and a solution that Student 1 has missed.

(b) Decide whether Student 2 is right.

Solution on page 1243

Exercise 3.4.35. STANDARD Let $A \in F^{m \times n}$. Prove or disprove each statement.

(a) If $A\mathbf{x} = \mathbf{y}$ has exactly one solution for some $\mathbf{y} \in F^m$, then A_r has a pivot in every column.

(b) If $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then $A\mathbf{x} = \mathbf{y}$ is consistent for every $\mathbf{y}$.

(c) If A_r has a pivot in every column, and B is obtained from A by deleting some of its columns, then B_r has a pivot in every column.

(d) If A_r has a pivot in every row, and B is obtained from A by deleting some of its columns, then B_r has a pivot in every row.

Solution on page 1244

Exercise 3.4.36. STANDARD Let F be a finite field with q elements, and let $A \in F^{m\times n}$ have k pivots.

(a) Show that $A\mathbf{x} = \mathbf{y}$ is consistent for exactly q^k vectors $\mathbf{y} \in F^m$.

(b) Each consistent system $A\mathbf{x} = \mathbf{y}$ has q^{n-k} solutions. Explain why $q^k \cdot q^{n-k} = q^n$ is no coincidence.

Solution on page 1244

Exercise 3.4.37. STANDARD Let $A \in F^{m\times n}$ and $1 \le j \le n$. Show that x_j is a free variable of A if and only if there is a solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{0}$ with $x_j = 1$ and $x_l = 0$ for every $l > j$.

Solution on page 1244

Exercise 3.4.38. STANDARD Let F be a subfield of a field K (for example $\mathbb{Q} \subseteq \mathbb{R}$ or $\mathbb{R} \subseteq \mathbb{C}$). Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$.

(a) Show that if $A\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x} \in K^n$, then it has a solution in F^n.

(b) Show that if $A\mathbf{x} = \mathbf{y}$ has exactly one solution in K^n, then that solution lies in F^n.

(c) Show that if $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution in K^n, then it has one in F^n.

Solution on page 1245

Exercise 3.4.39. STANDARD Let $m \ge 1$, let $c_1, \ldots, c_m \in F$ be distinct, let $d \ge 0$, and let $W \in F^{m\times(d+1)}$ be the matrix with rows $\left(1, c_i, c_i^2, \ldots, c_i^d\right)$, $i = 1, \ldots, m$. Show that:

(a) W_r has a pivot in every column if and only if $m \ge d + 1$;

(b) W_r has a pivot in every row if and only if $m \le d + 1$.

Interpret both statements as statements about polynomials of degree at most d with prescribed values at $c_1, \ldots, c_m$.

Solution on page 1245

Exercise 3.4.40. STANDARD Let p be a prime.

(a) Show that every function $g : \mathbb{F}_p \to \mathbb{F}_p$ is the polynomial function of exactly one polynomial $f \in \mathbb{F}_p[x]_{\le p-1}$. Check this by counting both sides.

(b) Find that polynomial for the function g with $g(0) = 1$ and $g(c) = 0$ for $c \ne 0$.

Solution on page 1246

Exercise 3.4.41. STANDARD For each prime p, find the number of $(a_0, a_1, a_2) \in \mathbb{F}_p^3$ with

$$a_0 + a_1 + a_2 = 1, \quad a_0 + 2a_1 + 4a_2 = 2, \quad a_0 + 3a_1 + 9a_2 = 4.$$

Solution on page 1246

Exercise 3.4.42. STANDARD For $p = a_0 + a_1x + a_2x^2 + a_3x^3 \in \mathbb{R}[x]$, write $p' = a_1 + 2a_2x + 3a_3x^2$.

(a) Show that for all $u_0, v_0, u_1, v_1 \in \mathbb{R}$ there is exactly one $p \in \mathbb{R}[x]_{\le 3}$ with

$$p(0) = u_0, \quad p'(0) = v_0, \quad p(1) = u_1, \quad p'(1) = v_1.$$

(b) Find the four polynomials h_1, h_2, h_3, h_4 obtained when one of u_0, v_0, u_1, v_1 is 1 and the others are 0, and write the general p in terms of them.

Solution on page 1246

Exercise 3.4.43. STANDARD

(a) Find all $w_0, w_1, w_2 \in \mathbb{R}$ such that

$$\int_{-1}^{1} p(t)\,\mathrm{d}t = w_0 p(-1) + w_1 p(0) + w_2 p(1)$$

for every $p \in \mathbb{R}[x]_{\leq 2}$.

(b) Show that the resulting rule is exact for every $p \in \mathbb{R}[x]_{\leq 3}$ as well, but not for x^4.

Solution on page 1247

Exercise 3.4.44. HARD Let $A, B \in F^{m\times n}$ and $\mathbf{y}, \mathbf{w} \in F^m$.

(a) Show that if the systems $A\mathbf{x} = \mathbf{y}$ and $B\mathbf{x} = \mathbf{w}$ have the same solution set, and it is nonempty, then the augmented matrices $(A \mid \mathbf{y})$ and $(B \mid \mathbf{w})$ are row equivalent.

(b) Show by an example that (a) fails without the word "nonempty".

Solution on page 1247

Exercise 3.4.45. HARD Let $\mathbf{x}_0, \mathbf{u}_1, \ldots, \mathbf{u}_r \in F^n$.

(a) Describe a procedure that produces a linear system whose solution set is exactly $\mathbf{x}_0 + \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$, and prove that it works.

(b) Carry it out over $\mathbb{R}$ for $\mathbf{x}_0 = (1, 0, 2, 0)$, $\mathbf{u}_1 = (1, 1, 0, 0)$, and $\mathbf{u}_2 = (0, 0, 1, 1)$.

Solution on page 1248

Exercise 3.4.46. HARD **Secret sharing.** A secret $s \in \mathbb{F}_p$ is to be shared among N people, where $N < p$, so that any $t + 1$ of them together can recover it, but any t of them together learn nothing about it ($0 \leq t < N$). The dealer chooses $a_1, \ldots, a_t \in \mathbb{F}_p$ at random, forms $f = s + a_1 x + \cdots + a_t x^t$, and gives person i the **share** $f(i)$, for $i = 1, \ldots, N$.

(a) Show that any $t + 1$ people can compute s from their shares.

(b) Let $i_1, \ldots, i_t$ be t of the people, with shares $v_1, \ldots, v_t$. Show that for every $s' \in \mathbb{F}_p$ there is exactly one choice of $(a_1, \ldots, a_t) \in \mathbb{F}_p^t$ for which a dealer with secret s' would have produced exactly these shares. Explain why this means that the t people learn nothing about s.

Solution on page 1249

Exercise 3.4.47. CHALLENGE For $n \geq 0$, let $p_n \in \mathbb{Q}[x]_{\leq n}$ be the polynomial with $p_n(k) = 2^k$ for $k = 0, 1, \ldots, n$. Show that $p_n(n + 1) = 2^{n+1} - 1$.

Solution on page 1249

3.5 Rank and the Fundamental Theorems

The two propositions of Section 3.4.4 located the answers to the two basic questions about $A\mathbf{x} = \mathbf{y}$ in the pivots of $A_r = \operatorname{rref}(A)$. A pivot in every column means that solutions are unique (Proposition 3.4.15), and a pivot in every row means that there is a solution for every $\mathbf{y}$ (Proposition 3.4.16). Each row and each column holds at most one pivot, so both conditions are statements about a single number, the number of pivots. This number is the rank of A. With it we complete the Fundamental Theorems of Section 2.7: the rank supplies the sixth condition in each of them, and the inequality $\operatorname{rank} A \leq \min(m, n)$ supplies their constraints on the shape of A. In particular, invertible matrices are square. Along the way we construct

one-sided inverses directly from the pivots, show that every invertible matrix is a product of elementary matrices, and show that a one-sided inverse of a square matrix is automatically two-sided. Finally, allowing column operations as well as row operations, we bring every matrix to a normal form that depends only on its rank.

3.5.1 Rank

Definition 3.5.1 (Rank). Let $A \in F^{m\times n}$. The **rank** of A, written $\operatorname{rank} A$ or $\operatorname{rank}(A)$, is the number of pivots of A, that is, the number of pivots of $\operatorname{rref}(A)$ (Definition 3.3.9).

This is the number k of Section 3.4. Since $\operatorname{rref}(A)$ is determined by A, however the reduction is carried out (Corollary 3.3.18), so is the rank. Everything we know about pivots translates into statements about it.

Proposition 3.5.2 (First properties of rank). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$.

(a) $\operatorname{rank} A \le \min(m, n)$. Moreover, $\operatorname{rank} A = m$ if and only if A_r has a pivot in every row, and $\operatorname{rank} A = n$ if and only if A_r has a pivot in every column.

(b) A has exactly $n - \operatorname{rank} A$ free columns, and A_r has exactly $m - \operatorname{rank} A$ zero rows.

(c) $\operatorname{rank}(GA) = \operatorname{rank} A$ for every invertible $G \in F^{m\times m}$. In particular, row equivalent matrices have the same rank.

(d) $\operatorname{rank} A$ is the number of indices $j \in \{1, \dots, n\}$ with $\mathbf{a}_j \notin \operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_{j-1})$.

Proof. ROUTINE

Uses. Proposition 3.3.8 (Properties of the output), Definition 3.4.2 (Pivot and free variables), Corollary 3.3.18 (RREF and row equivalence), Definition 3.2.21 (Row equivalence), Corollary 3.2.11 (Products of elementary matrices), Corollary 3.3.19 (Pivot columns, intrinsically), Definition 1.3.10 (Span)
Hint 1. Where does Proposition 3.3.8 put the pivots? Count the rows and the columns they occupy.
Hint 2. For (c), compare $\operatorname{rref}(GA)$ with $\operatorname{rref}(A)$. For (d), which result describes the pivot columns without running the algorithm?
Sketch. The r pivots sit in positions $(1, p_1), \dots, (r, p_r)$ with $p_1 < \cdots < p_r$: one in each of the first r rows and one in each of r distinct columns. This gives (a) and (b). Part (c) is $\operatorname{rref}(GA) = \operatorname{rref}(A)$, and (d) is Corollary 3.3.19(a), which says that the pivot columns are exactly the columns that are not combinations of earlier ones.

Let $r = \operatorname{rank} A$ and let $p_1 < \cdots < p_r$ be the pivot columns. By Proposition 3.3.8, the pivots are the entries in positions $(1, p_1), \dots, (r, p_r)$ of A_r, rows $1, \dots, r$ of A_r are nonzero, and rows $r+1, \dots, m$ are zero.

(a) The pivots lie in r distinct rows and r distinct columns, so $r \le m$ and $r \le n$. Each row contains at most one pivot, so every one of the m rows contains a pivot if and only if $r = m$. The same argument applies to the n columns.

(b) By Definition 3.4.2, the columns other than $p_1, \dots, p_r$ are the free columns, and there are $n - r$ of them. The zero rows of A_r are rows $r+1, \dots, m$.

(c) By Corollary 3.3.18(a), $\operatorname{rref}(GA) = \operatorname{rref}(A)$, so GA and A have the same pivots. If $A \sim B$, then $B = E_k \cdots E_1 A$ for elementary matrices E_i (Definition 3.2.21), and the product $E_k \cdots E_1$ is invertible (Corollary 3.2.11).

(d) By Corollary 3.3.19(a), column j is a pivot column if and only if $\mathbf{a}_j$ is not a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_{j-1}$, that is, if and only if $\mathbf{a}_j \notin \operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_{j-1})$ (Definition 1.3.10; for $j = 1$ the span of the empty list is $\{\mathbf{0}\}$). So the indices in (d) are exactly the pivot columns, and there are r of them. □

Part (d) describes the rank without reference to the algorithm. Scan the columns of A from left to right and count those that are not combinations of the columns before them; the count is $\operatorname{rank} A$. For instance, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$ has rank 1: its first column is nonzero, and the other two are multiples of it.

Part (c) is about row operations. Column operations also preserve the rank, $\operatorname{rank}(AH) = \operatorname{rank} A$ for invertible H, but row reduction gives no direct handle on this, because $\operatorname{rref}(AH)$ and $\operatorname{rref}(A)$ are in general different matrices. We prove it in Exercise 3.5.41, and Section 4.3 gives two more proofs.

Example 3.5.3 (A rank that depends on the field). Let

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 2 & 4 & 1 & 3 & 7 \\ 2 & 4 & 2 & 0 & 6 \\ 0 & 0 & 3 & 0 & 6 \end{pmatrix}.$$

Over $\mathbb{Q}$, pass 1 adds -2 times row 1 to rows 2 and 3. Pass 2 finds the pivot $(2,3)$ and adds -2 and -3 times row 2 to rows 3 and 4. Pass 3 scales row 3 by $-1/6$ and clears column 4:

$$A \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 3 & 0 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & -6 & -6 \\ 0 & 0 & 0 & -9 & -9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A_r.$$

So over $\mathbb{Q}$, $\operatorname{rank} A = 3 < \min(4,5)$, the pivot columns are $1, 3, 4$, and the free columns are $2, 5$. In terms of Proposition 3.5.2(d), $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_5 = \mathbf{a}_1 + 2\mathbf{a}_3 + \mathbf{a}_4$ are the columns that bring nothing new (Corollary 3.3.19(b)).

Now read the same entries in $\mathbb{F}_3$. The first two passes are the same computation with every entry reduced modulo 3, and the entries -6 and -9 of the second matrix are 0 in $\mathbb{F}_3$. So after pass 2 the last two rows are zero, the algorithm halts, and $\operatorname{rank} A = 2$ over $\mathbb{F}_3$. Over $\mathbb{F}_p$ with $p \geq 5$ the computation over $\mathbb{Q}$ goes through unchanged, since the only scalar we divided by is -6, which is nonzero in $\mathbb{F}_p$; so the rank is 3 there.

So the rank of a matrix with integer entries can depend on the field in which we read them. Enlarging the field never changes it: if F is a subfield of K and $A \in F^{m\times n}$, then $\operatorname{rref}(A)$ is the same over F and over K (Exercise 3.3.54), so a rational matrix has the same rank over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 2 & 4 & 1 & 3 & 7 \\ 2 & 4 & 2 & 0 & 6 \\ 0 & 0 & 3 & 0 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The matrix of Example 3.5.3 over $\mathbb{Q}$. At the end the summary reports the rank, the pivot columns, and the free columns. Use "Edit" to switch the field to $\mathbb{F}_3$ and watch the third pivot disappear, or to $\mathbb{F}_2$, where the rank is 3 again for different reasons. (Interactive in the web version.)

Exercise 3.5.4. ROUTINE Find the rank of each matrix.

(a) $\begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \\ 3 & -3 & 6 \end{pmatrix}$ over $\mathbb{Q}$.

(b) The Vandermonde matrix $V(1,2,3) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$ (Definition 3.4.22), over $\mathbb{Q}$ and over $\mathbb{F}_2$.

(c) The $n \times n$ matrix all of whose entries are 1, over any field.

Solution on page 1250

Exercise 3.5.5. STANDARD Let $n \geq 2$, and let $A \in F^{n\times n}$ be the matrix with entry $i + j$ in position (i, j) (that is, the element $(i + j) * 1$ of F). Show that $\operatorname{rank} A = 2$, whatever the field F.

Solution on page 1250

The rank also answers the consistency question for a single right-hand side. Adjoining $\mathbf{y}$ to A as an extra column can create at most one new pivot, and it does so exactly when the system has no solution.

Proposition 3.5.6 (Consistency and rank). Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$. Then

$$\operatorname{rank} A \leq \operatorname{rank}(A \mid \mathbf{y}) \leq \operatorname{rank} A + 1,$$

and $A\mathbf{x} = \mathbf{y}$ is consistent if and only if $\operatorname{rank}(A \mid \mathbf{y}) = \operatorname{rank} A$.

Proof. ROUTINE

Uses. Corollary 3.3.19 (Pivot columns, intrinsically), Corollary 3.4.5 (Consistency from the augmented matrix), Proposition 3.5.2 (First properties of rank)
Hint 1. Describe the pivot columns of $(A \mid \mathbf{y})$ with Corollary 3.3.19(a). Does whether column $j \leq n$ is a pivot column depend on the last column?
Hint 2. So the two matrices have the same pivot columns among $1, \ldots, n$. When is column $n + 1$ a pivot column?
Sketch. By Corollary 3.3.19(a), column j of a matrix is a pivot column exactly when it is not a combination of the earlier columns, a condition involving only columns $1, \ldots, j$. So for $j \leq n$, column j is a pivot column of $(A \mid \mathbf{y})$ if and only if it is one of A. Hence $\operatorname{rank}(A \mid \mathbf{y})$ is $\operatorname{rank} A + 1$ if the last column is a pivot column and $\operatorname{rank} A$ otherwise, and by Corollary 3.4.5 the last column is a pivot column exactly when the system is inconsistent.

The first n columns of $(A \mid \mathbf{y})$ are $\mathbf{a}_1, \ldots, \mathbf{a}_n$. By Corollary 3.3.19(a), for $j \leq n$, column j is a pivot column of $(A \mid \mathbf{y})$ if and only if $\mathbf{a}_j$ is not a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$, which is also the condition for column j to be a pivot column of A. So the two matrices have the same pivot columns among $1, \ldots, n$. Hence $\operatorname{rank}(A \mid \mathbf{y}) = \operatorname{rank} A$ if column $n + 1$ is not a pivot column of $(A \mid \mathbf{y})$, and $\operatorname{rank}(A \mid \mathbf{y}) = \operatorname{rank} A + 1$ if it is. By Corollary 3.4.5, column $n + 1$ is a pivot column if and only if $A\mathbf{x} = \mathbf{y}$ is inconsistent. □

In other words, $A\mathbf{x} = \mathbf{y}$ is consistent exactly when $\mathbf{y}$ brings nothing new beyond the columns of A. This criterion is often called the Rouché–Capelli theorem. It is convenient when ranks are already known, but in practice one row reduces $(A \mid \mathbf{y})$ and looks at the last column, which is what the proof does.

3.5.2 The Fundamental Theorems, completed

In Section 2.7 we stated three Fundamental Theorems for $T_A : F^n \to F^m$, each a list of equivalent conditions (I)–(V), and left open a sixth condition (the lecture marked it "secret") and the constraint each theorem puts on the shape of A. We can now fill in both.

Intuition (What the rank counts). Think of $A_r\mathbf{x} = G\mathbf{y}$ as the system $A\mathbf{x} = \mathbf{y}$ after elimination. Each of the $\operatorname{rank} A$ nonzero rows of A_r determines one pivot variable. The other $n - \operatorname{rank} A$ unknowns are free, and each free variable is an independent direction in which a solution can be moved without changing $A\mathbf{x}$; this is the failure of uniqueness. The remaining $m - \operatorname{rank} A$ rows of A_r are zero, and each produces a consistency condition $(G\mathbf{y})_i = 0$ on the right-hand side; this is the failure of existence. So $n - \operatorname{rank} A$ measures how far T_A is from injective, and $m - \operatorname{rank} A$ how far it is from surjective. Since $\operatorname{rank} A \leq$

$\min(m, n)$, a matrix with more columns than rows always has free variables, and a matrix with more rows than columns always has zero rows.

One-sided inverses from the pivots

Condition (V) of the first two theorems asks for a linear map: a left inverse of T_A, or a right inverse. That (V) implies the other conditions is quick (Proposition 2.5.4, Proposition 2.5.5). The converse requires producing the inverse, and row reduction produces it explicitly. First, a remark that lets us work with matrices instead of maps.

Remark 3.5.7 (Inverses of T_A and inverses of A). For $C \in F^{n\times m}$ we have $T_C \circ T_A = T_{CA}$ (Theorem 2.4.2(b)), and $T_{I_n} = \mathrm{id}_{F^n}$ (Proposition 2.3.13). Every linear map $S : F^m \to F^n$ is T_C for exactly one $C \in F^{n\times m}$ (Proposition 2.3.17). So $S = T_C$ is a left inverse of T_A if and only if $T_{CA} = T_{I_n}$, that is, if and only if $CA = I_n$, by the uniqueness in Proposition 2.3.17. Likewise T_C is a right inverse of T_A if and only if $AC = I_m$. Such a C is a left inverse, respectively a right inverse, of the matrix A (Definition 2.5.1, Definition 2.5.2), and we pass freely between the two readings.

Intuition (Choosing solutions linearly). A right inverse X of A is a rule that picks, for each $\mathbf{y}$, one solution $\mathbf{x} = X\mathbf{y}$ of $A\mathbf{x} = \mathbf{y}$, and does so linearly. When there are free variables, $A\mathbf{x} = \mathbf{y}$ has many solutions, and an arbitrary choice for each $\mathbf{y}$ need not be linear. But one choice is visibly linear: the particular solution, which sets every free variable to 0 and reads the pivot variables off $G\mathbf{y}$ (Definition 3.4.8). Writing this rule as a matrix gives the lecture's construction.

Proposition 3.5.8 (Right inverses from the pivots). Let $A \in F^{m\times n}$ with $\operatorname{rank} A = m$. Then A has a right inverse. Explicitly, let $p_1 < \cdots < p_m$ be the pivot columns of A, let $G \in F^{m\times m}$ be invertible with $GA = A_r$, and let $B \in F^{n\times m}$ be the matrix with entry 1 in position (p_i, i) for $i = 1, \ldots, m$ and 0 elsewhere. Then $A_rB = I_m$ and

$$A(BG) = I_m.$$

Moreover, for each $\mathbf{y} \in F^m$, the vector $BG\mathbf{y}$ is the particular solution of $A\mathbf{x} = \mathbf{y}$.

Proof. ROUTINE

Uses. Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Definition 3.4.8 (Basic solutions)
Hint 1. Compute A_rB one column at a time. Column i of B is a standard basis vector of F^n. Which one?
Hint 2. Column i of A_rB is $A_r\mathbf{e}_{p_i}$, which is column p_i of A_r. What is that column?
Hint 3. Write $A = G^{-1}A_r$ and compute ABG.
Sketch. Column i of B is $\mathbf{e}_{p_i}$, so column i of A_rB is column p_i of A_r, which is $\mathbf{e}_i$; hence $A_rB = I_m$. Then $ABG = G^{-1}A_rBG = G^{-1}G = I_m$. Finally, $B\mathbf{z}$ has z_i in position p_i and 0 in every free position, so $BG\mathbf{y}$ is the particular solution.

By construction, column i of B is $\mathbf{e}_{p_i} \in F^n$. By Definition 2.4.1 and Definition 2.3.8, column i of A_rB is $A_r\mathbf{e}_{p_i}$, which is column p_i of A_r. Since A_r is in RREF with pivot columns $p_1, \ldots, p_m$, this column is $\mathbf{e}_i \in F^m$ (Lemma 3.3.16(b)). So $A_rB = I_m$.

Since $GA = A_r$ and G is invertible, $A = G^{-1}A_r$. Hence

$$A(BG) = G^{-1}A_rBG = G^{-1}I_mG = I_m,$$

as desired.

For the last claim, let $\mathbf{z} = G\mathbf{y}$. The vector $B\mathbf{z} = z_1\mathbf{e}_{p_1} + \cdots + z_m\mathbf{e}_{p_m}$ has z_i in position p_i for each i, and 0 in every other position, that is, in every free position. Since $\operatorname{rank} A = m$, there are no consistency conditions, and by Definition 3.4.8(b) this is the particular solution of $A\mathbf{x} = \mathbf{y}$. □

So the ith column of BG is the particular solution of $A\mathbf{x} = \mathbf{e}_i$. In other words, the right inverse is found by solving the m systems $A\mathbf{x} = \mathbf{e}_1, \ldots, A\mathbf{x} = \mathbf{e}_m$, just as the inverse of a square matrix is (compare the intuition

after Algorithm 3.3.24). Any other choice of solutions would also give a right inverse; the particular solutions are simply the ones the reduced matrix hands us.

For left inverses the roles of rows and columns are exchanged. If rank $A = n$ there are no free variables, and the reduced system $A_r\mathbf{x} = G\mathbf{y}$ lists the unknowns $x_1, \dots, x_n$ themselves, so $\mathbf{x}$ can be read off from $G\mathbf{y}$. We give this argument, and a second one that never row reduces and instead extends the columns of A to a basis.

Proposition 3.5.9 (Left inverses). Let $A \in F^{m\times n}$ with rank $A = n$. Then A has a left inverse.

Proof (from the pivots). ROUTINE

Uses. Proposition 3.5.2 (First properties of rank), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Definition 2.4.1 (Matrix product)
Hint 1. If every column of A_r is a pivot column, what are $p_1, \dots, p_n$, and what does A_r look like?
Hint 2. $A_r = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. Find an $n \times m$ matrix C with $CA_r = I_n$, and then use $GA = A_r$.
Sketch. Every column is a pivot column, so $p_s = s$ and column s of A_r is $\mathbf{e}_s$: $A_r = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. With $C = \begin{pmatrix} I_n & 0 \end{pmatrix} \in F^{n\times m}$, $CA_r = I_n$, so $(CG)A = CA_r = I_n$. The matrix CG consists of the first n rows of G.

Let $G \in F^{m\times m}$ be invertible with $GA = A_r$ (Theorem 3.3.10). By Proposition 3.5.2(a), every column of A_r is a pivot column. So $p_1 < \cdots < p_n$ are n increasing numbers in $\{1, \dots, n\}$, which forces $p_s = s$, and by Lemma 3.3.16(b) column s of A_r is $\mathbf{e}_s \in F^m$. In block form,

$$A_r = \begin{pmatrix} I_n \\ 0 \end{pmatrix},$$

where the zero block has $m - n$ rows (and is absent if $m = n$). Let $C = \begin{pmatrix} I_n & 0 \end{pmatrix} \in F^{n\times m}$. For $s \le n$, column s of CA_r is $C\mathbf{e}_s$, which is column s of C, namely $\mathbf{e}_s \in F^n$. So $CA_r = I_n$, and

$$(CG)A = C(GA) = CA_r = I_n.$$

Thus CG is a left inverse of A. By Definition 2.4.1, row i of CG is row i of C times G, which is row i of G; so CG consists of the first n rows of G. □

Proof (by extending the columns to a basis). STANDARD

Uses. Proposition 3.4.15 (A pivot in every column), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.5.14 (Appending a vector), Proposition 1.4.14 (The span is the smallest subspace containing a list), Definition 1.6.1 (Basis), Theorem 2.2.1 (Linear maps are determined by a basis), Remark 3.5.7 (Inverses of T_A and inverses of A)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
Hint 1. The columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ of A are linearly independent. A left inverse S must send $\mathbf{a}_j = A\mathbf{e}_j$ to $\mathbf{e}_j$. On what kind of list can we prescribe the values of a linear map freely?
Hint 2. Extend $\mathbf{a}_1, \dots, \mathbf{a}_n$ to a basis of F^m by going through $\mathbf{e}_1, \dots, \mathbf{e}_m$ and appending each one that is not yet in the span (Proposition 1.5.14).
Hint 3. Define S on that basis by $S(\mathbf{a}_j) = \mathbf{e}_j$ and $S = \mathbf{0}$ on the appended vectors, and check $S \circ T_A = \mathrm{id}$ on the standard basis of F^n.
Sketch. The columns are independent (a pivot in every column). Appending each $\mathbf{e}_i$ that is not in the span so far keeps the list independent and eventually puts every $\mathbf{e}_i$ in its span, so it becomes a basis $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_t}$ of F^m. Let S be the linear map with $S(\mathbf{a}_j) = \mathbf{e}_j$ and $S(\mathbf{e}_{i_s}) = \mathbf{0}$. Then $S(T_A\mathbf{e}_j) = S(\mathbf{a}_j) = \mathbf{e}_j$, so $S \circ T_A$ and id agree on a basis and are equal.

Since rank $A = n$, A_r has a pivot in every column (Proposition 3.5.2(a)), so $A\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$ (Proposition 3.4.15). As $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$, this says that the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ are linearly independent (Theorem 1.5.5).

Extending to a basis. Let $\mathcal{L}_0$ be the list $\mathbf{a}_1, \dots, \mathbf{a}_n$. For $i = 1, \dots, m$, let $\mathcal{L}_i = \mathcal{L}_{i-1}$ if $\mathbf{e}_i \in \operatorname{span}(\mathcal{L}_{i-1})$, and let $\mathcal{L}_i$ be $\mathcal{L}_{i-1}$ with $\mathbf{e}_i$ appended otherwise. By Proposition 1.5.14 and induction on i, each $\mathcal{L}_i$ is linearly independent. Each $\mathbf{e}_i$ lies in $\operatorname{span}(\mathcal{L}_i)$, either because it was already in $\operatorname{span}(\mathcal{L}_{i-1}) \subseteq \operatorname{span}(\mathcal{L}_i)$ or because it was appended, and $\operatorname{span}(\mathcal{L}_i) \subseteq \operatorname{span}(\mathcal{L}_m)$ since $\mathcal{L}_i$ is a sublist of $\mathcal{L}_m$. So $\operatorname{span}(\mathcal{L}_m)$ is a subspace of F^m containing $\mathbf{e}_1, \dots, \mathbf{e}_m$, hence containing every linear combination of them (Proposition 1.4.14), that is, all of F^m. Thus

$$\mathcal{L}_m = \mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_t}$$

is linearly independent and spans F^m: it is a basis of F^m (Definition 1.6.1).

Defining S. By Theorem 2.2.1 there is a linear map $S : F^m \to F^n$ with $S(\mathbf{a}_j) = \mathbf{e}_j$ for $j = 1, \ldots, n$ and $S(\mathbf{e}_{i_s}) = \mathbf{0}$ for $s = 1, \ldots, t$. For each j, $(S \circ T_A)(\mathbf{e}_j) = S(A\mathbf{e}_j) = S(\mathbf{a}_j) = \mathbf{e}_j$. So the linear maps $S \circ T_A$ and id_{F^n} agree on the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, and by the uniqueness in Theorem 2.2.1 they are equal. By Remark 3.5.7, $S = T_C$ for a left inverse C of A. □

The first approach gives a formula: the first n rows of any invertible G with $GA = A_r$ form a left inverse of A. The second shows where the freedom lies. A left inverse is forced on the columns of A, which it must send to $\mathbf{e}_1, \ldots, \mathbf{e}_n$, but on the appended vectors $\mathbf{e}_{i_s}$ we could have prescribed any values at all. It also uses nothing about F^m beyond a finite spanning list, which is why it survives in abstract vector spaces (Exercise 3.5.40).

Example 3.5.10 (One-sided inverses, computed). *A right inverse.* Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$. Gauss–Jordan elimination performs $E_{1,-2,2}$ and then $E_{2,-3,1}$, and carrying I_2 along (Proposition 3.3.11) gives

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 2 & 4 & 7 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|cc} 1 & 2 & 0 & 7 & -3 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right).$$

So $\operatorname{rank} A = 2$, the pivot columns are $p_1 = 1$ and $p_2 = 3$, and $G = \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$. The matrix B has 1s in positions $(1,1)$ and $(3,2)$, so

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad BG = \begin{pmatrix} 7 & -3 \\ 0 & 0 \\ -2 & 1 \end{pmatrix}.$$

Check: the rows of A times the columns of BG give $7 - 6 = 1$, $-3 + 3 = 0$, $14 - 14 = 0$, and $-6 + 7 = 1$, so $A(BG) = I_2$. The middle row of BG is zero because x_2 is a free variable, set to 0 in every particular solution. Another right inverse is $BG + \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}(1 \;\; 1)$, since $(-2, 1, 0)$ lies in $\ker T_A$; right inverses are far from unique.

A left inverse. Let $A' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 2}$. Elimination performs $E_{1,-1,2}$, $E_{1,-2,3}$, and $E_{2,-1,3}$, ending at $A'_r = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, and the same operations turn I_3 into

$$G = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

So $\operatorname{rank} A' = 2$, and by the first approach the first two rows of G form a left inverse $C = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}$; indeed $CA' = I_2$. The second approach extends the columns $(1,1,2)$ and $(0,1,1)$ to a basis. The vector $\mathbf{e}_1$ is not in their span (a combination $x(1,1,2) + y(0,1,1) = (x, x+y, 2x+y)$ equal to $\mathbf{e}_1$ needs $x = 1$, $y = -1$, and then $2x + y = 1 \neq 0$), so the basis is $(1,1,2), (0,1,1), \mathbf{e}_1$. Solving for the coordinates of $\mathbf{e}_2$ and $\mathbf{e}_3$ in this basis,

$$\mathbf{e}_2 = -(1,1,2) + 2(0,1,1) + \mathbf{e}_1,$$
$$\mathbf{e}_3 = (1,1,2) - (0,1,1) - \mathbf{e}_1,$$

so $S(\mathbf{e}_1) = \mathbf{0}$, $S(\mathbf{e}_2) = (-1, 2)$, and $S(\mathbf{e}_3) = (1, -1)$:

$$S = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix}.$$

Check: S sends the columns $(1,1,2)$ and $(0,1,1)$ of A' to $(0-1+2, 0+2-2) = (1,0)$ and $(0-1+1, 0+2-1) = (0,1)$, so $SA' = I_2$. The two left inverses differ: $S - C = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. Each row of the difference is a multiple of $(-1,-1,1)$, the last row of G, which kills the columns of A'; Exercise 3.5.33 shows that this is the general pattern.

The row reducer below carries out the first computation. With I_2 carried along, the right block at the end is G.

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 2 & 4 & 7 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|cc} 1 & 2 & 0 & 7 & -3 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right)$$

Reducing $(A \mid I_2)$ for the matrix A of Example 3.5.10. The left block becomes A_r, with pivots in columns 1 and 3, and the right block becomes G. Placing the rows of G in rows $p_1 = 1$ and $p_2 = 3$ of a 3×2 matrix, with zeros in row 2, gives the right inverse BG. (Interactive in the web version.)

Exercise 3.5.11. ROUTINE Let $A = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 4}$. Find its rank and pivot columns, and use Proposition 3.5.8 to find a right inverse. Check your answer.

Solution on page 1251

The completed theorems

We restate each Fundamental Theorem in full. Conditions (I)–(V) are those of Section 2.7, and (VI) is new.

Theorem 3.5.12 (First Fundamental Theorem, completed). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is injective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has at most one solution $\mathbf{x} \in F^n$.
(III) The columns of A are linearly independent.
(IV) $\ker T_A = \{\mathbf{0}\}$.
(V) T_A has a left inverse: a linear map $S : F^m \to F^n$ with $S \circ T_A = \mathrm{id}_{F^n}$.
(VI) $\operatorname{rank} A = n$.

Moreover, if these hold, then $n \le m$.

Proof. ROUTINE

Uses. Theorem 2.7.2 (First Fundamental Theorem), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 3.5.2 (First properties of rank), Proposition 3.4.15 (A pivot in every column), Definition 2.6.5 (Kernel), Proposition 3.5.9 (Left inverses), Remark 3.5.7 (Inverses of T_A and inverses of A)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Theorem 2.7.2 already ties (I)–(IV) together and shows (V) ⇒ (I). Which of (I)–(IV) is closest to a statement about pivots?
Hint 2. (IV) says that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Combine Proposition 3.4.15 with Proposition 3.5.2(a).
Hint 3. For (VI) ⇒ (V), use Proposition 3.5.9. For the last claim, recall $\operatorname{rank} A \le m$.
Sketch. (I)–(IV) are equivalent and (V) implies (I) by Theorem 2.7.2. (IV) says $A\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$, which by Proposition 3.4.15 means a pivot in every column, that is, $\operatorname{rank} A = n$. (VI) gives a left inverse matrix by Proposition 3.5.9, hence a left inverse map. So (V) ⇒ (I) ⇒ (IV) ⇒ (VI) ⇒ (V). Finally $n = \operatorname{rank} A \le m$.

By Theorem 2.7.2, conditions (I)–(IV) are equivalent, and (V) implies (I) (this is Proposition 2.5.4). We show (IV) ⇒ (VI) ⇒ (V), which closes the cycle (V) ⇒ (I) ⇒ (IV) ⇒ (VI) ⇒ (V). In fact we show (IV) ⇔ (VI).

(IV) ⇔ (VI). By Definition 2.6.5, $\ker T_A = \{\mathbf{0}\}$ says that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. By Proposition 3.4.15 this holds if and only if A_r has a pivot in every column, which by Proposition 3.5.2(a) means $\operatorname{rank} A = n$.

(VI) ⇒ (V). By Proposition 3.5.9 there is $C \in F^{n\times m}$ with $CA = I_n$, and by Remark 3.5.7 the linear map $S = T_C$ satisfies $S \circ T_A = \mathrm{id}_{F^n}$.

Finally, if the conditions hold, then $n = \operatorname{rank} A \le m$ by Proposition 3.5.2(a). □

Theorem 3.5.13 (Second Fundamental Theorem, completed). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is surjective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has at least one solution $\mathbf{x} \in F^n$.
(III) The columns of A span F^m.
(IV) $\operatorname{im} T_A = F^m$.
(V) T_A has a right inverse: a linear map $S : F^m \to F^n$ with $T_A \circ S = \mathrm{id}_{F^m}$.
(VI) $\operatorname{rank} A = m$.

Moreover, if these hold, then $n \ge m$.

Proof. ROUTINE

Uses. Theorem 2.7.4 (Second Fundamental Theorem), Proposition 2.5.5 (Right inverses give existence), Proposition 3.5.2 (First properties of rank), Proposition 3.4.16 (A pivot in every row), Proposition 3.5.8 (Right inverses from the pivots), Remark 3.5.7 (Inverses of T_A and inverses of A)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Mirror the proof of Theorem 3.5.12. Which of (I)–(IV) is a statement about pivots in rows?
Hint 2. (II) is the second condition of Proposition 3.4.16. For (VI) ⇒ (V), use Proposition 3.5.8.
Sketch. (I)–(IV) are equivalent and (V) implies (I) by Theorem 2.7.4. (II) holds iff A_r has a pivot in every row (Proposition 3.4.16), iff $\operatorname{rank} A = m$. (VI) gives the right inverse BG of Proposition 3.5.8. So (V) ⇒ (I) ⇒ (II) ⇒ (VI) ⇒ (V), and $m = \operatorname{rank} A \le n$.

By Theorem 2.7.4, conditions (I)–(IV) are equivalent, and (V) implies (I) (this is Proposition 2.5.5). We show (II) ⇔ (VI) and (VI) ⇒ (V), which closes the cycle (V) ⇒ (I) ⇒ (II) ⇒ (VI) ⇒ (V).

(II) ⇔ (VI). By Proposition 3.4.16, (II) holds if and only if A_r has a pivot in every row, which by Proposition 3.5.2(a) means $\operatorname{rank} A = m$.

(VI) ⇒ (V). By Proposition 3.5.8 there is $X = BG \in F^{n\times m}$ with $AX = I_m$, and by Remark 3.5.7 the linear map $S = T_X$ satisfies $T_A \circ S = \mathrm{id}_{F^m}$.

Finally, if the conditions hold, then $m = \operatorname{rank} A \le n$ by Proposition 3.5.2(a). □

Theorem 3.5.14 (Third Fundamental Theorem, completed). Let $A \in F^{m\times n}$. The following are equivalent.

(I) T_A is bijective.
(II) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has exactly one solution $\mathbf{x} \in F^n$.
(III) The columns of A form a basis of F^m.
(IV) $\ker T_A = \{\mathbf{0}\}$ and $\operatorname{im} T_A = F^m$.
(V) T_A is invertible.
(VI) $\operatorname{rank} A = m = n$.

Moreover, if these hold, then $m = n$.

Proof. ROUTINE

Uses. Theorem 2.7.8 (Third Fundamental Theorem), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Proposition 2.7.7 (Bijective linear maps are invertible)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. A bijection is a map that is both injective and surjective.
Hint 2. Apply the two previous theorems.
Sketch. (I)–(V) are equivalent by Theorem 2.7.8. T_A is bijective iff it is injective and surjective, iff $\operatorname{rank} A = n$ and $\operatorname{rank} A = m$ by Theorem 3.5.12 and Theorem 3.5.13.

By Theorem 2.7.8, conditions (I)–(V) are equivalent; the step from (I) to (V) is that the inverse of a bijective linear map is linear (Proposition 2.7.7). For (I) ⇔ (VI): T_A is bijective if and only if it is injective and surjective, which by Theorem 3.5.12 and Theorem 3.5.13 holds if and only if rank $A = n$ and rank $A = m$. In that case $m = \operatorname{rank} A = n$. □

So the sixth conditions are rank $A = n$, rank $A = m$, and rank $A = m = n$, and the "moreover" parts follow from rank $A \leq \min(m, n)$. By Remark 3.5.7, condition (V) can be read for matrices: A has a left inverse, a right inverse, or an inverse, respectively. With this reading, the "moreover" parts give the fact that the lecture singles out: invertible matrices are square.

Corollary 3.5.15 (Invertible matrices are square). Let $A \in F^{m\times n}$.

(a) If A has a left inverse, then $n \leq m$.

(b) If A has a right inverse, then $n \geq m$.

(c) If A is invertible, that is, $AB = I_m$ and $BA = I_n$ for some $B \in F^{n\times m}$, then $m = n$.

Proof. ROUTINE

Uses. Remark 3.5.7 (Inverses of T_A and inverses of A), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed)
Hint 1. Translate "A has a left inverse" into condition (V) of one of the theorems.
Hint 2. Apply the "moreover" parts.
Sketch. A left inverse C of A gives the left inverse T_C of T_A, so (V) of Theorem 3.5.12 holds and $n \leq m$. Right inverses give $n \geq m$ by Theorem 3.5.13, and an inverse is both.

(a) If $CA = I_n$, then T_C is a left inverse of T_A (Remark 3.5.7), so condition (V) of Theorem 3.5.12 holds, and $n \leq m$.

(b) If $AC = I_m$, then T_C is a right inverse of T_A, so condition (V) of Theorem 3.5.13 holds, and $n \geq m$.

(c) An inverse B is both a left and a right inverse, so $n \leq m$ and $n \geq m$ by (a) and (b). □

Here is what the theorems say about a list of vectors in F^m, read through the matrix that has them as columns.

Corollary 3.5.16 (Lists of vectors in F^m). Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in F^m$, and let $A \in F^{m\times n}$ be the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

(a) The list is linearly independent if and only if rank $A = n$. In that case $n \leq m$.

(b) The list spans F^m if and only if rank $A = m$. In that case $n \geq m$.

(c) If $n = m$, the list is linearly independent if and only if it spans F^m, if and only if it is a basis of F^m.

Proof. ROUTINE

Uses. Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Definition 1.6.1 (Basis)
Hint 1. Which conditions of the completed theorems are about the columns of A?
Hint 2. For (c), when $n = m$, the conditions rank $A = n$ and rank $A = m$ coincide.
Sketch. (a) is (III) ⇔ (VI) of Theorem 3.5.12 with its "moreover", and (b) is the same for Theorem 3.5.13. If $n = m$, both say rank $A = m$; a basis is a list that is both independent and spanning.

(a) By (III) ⇔ (VI) of Theorem 3.5.12, the columns of A are linearly independent if and only if rank $A = n$, and then $n \leq m$.

(b) By (III) ⇔ (VI) of Theorem 3.5.13, the columns of A span F^m if and only if rank $A = m$, and then $n \geq m$.

(c) If $n = m$, the conditions in (a) and (b) are the same condition rank $A = m$. So the list is independent if and only if it spans F^m, and in that case it is both, which is the definition of a basis (Definition 1.6.1). □

In particular a basis of F^m, being independent and spanning, has $n \le m$ and $n \ge m$ vectors: every basis of F^m has exactly m vectors. This is how Section 4.1 begins (Theorem 4.1.1), on the way to defining the dimension of a vector space. Part (c) is a first instance of a principle that Section 4.2 makes general: for a list whose length is right, independence and spanning are the same condition. Part (a) also explains Corollary 3.4.19(a) from the side of the columns: more than m vectors in F^m are always dependent.

Remark 3.5.17 (Rank of a linear map). For a linear map $T : V \to W$ and bases $\mathcal{B}$ of V and $\mathcal{C}$ of W, it is tempting to define $\operatorname{rank} T$ as $\operatorname{rank} [T]_{\mathcal{C}\mathcal{B}}$. This is a good definition, but only because the rank of $[T]_{\mathcal{C}\mathcal{B}}$ does not depend on the choice of bases, and that needs ideas we do not have yet. Section 4.2 defines the rank of T intrinsically, as the dimension of $\operatorname{im} T$ (Definition 4.2.14), and shows that it equals $\operatorname{rank} [T]_{\mathcal{C}\mathcal{B}}$ for every choice of bases.

Exercise 3.5.18. ROUTINE For each item, either give an example or explain why none exists.

(a) $A \in \mathbb{R}^{3\times 5}$ with T_A injective.

(b) $A \in \mathbb{R}^{3\times 5}$ with T_A surjective.

(c) $A \in \mathbb{R}^{5\times 3}$ with $\operatorname{rank} A = 3$ and a right inverse.

(d) $A \in \mathbb{R}^{4\times 4}$ with $\operatorname{rank} A = 4$ that is not invertible.

Solution on page 1251

3.5.3 Invertible matrices are products of elementary matrices

Elementary matrices are invertible, and so are their products (Corollary 3.2.11). The converse says that the elementary matrices generate all invertible matrices: there are no others.

Theorem 3.5.19 (Invertible matrices are products of elementary matrices). Every invertible matrix $A \in F^{n\times n}$ is a product of elementary matrices. Explicitly, if Gauss–Jordan elimination on A performs the row operations with elementary matrices $E_1, \dots, E_t$, in this order, then

$$A = E_1^{-1} E_2^{-1} \cdots E_t^{-1},$$

and each E_i^{-1} is elementary. Consequently, a square matrix is invertible if and only if it is a product of elementary matrices.

Proof. ROUTINE

Uses. Theorem 3.3.23 (Invertibility and RREF), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.10 (Inverses of elementary matrices)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. What is $\operatorname{rref}(A)$ for an invertible A?
Hint 2. $GA = I_n$ with $G = E_t \cdots E_1$. Solve for A.
Sketch. By Theorem 3.3.23, $\operatorname{rref}(A) = I_n$, so $E_t \cdots E_1 A = I_n$ and $A = (E_t \cdots E_1)^{-1} = E_1^{-1} \cdots E_t^{-1}$, a product of elementary matrices by Proposition 3.2.10. The converse is Corollary 3.2.11.

Let $G = E_t \cdots E_1$, so that $GA = \operatorname{rref}(A)$ (Theorem 3.3.10). Since A is invertible, $\operatorname{rref}(A) = I_n$ by Theorem 3.3.23, so $GA = I_n$. By Corollary 3.2.11, G is invertible with $G^{-1} = E_1^{-1} \cdots E_t^{-1}$, and multiplying $GA = I_n$ on the left by G^{-1} gives

$$A = G^{-1} = E_1^{-1} E_2^{-1} \cdots E_t^{-1}.$$

Each E_i^{-1} is elementary by Proposition 3.2.10. (If $t = 0$, then $A = I_n$, the empty product.) Conversely, every product of elementary matrices is invertible by Corollary 3.2.11. □

Example 3.5.20 (Factoring an invertible matrix). Let $A = \begin{pmatrix} 3 & 6 \\ 1 & 3 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$. Elimination scales row 1 by $1/3$, adds -1 times row 1 to row 2, and adds -2 times row 2 to row 1:

$$A \xrightarrow{E_{1,1/3}} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \xrightarrow{E_{1,-1,2}} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \xrightarrow{E_{2,-2,1}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Inverting each operation and reversing the order (Proposition 3.2.10),

$$A = E_{1,3}E_{1,1,2}E_{2,2,1} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Read from the right, this builds A out of I_2: $E_{2,2,1}$ turns I_2 into $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, then adding row 1 to row 2 gives $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, and scaling row 1 by 3 gives A. In the plane, T_A is a horizontal shear, followed by a vertical shear, followed by a stretch along the x_1-axis (Example 3.2.33).

Combined with Section 3.3, the theorem gives several descriptions of the same class of matrices.

Corollary 3.5.21 (Invertibility and row equivalence).

(a) For $A \in F^{n\times n}$, the following are equivalent: A is invertible; A is a product of elementary matrices; $A \sim I_n$; $\operatorname{rref}(A) = I_n$.

(b) For $A, B \in F^{m\times n}$, we have $A \sim B$ if and only if $B = GA$ for some invertible $G \in F^{m\times m}$.

Proof. ROUTINE

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Definition 3.2.21 (Row equivalence), Proposition 3.2.22 (Row equivalence is an equivalence relation), Theorem 3.3.23 (Invertibility and RREF), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. By Definition 3.2.21, $A \sim B$ means $B = E_k \cdots E_1 A$. Compare this with "$B = GA$ for a product G of elementary matrices".
Hint 2. For (a), $A \sim I_n$ means $I_n = GA$ for a product G of elementary matrices; equivalently (by symmetry) $A = G'I_n$.
Sketch. By Definition 3.2.21, $A \sim B$ iff $B = GA$ with G a product of elementary matrices, and by Theorem 3.5.19 these G are exactly the invertible matrices; this is (b). For (a), $A \sim I_n$ iff $I_n \sim A$ iff $A = GI_n = G$ for an invertible G, and $\operatorname{rref}(A) = I_n$ is Theorem 3.3.23.

(b) By Definition 3.2.21, $A \sim B$ if and only if $B = GA$ for some product G of elementary matrices. By Theorem 3.5.19 and Corollary 3.2.11, the products of elementary matrices in $F^{m\times m}$ are exactly the invertible matrices.

(a) By Theorem 3.5.19, A is invertible if and only if it is a product of elementary matrices. By (b) and the symmetry of $\sim$ (Proposition 3.2.22), $A \sim I_n$ if and only if $A = GI_n = G$ for some invertible G, that is, if and only if A is invertible. Finally, A is invertible if and only if $\operatorname{rref}(A) = I_n$ by Theorem 3.3.23. □

The theorem is what makes the method of Technique 5.4.6 work: a statement about all invertible matrices that is easy to check for swaps, scalings, and replacements, and that survives products, holds for every invertible matrix. Its main use is in Section 5.4, where it proves $\det(AB) = \det A \det B$. Here is a first, simple instance, which turns statements about GAH into statements about row and column operations.

Corollary 3.5.22 (Row and column operations). Let $A, B \in F^{m\times n}$. Then $B = GAH$ for some invertible $G \in F^{m\times m}$ and $H \in F^{n\times n}$ if and only if B can be obtained from A by a finite sequence of elementary row operations and elementary column operations.

Proof. ROUTINE

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.28 (Column operations are right multiplications)
Techniques. Reduce to elementary matrices (Technique 5.4.6)
Hint 1. A row operation is left multiplication by an elementary matrix, and a column operation is right multiplication by one.
Hint 2. For $\Rightarrow$, factor G and H into elementary matrices.

Hint 3. Row operations and column operations commute: $(EX)E' = E(XE')$. So the operations may be performed in any interleaving.
Sketch. A row operation is $X \mapsto EX$ and a column operation is $X \mapsto XE'$ (Proposition 3.2.7, Proposition 3.2.28); since $(EX)E' = E(XE')$, any sequence of them turns A into $(E_s \cdots E_1)A(E'_1 \cdots E'_t)$, and both products are invertible. Conversely, factor $G = E_s \cdots E_1$ and $H = E'_1 \cdots E'_t$ by Theorem 3.5.19 and perform the corresponding operations.

By Proposition 3.2.7, performing a row operation on a matrix X replaces it by EX for an elementary $E \in F^{m\times m}$, and by Proposition 3.2.28, performing a column operation replaces it by XE' for an elementary $E' \in F^{n\times n}$. By associativity, $(EX)E' = E(XE')$, so in any sequence of such operations the row operations and column operations can be collected separately. A sequence with row operations $E_1, \ldots, E_s$ and column operations $E'_1, \ldots, E'_t$ (each list in the order performed) therefore turns A into

$$(E_s \cdots E_1)A(E'_1 \cdots E'_t),$$

and both products are invertible by Corollary 3.2.11.

Conversely, suppose $B = GAH$ with G, H invertible. By Theorem 3.5.19, $G = E_s \cdots E_1$ and $H = E'_1 \cdots E'_t$ for elementary matrices E_i and E'_j. Performing on A the row operations of $E_1, \ldots, E_s$ and then the column operations of $E'_1, \ldots, E'_t$ produces $(E_s \cdots E_1)A(E'_1 \cdots E'_t) = GAH = B$. □

So the theorem lets us pass between "multiply by an invertible matrix" and "perform elementary operations" in both directions. The first is convenient in algebra, the second in computation and in proofs by induction on the number of operations.

3.5.4 Square matrices

For a square matrix, $m = n$, so the sixth conditions of the three Fundamental Theorems all read $\operatorname{rank} A = n$. Injectivity, surjectivity, and bijectivity of T_A are then the same condition. The most useful consequence is that a square matrix needs only a one-sided inverse to be invertible.

Theorem 3.5.23 (One-sided inverses of square matrices). Let $A, B \in F^{n\times n}$. If $BA = I_n$, then $AB = I_n$. Consequently, if a square matrix A has a left inverse or a right inverse B, then A is invertible and $B = A^{-1}$.

Proof. STANDARD

Uses. Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.14 (Third Fundamental Theorem, completed), Remark 3.5.7 (Inverses of T_A and inverses of A), Definition 2.5.12 (Invertible)
Hint 1. $BA = I_n$ says that A has a left inverse. What does Theorem 3.5.12 then say about $\operatorname{rank} A$, and what does that mean when $m = n$?
Hint 2. A is invertible by Theorem 3.5.14. Compute $B = B\big(AA^{-1}\big)$.
Hint 3. For a right inverse, swap the roles: $AB = I_n$ says that A is a left inverse of B.
Sketch. If $BA = I_n$, then T_A has a left inverse, so $\operatorname{rank} A = n$ by Theorem 3.5.12. As A is square, $\operatorname{rank} A = n = m$, so A is invertible by Theorem 3.5.14. Then $B = B\big(AA^{-1}\big) = (BA)A^{-1} = A^{-1}$, so $AB = AA^{-1} = I_n$. If instead $AB = I_n$, apply this to the pair (B, A) to get $BA = I_n$.

Suppose $BA = I_n$. By Remark 3.5.7, T_B is a left inverse of T_A, so condition (V) of Theorem 3.5.12 holds and $\operatorname{rank} A = n$. Since A is $n \times n$, this says $\operatorname{rank} A = n = m$, so by Theorem 3.5.14 and Remark 3.5.7, A has an inverse A^{-1}, with $AA^{-1} = A^{-1}A = I_n$. Then

$$B = BI_n = B\big(AA^{-1}\big) = (BA)A^{-1} = I_nA^{-1} = A^{-1},$$

and therefore $AB = AA^{-1} = I_n$.

For the second sentence: if $BA = I_n$, we have just shown that A is invertible with $A^{-1} = B$. If $AB = I_n$, then the first sentence, applied with the roles of A and B exchanged, gives $BA = I_n$, so again $B = A^{-1}$. □

Intuition (Why squareness matters). The proof turned a statement about uniqueness into a statement about existence. $BA = I_n$ gives $A\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = BA\mathbf{x} = \mathbf{0}$, so there are no free variables; a square matrix

without free variables has as many pivots as rows, so there are no zero rows either, and every equation $A\mathbf{x} = \mathbf{y}$ is solvable. This is Technique 3.4.21 again. Both hypotheses are needed. For $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ we have $AB = (1) = I_1$, but $BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq I_2$. And the analogue for linear maps of an infinite-dimensional space fails: on $F[x]$, let S and L be the linear maps with $S\left(x^k\right) = x^{k+1}$ for all $k \geq 0$, and $L\left(x^k\right) = x^{k-1}$ for $k \geq 1$ and $L(1) = \mathbf{0}$ (Proposition 2.2.20). Then $L \circ S = \mathrm{id}$, but $(S \circ L)(1) = \mathbf{0}$, so $S \circ L \neq \mathrm{id}$.

The theorem halves the work of checking that a matrix is an inverse. It also gives the converse of the familiar fact that products of invertible matrices are invertible.

Corollary 3.5.24 (Products of square matrices). Let $A, B \in F^{n\times n}$. Then AB is invertible if and only if both A and B are invertible.

Proof. ROUTINE

Uses. Theorem 3.5.23 (One-sided inverses of square matrices), Definition 2.5.12 (Invertible)
Hint 1. If AB has inverse C, then $A(BC) = I_n$. What does that say about A?
Hint 2. For B, regroup $C(AB) = I_n$.
Sketch. If $C = (AB)^{-1}$, then $A(BC) = I_n$ and $(CA)B = I_n$, so A has a right inverse and B a left inverse; both are invertible by Theorem 3.5.23. Conversely, $(AB)\left(B^{-1}A^{-1}\right) = I_n$.

Suppose AB is invertible, with inverse C. Then $A(BC) = (AB)C = I_n$, so A has a right inverse, and $(CA)B = C(AB) = I_n$, so B has a left inverse. By Theorem 3.5.23, A and B are invertible.

Conversely, if A and B are invertible, then $(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AA^{-1} = I_n$, so AB has a right inverse and is invertible by Theorem 3.5.23. □

Again squareness is essential: with $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the product $AB = I_1$ is invertible, but neither factor is, since neither is square. What survives without squareness is that A has a right inverse and B a left inverse, as the first two lines of the proof show (compare Section 2.5).

We collect what we know about square matrices. Every item below has appeared already; the point is that they are all one condition.

Corollary 3.5.25 (The invertible matrix theorem). Let $A \in F^{n\times n}$. The following are equivalent.

(a) A is invertible.
(b) A has a left inverse.
(c) A has a right inverse.
(d) $\operatorname{rank} A = n$.
(e) $\operatorname{rref}(A) = I_n$.
(f) A is a product of elementary matrices.
(g) Any one of the conditions (I)–(VI) of Theorem 3.5.12, Theorem 3.5.13, or Theorem 3.5.14 holds.

In particular, $T_A : F^n \to F^n$ is injective if and only if it is surjective.

Proof. ROUTINE

Uses. Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Theorem 3.5.14 (Third Fundamental Theorem, completed), Theorem 3.5.23 (One-sided inverses of square matrices), Theorem 3.3.23 (Invertibility and RREF), Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Remark 3.5.7 (Inverses of T_A and inverses of A)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Show that each condition is equivalent to (d).
Hint 2. When $m = n$, what do the sixth conditions of the three theorems say?
Sketch. With $m = n$, condition (VI) of each of the three theorems reads $\operatorname{rank} A = n$, so every condition (I)–(VI) of each theorem is equivalent to (d). (a), (b), (c) are (V) of the three theorems, read for matrices; (e) and (f) are equivalent to (a) by Theorem 3.3.23 and Theorem 3.5.19.

Since $m = n$, condition (VI) of each of Theorem 3.5.12, Theorem 3.5.13, and Theorem 3.5.14 says rank $A = n$. So, by those theorems, every one of their conditions is equivalent to (d), which gives (d) ⇔ (g). By Remark 3.5.7, (b), (c), and (a) are condition (V) of the three theorems, read for matrices, so they are equivalent to (d) as well. (Alternatively, (b) ⇔ (a) ⇔ (c) is Theorem 3.5.23.) Finally, (a) ⇔ (e) is Theorem 3.3.23, and (a) ⇔ (f) is Theorem 3.5.19. The last sentence is the equivalence of (I) in Theorem 3.5.12 with (I) in Theorem 3.5.13. □

This contains Corollary 3.4.20, which was proved by counting pivots directly. The equivalence of injectivity and surjectivity for $T_A : F^n \to F^n$ extends to every linear map $T : V \to V$ of a finite-dimensional space, once dimension is available (Section 4.2).

Exercise 3.5.26. ROUTINE Find the error in the following argument, and show that its conclusion is false.

Claim. If $A \in F^{m\times n}$ and $B \in F^{n\times m}$ satisfy $AB = I_m$, then $BA = I_n$.

Proof. Since $AB = I_m$, we have $ABA = I_m A = A = AI_n$. Cancelling A on the left gives $BA = I_n$.

Solution on page 1252

3.5.5 Rank normal form

Row operations bring A to A_r, and no further: A_r is the simplest matrix row equivalent to A. If column operations are allowed too, more can be cleared. Each free column of A_r is a combination of the pivot columns to its left (Lemma 3.3.16(d)), and the pivot columns are $\mathbf{e}_1, \dots, \mathbf{e}_k$, so column operations can subtract those combinations and leave zero. After moving the pivot columns to the front, what remains is determined by the rank alone.

Theorem 3.5.27 (Rank normal form). Let $A \in F^{m\times n}$ and $k = \operatorname{rank} A$, and let $N_k \in F^{m\times n}$ be the matrix

$$N_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix},$$

with 1 in positions $(1,1), \dots, (k,k)$ and 0 elsewhere. There are invertible matrices $G \in F^{m\times m}$ and $H \in F^{n\times n}$ such that

$$GAH = N_k.$$

Here G can be any invertible matrix with $GA = A_r$.

Proof (column operations on the reduced form). STANDARD

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Proposition 3.3.8 (Properties of the output), Definition 2.4.1 (Matrix product), Theorem 2.4.34 (Block multiplication), Theorem 3.5.14 (Third Fundamental Theorem, completed), Remark 3.5.7 (Inverses of T_A and inverses of A), Definition 2.5.12 (Invertible)
Techniques. Reduce to RREF (Technique 3.3.22), Block decomposition (Technique 2.4.35)
Hint 1. Start from $GA = A_r$. It remains to find an invertible H with $A_r H = N_k$, that is, to clear A_r by column operations.
Hint 2. First reorder the columns so that the pivot columns come first. In block form, the result is $\begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}$ for some $k \times (n-k)$ matrix X.
Hint 3. Multiply on the right by $\begin{pmatrix} I_k & -X \\ 0 & I_{n-k} \end{pmatrix}$. Why is that matrix invertible?
Sketch. Let P have columns $\mathbf{e}_{p_1}, \dots, \mathbf{e}_{p_k}, \mathbf{e}_{f_1}, \dots, \mathbf{e}_{f_\ell}$, a reordering of the standard basis, so P is invertible. Then $A_r P$ lists the pivot columns $\mathbf{e}_1, \dots, \mathbf{e}_k$ first, and all columns vanish below row k, so $A_r P = \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}$. With $U = \begin{pmatrix} I_k & -X \\ 0 & I_\ell \end{pmatrix}$, invertible with inverse $\begin{pmatrix} I_k & X \\ 0 & I_\ell \end{pmatrix}$, block multiplication gives $A_r PU = N_k$. Take $H = PU$.

Let G be invertible with $GA = A_r$ (Theorem 3.3.10), with pivot columns $p_1 < \cdots < p_k$ and free columns $f_1 < \cdots < f_\ell$, where $\ell = n - k$. It suffices to find an invertible H with $A_r H = N_k$, since then $GAH = A_r H = N_k$.

Moving the pivot columns to the front. Let $P \in F^{n\times n}$ be the matrix with columns $\mathbf{e}_{p_1}, \dots, \mathbf{e}_{p_k}, \mathbf{e}_{f_1}, \dots, \mathbf{e}_{f_\ell}$. These are $\mathbf{e}_1, \dots, \mathbf{e}_n$ in another order, so they form a basis of F^n, and P is invertible by Theorem 3.5.14 ((III) ⇒

(V)) and Remark 3.5.7. Column j of A_rP is A_r times column j of P (Definition 2.4.1). For $j \le k$ this is $A_r\mathbf{e}_{p_j}$, column p_j of A_r, which is $\mathbf{e}_j$ (Lemma 3.3.16(b)). For $j = k+i$ it is column f_i of A_r. Rows $k+1, \dots, m$ of A_r are zero (Proposition 3.3.8(a)), so every column of A_rP vanishes below row k. In block form, with row sizes $k, m-k$ and column sizes k, ℓ,

$$A_rP = \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix},$$

where column i of $X \in F^{k\times\ell}$ consists of the first k entries of column f_i of A_r.

Clearing X. Let

$$U = \begin{pmatrix} I_k & -X \\ 0 & I_\ell \end{pmatrix}, \quad U' = \begin{pmatrix} I_k & X \\ 0 & I_\ell \end{pmatrix} \in F^{n\times n}.$$

By block multiplication (Theorem 2.4.34),

$$UU' = \begin{pmatrix} I_k & X - X \\ 0 & I_\ell \end{pmatrix} = I_n,$$

$$U'U = \begin{pmatrix} I_k & -X + X \\ 0 & I_\ell \end{pmatrix} = I_n,$$

so U is invertible (Definition 2.5.12), and

$$\begin{aligned} A_rPU &= \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_k & -X \\ 0 & I_\ell \end{pmatrix} \\ &= \begin{pmatrix} I_k & -X + X \\ 0 & 0 \end{pmatrix} = N_k. \end{aligned}$$

So $H = PU$ is invertible, as a product of invertible matrices, and $GAH = N_k$. (If $k = 0$, then $A_r = 0 = N_0$ and $H = I_n$ works; if $\ell = 0$, then $P = I_n$ and X and U are not needed. Blocks of size zero are simply absent.) □

Proof (by row reducing the transpose). STANDARD

> *Uses.* Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Proposition 3.3.8 (Properties of the output), Definition 2.4.21 (Transpose), Proposition 2.4.23 (Transpose of a product), Proposition 3.5.9 (Left inverses), Proposition 3.5.2 (First properties of rank), Corollary 3.5.16 (Lists of vectors in F^m)
> *Techniques.* Transpose to trade rows for columns (Technique 3.2.27), Reduce to RREF (Technique 3.3.22)
> *Hint 1.* Column operations on A_r are row operations on A_r^{T}. So row reduce $A_r^{\mathsf{T}} \in F^{n\times m}$.
> *Hint 2.* The first k columns of A_r^{T} are the nonzero rows of A_r, transposed, and the others are zero. Why are those k columns linearly independent?
> *Hint 3.* Row s of A_r has a 1 in column p_s, while every other row has a 0 there.
> *Sketch.* The first k columns of A_r^{T} form an $n \times k$ matrix M whose column s is row s of A_r; its p_t-th entry is 1 if $s = t$ and 0 otherwise, so the columns are independent and $\operatorname{rank} M = k$. As in the proof of Proposition 3.5.9, some invertible K has $KM = \begin{pmatrix} I_k \\ 0 \end{pmatrix}$, so $KA_r^{\mathsf{T}} = N_k^{\mathsf{T}}$ (the remaining columns are zero). Transposing, $A_rK^{\mathsf{T}} = N_k$, and K^{T} is invertible.

Let G be invertible with $GA = A_r$, and let $p_1 < \cdots < p_k$ be the pivot columns. By Proposition 3.3.8, rows $k+1, \dots, m$ of A_r are zero, so the transpose $A_r^{\mathsf{T}} \in F^{n\times m}$ (Definition 2.4.21) has the block form

$$A_r^{\mathsf{T}} = (M \;\; 0),$$

where $M \in F^{n\times k}$ has as its column s the transpose of row s of A_r, and the zero block has $m-k$ columns.

The columns of M are independent. Column p_t of A_r is $\mathbf{e}_t$ (Lemma 3.3.16(b)), so row s of A_r has entry 1 in column p_t if $s = t$, and 0 if $s \neq t$. Hence if $c_1, \dots, c_k \in F$ and $c_1\mathbf{m}_1 + \cdots + c_k\mathbf{m}_k = \mathbf{0}$, where $\mathbf{m}_s$ is column s of M, then the entry in position p_t gives $c_t = 0$, for each t. So the columns of M are linearly independent, and $\operatorname{rank} M = k$ by Corollary 3.5.16(a).

Reducing M. By Theorem 3.3.10 there is an invertible $K \in F^{n\times n}$ with $KM = \operatorname{rref}(M)$, and since $\operatorname{rank} M = k$ is the number of columns of M, the proof of Proposition 3.5.9 (first approach) shows that $\operatorname{rref}(M) = \begin{pmatrix} I_k \\ 0 \end{pmatrix}$. Therefore

$$KA_r^{\mathsf{T}} = (KM \;\; K0) = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} = N_k^{\mathsf{T}} \in F^{n\times m}.$$

Transposing with Proposition 2.4.23, $A_r K^{\mathsf{T}} = \left(KA_r^{\mathsf{T}}\right)^{\mathsf{T}} = N_k$. The matrix $H = K^{\mathsf{T}}$ is invertible, with inverse $\left(K^{-1}\right)^{\mathsf{T}}$, since $K^{\mathsf{T}}\left(K^{-1}\right)^{\mathsf{T}} = \left(K^{-1}K\right)^{\mathsf{T}} = I_n$ and similarly in the other order. So $GAH = A_r H = N_k$. □

The first proof writes H down: a reordering of the columns followed by subtracting multiples of the first k columns from the others. The second explains why the transpose of A_r is no harder to reduce than A_r itself: its nonzero columns are independent, because the pivot structure of A_r makes the nonzero *rows* of A_r independent. By Corollary 3.5.22, the theorem says that every matrix of rank k can be brought to N_k by elementary row and column operations.

Example 3.5.28 (A rank normal form). For the matrix A of Example 3.3.3 we found

$$A_r = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ -1 & -1/2 & 1 \end{pmatrix},$$

with pivot columns $1, 3$ and free columns $2, 4$, so $k = 2$. Following the first proof, P has columns $\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4$, and

$$A_r P = \begin{pmatrix} 1 & 0 & 2 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X = \begin{pmatrix} 2 & -2 \\ 0 & 2 \end{pmatrix},$$

$$U = \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and $H = PU$ is U with rows 2 and 3 exchanged,

$$H = \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$GAH = A_r H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = N_2.$$

The last two columns of H, $(-2, 1, 0, 0)$ and $(2, 0, -2, 1)$, are the basic solutions $\mathbf{s}_1, \mathbf{s}_2$ of Example 3.4.10. This is no accident: in general, column $k + i$ of $H = PU$ is $\mathbf{e}_{f_i} - \sum_s x_{si}\mathbf{e}_{p_s}$, which is the basic solution $\mathbf{s}_i$ (Definition 3.4.8).

What does the normal form say about the map T_A? Read $GAH = N_k$ as $AH = G^{-1}N_k$. The columns of H and of G^{-1} form bases of F^n and of F^m, and in these bases T_A is as simple as a linear map can be.

Corollary 3.5.29 (Bases adapted to T_A). Let $A \in F^{m\times n}$ have rank k. There are a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of F^n and a basis $\mathbf{c}_1, \ldots, \mathbf{c}_m$ of F^m such that

$$A\mathbf{b}_j = \mathbf{c}_j \quad (1 \le j \le k), \quad A\mathbf{b}_j = \mathbf{0} \quad (k < j \le n).$$

That is, with $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_m$, we have $[T_A]_{\mathcal{C}\mathcal{B}} = N_k$. Moreover

$$\ker T_A = \operatorname{span}(\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n),$$
$$\operatorname{im} T_A = \operatorname{span}(\mathbf{c}_1, \ldots, \mathbf{c}_k).$$

Proof. STANDARD

Uses. Theorem 3.5.27 (Rank normal form), Theorem 3.5.14 (Third Fundamental Theorem, completed), Remark 3.5.7 (Inverses of T_A and inverses of A), Definition 2.4.1 (Matrix product), Definition 2.3.1 (Matrix of a linear map), Corollary 1.5.17 (Sublists), Definition 1.3.10 (Span)
Hint 1. Take $\mathbf{b}_j$ to be the columns of H and $\mathbf{c}_i$ the columns of G^{-1}. Why are these bases?
Hint 2. Compare column j of AH with column j of $G^{-1}N_k$.
Hint 3. For the kernel, write $\mathbf{x} = \sum_j x_j\mathbf{b}_j$ and use that $\mathbf{c}_1, \ldots, \mathbf{c}_k$ are independent.
Sketch. With $GAH = N_k$, let $\mathbf{b}_j$ be the columns of H and $\mathbf{c}_i$ those of G^{-1}; both lists are bases by Theorem 3.5.14, since H and G^{-1} are invertible. Column j of $AH = G^{-1}N_k$ is $A\mathbf{b}_j$ on the left and $G^{-1}\mathbf{e}_j = \mathbf{c}_j$ (for $j \le k$) or $\mathbf{0}$ (for $j > k$) on the right. Then $A\sum_j x_j\mathbf{b}_j = \sum_{j\le k} x_j\mathbf{c}_j$, which is $\mathbf{0}$ iff $x_1 = \cdots = x_k = 0$ and ranges over $\operatorname{span}(\mathbf{c}_1, \ldots, \mathbf{c}_k)$.

By Theorem 3.5.27 there are invertible G and H with $GAH = N_k$, so $AH = G^{-1}N_k$. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be the columns of H and $\mathbf{c}_1, \ldots, \mathbf{c}_m$ the columns of G^{-1}. Since T_H and $T_{G^{-1}}$ are invertible (Remark 3.5.7), the columns of H form a basis of F^n and those of G^{-1} a basis of F^m, by Theorem 3.5.14 ((V) $\Rightarrow$ (III)).

By Definition 2.4.1, column j of AH is $A\mathbf{b}_j$, and column j of $G^{-1}N_k$ is G^{-1} times column j of N_k. That column is $\mathbf{e}_j \in F^m$ if $j \le k$ and $\mathbf{0}$ if $j > k$, so column j of $G^{-1}N_k$ is $\mathbf{c}_j$ if $j \le k$ and $\mathbf{0}$ if $j > k$. This proves the displayed equations. By Definition 2.3.1, column j of $[T_A]_{\mathcal{C}\mathcal{B}}$ is the coordinate vector of $A\mathbf{b}_j$ in the basis $\mathcal{C}$, which is $\mathbf{e}_j$ for $j \le k$ and $\mathbf{0}$ for $j > k$; so $[T_A]_{\mathcal{C}\mathcal{B}} = N_k$.

Every $\mathbf{x} \in F^n$ is $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$ for unique scalars x_j, and by linearity

$$A\mathbf{x} = x_1\mathbf{c}_1 + \cdots + x_k\mathbf{c}_k.$$

Since $\mathbf{c}_1, \ldots, \mathbf{c}_k$ is linearly independent (a sublist of a basis, Corollary 1.5.17), $A\mathbf{x} = \mathbf{0}$ if and only if $x_1 = \cdots = x_k = 0$, that is, if and only if $\mathbf{x} \in \operatorname{span}(\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n)$. And as $\mathbf{x}$ ranges over F^n, the scalars $x_1, \ldots, x_k$ range over all of F, so $A\mathbf{x}$ ranges over $\operatorname{span}(\mathbf{c}_1, \ldots, \mathbf{c}_k)$. □

So every linear map $F^n \to F^m$ of rank k is, in suitable coordinates, the map $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_k, 0, \ldots, 0)$: it keeps k coordinates and kills the rest. Choosing bases to make a map simple is a theme of the rest of the book (Section 4.5, and diagonalization in Section 7.4). The corollary also previews Section 4.2. In the first proof of Theorem 3.5.27, the last $n-k$ columns of H are the basic solutions, and the first k columns of G^{-1} are $A\mathbf{e}_{p_1}, \ldots, A\mathbf{e}_{p_k}$, the pivot columns of A. So the kernel is spanned by $n-k$ vectors, one for each free variable, and the image by the k pivot columns of A. Once dimension is defined, this says $\dim\ker T_A + \dim\operatorname{im} T_A = n$.

Is k the only number for which A has a normal form N_k? It is: if $GAH = N_j$ for some invertible G and H, then $j = \operatorname{rank} A$. This amounts to $\operatorname{rank}(AH) = \operatorname{rank} A$ for invertible H, and is proved in Exercise 3.5.41. It follows that two matrices $A, B \in F^{m\times n}$ satisfy $B = GAH$ for some invertible G, H if and only if they have the same rank: rank is the only invariant of a matrix under row and column operations together.

Exercise 3.5.30. ROUTINE Let $A = \begin{pmatrix} 1 & i & 1+i \\ i & -1 & -1+i \end{pmatrix} \in \mathbb{C}^{2\times 3}$. Find invertible G and H with $GAH = N_k$, where $k = \operatorname{rank} A$.

Solution on page 1252

Exercises

Exercise 3.5.31. STANDARD For a scalar t, let

$$A_t = \begin{pmatrix} t & 1 & 1 \\ 1 & t & 1 \\ 1 & 1 & t \end{pmatrix}.$$

(a) Find $\operatorname{rank} A_t$ for every $t \in \mathbb{R}$.

(b) Find $\operatorname{rank} A_t$ for every $t \in \mathbb{F}_3$. Which rank that occurs over $\mathbb{R}$ never occurs over $\mathbb{F}_3$, and why?

Solution on page 1252

Exercise 3.5.32. ROUTINE Over $\mathbb{F}_5$, let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{pmatrix}$. Show that $\operatorname{rank} A = 2$, find a left inverse of A with the first approach to Proposition 3.5.9, and find a second left inverse. How many left inverses does A have? (Use Exercise 3.5.33 for the count.)

Solution on page 1253

Exercise 3.5.33. STANDARD Let $A \in F^{m\times n}$, and let $G \in F^{m\times m}$ be invertible with $GA = A_r$.

(a) Suppose $\operatorname{rank} A = m$, and let X_0 be a right inverse of A. Show that the right inverses of A are exactly the matrices $X_0 + K$ with $AK = 0$, that is, with every column of K in $\ker T_A$.

(b) Suppose $\operatorname{rank} A = n$, and let C_0 consist of the first n rows of G. Show that $C \in F^{n\times m}$ is a left inverse of A if and only if every row of $C - C_0$ is a linear combination of the last $m - n$ rows of G.

(c) Let F be finite with q elements. Show that a matrix of rank m has exactly $q^{m(n-m)}$ right inverses, and a matrix of rank n has exactly $q^{n(m-n)}$ left inverses.

(d) Over an arbitrary field, show that a matrix with a one-sided inverse has exactly one such inverse if and only if it is square.

Solution on page 1254

Exercise 3.5.34. STANDARD Prove or disprove each statement. Matrices are over a field F and have the indicated sizes.

(a) If $A \in F^{m\times n}$ has rank n and $B \in F^{n\times p}$ has rank p, then AB has rank p.

(b) If $A \in F^{m\times n}$ has rank m and $B \in F^{n\times p}$ has rank n, then AB has rank m.

(c) For every $A \in F^{n\times n}$, $\operatorname{rank}(A^2) = \operatorname{rank} A$.

(d) If $A \in F^{m\times n}$ has rank n, then A^{T} has rank n.

(e) The matrices G and H in Theorem 3.5.27 are determined by A.

(f) If $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $m > n$, then AB is not invertible.

Solution on page 1255

Exercise 3.5.35. STANDARD *Claim.* Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $BA = I_n$. Then $AB = I_m$.

Proof. Let $P = AB$. Then $P^2 = A(BA)B = AI_nB = AB = P$. Multiplying $P^2 = P$ on the left by P^{-1} gives $P = I_m$.

(a) Find the error.

(b) Give a counterexample to the claim.

(c) Show that whenever $BA = I_n$, the matrix $P = AB$ satisfies $P^2 = P$, and that $P = I_m$ if and only if $m = n$.

Solution on page 1255

Exercise 3.5.36. STANDARD Let $A \in F^{m\times n}$, $B \in F^{p\times q}$, and $C \in F^{m\times q}$.

(a) Show that $\operatorname{rank}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \operatorname{rank} A + \operatorname{rank} B$.

(b) Show that $\operatorname{rank}\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \geq \operatorname{rank} A + \operatorname{rank} B$, and give an example where the inequality is strict.

Solution on page 1255

Exercise 3.5.37. STANDARD Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$. Show that exactly one of the following holds:

(i) $A\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x} \in F^n$;

(ii) there is $\mathbf{z} \in F^m$ with $\mathbf{z}^{\mathsf{T}}A = \mathbf{0}$ and $\mathbf{z}^{\mathsf{T}}\mathbf{y} \neq 0$.

Then find such a $\mathbf{z}$ for the inconsistent system of Example 3.4.6, with $\mathbf{y} = (4, 2, 6)$.

Solution on page 1256

Exercise 3.5.38. STANDARD Let $A, B, C \in F^{n\times n}$ with $ABC = I_n$. Prove or disprove:

(a) $BCA = I_n$;

(b) $CAB = I_n$;

(c) $BAC = I_n$.

Solution on page 1256

Exercise 3.5.39. HARD Let $A, B \in F^{n\times n}$ satisfy $AB = A + B$. Show that $AB = BA$.

Solution on page 1257

Exercise 3.5.40. STANDARD Let V and W be vector spaces over F, and let $T : V \to W$ be linear.

(a) Suppose V has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and W is spanned by a finite list $\mathbf{w}_1, \ldots, \mathbf{w}_p$. Show that if T is injective, then T has a left inverse.

(b) Suppose W has a basis $\mathbf{c}_1, \ldots, \mathbf{c}_m$. Show that if T is surjective, then T has a right inverse. (No assumption on V is needed.)

Solution on page 1257

Exercise 3.5.41. HARD This exercise shows that the number k in Theorem 3.5.27 is determined by A, and draws the consequences. Throughout, $N_j \in F^{m\times n}$ is as in Theorem 3.5.27.

(a) Let $Q \in F^{n\times n}$ be invertible and $k \leq \min(m, n)$. Show that $\operatorname{rank}(N_k Q) = k$.

(b) Let $A \in F^{m\times n}$. Show that if $GAH = N_j$ for some invertible G and H, then $j = \operatorname{rank} A$.

(c) Show that $\operatorname{rank}(AH) = \operatorname{rank} A$ for every invertible $H \in F^{n\times n}$.

(d) Show that $A, B \in F^{m\times n}$ satisfy $B = GAH$ for some invertible G, H if and only if $\operatorname{rank} A = \operatorname{rank} B$.

Solution on page 1258

Exercise 3.5.42. STANDARD A matrix $P \in F^{n\times n}$ is **idempotent** if $P^2 = P$. Show that every $A \in F^{n\times n}$ can be written as $A = SP$ with S invertible and P idempotent, and also as $A = P'S'$ with P' idempotent and S' invertible.

Solution on page 1259

Exercise 3.5.43. HARD Let $A \in F^{m\times n}$. A matrix $X \in F^{n\times m}$ is a **generalized inverse** of A if $AXA = A$ and $XAX = X$.

(a) Show that every A has a generalized inverse.

(b) Show that if A has a left inverse or a right inverse, then every left inverse or right inverse of A is a generalized inverse.

(c) Let X satisfy $AXA = A$. Show that whenever $A\mathbf{x} = \mathbf{y}$ is consistent, $\mathbf{x} = X\mathbf{y}$ is a solution.

Solution on page 1259

Exercise 3.5.44. HARD Let $A \in F^{n\times n}$. Show that A is not invertible if and only if there is a nonzero $B \in F^{n\times n}$ with $AB = BA = 0$.

Solution on page 1260

Exercise 3.5.45. HARD

(a) Let F be a field with at least three elements. Show that every $A \in F^{n\times n}$ is the sum of two invertible matrices.

(b) Show that (a) fails for $F = \mathbb{F}_2$ and $n = 1$.

Solution on page 1260

Exercise 3.5.46. CHALLENGE Show that for $n \geq 2$, every $A \in \mathbb{F}_2^{n\times n}$ is the sum of two invertible matrices. (Compare Exercise 3.5.45.)

Solution on page 1260

3.6 Proof patterns

Nearly every proof in this chapter ran through Gauss–Jordan elimination, but it used the algorithm in only a handful of ways: as a reduction to a simpler matrix, as an object whose correctness had to be proved, as a way to settle existence questions, and as a table from which structure can be read. This section sorts the chapter's proofs by the question they answer, compares the tools on examples from the chapter, and says when each tool is the wrong one. The exercises at the end practice one move at a time, and the chapter's problems (Problems for Chapter 3) mix them.

3.6.1 Proving a statement about every matrix

The chapter has two reductions. Technique 3.3.22 replaces A by $\operatorname{rref}(A) = GA$ with G invertible (Theorem 3.3.10). Technique 5.4.6, which is named in Section 5.4 but first used here, writes an invertible matrix as a product of elementary matrices (Theorem 3.5.19) and checks the three types one at a time. The rank normal form $GAH = N_k$ (Theorem 3.5.27) is a two-sided version of the first. To choose between them, ask what the property you want has to survive.

Reducing to RREF needs a property P with $P(A) \Leftrightarrow P(GA)$ for every invertible G. The chapter found many: the kernel (Corollary 3.2.23), the linear relations among the columns (Corollary 3.2.17), and

everything defined from these, such as which columns are pivot columns (Corollary 3.3.19), the rank (Proposition 3.5.2(c)), and, for square A, invertibility. Solution sets survive only if the right-hand side moves too: $A\mathbf{x} = \mathbf{y}$ becomes $\operatorname{rref}(A)\mathbf{x} = G\mathbf{y}$ (Proposition 3.2.14). Once invariance is checked, the question is about a matrix whose pivot columns are $\mathbf{e}_1, \ldots, \mathbf{e}_r$ and whose zero rows are visible, and it usually answers itself. This is how Theorem 3.3.23, Proposition 3.4.15, and Corollary 3.3.20 were proved.

Reducing to elementary matrices needs something else: a statement about invertible matrices that survives products. Then two checks suffice, one computation for each type and one step showing that the statement passes from A to EA (or AE). Invariance under $A \mapsto GA$ plays no role. This is how Corollary 3.5.22 turned "multiply by invertible matrices" into "perform row and column operations", and how Section 5.4 proves $\det(AB) = \det A \det B$. The method only reaches invertible matrices, because products of elementary matrices are invertible. A statement about all square matrices needs a second argument for the others. Often both sides of the claim simply vanish there; otherwise one reduces to RREF, writes $A = G^{-1}\operatorname{rref}(A)$, and uses the zero row of $\operatorname{rref}(A)$ (Exercise 3.6.10). Forgetting this case is the error in Exercise 5.4.35.

Example 3.6.1 (Which reduction fits). *Commuting with every invertible matrix.* Which $M \in F^{n\times n}$ satisfy $MX = XM$ for every invertible X? Reducing to RREF is useless here, since M commuting with X says nothing about M commuting with GX. But the property survives products: if $MX = XM$ and $MY = YM$, then $MXY = XMY = XYM$. So by Theorem 3.5.19, M commutes with every invertible matrix if and only if it commutes with every elementary one, and for $n \geq 2$ the type 3 matrices alone already force $M = \lambda I_n$ (Exercise 3.2.52; for $n = 1$ every M works).

Rank under row and column operations. That $\operatorname{rank}(GA) = \operatorname{rank} A$ for invertible G is one line by reducing to RREF: $\operatorname{rref}(GA) = \operatorname{rref}(A)$ (Corollary 3.3.18(a)). The same attempt for AH fails. With $A = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $H = E_{12}$ we get $AH = \begin{pmatrix} 1 & 0 \end{pmatrix}$, and A and AH are both in RREF, so $\operatorname{rref}(AH) \neq \operatorname{rref}(A)$, although both have rank 1. Right multiplication does not preserve the reduced form, so the reduction cannot see that $\operatorname{rank}(AH) = \operatorname{rank} A$. A different argument is needed: the uniqueness of the rank normal form (Exercise 3.5.41), or transposition, which turns the column operations into row operations of A^{T} and so reduces the question to the equality $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$ of Section 4.3.

3.6.2 Proving that an algorithm is correct

An algorithm proves that something exists once two things are shown about every run (Technique 3.3.5). For *termination*, find a quantity that moves in one direction and cannot do so forever: in Gauss–Jordan elimination, the number k of pivoted rows grows by 1 in each pass and never exceeds m (Proposition 3.3.6). For *correctness*, find an invariant that the starting state satisfies and every step preserves, and that, combined with the reason for halting, gives the property of the output.

The invariant is where the work is, and its typical feature is visible in Lemma 3.3.7. The property we want at the end, conditions (R1)–(R4), is not preserved pass by pass: halfway through, the unpivoted rows are an arbitrary mess. The invariant has to say something about that mess, and part (e) does: the unpivoted rows vanish in every column up to the last pivot. Part (e) is what forces each new pivot to lie to the right of the old ones, which is (b), and it, together with the vanishing of the new pivot row to the left of its pivot, is exactly what choosing the *leftmost* column in step 1 of Algorithm 3.3.1 keeps true. Drop that choice and both the proof and the algorithm break (Exercise 3.3.47). The same pattern recurs in the exercises: Euclid's algorithm by row operations has the monovariant $|b|$ and the invariant "the common divisors of the two entries" (Exercise 3.2.54), and elimination on a strictly diagonally dominant matrix needs no swap because the block still to be processed stays diagonally dominant (Exercise 3.3.56). Even Theorem 3.3.12 was proved by running the algorithm, on a matrix of the right shape, with the invariant "nothing has changed yet". Exercise 3.6.11 practices the technique on an algorithm that has nothing to do with elimination.

An algorithm proves that *some* output with the guaranteed properties exists. It does not prove that this output is the only object with those properties, and a run that made different choices could, as far as the invariant knows, stop somewhere else. That is why the uniqueness of the reduced row echelon form needed its own theorem (Theorem 3.3.17), and why, before it, we could not say that another way of reducing A finds the same pivots, or that $\operatorname{rank}(GA) = \operatorname{rank} A$ (Remark 3.3.2). For echelon forms that are not reduced, uniqueness really fails (Example 3.3.29).

3.6.3 Turning an existence question into a linear system

"Is there a polynomial, a combination, a map, a set of weights, such that ...?" The move of Technique 3.1.9 is to write the unknown object with unknown scalars, translate each requirement into an equation in them, and check that the equations are linear. The question has become the consistency of a linear system, and the set of all answers has the shape of Technique 3.1.18: one answer plus everything that solves the homogeneous version. The chapter used this for membership in a span (Example 3.1.10), interpolation (Theorem 3.4.24), chemical balancing (Exercise 3.1.39), quadrature weights (Exercise 3.4.43), and antiderivatives, where the homogeneous solutions are the constants (Example 3.1.20).

Example 3.6.2 (One question, three answers). Which polynomials take the values $2, 1, 4$ at $-1, 0, 1$? The answer depends on how many coefficients we allow, that is, on the number of unknowns against the three equations (Exercise 3.1.37).

- In $\mathbb{R}[x]_{\leq 1}$, write $p = a + bx$. The conditions $a - b = 2$, $a = 1$, $a + b = 4$ give $b = -1$ and $b = 3$, so there is no solution: three equations, two unknowns, and an inconsistent system.
- In $\mathbb{R}[x]_{\leq 2}$ there are three unknowns and exactly one solution, $1 + x + 2x^2$.
- In $\mathbb{R}[x]_{\leq 3}$ there are four unknowns, and the solutions are $1 + x + 2x^2 + t\big(x^3 - x\big)$ for $t \in \mathbb{R}$: the particular solution from the previous case plus the cubics that vanish at $-1, 0, 1$.

The count of unknowns against equations predicts what can happen (Corollary 3.4.19), but only the system decides. The square case has exactly one solution here because of a structural fact, which is the subject of the next subsection.

The technique is the wrong tool when the requirements are not linear in the unknowns. Asking for a matrix X with $X^2 = A$ gives quadratic equations in the entries of X, and elimination does not apply. Side conditions such as positivity or integrality are not linear either: the linear system only finds the subspace of all balancings in Exercise 3.1.39, and a separate divisibility argument picks the smallest positive integer one; the nonnegative flows of Exercise 3.1.38 come from an inequality on the parameter. Also remember that the system lives over a field, and solving it divides by scalars: in Example 3.1.10 the same equations are solvable over $\mathbb{R}$ and inconsistent over $\mathbb{F}_2$, because one step divided by 2.

3.6.4 When uniqueness gives existence

Uniqueness and existence behave differently. Uniqueness is a property of the homogeneous problem alone, for a matrix of any size: a consistent system has exactly one solution if and only if $\ker T_A = \{\mathbf{0}\}$ (Corollary 3.1.22). Existence for every right-hand side is a property of the zero rows of $\operatorname{rref}(A)$ (Proposition 3.4.16). For a square matrix the two conditions coincide, since both say $\operatorname{rank} A = n$ (Corollary 3.4.20). So for square problems the homogeneous statement buys existence for free (Technique 3.4.21). The homogeneous statement says that a single object is zero, and it is usually proved by structure rather than computation: a polynomial with too many roots (Lemma 3.4.23), a maximum principle (Exercise 3.1.47), or the observation that $BA = I_n$ leaves no room for free variables (Theorem 3.5.23).

Example 3.6.3 (Existence, with and without a formula). Interpolation (Theorem 3.4.24) was proved without solving anything: $n + 1$ conditions on $n + 1$ coefficients, and a polynomial of degree at most n with $n + 1$ roots is zero. The Lagrange polynomials (Definition 3.4.26) were then *defined* through the theorem,

so Lagrange's formula did not give a second existence proof. That needs the closed form $L_i = \prod_{j \neq i}(x - c_j)/(c_i - c_j)$, which is a construction, and which waits for polynomial multiplication in Section 7.1. The construction gives more, a formula, but only when a formula is available. For the plate of Exercise 3.1.47 there is none, and uniqueness is the natural route to existence.

There are three ways to misuse the technique. First, on a non-square system it proves nothing: $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ has trivial kernel and $Ax = (1, 0)$ has no solution (Exercise 3.1.34). Second, counting conditions is not the proof; the homogeneous problem must actually have only the zero solution, and when it does not, existence fails too for some data. In the cyclic system of Exercise 3.1.46 with n even, the homogeneous solution $(1, -1, \ldots, 1, -1)$ comes with the condition that the alternating sum of the right-hand sides vanish, since the alternating sum of the left sides is 0. Exercise 3.6.14 is a smaller instance. Third, the technique depends on finite size. On $F[x]$, the shift S with $S\left(x^k\right) = x^{k+1}$ has a left inverse but is not surjective (the intuition after Theorem 3.5.23), and Section 4.2 replaces "square" by $\dim V = \dim W$.

3.6.5 Reading structure off the reduced form

Once $GA = R = \operatorname{rref}(A)$ is computed, a great deal can be read off without further work:

- a pivot in every column means at most one solution for each $\mathbf{y}$, independent columns, and injectivity of T_A (Proposition 3.4.15, Theorem 3.5.12); a pivot in every row means solutions for every $\mathbf{y}$, spanning columns, and surjectivity (Proposition 3.4.16, Theorem 3.5.13);
- the pivot columns are the columns of A that are not combinations of earlier ones, and the entries of R in a free column are the coefficients that build that column of A from the pivot columns (Corollary 3.3.19);
- the basic solutions, one for each free column, describe $\ker T_A$ (Theorem 3.4.9);
- the last $m - \operatorname{rank} A$ rows of G give the consistency conditions, so they describe $\operatorname{im} T_A$ (Corollary 3.4.4), and they are also the relations among the rows of A (Exercise 3.3.57).

Everything in the first three items is about relations among the columns, and R carries it alone. The image is different: it needs G, or the original columns of A.

Example 3.6.4 (Everything about one matrix). For the matrix A of Example 3.3.3, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_4$,

$$A = \begin{pmatrix} 0 & 0 & 2 & 4 \\ 2 & 4 & 2 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix},$$

$$R = \begin{pmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$G = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ -1 & -1/2 & 1 \end{pmatrix}.$$

The rank is 2, with pivot columns 1, 3 and free columns 2, 4. From the entries of R, $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_4 = -2\mathbf{a}_1 + 2\mathbf{a}_3$. The basic solutions $(-2, 1, 0, 0)$ and $(2, 0, -2, 1)$ describe $\ker T_A$ (Example 3.4.10). The last row of G gives the single consistency condition $-y_1 - y_2/2 + y_3 = 0$, so $\operatorname{im} T_A$ is the plane $2y_3 = 2y_1 + y_2$ (Example 3.4.6), spanned by the pivot columns $\mathbf{a}_1 = (0, 2, 1)$ and $\mathbf{a}_3 = (2, 2, 3)$ of A.

The classic error is to take the pivot columns of R instead. They are $\mathbf{e}_1$ and $\mathbf{e}_2$, and $\mathbf{e}_1 = (1, 0, 0)$ is not in $\operatorname{im} T_A$, since $2 \cdot 0 \neq 2 \cdot 1 + 0$. The columns of R span $\operatorname{im} T_R$, which is $\operatorname{im} T_A$ moved by G (Non-example 3.2.19). The fix is to read *which* columns are pivot columns from R and to take those columns *from* A. Exercise 3.6.6 and Exercise 3.6.7 practice the distinction.

3.6.6 Row operations and column operations

Row operations multiply on the left by an invertible G, and column operations multiply on the right by an invertible H. They preserve different things (Table 3.6.1).

Property of A	Row operations, $A \mapsto GA$	Column operations, $A \mapsto AH$
$\ker T_A$	kept (Corollary 3.2.23)	moved by H^{-1}
relations among the columns	kept (Corollary 3.2.17)	changed
$\operatorname{im} T_A$, the span of the columns	moved by G	kept (Proposition 3.2.30)
relations among the rows	changed	kept (Exercise 3.6.8)
$\operatorname{rref}(A)$ and the pivot columns	kept (Corollary 3.3.18)	changed
$\operatorname{rank} A$	kept (Proposition 3.5.2)	kept (Exercise 3.5.41)

Table 3.6.1. What the two kinds of operations preserve. "Moved by G" means that the new image is $\{G\mathbf{y} \mid \mathbf{y} \in \operatorname{im} T_A\}$, and similarly for the kernel, which becomes $\{H^{-1}\mathbf{x} \mid \mathbf{x} \in \ker T_A\}$. Transposing turns column operations on A into row operations on A^{T} (Technique 3.2.27); for instance, the relations among the rows of A form the kernel of $T_{A^{\mathsf{T}}}$, which is why column operations keep them.

Example 3.6.5 (One matrix, both operations). Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, so $\ker T_A = \operatorname{span}((1,-1))$ and $\operatorname{im} T_A = \operatorname{span}((1,1))$. The row operation $E_{1,-1,2}$ gives $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, with the same kernel and the image $\operatorname{span}((1,0))$. The column operation "add -1 times column 1 to column 2", which is right multiplication by $H = E_{2,-1,1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, gives $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, with the same image and the kernel $\operatorname{span}((0,1))$. Indeed $H^{-1}(1,-1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = (0,-1)$, as the table predicts.

So row operations are the tool for solving systems and for finding relations among the columns, and column operations are the tool for questions about the span of the columns. Using column operations to solve $A\mathbf{x} = \mathbf{y}$ changes the unknowns: $AH\mathbf{x}' = \mathbf{y}$ is solved by $\mathbf{x}'$ exactly when $A\mathbf{x} = \mathbf{y}$ is solved by $\mathbf{x} = H\mathbf{x}'$, so the answer must be translated back. Using row operations on a question about the span of the columns moves the span. This is why the bad right-hand side in the proof of Proposition 3.4.16 is $G^{-1}\mathbf{e}_m$ and not $\mathbf{e}_m$ (Remark 3.4.17). With both kinds of operations allowed, only the rank survives, and every matrix of rank k becomes N_k (Theorem 3.5.27).

3.6.7 Reading the chapter's proofs

Uniqueness of the reduced form, two ways. Writing $R' = GR$ column by column suggests the first proof of Theorem 3.3.17: track G, and show pivot by pivot, from the left, that the pivot columns agree and that G fixes $\mathbf{e}_1, \ldots, \mathbf{e}_r$. It needs invertibility of G in a concrete form, the independence of its columns, to rule out a pivot that appears too early in one matrix. The second proof forgets G. It notices that everything about a matrix in RREF, which columns are pivot columns and what the other columns are, is determined by the linear relations among its columns (Lemma 3.3.16), and that row operations preserve those relations. The comparison teaches a habit: after a proof, list the hypotheses it actually used. The second proof used one consequence of $R' = GR$, namely that R and R' have the same kernel, so it proves the stronger statement Corollary 3.3.20 at no extra cost. The first proof, in turn, says something the second does not: G is rigid on $\mathbf{e}_1, \ldots, \mathbf{e}_r$ and free elsewhere (Exercise 3.3.46).

Constructive and structural proofs. Several results got one proof of each kind. Proposition 3.4.15 was proved by setting one free variable to 1, which produces a nonzero solution, and by the linear dependence lemma, which explains what a missing pivot means. Proposition 3.5.9 was proved from the pivots, which gives a formula, and by extending the columns to a basis, which works for linear maps between abstract vector

spaces as long as the domain has a finite basis and the target a finite spanning list (Exercise 3.5.40). Theorem 3.5.27 was proved by clearing columns and by row reducing the transpose. The computational proof usually produces a formula or a witness, and the structural one usually survives the passage to abstract vector spaces.

Do it to the identity. To find the matrix of an operation that acts on each column by the same linear map, perform the operation on I_m (Lemma 3.2.6, the first proof of Proposition 3.2.7). The same idea, run along with the algorithm, records the whole elimination: carrying I_m next to A produces G (Proposition 3.3.11), and carrying I_n next to a square A produces A^{-1} (Algorithm 3.3.24).

3.6.8 Common errors

Each of these appears in the chapter, most of them as find-the-error exercises.

- *Simultaneous row operations.* Replacing two equations at once, where the second replacement uses an equation the first one changed, can gain solutions (Exercise 3.1.35). So can scaling a row by 0, which is why it is not an elementary operation.
- *The order of the factors.* The first row operation is the rightmost factor of $G = E_t \cdots E_1$ (Example 3.2.8); for column operations it is the leftmost (Remark 3.2.29).
- *Row operations do not preserve the image.* $A\mathbf{x} = \mathbf{y}$ is equivalent to $GA\mathbf{x} = G\mathbf{y}$, not to $GA\mathbf{x} = \mathbf{y}$ (Exercise 3.2.43, Remark 3.4.17), and the columns of $\operatorname{rref}(A)$ do not span $\operatorname{im} T_A$ (Exercise 3.6.6).
- *Uniqueness is not existence.* A trivial kernel gives at most one solution, not one (Exercise 3.1.34). Checking that some vectors are solutions does not show that they are all of them (Exercise 3.4.34, Remark 3.4.12). Counting equations and unknowns does not replace proving uniqueness (Exercise 3.6.14).
- *The shape bounds the rank only from above.* $\operatorname{rank} A \leq \min(m, n)$, and equality is a condition to check, not a consequence of the shape (Problem 3.7).
- *One-sided inverses of non-square matrices.* Cancelling a matrix that has only a right inverse, or inverting AB when A is not square, is invalid (Exercise 3.5.26, Exercise 3.5.35).
- *Rigidity overstated.* $GR = R$ with R in RREF does not force $G = I_m$ (Exercise 3.3.46).
- *Reducing to elementary matrices beyond its reach.* Only invertible matrices are products of elementary matrices; the other case needs its own argument (Exercise 5.4.35, Problem 3.7).
- *Hidden division.* Pivots, ranks, and inverses depend on the field, because elimination divides (Example 3.5.3, Exercise 3.3.27).
- *Pictures from the plane.* Three planes in $\mathbb{R}^3$, no two of them parallel, can still have no common point (Exercise 3.1.36).

Exercises

Exercise 3.6.6. ROUTINE Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 3 & 6 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}, \quad \operatorname{rref}(A) = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

A student argues: "Row operations do not change the linear relations among the columns, and the pivot columns of $\operatorname{rref}(A)$ are independent. So $\operatorname{im} T_A$ has the independent spanning list $(1, 0, 0), (0, 1, 0)$, the pivot columns of $\operatorname{rref}(A)$."

(a) Show that $(1, 0, 0) \notin \operatorname{im} T_A$, and find the error.

(b) Find two vectors that span $\operatorname{im} T_A$, and an equation for $\operatorname{im} T_A$.

Solution on page 1261

Exercise 3.6.7. ROUTINE A matrix $A \in \mathbb{R}^{3\times 4}$ has

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and its first and third columns are $\mathbf{a}_1 = (1, 1, 2)$ and $\mathbf{a}_3 = (0, 1, -1)$.

(a) Find A.

(b) Find vectors spanning $\ker T_A$, and an equation describing $\operatorname{im} T_A$.

Solution on page 1262

Exercise 3.6.8. STANDARD Let $A \in F^{m\times n}$ and let $H \in F^{n\times n}$ be invertible. Decide which of the following are always the same for A and for AH, with a proof or a counterexample.

(a) The image $\operatorname{im} T_A$.
(b) The kernel $\ker T_A$.
(c) The set of row vectors $\mathbf{z} \in F^{1\times m}$ with $\mathbf{z}A = \mathbf{0}$ (the relations among the rows).
(d) The reduced row echelon form.
(e) The rank.
(f) Whether there is a left inverse.

Solution on page 1262

Exercise 3.6.9. HARD Let $A \in F^{m\times n}$ and $B \in F^{p\times n}$.

(a) Show that $\ker T_A \subseteq \ker T_B$ if and only if $B = XA$ for some $X \in F^{p\times m}$.

(b) Deduce that $\ker T_A \subseteq \ker T_B$ if and only if every row of B is a linear combination of the rows of A.

Solution on page 1263

Exercise 3.6.10. STANDARD For $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, let $D(X) = ad - bc$. Prove that $D(XY) = D(X)D(Y)$ for all $X, Y \in F^{2\times 2}$, by reducing to elementary matrices: first for invertible X, then separately for X that are not invertible. You may use Exercise 3.2.42.

Solution on page 1264

Exercise 3.6.11. STANDARD Start with a real $m \times n$ matrix. A **move** chooses a row or a column whose entries have a negative sum and multiplies every entry of it by -1.

(a) Show that every sequence of moves is finite, whatever choices are made.

(b) Deduce that some finite sequence of moves produces a matrix all of whose row sums and column sums are nonnegative.

(c) Express the final matrix as $D_1 A D_2$, where A is the starting matrix and D_1, D_2 are products of elementary matrices of one type.

Solution on page 1264

Exercise 3.6.12. STANDARD

(a) Find all $p \in \mathbb{Q}[x]_{\leq 4}$ such that $p(t+1) - p(t) = t^3$ for every $t \in \mathbb{Q}$.

(b) Deduce that $0^3 + 1^3 + \cdots + (N-1)^3 = (N(N-1)/2)^2$ for every integer $N \geq 1$.

Solution on page 1265

Exercise 3.6.13. STANDARD Let $n \geq 3$. Read indices modulo n, so that $x_0 = x_n$ and $x_{n+1} = x_1$.

(a) Show that for every $\mathbf{y} \in \mathbb{R}^n$ there is exactly one $\mathbf{x} \in \mathbb{R}^n$ with

$$4x_i - x_{i-1} - x_{i+1} = y_i \quad (i = 1, \ldots, n).$$

(b) Now replace 4 by 2. Show that the system $2x_i - x_{i-1} - x_{i+1} = y_i$ $(i = 1, \ldots, n)$ has a solution if and only if $y_1 + \cdots + y_n = 0$. You may use Exercise 3.5.37.

Solution on page 1266

Exercise 3.6.14. ROUTINE For $p = a + bx + cx^2 \in \mathbb{R}[x]_{\leq 2}$, write $p' = b + 2cx$. Find the error in the following argument, and decide for which $(u, v, w) \in \mathbb{R}^3$ the conclusion holds.

Claim. For all $u, v, w \in \mathbb{R}$ there is exactly one $p \in \mathbb{R}[x]_{\leq 2}$ with $p(0) = u$, $p(1) = v$, and $p'(1/2) = w$.

Proof. The three conditions are linear equations in the three unknown coefficients of p. The system is square, so by Technique 3.4.21 it has exactly one solution.

Solution on page 1267

Problems for Chapter 3

These problems combine the sections of Chapter 3, and some of them combine it with Chapters 1 and 2. The first four are review prompts: state a main result, prove it without looking, and then answer a question that the text did not ask. The rest are ordered roughly by difficulty. Problems of kind "competition" are harder than anything the course expects. Section 3.6 surveys the techniques they use.

Review: state and prove

Problem 3.1. STANDARD

(a) State the theorem that describes the solution set of a consistent linear equation $T(\mathbf{x}) = \mathbf{y}$, where $T : V \to W$ is linear, and prove it.

(b) Let $A \in F^{m \times n}$. Show that the set of $\mathbf{y} \in F^m$ for which $A\mathbf{x} = \mathbf{y}$ has *exactly one* solution is either empty or all of $\operatorname{im} T_A$, and decide which in terms of $\ker T_A$.

Solution on page 1267

Problem 3.2. STANDARD

(a) State the theorem on the uniqueness of the reduced row echelon form, and prove it by comparing the linear relations among the columns.

(b) Show that the analogous statement for column operations is false: find matrices $R \neq R'$, both in RREF, and an invertible H with $R' = RH$.

(c) Show that nevertheless, if R is in RREF, H is invertible, and RH is also in RREF, then R and RH have the same number of pivots.

Solution on page 1268

Problem 3.3. STANDARD

(a) State the criterion for invertibility of a square matrix in terms of its reduced row echelon form, prove it, and explain why Algorithm 3.3.24 is correct.

(b) Suppose that while row reducing $(A \mid I_n)$, at some intermediate stage, row i of the left block is zero. Show that row i of the right block is then a nonzero row vector $\mathbf{z}$ with $\mathbf{z}A = \mathbf{0}$, and deduce directly from this, without using (a), that A is not invertible.

(c) Carry out (b) for the integer matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 2 & 9 \end{pmatrix}$, read over $\mathbb{F}_5$. Is A invertible over $\mathbb{Q}$?

Solution on page 1269

Problem 3.4. STANDARD

(a) State the rank normal form theorem and prove it.

(b) Deduce that every $A \in F^{m\times n}$ of rank $k \geq 1$ is a sum of k matrices of rank 1.

(c) Deduce that every $A \in F^{m\times n}$ of rank $k \geq 1$ can be written as $A = CS$ with $C \in F^{m\times k}$ and $S \in F^{k\times n}$.

Solution on page 1270

Synthesis problems

Problem 3.5. ROUTINE Let

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 & 3 \\ 2 & 4 & 1 & 4 & 1 \\ 1 & 2 & 1 & 3 & 3 \end{pmatrix}.$$

(a) Find $\operatorname{rref}(A)$ and $\operatorname{rank} A$ over $\mathbb{Q}$ and over $\mathbb{F}_5$. Find a relation among the rows of A over $\mathbb{F}_5$ that explains the difference.

(b) Over $\mathbb{F}_5$, write each free column of A as a combination of the pivot columns, and find vectors spanning $\ker T_A$.

(c) Over $\mathbb{F}_5$, for which $\mathbf{y}$ is $A\mathbf{x} = \mathbf{y}$ consistent, and how many such $\mathbf{y}$ are there? Find all solutions of $A\mathbf{x} = (1, 1, 0)$, and count them.

(d) Over $\mathbb{Q}$, find a right inverse of A. Does A have a right inverse over $\mathbb{F}_5$?

Solution on page 1271

Problem 3.6. ROUTINE Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. For each of the fields $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_2$, $\mathbb{F}_3$, and $\mathbb{F}_5$, find every scalar λ for which $A - \lambda I_2$ is not invertible, and for each such λ find all $\mathbf{x}$ with $A\mathbf{x} = \lambda\mathbf{x}$.

Solution on page 1272

Problem 3.7. ROUTINE Find the error in each argument, and give a counterexample to its conclusion.

(a) *Claim.* Every $A \in F^{m\times n}$ with $m \leq n$ has a right inverse.

Proof. Run Gauss–Jordan elimination on A. Each pivot uses up one row and one column, and there are at least as many columns as rows, so every row receives a pivot. Hence $\operatorname{rank} A = m$, and Proposition 3.5.8 gives a right inverse.

(b) *Claim.* $\operatorname{rank}(AB) = \operatorname{rank} B$ for all $A \in F^{m\times m}$ and $B \in F^{m\times n}$.

Proof. By Theorem 3.5.19, $A = E_1E_2\cdots E_t$ with each E_i elementary. Multiplying B on the left by E_t, then by E_{t-1}, and so on, performs row operations on B, and row operations do not change the rank (Proposition 3.5.2(c)).

Solution on page 1273

Problem 3.8. STANDARD Prove or disprove each statement.

(a) If $A \in F^{m\times n}$ and $A\mathbf{x} = \mathbf{y}$ has exactly one solution for some $\mathbf{y}$, then $n \le m$.

(b) If $A, B \in F^{m\times n}$ and, for every $\mathbf{y} \in F^m$, the systems $A\mathbf{x} = \mathbf{y}$ and $B\mathbf{x} = \mathbf{y}$ have the same solution set, then $A = B$.

(c) If $A \in F^{m\times n}$ and $B \in F^{p\times n}$ satisfy $\ker T_A = \ker T_B$, then $\operatorname{rank} A = \operatorname{rank} B$, even when $m \neq p$.

(d) If $A \in F^{m\times n}$ and $H \in F^{n\times n}$ is invertible, then A and AH have the same pivot columns.

(e) If column j of $\operatorname{rref}(A)$ is zero, then column j of A is zero.

(f) The matrices of rank at most 1 form a subspace of $F^{2\times 2}$.

Solution on page 1273

Problem 3.9. STANDARD Let F be a finite field with q elements, let $n \ge 0$ and $k \ge 1$, let $c_1, \dots, c_k \in F$ be distinct, and let $d_1, \dots, d_k \in F$.

(a) Suppose $k \le n+1$. Show that exactly q^{n+1-k} polynomials $p \in F[x]_{\le n}$ satisfy $p(c_i) = d_i$ for $i = 1, \dots, k$.

(b) Suppose $k > n+1$. Show that at most one $p \in F[x]_{\le n}$ satisfies these conditions, and that for some choice of the d_i there is none.

(c) How many $p \in \mathbb{F}_7[x]_{\le 4}$ satisfy $p(1) = 2$, $p(2) = 3$, and $p(4) = 0$? How many of these also satisfy $p(0) = 0$?

Solution on page 1274

Problem 3.10. STANDARD Let $\mathbf{a}_1, \dots, \mathbf{a}_n \in F^m$ be linearly independent, and let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$.

(a) Show that the pivot columns of the $m \times (n+m)$ matrix $(A \mid I_m)$ are $1, \dots, n$ together with some columns $n + i_1 < \cdots < n + i_t$, and that $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_t}$ is a basis of F^m. Show also that $n + t = m$.

(b) Carry this out for $\mathbf{a}_1 = (2, 1, 0, 0)$ and $\mathbf{a}_2 = (1, 1, 1, 0)$ in $\mathbb{Q}^4$, and express the standard basis vectors left out as combinations of the basis you found.

(c) In the second proof of Proposition 3.5.9, the columns of A were extended to a basis by appending, for $i = 1, \dots, m$ in turn, each $\mathbf{e}_i$ not in the span of the list built so far. Show that this appends exactly $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_t}$.

Solution on page 1274

Problem 3.11. STANDARD Let $m, n \ge 1$, and let $\mathbf{u}_1, \dots, \mathbf{u}_k \in F^n$ and $\mathbf{w}_1, \dots, \mathbf{w}_k \in F^m$.

(a) Show that there is a linear map $T : F^n \to F^m$ with $T(\mathbf{u}_j) = \mathbf{w}_j$ for $j = 1, \dots, k$ if and only if every linear relation among the $\mathbf{u}_j$ also holds among the $\mathbf{w}_j$: whenever $c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$, also $c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k = \mathbf{0}$.

(b) Suppose such a T exists. Show that it is unique if and only if $\mathbf{u}_1, \dots, \mathbf{u}_k$ span F^n.

(c) Find all such T for $\mathbf{u}_1 = (1, 0, 1)$, $\mathbf{u}_2 = (1, 1, 0)$, $\mathbf{u}_3 = (3, 1, 2)$ and $\mathbf{w}_1 = (1, 1)$, $\mathbf{w}_2 = (0, 2)$, $\mathbf{w}_3 = (2, 4)$ over $\mathbb{Q}$. What changes if $\mathbf{w}_3 = (2, 5)$?

Solution on page 1275

Problem 3.12. STANDARD For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, define $C_A : F^{2\times 2} \to F^{2\times 2}$ by $C_A(X) = AX - XA$.

(a) Show that C_A is linear, and find its matrix M with respect to the basis $\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}, \mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}, \mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}, \mathbf{e}_2\mathbf{e}_2^{\mathsf{T}}$ of matrix units.

(b) Show that $\operatorname{rank} M = 0$ if A is a scalar multiple of I_2, and $\operatorname{rank} M = 2$ otherwise.

(c) For $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, find all X with $AX = XA$, and all matrices of the form $AX - XA$.

Solution on page 1276

Problem 3.13. STANDARD For $p = a_0 + a_1x + a_2x^2 + a_3x^3 \in \mathbb{R}[x]_{\le 3}$, write $p' = a_1 + 2a_2x + 3a_3x^2$ and $p'' = 2a_2 + 6a_3x$. Given values $d_0, d_1, d_2 \in \mathbb{R}$, a **natural cubic spline** through $(-1, d_0)$, $(0, d_1)$, $(1, d_2)$ is a pair (p, q) of polynomials in $\mathbb{R}[x]_{\le 3}$ (the pieces on $[-1, 0]$ and on $[0, 1]$) such that

$$p(-1) = d_0, \quad p(0) = q(0) = d_1, \quad q(1) = d_2, \quad p'(0) = q'(0), \quad p''(0) = q''(0), \quad p''(-1) = q''(1) = 0.$$

(a) Show that these conditions form a square linear system in the coefficients of p and q.

(b) Show that for all d_0, d_1, d_2 there is exactly one natural cubic spline.

(c) Find the spline for $(d_0, d_1, d_2) = (0, 1, 0)$. Use (b) to shorten the computation.

Solution on page 1277

Problem 3.14. STANDARD Let $A \in F^{k\times k}$ be invertible, $B \in F^{k\times \ell}$, $C \in F^{\ell\times k}$, $D \in F^{\ell\times \ell}$, and

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad S = D - CA^{-1}B \in F^{\ell\times \ell}.$$

The matrix S is the **Schur complement** of A in M.

(a) Find invertible block matrices P and Q with $PMQ = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}$.

(b) Deduce that $\operatorname{rank} M = k + \operatorname{rank} S$. In particular, M is invertible if and only if S is.

(c) For which $t \in \mathbb{Q}$ does the following matrix have rank 4?

$$M_t = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & t & 1 \\ 0 & 1 & 1 & t \end{pmatrix}$$

Solution on page 1278

Problem 3.15. HARD An electrical network has nodes $0, 1, \ldots, n$. Some pairs of nodes are joined by a resistor; let $c_{ij} = c_{ji} > 0$ be the conductance (the reciprocal of the resistance) if i and j are joined, and $c_{ij} = 0$ otherwise, with $c_{ii} = 0$. The network is **connected**: any node can be reached from any other along resistors. Node 0 is grounded, so its potential is $v_0 = 0$. An external source injects a current s_i into node i for $i = 1, \ldots, n$. By Ohm's law, the current from i to j through their resistor is $c_{ij}(v_i - v_j)$, and Kirchhoff's current law at node i says

$$\sum_{j=0}^{n} c_{ij}(v_i - v_j) = s_i \quad (i = 1, \ldots, n).$$

(a) Show that for every $\mathbf{s} \in \mathbb{R}^n$ there is exactly one vector of potentials $\mathbf{v} = (v_1, \ldots, v_n)$ satisfying these equations.

(b) Show that the current flowing into the ground node through the resistors is automatically $s_1 + \cdots + s_n$.

(c) In the network of Figure 3.1, one ampere is injected at node 2. Find the potentials, the current through the diagonal resistor, and the **effective resistance** between nodes 2 and 0, which is v_2 divided by the injected current.

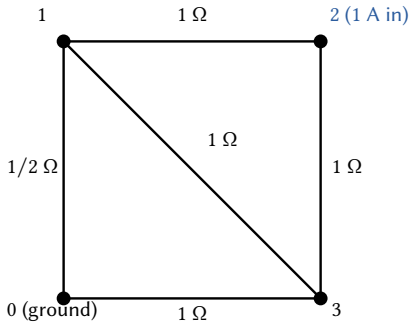


Figure 3.1. Four nodes on a square with one diagonal. The resistor between nodes 0 and 1 has resistance $1/2$ Ω (conductance 2); the other four have resistance 1 Ω.

Solution on page 1279

Problem 3.16. HARD For $n \geq 1$, let $P_n \in F^{n \times n}$ have entry 1 in every position (i, j) with $|i - j| \leq 1$, and 0 elsewhere.

(a) Show that, over every field F, P_n is invertible if and only if n is not congruent to 2 modulo 3.

(b) Over $\mathbb{F}_2$, P_n is the matrix of "lights out" on a row of n lamps, where pressing a lamp toggles it and its immediate neighbors (compare Exercise 3.1.40). Suppose $n \equiv 2 \bmod 3$. Show that a pattern $\mathbf{y} \in \mathbb{F}_2^n$ of lit lamps can be switched off if and only if the number of lit lamps in positions not divisible by 3 is even. You may use Exercise 3.5.37.

Solution on page 1280

Problem 3.17. HARD Let F be a field and $1 \leq r \leq \min(m, n)$.

(a) Let $A \in F^{m \times n}$ have rank r, and let $S \in F^{r \times n}$ consist of the r nonzero rows of $\operatorname{rref}(A)$. Show that $A = CS$ for exactly one $C \in F^{m \times r}$, and that this C has rank r. Conversely, show that if $S \in F^{r \times n}$ is in RREF with no zero row and $C \in F^{m \times r}$ has rank r, then CS has rank r and $\operatorname{rref}(CS) = \begin{pmatrix} S \\ 0 \end{pmatrix}$.

(b) Let F have q elements, and let $N(r, n)$ be the number of $r \times n$ matrices in RREF with no zero row (compare Exercise 3.3.58). Show that the number of matrices in $F^{m \times n}$ of rank r is

$$N(r, n)(q^m - 1)(q^m - q)\cdots\left(q^m - q^{r-1}\right).$$

(c) Deduce that

$$N(r, n) = \frac{(q^n - 1)(q^n - q)\cdots\left(q^n - q^{r-1}\right)}{(q^r - 1)(q^r - q)\cdots\left(q^r - q^{r-1}\right)},$$

and check this against Exercise 3.3.48 for $r = 2$, $n = 3$.

Solution on page 1281

Problem 3.18. HARD For $A \in \mathbb{Z}^{n\times n}$, let $\overline{A} \in \mathbb{F}_2^{n\times n}$ be the matrix of remainders of its entries modulo 2.

(a) Show that if $\overline{A}$ is invertible over $\mathbb{F}_2$, then A is invertible as a real matrix.

(b) Deduce that an integer matrix whose diagonal entries are odd and whose other entries are even is invertible over $\mathbb{R}$.

(c) Show that an integer matrix whose diagonal entries are even and whose other entries are odd is invertible over $\mathbb{R}$ when n is even, and give a non-invertible example with $n = 3$.

(d) Show that the converse of (a) is false.

Solution on page 1282

Problem 3.19. CHALLENGE Let F be a field, let $a_1, \ldots, a_n \in F$ be distinct, let $b_1, \ldots, b_n \in F$ be distinct, and suppose $a_i + b_j \neq 0$ for all i, j.

(a) Show that the **Cauchy matrix** $C \in F^{n\times n}$ with entries $c_{ij} = 1/\big(a_i + b_j\big)$ is invertible.

(b) Deduce that the **Hilbert matrix** $H_n \in \mathbb{Q}^{n\times n}$ with entries $1/(i + j - 1)$ is invertible.

Solution on page 1283

CHAPTER 4

Dimension

Dimension, rank–nullity, rank inequalities, sums and direct sums, change of basis, quotients, and affine geometry.

4.1 Dimension

The standard bases of F^n, $F^{m\times n}$, and $F[x]_{\le n}$ have n, mn, and $n+1$ entries. Each of these spaces has many other bases (Section 1.6), and this section begins by showing that all of them have the same length. That length is the dimension of the space. We prove it twice: with the matrix machinery of Section 3.5, as the lecture does, and with the exchange lemma, which works inside V and says more, namely that an independent list is never longer than a spanning list.

Dimension then turns many questions into counting. A list of the right length is a basis as soon as it is independent or spanning, every independent list extends to a basis, and a subspace is no bigger than the space around it. The counting has consequences well beyond bases. When a chain $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots$ ends in $\mathbf{0}$, its nonzero vectors are independent, so an operator on an n-dimensional space that is eventually 0 is already 0 after n steps. Any n^2+1 matrices of size $n\times n$ satisfy a linear relation, so every square matrix satisfies a polynomial equation. And over a field with q elements, a space of dimension n has exactly q^n vectors.

Throughout, V is a vector space over a field F.

4.1.1 Bases of F^m

Intuition (Why the number of vectors should be fixed). A basis $\mathcal{B} = \mathbf{v}_1, \dots, \mathbf{v}_n$ of V names every vector by its coordinates $[\mathbf{v}]_{\mathcal{B}} \in F^n$, and the coordinate map $V \to F^n$ is a bijection that respects linear combinations (Proposition 1.6.16, Proposition 1.6.19). So a basis of length n makes V a copy of F^n with the vectors renamed. If V also had a basis of length m, it would be a copy of F^m as well, and translating from one set of names to the other would give a bijection $F^m \to F^n$ that respects linear combinations, that is, an invertible linear map. Every linear map $F^m \to F^n$ is T_A for a matrix $A \in F^{n\times m}$ (Proposition 2.3.17), and invertible matrices are square (Corollary 3.5.15). So $m = n$. The lecture runs this argument in the simplest case, $V = F^m$, where no translation is needed.

Theorem 4.1.1 (Bases of F^m). Every basis of F^m has exactly m vectors.

Proof (via the Third Fundamental Theorem). ROUTINE

Uses. Theorem 3.5.14 (Third Fundamental Theorem, completed), Definition 1.6.1 (Basis)
Techniques. Translate between maps and matrices (Technique 2.3.27)

Hint 1. Put the vectors of a basis $\mathcal{B}$ of F^m into a matrix as its columns. What is its size?
Hint 2. Condition (III) of Theorem 3.5.14 is a statement about the columns of a matrix. What does the "moreover" part of that theorem conclude?
Sketch. If $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ is a basis of F^m, the $m \times n$ matrix with columns $\mathbf{v}_1, ..., \mathbf{v}_n$ satisfies condition (III) of Theorem 3.5.14, so $m = n$ by the "moreover" part.

Let $\mathcal{B} = \mathbf{v}_1, ..., \mathbf{v}_n$ be a basis of F^m, and let

$$A = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \in F^{m \times n}$$

be the matrix whose columns are $\mathbf{v}_1, ..., \mathbf{v}_n$. Its columns form a basis of F^m, which is condition (III) of Theorem 3.5.14. By the "moreover" part of that theorem, $m = n$. □

The proof is two lines because the work was done in Section 3.5. The "moreover" part of Theorem 3.5.14 comes from the inequality $\operatorname{rank} A \leq \min(m, n)$, which holds because the pivots of $\operatorname{rref}(A)$ occupy distinct rows and distinct columns (Proposition 3.5.2). So the theorem is really a count of pivots. To prove the corresponding statement in an arbitrary vector space without first choosing coordinates, we need a count that happens inside V. That count is the exchange lemma.

4.1.2 The exchange lemma

Intuition (Trading spanning vectors for independent ones). Let $\mathbf{u}_1, ..., \mathbf{u}_m$ be independent and let $\mathbf{w}_1, ..., \mathbf{w}_n$ span V. Since the $\mathbf{w}_i$ span, $\mathbf{u}_1$ is a linear combination of them, and since $\mathbf{u}_1 \neq \mathbf{0}$, some $\mathbf{w}_i$ appears with a nonzero coefficient. Solving for that $\mathbf{w}_i$ shows that it lies in the span of $\mathbf{u}_1$ and the other $\mathbf{w}$'s, so we can trade it for $\mathbf{u}_1$ and still have a spanning list of length n. This is the exchange of Exercise 1.3.34 and Exercise 1.5.33. Now trade a second $\mathbf{w}$ for $\mathbf{u}_2$, and so on. The independence of the $\mathbf{u}$'s guarantees that each trade removes a $\mathbf{w}$, and never one of the $\mathbf{u}$'s brought in earlier. If there were more $\mathbf{u}$'s than $\mathbf{w}$'s, we would run out of $\mathbf{w}$'s to trade.

Lemma 4.1.2 (Exchange lemma). Let $\mathbf{u}_1, ..., \mathbf{u}_m$ be a linearly independent list in V, and let $\mathbf{w}_1, ..., \mathbf{w}_n$ be a list that spans V. Then $m \leq n$. More precisely, the $\mathbf{w}_i$ can be renumbered so that

$$\mathbf{u}_1, ..., \mathbf{u}_m, \mathbf{w}_{m+1}, ..., \mathbf{w}_n$$

spans V.

In words: an independent list is never longer than a spanning list, and the independent list can replace an equal number of entries of the spanning list without losing the span. The result is often called the Steinitz exchange lemma. The first proof makes the trades of the intuition one at a time, and an induction keeps track of them.

Proof (by exchange). STANDARD

Uses. Lemma 1.5.10 (Linear dependence lemma), Corollary 1.5.13 (Independence, one entry at a time), Corollary 1.5.17 (Sublists), Proposition 1.5.18 (Operations that preserve independence), Corollary 1.3.20 (Operations that preserve the span)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Prove by induction on $k = 0, 1, ..., m$: $k \leq n$, and after renumbering the $\mathbf{w}_i$, the list $\mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{w}_{k+1}, ..., \mathbf{w}_n$ spans V.
Hint 2. In the inductive step, insert $\mathbf{u}_{k+1}$ right after $\mathbf{u}_k$. Why is the longer list dependent, and what does Lemma 1.5.10 then let you delete?
Hint 3. The deleted entry is in the span of the entries before it. It cannot be one of $\mathbf{u}_1, ..., \mathbf{u}_{k+1}$, which come first and form an independent list (Corollary 1.5.13). So it is a $\mathbf{w}$, and in particular some $\mathbf{w}$ was still there.
Sketch. Induct on k. Suppose $\mathcal{B}_k = \mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{w}_{k+1}, ..., \mathbf{w}_n$ spans V. Inserting $\mathbf{u}_{k+1} \in \operatorname{span}(\mathcal{B}_k)$ after $\mathbf{u}_k$ gives a dependent list that still spans. By Lemma 1.5.10, some entry lies in the span of the entries before it, and deleting it keeps the span. It is not among the independent $\mathbf{u}_1, ..., \mathbf{u}_{k+1}$ at the front, so it is some $\mathbf{w}_i$ with $i > k$. Hence $k + 1 \leq n$, and after renumbering, deleting it gives $\mathcal{B}_{k+1}$. At $k = m$ we get $m \leq n$.

We prove by induction on $k \in \{0, 1, ..., m\}$ the statement

$P(k)$: $k \le n$, and the $\mathbf{w}_i$ can be renumbered so that $\mathcal{B}_k = \mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{w}_{k+1}, ..., \mathbf{w}_n$ spans V.

For $k = m$ this is the lemma.

Base case. $P(0)$ says that $0 \le n$ and that $\mathbf{w}_1, ..., \mathbf{w}_n$ spans V, which is the hypothesis.

Inductive step. Let $0 \le k < m$ and assume $P(k)$, with the $\mathbf{w}_i$ numbered so that $\mathcal{B}_k$ spans V. Let

$$\mathcal{L} = \mathbf{u}_1, ..., \mathbf{u}_k, \mathbf{u}_{k+1}, \mathbf{w}_{k+1}, ..., \mathbf{w}_n$$

be $\mathcal{B}_k$ with $\mathbf{u}_{k+1}$ inserted after $\mathbf{u}_k$. The list $\mathcal{L}$ spans V, since it contains every entry of $\mathcal{B}_k$ (Corollary 1.3.20). It is linearly dependent: $\mathbf{u}_{k+1} \in V = \operatorname{span}(\mathcal{B}_k)$, so the reordering $\mathcal{B}_k, \mathbf{u}_{k+1}$ of $\mathcal{L}$ is dependent by Corollary 1.5.13, and reordering does not affect independence (Proposition 1.5.18(a)).

By Lemma 1.5.10, some entry of $\mathcal{L}$ lies in the span of the entries before it, and deleting that entry does not change the span. The first $k+1$ entries of $\mathcal{L}$ form the list $\mathbf{u}_1, ..., \mathbf{u}_{k+1}$, which is independent as a sublist of an independent list (Corollary 1.5.17). By Corollary 1.5.13, none of them lies in the span of the entries before it. So the entry in question is $\mathbf{w}_i$ for some i with $k+1 \le i \le n$. In particular $k+1 \le n$. Renumber $\mathbf{w}_{k+1}, ..., \mathbf{w}_n$ so that this entry becomes $\mathbf{w}_{k+1}$. Deleting it from $\mathcal{L}$ leaves

$$\mathcal{B}_{k+1} = \mathbf{u}_1, ..., \mathbf{u}_{k+1}, \mathbf{w}_{k+2}, ..., \mathbf{w}_n,$$

whose span is $\operatorname{span}(\mathcal{L}) = V$. This is $P(k+1)$. □

Each step of the induction removes a $\mathbf{w}$ that is still present, so the steps assign to each $\mathbf{u}$ a different $\mathbf{w}$, and $m \le n$ is the pigeonhole principle. The second proof replaces the trades by a single linear system. It shows that $m \le n$, but not the renumbering statement.

Proof (via a homogeneous system). STANDARD

> *Uses.* Definition 1.3.12 (Spanning list), Corollary 3.4.19 (More unknowns than equations), Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 2.3.11 (Entries of $A\mathbf{x}$)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* Write each $\mathbf{u}_j$ as a linear combination of $\mathbf{w}_1, ..., \mathbf{w}_n$, and collect the coefficients into a matrix A. What is its size?
> *Hint 2.* Suppose $m > n$. Then the system $A\mathbf{x} = \mathbf{0}$ has more unknowns than equations. Turn a nonzero solution into a linear relation among the $\mathbf{u}_j$.
> *Sketch.* Write $\mathbf{u}_j = \sum_i a_{ij}\mathbf{w}_i$ and let $A = (a_{ij}) \in F^{n \times m}$. If $m > n$, some $\mathbf{x} \ne \mathbf{0}$ has $A\mathbf{x} = \mathbf{0}$ (Corollary 3.4.19). Regrouping, $\sum_j x_j \mathbf{u}_j = \sum_i (A\mathbf{x})_i \mathbf{w}_i = \mathbf{0}$, a nontrivial relation among the independent $\mathbf{u}_j$.

Suppose, to the contrary, that $m > n$. Since $\mathbf{w}_1, ..., \mathbf{w}_n$ spans V, for each j there are scalars $a_{1j}, ..., a_{nj}$ with

$$\mathbf{u}_j = a_{1j}\mathbf{w}_1 + \cdots + a_{nj}\mathbf{w}_n.$$

Let $A = (a_{ij}) \in F^{n \times m}$. It has more columns than rows, so by Corollary 3.4.19(a) there is $\mathbf{x} = (x_1, ..., x_m) \ne \mathbf{0}$ with $A\mathbf{x} = \mathbf{0}$. Regrouping the double sum by repeated use of Lemma 1.3.9,

$$\sum_{j=1}^{m} x_j \mathbf{u}_j = \sum_{j=1}^{m} x_j \sum_{i=1}^{n} a_{ij}\mathbf{w}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} x_j \right) \mathbf{w}_i = \sum_{i=1}^{n} (A\mathbf{x})_i \mathbf{w}_i = \mathbf{0},$$

where $(A\mathbf{x})_i = \sum_j a_{ij} x_j$ is the i-th entry of $A\mathbf{x}$ (Proposition 2.3.11). Since some $x_j \ne 0$, this is a nontrivial linear relation of $\mathbf{u}_1, ..., \mathbf{u}_m$, contradicting their independence (Theorem 1.5.5). So $m \le n$. □

Example 4.1.3 (The exchange, step by step). In $\mathbb{R}^3$, let $\mathbf{u}_1 = (1, 1, 0)$ and $\mathbf{u}_2 = (0, 1, 1)$, which are independent since neither is a multiple of the other, and let $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 = \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Step 1. Insert $\mathbf{u}_1$ at the front: $\mathbf{u}_1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. The relation $\mathbf{u}_1 - \mathbf{e}_1 - \mathbf{e}_2 = \mathbf{0}$ shows that $\mathbf{e}_2 \in \operatorname{span}(\mathbf{u}_1, \mathbf{e}_1)$, and $\mathbf{e}_2$ is the first entry with this property, since $\mathbf{e}_1 \notin \operatorname{span}(\mathbf{u}_1)$. Delete it: $\mathbf{u}_1, \mathbf{e}_1, \mathbf{e}_3$ spans $\mathbb{R}^3$.

Step 2. Insert $\mathbf{u}_2$: $\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1, \mathbf{e}_3$. Now $\mathbf{u}_2 = \mathbf{u}_1 - \mathbf{e}_1 + \mathbf{e}_3$, so $\mathbf{e}_3 \in \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1)$. Here $\mathbf{e}_3$ is the only entry in the span of the earlier ones: $\mathbf{u}_1$ and $\mathbf{u}_2$ cannot be, and $\mathbf{e}_1 = (1, 0, 0) \notin \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$, whose vectors $a(1, 1, 0) +$

$b(0,1,1) = (a, a+b, b)$ have middle coordinate equal to the sum of the others. Delete $\mathbf{e}_3$: the list $\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1$ spans $\mathbb{R}^3$.

Two independent vectors pushed out two of the three standard basis vectors, and the one that survived, $\mathbf{e}_1$, was enough to complete $\mathbf{u}_1, \mathbf{u}_2$ to a spanning list of $\mathbb{R}^3$. (So would $\mathbf{e}_2$ or $\mathbf{e}_3$ have been; the exchange simply found one.)

With the lemma, the theorem on F^m has a proof that uses nothing but the standard basis.

Proof of Theorem 4.1.1 (via the exchange lemma). ROUTINE

> *Uses.* Lemma 4.1.2 (Exchange lemma), Proposition 1.6.7 (Standard bases), Proposition 1.6.2 (Bases are independent spanning lists)
> *Hint 1.* Compare a basis $\mathcal{B}$ of F^m of length n with the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$. Both lists are independent and both span.
> *Hint 2.* Apply Lemma 4.1.2 twice, with the roles of the two lists swapped.
> *Sketch.* $\mathcal{B}$ is independent and $\mathbf{e}_1, \ldots, \mathbf{e}_m$ spans F^m, so $n \le m$. The standard basis is independent and $\mathcal{B}$ spans, so $m \le n$.

Let $\mathcal{B}$ be a basis of F^m of length n. The standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ is a basis of F^m (Proposition 1.6.7). By Proposition 1.6.2 both lists are independent and both span F^m. Applying Lemma 4.1.2 to the independent list $\mathcal{B}$ and the spanning list $\mathbf{e}_1, \ldots, \mathbf{e}_m$ gives $n \le m$. Applying it to the independent list $\mathbf{e}_1, \ldots, \mathbf{e}_m$ and the spanning list $\mathcal{B}$ gives $m \le n$. Hence $n = m$. □

> *Intuition (Matrix machinery versus pure exchange).* Both proofs of Theorem 4.1.1 are counting arguments, and they count different things. The first counts pivots: $\operatorname{rank} A \le \min(m, n)$ because each pivot of $\operatorname{rref}(A)$ has a row and a column to itself. Behind it stands all of Gauss–Jordan elimination, and in return the proof is a computation: row reduce, and the pivots show which columns are independent and whether they span. The second counts trades: each $\mathbf{u}$ pushes out a $\mathbf{w}$ that has not been pushed out before. It never mentions matrices or coordinates, and the only tool it uses is the linear dependence lemma. That is why it works verbatim in $F[x]_{\le n}$, in spaces of functions, and in subspaces cut out by equations, where there is no matrix to reduce until a basis has been chosen. The price is that it is not a computation. It asserts that some $\mathbf{w}$ can be removed, and finding one means testing membership in spans, which in practice is again a row reduction.
>
> The second proof of the exchange lemma sits between the two. It uses coordinates, but only with respect to a spanning list, which need not be a basis, and the single fact about linear systems that it needs is that a homogeneous system with more unknowns than equations has a nonzero solution.

4.1.3 Dimension

The theorem on F^m extends to every finite-dimensional space, again in two ways. The lecture transports the question to F^m by coordinates. The exchange lemma answers it in V directly.

> **Theorem 4.1.4** (All bases have the same length). Let V be finite-dimensional. Any two bases of V have the same length.

Proof (via coordinates). ROUTINE

> *Uses.* Proposition 1.6.20 (Coordinates translate questions about V into F^n), Theorem 4.1.1 (Bases of F^m)
> *Hint 1.* Take bases $\mathcal{B}$ of length m and $\mathcal{C}$ of length n. Use $\mathcal{B}$-coordinates to move the list $\mathcal{C}$ into F^m.
> *Hint 2.* By Proposition 1.6.20(c), the coordinate vectors of a basis of V form a basis of F^m. How long is every basis of F^m?
> *Sketch.* If $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_n$ are bases, then $[\mathbf{w}_1]_{\mathcal{B}}, \ldots, [\mathbf{w}_n]_{\mathcal{B}}$ is a basis of F^m by Proposition 1.6.20(c), so $n = m$ by Theorem 4.1.1.

Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathcal{C} = \mathbf{w}_1, \ldots, \mathbf{w}_n$ be bases of V. Since $\mathcal{C}$ is a basis of V, Proposition 1.6.20(c), applied with the basis $\mathcal{B}$, shows that $[\mathbf{w}_1]_{\mathcal{B}}, \ldots, [\mathbf{w}_n]_{\mathcal{B}}$ is a basis of F^m. It has n vectors, so $n = m$ by Theorem 4.1.1.□

Proof (via the exchange lemma). ROUTINE

> *Uses.* Lemma 4.1.2 (Exchange lemma), Proposition 1.6.2 (Bases are independent spanning lists)

Hint 1. A basis is both independent and spanning.
Hint 2. Apply Lemma 4.1.2 with $\mathcal{B}$ as the independent list and $\mathcal{C}$ as the spanning list, and then the other way around.
Sketch. $\mathcal{B}$ is independent and $\mathcal{C}$ spans, so $m \le n$; symmetrically $n \le m$.

Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V of lengths m and n. By Proposition 1.6.2, $\mathcal{B}$ is independent and $\mathcal{C}$ spans V, so $m \le n$ by Lemma 4.1.2. Exchanging the roles of $\mathcal{B}$ and $\mathcal{C}$ gives $n \le m$. □

The first proof reduces the question about V to the same question about F^m, and the dictionary of Proposition 1.6.20 guarantees that nothing is lost in translation. The second proof needs no dictionary. Both are worth knowing: transporting a problem to F^n by coordinates is how most computations in abstract spaces are done, and arguing inside V is how most proofs about them are written.

By Corollary 1.6.24 a finite-dimensional space has a basis, and by Theorem 4.1.4 all of its bases have the same length. So the following definition makes sense.

Definition 4.1.5 (Dimension). Let V be a finite-dimensional vector space. The **dimension** of V, written $\dim V$, is the length of any basis of V. If $\dim V = n$, we say that V is **n-dimensional**. If V is infinite-dimensional, we write $\dim V = \infty$.

The symbol ∞ is shorthand and not a number we compute with. It lets us write statements such as "$\dim V < \infty$" for "V is finite-dimensional". Since $\{\mathbf{0}\}$ has the basis $()$ and every other space has a nonzero vector, which no basis of length 0 can represent, $\dim V = 0$ exactly when $V = \{\mathbf{0}\}$.

Remark 4.1.6 (The field is part of the data). Whether a list is a basis depends on which scalars are allowed (Example 1.6.5), so the dimension depends on the field. When the field matters we write $\dim_F V$. For example, $\dim_{\mathbb{C}} \mathbb{C} = 1$, with basis 1, while $\dim_{\mathbb{R}} \mathbb{C} = 2$, with basis $1, i$. The set and the addition are the same; only the scalars differ.

Example 4.1.7 (Dimensions of the standard spaces).

(a) By Proposition 1.6.7, $\dim F^n = n$, $\dim F^{m\times n} = mn$, and $\dim F[x]_{\le n} = n + 1$.

(b) The plane $P = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ has the basis $(1, 0, -1), (0, 1, -1)$ (Example 1.6.3(b)), so $\dim P = 2$.

(c) If V is a vector space over $\mathbb{C}$ with $\dim_{\mathbb{C}} V = n$ and basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then $\mathbf{v}_1, i\mathbf{v}_1, \ldots, \mathbf{v}_n, i\mathbf{v}_n$ is a basis of V over $\mathbb{R}$. It is independent over $\mathbb{R}$ by Exercise 1.5.37, and it spans V over $\mathbb{R}$, since every $\mathbf{v} \in V$ is $\sum_j (a_j + b_j i)\mathbf{v}_j = \sum_j (a_j \mathbf{v}_j + b_j (i\mathbf{v}_j))$ with $a_j, b_j \in \mathbb{R}$. So $\dim_{\mathbb{R}} V = 2n$; for instance $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$.

(d) $F[x]$ is infinite-dimensional (Proposition 1.6.25), and so is $\mathbb{R}$ over $\mathbb{Q}$ (Exercise 1.3.45). In both cases $\dim = \infty$. For $F[x]$ the set $\{1, x, x^2, \ldots\}$ is a basis in the infinite sense of Aside (Bases of infinite-dimensional spaces). For $\mathbb{R}$ over $\mathbb{Q}$ every such basis is uncountable, and none can be described explicitly; their existence rests on the axiom of choice (Section 15.1).

(e) Linear maps $F^n \to F^m$ correspond to matrices in $F^{m\times n}$ (Theorem 2.3.20), and the correspondence respects sums and scalar multiples (Proposition 2.4.12). We will see in Section 4.5 that this makes $\dim \mathrm{L}(V, W) = (\dim V)(\dim W)$ for all finite-dimensional V and W.

Non-example 4.1.8 (What the dimension does not count).

(a) Every vector of the plane P in (b) above has three coordinates, but $\dim P = 2$. The dimension counts the vectors in a basis of P, not the coordinates used to write down its elements.

(b) The list $(1, 0), (0, 1), (1, 1)$ spans $\mathbb{R}^2$ but $\dim \mathbb{R}^2 = 2$. A spanning list can be longer than the dimension, and Corollary 4.1.11 below shows that it can never be shorter.

(c) $\mathbb{C}$ and $\mathbb{R}^2$ can be identified as sets with the same addition (Example 1.6.5), yet $\dim_{\mathbb{C}} \mathbb{C} = 1$ and $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$. Without the field, "the dimension of $\mathbb{C}$" is ambiguous.

Exercise 4.1.9. ROUTINE Find a basis and the dimension of $U = \{(a, b, a+b, a-b) \mid a, b \in F\} \subseteq F^4$.

Solution on page 1284

4.1.4 Lists of the right length

Once $\dim V$ is known, the length of a list already says a good deal about it. The lecture proves the next theorem with coordinates and the First and Third Fundamental Theorems. The exchange lemma gives a second proof.

Theorem 4.1.10 (Independent lists are short). Let V be finite-dimensional with $\dim V = n$, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a linearly independent list in V.

(a) $k \leq n$.

(b) If $k = n$, then $\mathcal{L}$ is a basis of V.

Proof (via coordinates and the Fundamental Theorems). ROUTINE

Uses. Proposition 1.6.20 (Coordinates translate questions about V into F^n), Corollary 3.5.16 (Lists of vectors in F^m), Definition 4.1.5 (Dimension)
Hint 1. Fix a basis $\mathcal{B}$ of V, of length n, and pass to the coordinate vectors $[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}} \in F^n$. Are they independent?
Hint 2. Corollary 3.5.16 describes independent and spanning lists in F^n. Apply (a) of it, and for $k = n$ apply (c), then translate back with Proposition 1.6.20.
Sketch. With a basis $\mathcal{B}$ of length n, the coordinate vectors $[\mathbf{v}_i]_{\mathcal{B}}$ form an independent list of length k in F^n (Proposition 1.6.20(a)), so $k \leq n$ by Corollary 3.5.16(a). If $k = n$, they form a basis of F^n by Corollary 3.5.16(c), so $\mathcal{L}$ is a basis of V by Proposition 1.6.20(c).

Fix a basis $\mathcal{B}$ of V; it has length n (Definition 4.1.5). By Proposition 1.6.20(a), the list $[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_k]_{\mathcal{B}}$ is linearly independent in F^n.

(a) By Corollary 3.5.16(a), applied to the $n \times k$ matrix with these columns, $k \leq n$.

(b) If $k = n$, then by Corollary 3.5.16(c) the independent list $[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}$ is a basis of F^n. By Proposition 1.6.20(c), $\mathcal{L}$ is a basis of V. □

Proof (via the exchange lemma). ROUTINE

Uses. Lemma 4.1.2 (Exchange lemma), Proposition 1.5.14 (Appending a vector), Proposition 1.6.2 (Bases are independent spanning lists), Definition 4.1.5 (Dimension)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For (a), a basis of V is a spanning list of length n.
Hint 2. For (b), suppose $\mathcal{L}$ does not span V. Append a vector outside its span and count.
Sketch. (a) A basis of V spans and has length n, so $k \leq n$ by Lemma 4.1.2. (b) If an independent $\mathcal{L}$ of length n did not span, appending $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$ would give an independent list of length $n+1$ (Proposition 1.5.14), contradicting (a).

(a) A basis of V spans V and has length n, so $k \leq n$ by Lemma 4.1.2.

(b) Suppose $k = n$. If $\mathcal{L}$ did not span V, there would be some $\mathbf{w} \in V$ with $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and $\mathcal{L}, \mathbf{w}$ would be an independent list of length $n+1$ by Proposition 1.5.14, contradicting (a). So $\mathcal{L}$ spans V, and being independent, it is a basis (Proposition 1.6.2). □

The contrapositive of (a) is used constantly: in an n-dimensional space, every list of more than n vectors is linearly dependent. For instance, any four vectors in $\mathbb{R}^3$ are dependent, which is what the span explorer of Section 1.5.3 showed for three vectors in $\mathbb{R}^2$. Spanning lists satisfy the mirror-image statements.

Corollary 4.1.11 (Spanning lists are long). Let V be finite-dimensional with $\dim V = n$, and let $\mathcal{S}$ be a list of length l that spans V.

(a) $l \geq n$.

(b) If $l = n$, then $\mathcal{S}$ is a basis of V.

Proof. ROUTINE

> *Uses.* Theorem 1.6.10 (Every spanning list contains a basis), Theorem 4.1.4 (All bases have the same length), Definition 4.1.5 (Dimension)
> *Hint 1.* Every spanning list contains a basis as a sublist (Theorem 1.6.10). How long is that basis?
> *Hint 2.* For (b), a sublist with as many entries as the whole list is the whole list.
> *Sketch.* $\mathcal{S}$ has a sublist that is a basis, of length n (Theorem 4.1.4), so $n \le l$. If $l = n$, that sublist has all l entries, so it is $\mathcal{S}$.

By Theorem 1.6.10, some sublist $\mathcal{S}'$ of $\mathcal{S}$ is a basis of V, and by Theorem 4.1.4 it has length n.

(a) A sublist is obtained by deleting entries, so $n = |\mathcal{S}'| \le l$.

(b) If $l = n$, then $\mathcal{S}'$ is obtained from $\mathcal{S}$ by deleting $l - n = 0$ entries, so $\mathcal{S}' = \mathcal{S}$, and $\mathcal{S}$ is a basis. □

Part (a) also follows from Lemma 4.1.2, applied to a basis of V and $\mathcal{S}$. Together, the two results say that for a list of length exactly $\dim V$, being independent, spanning, and being a basis are the same condition. So to check that a list of the right length is a basis, check whichever of the two properties is easier, and the other comes for free. In F^m this is Corollary 3.5.16(c), now available in every finite-dimensional space.

Example 4.1.12 (Two bases for the price of one property).

(a) Let $c \in F$ and $n \ge 0$. The polynomials $1, x-c, (x-c)^2, \ldots, (x-c)^n$ are nonzero with distinct degrees $0, 1, \ldots, n$, so they are independent (Proposition 1.5.24). There are $n+1 = \dim F[x]_{\le n}$ of them, so they form a basis of $F[x]_{\le n}$ by Theorem 4.1.10(b). We never had to show that every polynomial of degree at most n is a combination of powers of $x - c$. The coordinates of p in this basis are the coefficients of its expansion in powers of $x - c$, which over $\mathbb{R}$ are the Taylor coefficients at c; Exercise 1.6.18 computed them for $n = 2$ and $c = 1$.

(b) In $\mathbb{C}^{2\times 2}$, consider

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

If $aI_2 + bX + cY + dZ = 0$, then

$$\begin{pmatrix} a+d & b-ic \\ b+ic & a-d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Adding and subtracting the diagonal entries gives $2a = 2d = 0$, and the off-diagonal entries give $2b = 0$ and $2ic = 0$. So $a = b = c = d = 0$, the list is independent, and since $\dim \mathbb{C}^{2\times 2} = 4$ it is a basis. (These are the Pauli matrices of quantum mechanics.)

> *Intuition (Halving the work, and where the halving comes from).* For a list of length $\dim V$, independence and spanning are equivalent. In F^n this was the square case of the Fundamental Theorems: for a square matrix, "pivot in every column" and "pivot in every row" both say that there are n pivots. In general it is the exchange lemma: an independent list that did not span could be lengthened, and a spanning list that was dependent could be shortened, and either would produce a list whose length is on the wrong side of $\dim V$. The count only works when the length is *exactly* right. A list of length less than $\dim V$ is never spanning, one of length more is never independent, and these are the shape constraints of Corollary 3.5.16(a) and (b) in disguise.

Dimension also gives a convenient test for infinite dimension. It is the test promised after Proposition 1.6.25.

Corollary 4.1.13 (Detecting infinite dimension). A vector space V is infinite-dimensional if and only if for every $k \ge 0$ it contains a linearly independent list of length k.

Proof. ROUTINE

Uses. Theorem 4.1.10 (Independent lists are short), Proposition 1.5.14 (Appending a vector), Definition 1.6.23 (Finite-dimensional), Proposition 1.5.15 (Short lists, zeros, and repeats)
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. For (⇐), if V were finite-dimensional with $\dim V = n$, which length of independent list would be impossible?
Hint 2. For (⇒), build the lists by induction on k. If V is infinite-dimensional, can an independent list span it?
Sketch. (⇐) If $\dim V = n$, an independent list of length $n+1$ contradicts Theorem 4.1.10(a). (⇒) Induct on k, starting from $()$. An independent list $\mathcal{L}$ of length k does not span V, since no list does, so appending some $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$ gives an independent list of length $k+1$ (Proposition 1.5.14).

(⇐) We prove the contrapositive. If V is finite-dimensional, say $\dim V = n$, then by Theorem 4.1.10(a) there is no independent list of length $n+1$ in V.

(⇒) Suppose V is infinite-dimensional. We show by induction on k that V contains an independent list of length k. For $k = 0$ take $()$ (Proposition 1.5.15(a)). If $\mathcal{L}$ is independent of length k, then $\mathcal{L}$ does not span V, because no list spans V (Definition 1.6.23). So there is $\mathbf{w} \in V$ with $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and $\mathcal{L}, \mathbf{w}$ is independent of length $k+1$ by Proposition 1.5.14. □

Example 4.1.14 (Function spaces are infinite-dimensional). Let $I \subseteq \mathbb{R}$ be an infinite set, for instance $\mathbb{R}$ or $[0,1]$. For each k, the functions $1, t, t^2, \ldots, t^k$ on I are linearly independent in $\mathbb{R}^I$: if $a_0 + a_1 t + \cdots + a_k t^k = 0$ for every $t \in I$, then the polynomial $a_0 + a_1 x + \cdots + a_k x^k$ has infinitely many roots, so it is the zero polynomial (Proposition 7.1.22), and all $a_i = 0$. By Corollary 4.1.13, $\mathbb{R}^I$ is infinite-dimensional. These functions are continuous, so the subspace $C[0,1]$ of $\mathbb{R}^{[0,1]}$ is infinite-dimensional as well.

Exercise 4.1.15. ROUTINE Show that $1 + x, (1+x)^2, x^2 + x^3, x^3$ is a basis of $F[x]_{\leq 3}$ for every field F.

Solution on page 1284

4.1.5 Extending to a basis

A spanning list can be thinned to a basis (Theorem 1.6.10). Remark 1.6.26 described the opposite process: grow an independent list by appending vectors outside its span. It warned that growing need not stop, as in $F[x]$. In a finite-dimensional space it does stop, because Theorem 4.1.10 caps the length of an independent list. This is the lecture's corollary, and the homework step of appending one vector (Proposition 1.5.14) is what drives it. We also record which vectors can be appended.

Theorem 4.1.16 (Extending to a basis). Let V be finite-dimensional with $\dim V = n$, and let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ be a linearly independent list in V. Then $\mathcal{L}$ extends to a basis of V: there are vectors $\mathbf{w}_1, \ldots, \mathbf{w}_{n-k} \in V$ such that

$$\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}_1, \ldots, \mathbf{w}_{n-k}$$

is a basis of V. Moreover, if $\mathcal{S}$ is any list that spans V, the $\mathbf{w}_j$ can be chosen among the entries of $\mathcal{S}$.

Proof. STANDARD

Uses. Theorem 4.1.10 (Independent lists are short), Proposition 1.5.14 (Appending a vector), Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 4.1.4 (All bases have the same length), Definition 4.1.5 (Dimension)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. By Theorem 4.1.10, $k \leq n$, and if $k = n$ there is nothing to do. Induct on $n - k$.
Hint 2. If $k < n$, then $\mathcal{L}$ does not span V. Why must some entry of $\mathcal{S}$ lie outside $\operatorname{span}(\mathcal{L})$?
Hint 3. Append that entry. The new list is independent (Proposition 1.5.14) and one entry longer, so the inductive hypothesis applies to it.
Sketch. By Theorem 4.1.10, $k \leq n$; induct on $d = n - k$. If $d = 0$, $\mathcal{L}$ is already a basis. If $d > 0$, $\mathcal{L}$ is not a basis (wrong length), so it does not span. If every entry of $\mathcal{S}$ were in $\operatorname{span}(\mathcal{L})$, then $V = \operatorname{span}(\mathcal{S}) \subseteq \operatorname{span}(\mathcal{L})$; so some entry $\mathbf{s}$ of $\mathcal{S}$ is not. Then $\mathcal{L}, \mathbf{s}$ is independent of length $k+1$, and the inductive hypothesis extends it by entries of $\mathcal{S}$.

Fix a spanning list $\mathcal{S}$ of V; one exists because V is finite-dimensional. By Theorem 4.1.10(a), every independent list in V has length at most n. We prove by induction on $d \geq 0$ that every independent list of length $n -$

d extends to a basis of V by appending d entries of $\mathcal{S}$. Taking $\mathcal{S}$ to be any spanning list proves the theorem, including the "moreover".

Base case. If $d = 0$, then $\mathcal{L}$ is independent of length n, so it is a basis by Theorem 4.1.10(b), and we append nothing.

Inductive step. Let $d \geq 1$, assume the claim for $d-1$, and let $\mathcal{L}$ be independent of length $k = n - d < n$. Then $\mathcal{L}$ is not a basis of V, since every basis has length n (Theorem 4.1.4), so $\mathcal{L}$ does not span V. If every entry of $\mathcal{S}$ were in $\operatorname{span}(\mathcal{L})$, then, since $\operatorname{span}(\mathcal{L})$ is a subspace, $V = \operatorname{span}(\mathcal{S}) \subseteq \operatorname{span}(\mathcal{L})$ by Proposition 1.4.14(b), which is false. So some entry $\mathbf{s}$ of $\mathcal{S}$ lies outside $\operatorname{span}(\mathcal{L})$. By Proposition 1.5.14, $\mathcal{L}, \mathbf{s}$ is independent, of length $k+1 = n-(d-1)$. By the inductive hypothesis, it extends to a basis of V by appending $d-1$ entries of $\mathcal{S}$. Together with $\mathbf{s}$, these are d entries of $\mathcal{S}$ appended to $\mathcal{L}$, as desired. □

The lecture remarks that its argument "may be formalized as an induction on $d-k$", where its d is the dimension; our $d = n - k$ is that quantity. Unwound, the proof says: test the entries of $\mathcal{S}$ one at a time and append each one that is not in the span of the list built so far. The list stays independent, and the process stops because its length cannot exceed n. Finite dimension enters exactly there, in the cap on the length. In $F[x]$ there is no spanning list to draw from, and appending vectors outside the span can go on forever (Remark 1.6.26). In infinite dimensions, an extension to a basis in the sense of Aside (Bases of infinite-dimensional spaces) needs Zorn's lemma (Section 15.1).

Unwound in this way, the procedure is the sifting of Section 1.6.3 applied to the list $\mathcal{L}, \mathcal{S}$. Since $\mathcal{L}$ is independent, no entry of $\mathcal{L}$ lies in the span of the entries before it (Corollary 1.5.13), so sifting keeps all of $\mathcal{L}$, and it keeps exactly the entries of $\mathcal{S}$ not already accounted for. In F^n, with $\mathcal{S} = \mathbf{e}_1, \ldots, \mathbf{e}_n$, sifting is row reduction: by Corollary 3.3.19, the pivot columns of a matrix are exactly the columns that are not combinations of the columns before them.

Example 4.1.17 (Extending by row reduction). Extend $\mathbf{v}_1 = (1,1,0,0)$ and $\mathbf{v}_2 = (1,1,1,1)$ to a basis of $\mathbb{Q}^4$. Row reduce the matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

The pivot columns are $1, 2, 3, 5$. So $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_3$ is a basis of $\mathbb{Q}^4$. The free columns explain the rejections: column 4 of the reduced matrix says $\mathbf{e}_2 = \mathbf{v}_1 - \mathbf{e}_1$, and column 6 says $\mathbf{e}_4 = -\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{e}_3$ (Corollary 3.3.19(b)).

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

The reduction of Example 4.1.17. The first two columns are $\mathbf{v}_1, \mathbf{v}_2$ and the last four are $\mathbf{e}_1, \ldots, \mathbf{e}_4$. The pivot columns are the columns that survive sifting. Use "Edit" to put your own independent vectors in the first columns, or switch to $\mathbb{F}_p$; as long as those vectors are independent over the chosen field, they always end up as pivot columns. (Interactive in the web version.)

Extension is most useful when the independent list is a basis of a subspace. Then the extended basis is adapted to the subspace: its first vectors describe the subspace, and the remaining ones describe "the rest" of V, on which we are free to do as we like.

Technique 4.1.18 (Extend a basis of a subspace). To study a subspace U of a finite-dimensional space V, choose a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of U and extend it to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$ of V (Theorem 4.1.16). The new vectors are independent of U in a strong sense: a combination of them lies in U only if it is $\mathbf{0}$. Since a linear map may be prescribed freely on a basis (Theorem 2.2.1), the extension lets us define maps that do one thing on U and anything we like on the $\mathbf{w}_j$. Since the extended list is a basis, it also lets us count: $\dim V - \dim U$ is the number r of new vectors.

Reach for it when you need a linear map with prescribed behavior on a subspace, when you need to compare a subspace with the whole space, or when you want coordinates in which a subspace looks like the span of the first few standard basis vectors. The canonical example is Proposition 4.1.19. Later the same move proves rank–nullity (Section 4.2), produces complements and the dimension formula for sums (Section 4.4), and starts the proof of triangularization (Section 7.5).

Used in. Proposition 4.1.19, Theorem 4.2.11, Proposition 4.2.26, Proposition 4.4.16, Proposition 4.4.18, Theorem 4.4.19, Exercise 4.4.38, Exercise 4.4.42, Exercise 4.4.50, Corollary 4.5.22, Theorem 4.6.27, Problem 4.1, Problem 4.13, Problem 7.11, Proposition 8.4.19.

Proposition 4.1.19 (Extending linear maps and separating vectors). Let V be finite-dimensional, and let U be a subspace of V.

(a) For every vector space W and every linear map $S : U \to W$, there is a linear map $T : V \to W$ with $T(\mathbf{u}) = S(\mathbf{u})$ for all $\mathbf{u} \in U$.

(b) For every $\mathbf{v} \in V$ with $\mathbf{v} \notin U$, there is a linear map $f : V \to F$ with $f(\mathbf{u}) = 0$ for all $\mathbf{u} \in U$ and $f(\mathbf{v}) = 1$.

The proof uses that U has a basis. That every subspace of a finite-dimensional space is finite-dimensional is Proposition 4.1.21(a) below, whose proof does not use this proposition.

Proof. STANDARD

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 1.5.14 (Appending a vector)
Techniques. Extend a basis of a subspace (Technique 4.1.18), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Take a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of U and extend it to a basis of V. A linear map on V can be prescribed freely on this basis.
Hint 2. For (a), send $\mathbf{u}_i$ to $S(\mathbf{u}_i)$ and the new basis vectors to $\mathbf{0}$. Why do T and S then agree on all of U?
Hint 3. For (b), first extend $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}$, which is independent because $\mathbf{v} \notin U$.
Sketch. Take a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of U. (a) Extend it to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$ of V and let T send $\mathbf{u}_i \mapsto S(\mathbf{u}_i)$, $\mathbf{w}_j \mapsto \mathbf{0}$ (Theorem 2.2.1). Then $T(\sum c_i \mathbf{u}_i) = \sum c_i S(\mathbf{u}_i) = S(\sum c_i \mathbf{u}_i)$. (b) $\mathbf{v} \notin \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, so $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}$ is independent; extend it to a basis and let f be 1 on $\mathbf{v}$ and 0 on every other basis vector.

By Proposition 4.1.21(a), U is finite-dimensional; let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be a basis of U. It is an independent list in V.

(a) By Theorem 4.1.16 there are $\mathbf{w}_1, \dots, \mathbf{w}_r$ such that $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$ is a basis of V. By Theorem 2.2.1 there is a linear map $T : V \to W$ with $T(\mathbf{u}_i) = S(\mathbf{u}_i)$ for all i and $T(\mathbf{w}_j) = \mathbf{0}$ for all j. Let $\mathbf{u} \in U$, and write $\mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$. Since T and S are both linear,

$$T(\mathbf{u}) = \sum_{i=1}^{k} c_i T(\mathbf{u}_i) = \sum_{i=1}^{k} c_i S(\mathbf{u}_i) = S(\mathbf{u}),$$

as desired.

(b) Since $U = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$ and $\mathbf{v} \notin U$, the list $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}$ is independent by Proposition 1.5.14. By Theorem 4.1.16 it extends to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}, \mathbf{w}_1, \dots, \mathbf{w}_r$ of V. By Theorem 2.2.1 there is a linear map f :

$V \to F$ with $f(\mathbf{v}) = 1$ and with value 0 on every other vector of this basis. For $\mathbf{u} = \sum_i c_i \mathbf{u}_i \in U$, linearity gives $f(\mathbf{u}) = \sum_i c_i f(\mathbf{u}_i) = 0$. □

Part (b) says that a vector outside a subspace can always be told apart from the subspace by a linear function. It is the starting point of duality (Section 13.1). Neither part is automatic in infinite dimensions without Zorn's lemma, which is needed even to extend a basis of U (Section 15.1).

Exercise 4.1.20. ROUTINE Extend $(1, 2, 1)$ to a basis of F^3 by appending standard basis vectors, (a) for $F = \mathbb{Q}$, and (b) for $F = \mathbb{F}_3$. Which standard basis vectors are appended?

Solution on page 1285

4.1.6 Dimensions of subspaces

Every subspace of F^3 ought to have dimension 0, 1, 2, or 3. Proving that a subspace U of a finite-dimensional space has dimension at most $\dim V$ looks like a direct application of Theorem 4.1.10, but there is a gap: the theorem applies to a basis of U, and we do not yet know that U has one. A basis of V is no help, since none of its vectors need lie in U. (In $\mathbb{R}^2$, the line $\operatorname{span}((1, 1))$ contains neither $\mathbf{e}_1$ nor $\mathbf{e}_2$.) The lecture closes the gap by growing an independent list inside U. The bound on independent lists in V forces the growth to stop, and it can only stop at a basis of U.

Proposition 4.1.21 (Dimensions of subspaces). Let V be finite-dimensional with $\dim V = n$, and let U be a subspace of V.

(a) U is finite-dimensional, and $\dim U \leq n$.

(b) $\dim U = n$ if and only if $U = V$.

Proof. STANDARD

Uses. Theorem 4.1.10 (Independent lists are short), Proposition 1.5.14 (Appending a vector), Proposition 1.6.2 (Bases are independent spanning lists), Proposition 1.4.14 (The span is the smallest subspace containing a list), Definition 1.4.1 (Subspace)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. An independent list in U is also an independent list in V, so it has length at most n. Choose an independent list in U that is as long as possible.
Hint 2. If that list did not span U, what could you append to it?
Hint 3. For (b), a basis of U of length n is an independent list of length n in V. Apply Theorem 4.1.10(b).
Sketch. Independent lists in U are independent in V, so they have length at most n. Take one, $\mathcal{L}$, of maximal length. If some $\mathbf{u} \in U$ were outside $\operatorname{span}(\mathcal{L})$, then $\mathcal{L}, \mathbf{u}$ would be a longer independent list in U. So $\mathcal{L}$ spans U and is a basis of U of length at most n. For (b), if $\dim U = n$, a basis of U is independent of length n in V, hence a basis of V, so $V = \operatorname{span}(\mathcal{L}) \subseteq U$.

A linear relation among vectors of U is computed with the operations of V (Definition 1.4.1), so a list in U is independent in U if and only if it is independent in V.

(a) By Theorem 4.1.10(a), every independent list in U has length at most n. The empty list is independent, so the lengths of the independent lists in U form a nonempty set of integers between 0 and n. Choose an independent list $\mathcal{L}$ in U whose length k is as large as possible. Suppose, to the contrary, that some $\mathbf{u} \in U$ is not in $\operatorname{span}(\mathcal{L})$. Then $\mathcal{L}, \mathbf{u}$ is an independent list in U (Proposition 1.5.14) of length $k + 1$, contradicting the choice of $\mathcal{L}$. So $\mathcal{L}$ spans U, and it is a basis of U (Proposition 1.6.2). Hence U is finite-dimensional, and $\dim U = k \leq n$.

(b) If $U = V$, then $\dim U = n$. Conversely, suppose $\dim U = n$, and let $\mathcal{L}$ be a basis of U. It is an independent list in V of length n, so it is a basis of V by Theorem 4.1.10(b). Since U is a subspace containing the entries of $\mathcal{L}$, Proposition 1.4.14 gives $V = \operatorname{span}(\mathcal{L}) \subseteq U$. Hence $U = V$. □

The lecture's proof is the same argument told as a process: grow an independent list in U, one vector at a time, and observe that the process cannot reach length $n+1$. Choosing a longest list in advance replaces the process by an extremal choice (Technique 1.5.9), and then there is no need to say "and so on".

Since every subspace is a vector space in its own right, (b) can be applied with W in place of V: if $U \subseteq W$ are subspaces of a finite-dimensional space and $\dim U = \dim W$, then $U = W$. This is how equality of subspaces is usually proved in practice. Show one containment, and compare dimensions. Equivalently, a proper subspace has strictly smaller dimension, so a subspace can be pinned down by squeezing its dimension between two bounds.

Example 4.1.22 (Squeezing a dimension). Let $U = \{A \in F^{2\times 2} \mid \operatorname{tr} A = 0\}$, the trace-zero matrices of Example 1.4.11. The matrices

$$U_{12}, \quad U_{21}, \quad U_{11} - U_{22}$$

lie in U and are independent, since $aU_{12} + bU_{21} + c(U_{11} - U_{22}) = \begin{pmatrix} c & a \\ b & -c \end{pmatrix}$ is 0 only if $a = b = c = 0$. So $\dim U \geq 3$, by Theorem 4.1.10(a) applied in U. On the other hand $U_{11} \notin U$, so $U \neq F^{2\times 2}$ and $\dim U < 4$ by Proposition 4.1.21(b). Hence $\dim U = 3$, and by Theorem 4.1.10(b) the three matrices form a basis of U. We never checked that they span. The same squeeze shows that the trace-zero matrices in $F^{n\times n}$ have dimension $n^2 - 1$ (Exercise 4.1.43).

Example 4.1.23 (The subspaces of F^3). Let U be a subspace of F^3. By Proposition 4.1.21, $\dim U \in \{0, 1, 2, 3\}$, and U has a basis of that length:

- $\dim U = 0$: $U = \{\mathbf{0}\}$;
- $\dim U = 1$: $U = \operatorname{span}(\mathbf{v})$ for some $\mathbf{v} \neq \mathbf{0}$, a line through $\mathbf{0}$;
- $\dim U = 2$: $U = \operatorname{span}(\mathbf{v}, \mathbf{w})$ for some independent $\mathbf{v}, \mathbf{w}$, a plane through $\mathbf{0}$;
- $\dim U = 3$: $U = F^3$, by Proposition 4.1.21(b).

Over $\mathbb{R}$ these are the subspaces of $\mathbb{R}^3$ that Section 1.4 anticipated: the origin, the lines and the planes through it, and $\mathbb{R}^3$. The proof of Theorem 1.4.25 for F^2 worked by cases on how many independent directions U contains. Dimension is that count, and it settles F^n for every n at once.

Exercise 4.1.24. ROUTINE Let $U = \{p \in \mathbb{R}[x]_{\leq 3} \mid p(2) = 0\}$. Find $\dim U$ and a basis of U, by a squeeze.

Solution on page 1285

4.1.7 Nilpotent chains

Some operators become 0 after finitely many applications. On F^4, the shift

$$N(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4, 0)$$

moves every entry one place up and pushes a 0 in at the bottom, so $N^4 = 0$. Following one vector,

$$\mathbf{e}_4 \xrightarrow{N} \mathbf{e}_3 \xrightarrow{N} \mathbf{e}_2 \xrightarrow{N} \mathbf{e}_1 \xrightarrow{N} \mathbf{0},$$

so $N^3 \neq 0$ but $N^4 = 0$. Over $\mathbb{R}$, differentiation on $\mathbb{R}[x]_{\leq 3}$ behaves the same way:

$$x^3 \xrightarrow{D} 3x^2 \xrightarrow{D} 6x \xrightarrow{D} 6 \xrightarrow{D} \mathbf{0}.$$

An operator $T \in \mathrm{L}(V)$ with $T^m = 0$ for some $m \geq 1$ is called **nilpotent**; Chapter 10 studies these operators in depth (Section 10.1). In both examples the chain of images ends at $\mathbf{0}$ after exactly $\dim V$ steps, and the vectors in the chain form a basis. That is no accident. The vectors in such a chain are always independent, and the reason is a move worth naming.

Technique 4.1.25 (Apply an operator to a relation to kill terms). To show that $\mathbf{v}_1, \dots, \mathbf{v}_k$ is independent, take a relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ and apply a linear map L that sends all but a few of the $\mathbf{v}_i$ to $\mathbf{0}$. Since $L(\mathbf{0}) = \mathbf{0}$, the result is a shorter relation, ideally a single term $a_j L(\mathbf{v}_j) = \mathbf{0}$ with $L(\mathbf{v}_j) \neq \mathbf{0}$, which forces $a_j = 0$. Then repeat with another map, or on the shorter relation that remains.

Reach for it when the vectors come with an operator that acts on them in simple and distinguishable ways: a power of T along a chain $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots$; differentiation or a shift on exponentials and geometric sequences (Example 1.5.27, Exercise 1.5.44); evaluation at a point for functions that vanish at all but one point; a dot product with one of the vectors when they are mutually orthogonal (Exercise 1.5.46). The canonical example is Proposition 4.1.26. The same move shows that eigenvectors for distinct eigenvalues are independent (Section 7.2), and it organizes the Jordan chains of Section 10.3.

Used in. Example 1.5.27, Exercise 1.5.47, Proposition 4.1.26, Corollary 4.1.27, Exercise 4.1.45, Exercise 4.1.51, Exercise 4.1.60, Proposition 4.4.29, Proposition 4.6.28, Exercise 4.8.5, Exercise 4.8.8, Exercise 4.8.14, Problem 4.3, Problem 4.5, Problem 4.8, Theorem 7.2.21, Proposition 7.2.37, Exercise 7.2.51, Exercise 7.7.9, Problem 7.1, Problem 7.3, Problem 7.4, Problem 7.12, Problem 7.15, Exercise 4.5.50.

Proposition 4.1.26 (Nilpotent chains are independent). Let $T \in \mathrm{L}(V)$, let $\mathbf{v} \in V$, and let $k \geq 0$ be such that

$$T^k\mathbf{v} \neq \mathbf{0} \quad \text{and} \quad T^{k+1}\mathbf{v} = \mathbf{0}.$$

Then $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots, T^k\mathbf{v}$ is linearly independent.

Intuition. The chain $\mathbf{v} \mapsto T\mathbf{v} \mapsto \cdots \mapsto T^k\mathbf{v} \mapsto \mathbf{0}$ has a last nonzero vector. Applying T^k to the chain kills every vector except the first, which it sends to that last nonzero vector $T^k\mathbf{v}$. So applying T^k to a relation isolates the coefficient of $\mathbf{v}$. Once that coefficient is known to be 0, applying T^{k-1} isolates the next one, and so on down the chain.

Proof. STANDARD

Uses. Theorem 1.5.5 (Independence is decided at zero), Notation 2.1.18, Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Proposition 1.2.35 (Zero products)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Take a relation $a_0\mathbf{v} + a_1T\mathbf{v} + \cdots + a_kT^k\mathbf{v} = \mathbf{0}$ and apply T^k. Which terms survive?
Hint 2. $T^k(T^j\mathbf{v}) = T^{k+j}\mathbf{v}$, and this is $\mathbf{0}$ for $j \geq 1$, since $T^{k+j}\mathbf{v} = T^{j-1}(T^{k+1}\mathbf{v})$. So only $a_0T^k\mathbf{v}$ is left.
Hint 3. Prove $a_i = 0$ by induction on i: if $a_0 = \cdots = a_{i-1} = 0$, apply T^{k-i} to the relation.
Sketch. Let $\sum_{j=0}^{k} a_jT^j\mathbf{v} = \mathbf{0}$. Since $T^l\mathbf{v} = \mathbf{0}$ for all $l \geq k+1$, applying T^k leaves $a_0T^k\mathbf{v} = \mathbf{0}$, so $a_0 = 0$. If $a_0 = \cdots = a_{i-1} = 0$, applying T^{k-i} to $\sum_{j\geq i} a_jT^j\mathbf{v} = \mathbf{0}$ leaves $a_iT^k\mathbf{v} = \mathbf{0}$, so $a_i = 0$.

First, $T^l\mathbf{v} = \mathbf{0}$ for every $l \geq k+1$, because $T^l\mathbf{v} = T^{l-k-1}(T^{k+1}\mathbf{v}) = T^{l-k-1}(\mathbf{0}) = \mathbf{0}$ (Notation 2.1.18, Proposition 2.1.4). Suppose that

$$a_0\mathbf{v} + a_1T\mathbf{v} + \cdots + a_kT^k\mathbf{v} = \mathbf{0}. \tag{4.1.1}$$

We show by induction on $i \in \{0, 1, \dots, k\}$ that $a_i = 0$. Let i be given and assume that $a_j = 0$ for all $j < i$, which is no assumption when $i = 0$. Then (4.1.1) reads $\sum_{j=i}^{k} a_jT^j\mathbf{v} = \mathbf{0}$. Apply the linear map T^{k-i} to both sides:

$$\mathbf{0} = \sum_{j=i}^{k} a_jT^{k-i+j}\mathbf{v} = a_iT^k\mathbf{v} + \sum_{j=i+1}^{k} a_jT^{k+(j-i)}\mathbf{v} = a_iT^k\mathbf{v},$$

since every exponent $k + (j - i)$ in the last sum is at least $k + 1$. As $T^k\mathbf{v} \neq \mathbf{0}$, Proposition 1.2.35 gives $a_i = 0$. By induction all $a_i = 0$, so the list is independent by Theorem 1.5.5. □

apply	$a_0\mathbf{v}$	$a_1T\mathbf{v}$	$a_2T^2\mathbf{v}$	$a_3T^3\mathbf{v}$	gives
T^3	$a_0T^3\mathbf{v}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$a_0 = 0$
T^2	·	$a_1T^3\mathbf{v}$	$\mathbf{0}$	$\mathbf{0}$	$a_1 = 0$
T	·	·	$a_2T^3\mathbf{v}$	$\mathbf{0}$	$a_2 = 0$
id	·	·	·	$a_3T^3\mathbf{v}$	$a_3 = 0$

Table 4.1.1. The proof of Proposition 4.1.26 for $k = 3$. Each row applies a power of T to the relation $a_0\mathbf{v} + a_1T\mathbf{v} + a_2T^2\mathbf{v} + a_3T^3\mathbf{v} = \mathbf{0}$. Terms pushed past $T^3\mathbf{v}$ become $\mathbf{0}$; a dot marks a term whose coefficient is already known to be 0. Exactly one term survives in each row, and it is a multiple of $T^3\mathbf{v} \neq \mathbf{0}$.

The proposition did not assume that V is finite-dimensional. Combined with the bound on independent lists, it bounds how long an operator on a finite-dimensional space can take to reach 0. The striking part is that the bound depends only on $\dim V$: nothing about the particular operator enters.

Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps). Let V be finite-dimensional with $\dim V = n$, and let $T \in \mathrm{L}(V)$ with $T^m = 0$ for some $m \geq 1$. Then $T^n = 0$. Similarly, if $A \in F^{n\times n}$ and $A^m = 0$ for some $m \geq 1$, then $A^n = 0$.

Proof. STANDARD

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Theorem 2.4.2 (The matrix of a composite), Definition 2.3.15 (The linear map of a matrix)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Fix $\mathbf{v} \neq \mathbf{0}$. Since $T^m\mathbf{v} = \mathbf{0}$, there is a largest k with $T^k\mathbf{v} \neq \mathbf{0}$. What does Proposition 4.1.26 say about $\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}$?
Hint 2. That list has $k+1$ entries in an n-dimensional space, so $k+1 \leq n$. Now compute $T^n\mathbf{v}$.
Sketch. For $\mathbf{v} \neq \mathbf{0}$, let k be the largest integer with $T^k\mathbf{v} \neq \mathbf{0}$; it exists, and $k < m$, because $T^0\mathbf{v} = \mathbf{v} \neq \mathbf{0}$ and $T^m\mathbf{v} = \mathbf{0}$. By Proposition 4.1.26 the chain $\mathbf{v}, \ldots, T^k\mathbf{v}$ is independent, so $k+1 \leq n$. Hence $T^n\mathbf{v} = T^{n-k-1}\big(T^{k+1}\mathbf{v}\big) = \mathbf{0}$. For matrices, apply this to T_A.

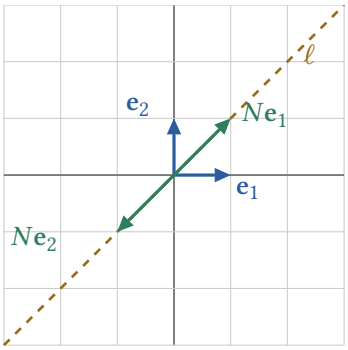
Let $\mathbf{v} \in V$. If $\mathbf{v} = \mathbf{0}$, then $T^n\mathbf{v} = \mathbf{0}$. Suppose $\mathbf{v} \neq \mathbf{0}$. The set of integers $j \geq 0$ with $T^j\mathbf{v} \neq \mathbf{0}$ contains 0, as $T^0\mathbf{v} = \mathbf{v}$, and contains no $j \geq m$, as $T^j\mathbf{v} = T^{j-m}(T^m\mathbf{v}) = \mathbf{0}$ for those. Let k be its largest element. Then $T^k\mathbf{v} \neq \mathbf{0}$ and $T^{k+1}\mathbf{v} = \mathbf{0}$, so by Proposition 4.1.26 the list $\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}$ is independent. It has $k+1$ entries, so $k+1 \leq n$ by Theorem 4.1.10(a). Therefore

$$T^n\mathbf{v} = T^{n-k-1}\big(T^{k+1}\mathbf{v}\big) = T^{n-k-1}(\mathbf{0}) = \mathbf{0}.$$

This holds for every $\mathbf{v}$, so $T^n = 0$.

For $A \in F^{n\times n}$, the linear map $T_A : F^n \to F^n$ satisfies $(T_A)^j = T_{A^j}$ for all $j \geq 1$, by Theorem 2.4.2 and induction on j. So $(T_A)^m = T_{A^m} = 0$, and by the first part $T_{A^n} = (T_A)^n = 0$. The columns of A^n are the vectors $A^n\mathbf{e}_j = T_{A^n}(\mathbf{e}_j) = \mathbf{0}$, so $A^n = 0$. □

The exponent n cannot be lowered in general: the shift N on F^4 above has $N^3 \neq 0$. And the corollary rules out operators that take longer. For instance, there is no $A \in \mathbb{R}^{3\times3}$ with $A^3 \neq 0$ but $A^4 = 0$, however cleverly its entries are chosen. In the plane, the corollary says that a nilpotent 2×2 matrix satisfies $A^2 = 0$.



The map of $N = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ (called A in the interactive version) flattens the plane onto the line ℓ through $(1,1)$, and it sends ℓ itself to $\mathbf{0}$, since $N(1,1) = \mathbf{0}$. So $N^2 = 0$. For any $\mathbf{v}$ off ℓ, such as $\mathbf{e}_1$, the chain $\mathbf{v}, N\mathbf{v}$ ends at $\mathbf{0}$; by Proposition 4.1.26 it is independent, and having length $2 = \dim \mathbb{R}^2$, it is a basis (Theorem 4.1.10(b)). In the interactive version, move the columns: a nonzero 2×2 matrix is nilpotent exactly when it flattens the plane onto a line and sends that line to $\mathbf{0}$. (Interactive in the web version.)

Over a field in which some integer is 0, the derivative dies much sooner than in characteristic 0, sooner than $\dim V$ steps.

Example 4.1.28 (Differentiation over $\mathbb{F}_p$). Let p be a prime and let D be the formal derivative on $\mathbb{F}_p[x]$ (Example 2.1.9). Then $D^p\big(x^j\big) = j(j-1)\cdots(j-p+1)x^{j-p}$ for $j \geq p$, and the product of the p consecutive integers $j, j-1, \ldots, j-p+1$ is divisible by p, so it is 0 in $\mathbb{F}_p$. For $j < p$, $D^p\big(x^j\big) = 0$ already. So $D^p = 0$ on all of $\mathbb{F}_p[x]$, an infinite-dimensional space.

For $p = 2$, on $\mathbb{F}_2[x]_{\leq 3}$ the chains are

$$x^3 \xrightarrow{D} x^2 \xrightarrow{D} \mathbf{0}, \quad x \xrightarrow{D} 1 \xrightarrow{D} \mathbf{0},$$

since $D\big(x^3\big) = 3x^2 = x^2$ and $D\big(x^2\big) = 2x = 0$. Here $D^2 = 0$ although $\dim \mathbb{F}_2[x]_{\leq 3} = 4$. The basis $x^3, x^2, x, 1$ splits into two chains of length 2 instead of one chain of length 4. How a nilpotent operator splits a space into chains is the subject of Section 10.3.

Exercise 4.1.29. ROUTINE For which n is there a matrix $A \in F^{n\times n}$ with $A^4 \neq 0$ and $A^5 = 0$? For each such n, give one.

Solution on page 1285

4.1.8 Too many vectors

Theorem 4.1.10 also works in the other direction. To prove that some linear relation exists, without being able to write one down, it is enough to produce more vectors than the dimension of a space that contains them.

Technique 4.1.30 (Too many vectors force a dependence). If a list of more than n vectors lies in a space of dimension at most n, it is linearly dependent (Theorem 4.1.10(a)), so it has a nontrivial linear relation. The relation is the output: it is a statement about the vectors that we did not have to guess. Natural sources of many vectors in a space of known dimension are powers $I, A, A^2, \ldots$ of a matrix, iterates $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ of a vector, and powers $1, \alpha, \alpha^2, \ldots$ of a number inside a space spanned by a few known quantities. Often the relation is then refined, for instance by choosing one with the smallest index of a nonzero coefficient (Technique 1.5.9).

Reach for it when you need an identity or an equation of a given shape and have no formula for it, but can count. The canonical example is Proposition 4.1.31. The iterates $\mathbf{v}, T\mathbf{v}, \ldots, T^n\mathbf{v}$ give eigenvalues

without determinants in Section 7.3; the powers of T lead to the minimal polynomial (Section 7.6); and the powers of a number lead to algebraic numbers (Section 14.2).
Used in. Exercise 2.5.49, Proposition 4.1.31, Exercise 4.1.33, Exercise 4.1.56, Exercise 4.1.57, Exercise 4.1.60, Exercise 4.8.4, Exercise 4.8.9, Problem 4.15, Proposition 7.2.35, Theorem 7.3.27, Proposition 7.6.4.

Proposition 4.1.31 (Every square matrix satisfies a polynomial).

(a) Let $A \in F^{n\times n}$ with $n \geq 1$. There are $c_0, c_1, \ldots, c_{n^2} \in F$, not all 0, with

$$c_0 I_n + c_1 A + c_2 A^2 + \cdots + c_{n^2} A^{n^2} = 0.$$

(b) Let V be finite-dimensional with $\dim V = n \geq 1$, and let $T \in \mathrm{L}(V)$. There are $c_0, \ldots, c_{n^2} \in F$, not all 0, with $c_0 \mathrm{id}_V + c_1 T + \cdots + c_{n^2} T^{n^2} = 0$.

Proof. ROUTINE

Uses. Example 4.1.7 (Dimensions of the standard spaces), Theorem 4.1.10 (Independent lists are short), Theorem 1.5.5 (Independence is decided at zero), Notation 2.4.16 (Powers of a matrix), Theorem 2.3.20 (Linear maps correspond to matrices)
Techniques. Too many vectors force a dependence (Technique 4.1.30), Translate between maps and matrices (Technique 2.3.27)
Hint 1. Count the matrices $I_n, A, A^2, \ldots, A^{n^2}$, and recall $\dim F^{n\times n}$.
Hint 2. For (b), apply (a) to $A = [T]_{\mathcal{B}\mathcal{B}}$ for a basis $\mathcal{B}$ of V, and use $\left[T^j\right]_{\mathcal{B}\mathcal{B}} = A^j$ (Notation 2.4.16).
Sketch. The list $I_n, A, \ldots, A^{n^2}$ has n^2+1 entries in $F^{n\times n}$, which has dimension n^2, so it is dependent, and a nontrivial relation is the claim. For (b), take the relation for $A = [T]_{\mathcal{B}\mathcal{B}}$; by Notation 2.4.16 it says $\left[\sum c_j T^j\right]_{\mathcal{B}\mathcal{B}} = 0 = [0]_{\mathcal{B}\mathcal{B}}$, and a map is determined by its matrix.

(a) The list $I_n, A, A^2, \ldots, A^{n^2}$ has n^2+1 entries, and it lies in $F^{n\times n}$, which has dimension n^2 (Example 4.1.7). By Theorem 4.1.10(a) it is linearly dependent, so by Theorem 1.5.5 it has a nontrivial linear relation. That relation is the claim.

(b) Fix a basis $\mathcal{B}$ of V and let $A = [T]_{\mathcal{B}\mathcal{B}} \in F^{n\times n}$. Take $c_0, \ldots, c_{n^2}$, not all 0, as in (a). By Notation 2.4.16,

$$\left[c_0 \mathrm{id}_V + c_1 T + \cdots + c_{n^2} T^{n^2}\right]_{\mathcal{B}\mathcal{B}} = c_0 I_n + c_1 A + \cdots + c_{n^2} A^{n^2} = 0 = [0]_{\mathcal{B}\mathcal{B}}.$$

By Theorem 2.3.20, a linear map $V \to V$ is determined by its matrix with respect to $\mathcal{B}$, so $c_0 \mathrm{id}_V + c_1 T + \cdots + c_{n^2} T^{n^2} = 0$. □

In the language of polynomials, A is a root of the nonzero polynomial $c_0 + c_1 x + \cdots + c_{n^2} x^{n^2}$. Evaluating polynomials at matrices and operators is made formal in Section 7.1 and Section 7.2, and Section 7.6 studies the polynomial of smallest degree that A satisfies. The bound n^2 is far from the truth: the Cayley–Hamilton theorem (Theorem 7.6.20) gives a nonzero polynomial of degree n. For now, the point is that the existence of *some* polynomial costs nothing beyond counting.

Example 4.1.32 ($\sqrt{2}+\sqrt{3}$ is a root of a rational polynomial). Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$ (Example 1.2.20), and let $\alpha = \sqrt{2}+\sqrt{3}$. Expanding,

$$\alpha^2 = 5 + 2\sqrt{6}, \quad \alpha^3 = 11\sqrt{2} + 9\sqrt{3}, \quad \alpha^4 = 49 + 20\sqrt{6}.$$

So $1, \alpha, \alpha^2, \alpha^3, \alpha^4$ all lie in $W = \mathrm{span}_{\mathbb{Q}}\left(1, \sqrt{2}, \sqrt{3}, \sqrt{6}\right)$, which has dimension at most 4 by Corollary 4.1.11(a). Five vectors in a space of dimension at most 4 are dependent: there are rationals $c_0, \ldots, c_4$, not all 0, with $c_0 + c_1\alpha + \cdots + c_4\alpha^4 = 0$. Before computing anything, we know that α is a root of a nonzero polynomial with rational coefficients and degree at most 4.

To find one, note that the even powers lie in $\mathrm{span}_{\mathbb{Q}}\left(1, \sqrt{6}\right)$, and look for $\alpha^4 + a\alpha^2 + b = 0$:

$$49 + 20\sqrt{6} + a\left(5 + 2\sqrt{6}\right) + b = (49 + 5a + b) + (20 + 2a)\sqrt{6}.$$

Taking $a = -10$ and $b = 1$ makes both brackets 0, so $\alpha^4 - 10\alpha^2 + 1 = 0$. Making the brackets 0 is sufficient whether or not $1, \sqrt{6}$ is independent over $\mathbb{Q}$. (It is: $a' + b'\sqrt{6} = 0$ with $b' \neq 0$ would make $\sqrt{6}$ rational.)

Exercise 4.1.33. ROUTINE Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. By Proposition 4.1.31, $I_2, A, \ldots, A^4$ satisfy a nontrivial relation. Find a nontrivial relation among I_2, A, A^2 alone, and use it to compute A^{10}.

Solution on page 1286

4.1.9 Dimension over a finite field

Over a finite field, dimension counts elements. In Exercise 1.5.43 and Exercise 1.6.38 we counted spans and bases in F^n by hand; with dimension available, the counts hold in every finite-dimensional space and come with a short proof.

Proposition 4.1.34 (Size of a vector space over a finite field). Let F be a finite field with q elements, and let V be a vector space over F. Then V is finite-dimensional if and only if V is a finite set, and in that case

$$|V| = q^{\dim V}.$$

Proof. ROUTINE

Uses. Proposition 1.6.16 (The coordinate map is a bijection), Definition 1.6.23 (Finite-dimensional), Definition 1.2.4 (The space F^n), Definition 4.1.5 (Dimension)
Hint 1. If V is a finite set, find a spanning list directly.
Hint 2. If $\dim V = n$, the coordinate map of a basis is a bijection $V \to F^n$. How many elements does F^n have?
Sketch. A finite V is spanned by the list of all its elements. Conversely, a basis of length $n = \dim V$ gives a bijection $V \to F^n$ (Proposition 1.6.16), and F^n has q^n elements, one for each choice of n entries.

If V is finite, the list of all elements of V spans V, so V is finite-dimensional. Conversely, suppose $\dim V = n$, and let $\mathcal{B}$ be a basis of V. By Proposition 1.6.16, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ is a bijection $V \to F^n$. An element of F^n is a list of n entries, each chosen from q elements, so $|F^n| = q^n$. Hence V is finite and $|V| = q^n$. □

So every subspace of $\mathbb{F}_p^n$ has a power of p as its number of elements, and a subspace with p^k elements has dimension k. For instance, no subspace of $\mathbb{F}_3^5$ has 30 elements, and a subspace of $\mathbb{F}_2^7$ with 16 elements is 4-dimensional.

Corollary 4.1.35 (Counting independent lists and bases). Let F be a finite field with q elements, and let V be a vector space over F with $\dim V = n$. For $0 \leq k \leq n$, the number of linearly independent lists of length k in V is

$$N_q(n,k) = (q^n - 1)(q^n - q)\left(q^n - q^2\right)\cdots\left(q^n - q^{k-1}\right).$$

In particular V has exactly $N_q(n,n)$ bases.

Proof. STANDARD

Uses. Proposition 4.1.34 (Size of a vector space over a finite field), Proposition 1.5.14 (Appending a vector), Corollary 1.5.17 (Sublists), Proposition 1.6.2 (Bases are independent spanning lists), Theorem 4.1.10 (Independent lists are short), Theorem 4.1.4 (All bases have the same length)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. An independent list of length $k+1$ is an independent list $\mathcal{L}$ of length k followed by a vector outside $\operatorname{span}(\mathcal{L})$. How many vectors does $\operatorname{span}(\mathcal{L})$ have?
Hint 2. $\mathcal{L}$ is a basis of $\operatorname{span}(\mathcal{L})$, so $|\operatorname{span}(\mathcal{L})| = q^k$ by Proposition 4.1.34. Induct on k. For bases, use Theorem 4.1.10(b).
Sketch. By Corollary 1.5.17 and Proposition 1.5.14, the independent lists of length $k+1$ are the lists $\mathcal{L}, \mathbf{w}$ with $\mathcal{L}$ independent of length k and $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$. Since $\mathcal{L}$ is a basis of $\operatorname{span}(\mathcal{L})$, that span has q^k elements, leaving $q^n - q^k$ choices of $\mathbf{w}$. Induction on k gives the product. The bases are exactly the independent lists of length n.

Let I_k be the number of independent lists of length k in V. We show $I_k = N_q(n,k)$ by induction on k. For $k = 0$, the only list of length 0 is (), which is independent, and $N_q(n,0) = 1$ is the empty product.

Let $0 \le k < n$. A list $\mathcal{L}, \mathbf{w}$ of length $k+1$ is independent if and only if $\mathcal{L}$ is independent and $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$: if it is independent, so is its sublist $\mathcal{L}$ (Corollary 1.5.17), and then $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$ by Proposition 1.5.14; the converse is Proposition 1.5.14. For an independent $\mathcal{L}$ of length k, the subspace $\operatorname{span}(\mathcal{L})$ has $\mathcal{L}$ as a basis (Proposition 1.6.2), so it has dimension k and q^k elements by Proposition 4.1.34. That leaves $q^n - q^k$ choices of $\mathbf{w}$ for each of the I_k choices of $\mathcal{L}$. So

$$I_{k+1} = I_k\left(q^n - q^k\right) = N_q(n,k)\left(q^n - q^k\right) = N_q(n,k+1).$$

Every basis of V has length n (Theorem 4.1.4) and is independent, and every independent list of length n is a basis (Theorem 4.1.10(b)). So the bases of V are the independent lists of length n, and there are $N_q(n,n)$ of them. □

By Theorem 3.5.14, a matrix in $F^{n\times n}$ is invertible if and only if its columns form a basis of F^n, and the columns determine the matrix. So there are exactly $N_q(n,n)$ invertible $n \times n$ matrices over a field with q elements. Over $\mathbb{F}_2$ this gives $3 \cdot 2 = 6$ invertible 2×2 matrices and $7 \cdot 6 \cdot 4 = 168$ invertible 3×3 matrices.

Counting subspaces BEYOND 21-242

Counting subspaces over finite fields is standard combinatorics, but it is not part of 21-242.

How many k-dimensional subspaces does $\mathbb{F}_2^3$ have? Each such subspace has many bases, and each independent list spans exactly one subspace. Counting the pairs (subspace, basis) in two ways answers the question, and the move is worth a name.

Technique 4.1.36 (Double counting). To count a set X that is hard to list, find a second set P of pairs (x,y) with $x \in X$ that is easy to count in another way. If every $x \in X$ occurs in the same number d of pairs, then $|P| = |X| \cdot d$, and $|X| = |P|/d$. Its cousin, the pigeonhole principle, compares two sets through a one-to-one assignment: the proof by exchange of Lemma 4.1.2 assigns to each $\mathbf{u}$ a different $\mathbf{w}$, which is why $m \le n$.

Reach for it when the objects to be counted are hard to enumerate but each one is described by many easy-to-count objects, such as subspaces by their bases. The canonical example is Proposition 4.1.37. Counting over finite fields returns in Section 14.3.

Used in. Proposition 4.1.37, Exercise 4.1.54, Exercise 4.4.52, Exercise 4.7.63, Problem 4.10, Problem 4.14, Exercise 7.3.52.

Proposition 4.1.37 (Counting subspaces). Let F be a finite field with q elements, let $\dim V = n$, and let $0 \le k \le n$. With N_q as in Corollary 4.1.35, the number of k-dimensional subspaces of V is

$$N_q \frac{n,k}{N_q}(k,k) = \frac{(q^n-1)(q^n-q)\cdots\left(q^n-q^{k-1}\right)}{\left(q^k-1\right)\left(q^k-q\right)\cdots\left(q^k-q^{k-1}\right)}.$$

Proof. STANDARD

Uses. Corollary 4.1.35 (Counting independent lists and bases), Theorem 1.6.10 (Every spanning list contains a basis), Definition 4.1.5 (Dimension)
Techniques. Double counting (Technique 4.1.36)
Hint 1. Count the pairs $(U, \mathcal{L})$, where U is a k-dimensional subspace of V and $\mathcal{L}$ is a basis of U.
Hint 2. Counted by U first: each U has $N_q(k,k)$ bases. Counted by $\mathcal{L}$ first: which lists $\mathcal{L}$ occur, and in how many pairs?
Sketch. Let P be the set of pairs $(U, \mathcal{L})$ with $\dim U = k$ and $\mathcal{L}$ a basis of U. Each k-dimensional U has $N_q(k,k)$ bases (Corollary 4.1.35 applied to U), so $|P| = s \cdot N_q(k,k)$, where s is the number of subspaces. Each independent list $\mathcal{L}$ of length k in V occurs in exactly one pair, with $U = \operatorname{span}(\mathcal{L})$, so $|P| = N_q(n,k)$.

Let s be the number of k-dimensional subspaces of V, and let P be the set of pairs $(U, \mathcal{L})$ in which U is a k-dimensional subspace of V and $\mathcal{L}$ is a basis of U. We count P in two ways.

By subspaces. A k-dimensional subspace U is a vector space over F of dimension k, so by Corollary 4.1.35 it has $N_q(k,k)$ bases. Hence $|P| = s \cdot N_q(k,k)$.

By lists. If $(U, \mathcal{L}) \in P$, then $\mathcal{L}$ is an independent list of length k in V, and $U = \operatorname{span}(\mathcal{L})$ is determined by $\mathcal{L}$. Conversely, every independent list $\mathcal{L}$ of length k in V is a basis of the subspace $\operatorname{span}(\mathcal{L})$ (Theorem 1.6.10), which then has dimension k, so $(\operatorname{span}(\mathcal{L}), \mathcal{L}) \in P$. So the pairs in P correspond to the independent lists of length k in V, and $|P| = N_q(n,k)$ by Corollary 4.1.35.

Comparing, $s \cdot N_q(k,k) = N_q(n,k)$. Since $N_q(k,k) \geq 1$ (a k-dimensional space has a basis), we may divide.□

Example 4.1.38 (The subspaces of $\mathbb{F}_2^3$). For $q = 2$ and $n = 3$: the number of 1-dimensional subspaces is $N_2(3,1)/N_2(1,1) = 7/1 = 7$, one for each nonzero vector, since a line over $\mathbb{F}_2$ is $\{\mathbf{0}, \mathbf{v}\}$. The number of 2-dimensional subspaces is

$$N_2 \frac{3,2}{N_2}(2,2) = \frac{7 \cdot 6}{3 \cdot 2} = 7.$$

Each plane has $2^2 = 4$ elements and contains 3 of the 7 lines. Taking the 7 lines as points and the 7 planes as lines, with "lies in" as incidence, gives the Fano plane of combinatorics.

Exercise 4.1.39. ROUTINE (a) Can a vector space over $\mathbb{F}_5$ have exactly 100 elements? (b) How many elements does a 3-dimensional subspace of $\mathbb{F}_5^7$ have? (c) How many bases does $\mathbb{F}_3^2$ have?

Solution on page 1286

Exercises

Exercise 4.1.40. ROUTINE Find a basis and the dimension of each subspace.

(a) $U_1 = \{\mathbf{x} \in F^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$.

(b) $U_2 = \{p \in F[x]_{\leq 3} \mid p(0) = p(1) = 0\}$.

(c) U_3, the upper triangular matrices in $F^{3\times 3}$, those with $a_{ij} = 0$ whenever $i > j$.

Solution on page 1286

Exercise 4.1.41. ROUTINE Let V be a vector space with $\dim V = 4$. Prove or disprove each statement.

(a) Every list of 5 vectors in V spans V.

(b) Every list of 3 vectors in V is linearly independent.

(c) If $\mathbf{v}_1, \ldots, \mathbf{v}_4$ spans V, then it is linearly independent.

(d) If U and W are subspaces of V with $\dim U = \dim W$, then $U = W$.

(e) If $U \subseteq W$ are subspaces of V with $\dim U = \dim W$, then $U = W$.

(f) $\mathbb{R}^2$ is a 2-dimensional subspace of $\mathbb{R}^3$.

Solution on page 1286

Exercise 4.1.42. ROUTINE Let $\mathbf{v}_1 = (0,0,1,2)$ and $\mathbf{v}_2 = (1,0,2,1)$. Extend $\mathbf{v}_1, \mathbf{v}_2$ to a basis of F^4 by sifting the list $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$, first for $F = \mathbb{Q}$ and then for $F = \mathbb{F}_3$. Explain why the answers differ.

Solution on page 1287

Exercise 4.1.43. ROUTINE Let $n \geq 1$. Show that the trace-zero matrices $U = \{A \in F^{n\times n} \mid \operatorname{tr} A = 0\}$ form a subspace of dimension $n^2 - 1$, and find a basis of U.

Solution on page 1287

Exercise 4.1.44. ROUTINE Let $A \in F^{n\times n}$ with $A^m = 0$ for some $m \geq 1$. Show that $I_n - A$ is invertible with inverse $I_n + A + A^2 + \cdots + A^{n-1}$, however large m is.

Solution on page 1288

Exercise 4.1.45. ROUTINE Let X be a set, and for $x \in X$ let $\delta_x \in F^X$ be the function with $\delta_x(x) = 1$ and $\delta_x(y) = 0$ for $y \neq x$.

(a) If $X = \{x_1, \ldots, x_n\}$ has n elements, show that $\delta_{x_1}, \ldots, \delta_{x_n}$ is a basis of F^X, so $\dim F^X = n$.

(b) If X is infinite, show that F^X is infinite-dimensional.

Solution on page 1288

Exercise 4.1.46. STANDARD Let $n \geq 1$. In $F^{n\times n}$, let S be the symmetric matrices ($a_{ij} = a_{ji}$ for all i, j) and K the skew-symmetric matrices ($a_{ij} = -a_{ji}$ for all i, j), as in Exercise 1.4.36.

(a) Show that $\dim S = n(n+1)/2$.

(b) Show that $\dim K = n(n-1)/2$ if $1 + 1 \neq 0$ in F, and $\dim K = n(n+1)/2$ if $1 + 1 = 0$.

Solution on page 1288

Exercise 4.1.47. STANDARD A matrix $A \in \mathbb{C}^{n\times n}$ is **Hermitian** if $a_{ji} = \overline{a_{ij}}$ for all i, j, where $\overline{z}$ is the complex conjugate of z (Definition B.1.1). Let H be the set of Hermitian matrices.

(a) Show that H is a subspace of $\mathbb{C}^{n\times n}$ regarded as a vector space over $\mathbb{R}$, but not a subspace of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$.

(b) Show that $\dim_{\mathbb{R}} H = n^2$, while $\dim_{\mathbb{R}} \mathbb{C}^{n\times n} = 2n^2$.

Solution on page 1289

Exercise 4.1.48. STANDARD Find the error in the following argument, and decide whether the claim is true.

Claim. Let $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V, and let U be a subspace of V. Then the entries of $\mathcal{B}$ that lie in U form a basis of U.

Proof. Let $\mathcal{B}'$ be the sublist of entries of $\mathcal{B}$ that lie in U. It is independent, as a sublist of an independent list. Let $\mathbf{u} \in U$, and write $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Since $\mathbf{u} \in U$ and U is a subspace, each term $a_i\mathbf{v}_i$ lies in U. So $a_i = 0$ whenever $\mathbf{v}_i \notin U$, and $\mathbf{u} \in \operatorname{span}(\mathcal{B}')$. Hence $\mathcal{B}'$ spans U.

Solution on page 1290

Exercise 4.1.49. STANDARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be linearly independent in V, with $m \geq 1$, and let $\mathbf{w} \in V$.

(a) Show that $\mathbf{v}_1 - \mathbf{w}, \ldots, \mathbf{v}_m - \mathbf{w}$ is linearly dependent if and only if $\mathbf{w} = t_1\mathbf{v}_1 + \cdots + t_m\mathbf{v}_m$ for some scalars with $t_1 + \cdots + t_m = 1$.

(b) Show that $\dim \operatorname{span}(\mathbf{v}_1 - \mathbf{w}, \ldots, \mathbf{v}_m - \mathbf{w}) \geq m - 1$ in all cases.

Solution on page 1290

Exercise 4.1.50. STANDARD Let $\dim V = n$.

(a) Show that if $U_0 \subsetneq U_1 \subsetneq \cdots \subsetneq U_k$ are subspaces of V, then $k \leq n$.

(b) Show that such a chain with $k = n$ exists, and that in every such chain $\dim U_i = i$ for all i. In particular $U_0 = \{\mathbf{0}\}$ and $U_n = V$.

Solution on page 1290

Exercise 4.1.51. STANDARD Let $n \geq 0$ and let F be a field in which $n! = 1 \cdot 2 \cdots n$ is nonzero, such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{F}_p$ with $p > n$. Let D be the formal derivative on $F[x]$ (Example 2.1.9).

(a) Let $p \in F[x]$ have degree n. Show that $p, Dp, D^2p, \ldots, D^np$ is a basis of $F[x]_{\leq n}$.

(b) Show that (a) fails for $F = \mathbb{F}_2$, $n = 2$, and $p = x^2$.

Solution on page 1291

Exercise 4.1.52. STANDARD Let $\dim V = n \geq 1$, and let $T \in \mathrm{L}(V)$ satisfy $T^{n-1} \neq 0$ and $T^n = 0$.

(a) Show that there is $\mathbf{v} \in V$ such that $\mathbf{v}, T\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is a basis of V.

(b) Find the matrix of T with respect to the basis $\mathcal{B} = T^{n-1}\mathbf{v}, \ldots, T\mathbf{v}, \mathbf{v}$.

Solution on page 1291

Exercise 4.1.53. STANDARD

(a) What fraction of the matrices in $\mathbb{F}_2^{3\times 3}$ are invertible?

(b) Show that for every $n \geq 1$, more than a quarter of the matrices in $\mathbb{F}_2^{n\times n}$ are invertible.

Solution on page 1292

Exercise 4.1.54. STANDARD BEYOND 21-242

(a) For $0 \leq k \leq 4$, find the number of k-dimensional subspaces of $\mathbb{F}_2^4$.

(b) Let F have q elements and $\dim V = n$. Show that V has as many k-dimensional subspaces as $(n-k)$-dimensional ones.

Solution on page 1292

Exercise 4.1.55. HARD Let $\dim V = n \geq 1$, and let $N \in \mathrm{L}(V)$ satisfy $N^{n-1} \neq 0$ and $N^n = 0$. Show that $S \in \mathrm{L}(V)$ commutes with N, that is, $SN = NS$, if and only if

$$S = c_0\mathrm{id}_V + c_1N + \cdots + c_{n-1}N^{n-1}$$

for some $c_0, \ldots, c_{n-1} \in F$.

Solution on page 1293

Exercise 4.1.56. HARD Let V be finite-dimensional, and let $T \in \mathrm{L}(V)$ be injective.

(a) Show that T is invertible, and that $T^{-1} = b_0\mathrm{id}_V + b_1T + \cdots + b_dT^d$ for some $d \geq 0$ and scalars $b_0, \ldots, b_d$.

(b) Let $M \in \mathrm{L}(F[x])$ be multiplication by x (Example 2.1.9). Show that M is injective but not surjective. Which step of (a) fails for M?

Solution on page 1293

Exercise 4.1.57. HARD Let $A_1, \ldots, A_{n+1}$ be nonempty subsets of $\{1, \ldots, n\}$. Show that there are disjoint nonempty sets $I, J \subseteq \{1, \ldots, n+1\}$ with

$$\bigcup_{i\in I} A_i = \bigcup_{j\in J} A_j.$$

Solution on page 1294

Exercise 4.1.58. HARD Show that for every field F and every $n \geq 1$, the space $F^{n\times n}$ has a basis consisting of invertible matrices. Write one down for $\mathbb{F}_2^{2\times 2}$.

Solution on page 1294

Exercise 4.1.59. HARD Let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_n$ be bases of V, and fix i. Show that there is an index j such that both of the following are bases of V:

(a) $\mathcal{B}$ with $\mathbf{b}_i$ replaced by $\mathbf{c}_j$;

(b) $\mathcal{C}$ with $\mathbf{c}_j$ replaced by $\mathbf{b}_i$.

Solution on page 1295

Exercise 4.1.60. CHALLENGE Let F be a field of characteristic 0 (Definition 1.1.28), such as $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, and let $A, B \in F^{n\times n}$ satisfy $AB - BA = A$. Show that $A^n = 0$.

Solution on page 1295

4.2 Rank–nullity

A linear map $T : V \to W$ collapses $\ker T$ to $\mathbf{0}$ and reaches $\operatorname{im} T$. In Section 2.6 we measured the two separately; with dimension available we can compare them. The answer is the rank–nullity theorem: if V is finite-dimensional, then

$$\dim \ker T + \dim \operatorname{im} T = \dim V.$$

Informally, each of the $\dim V$ independent directions in V is either collapsed by T or carried over to W. For a matrix, the statement is that every column is either a pivot column or a free column.

We follow the lecture's route. Restricting a map to a subspace (Section 4.2.1) lets us compute $\dim \operatorname{im} T_A = \operatorname{rank} A$ (Section 4.2.3), and counting free columns then proves rank–nullity for matrices. A second proof, which extends a basis of $\ker T$, works directly for every linear map (Section 4.2.4). The rest of the section draws consequences. Among them are the fact that the rank of a matrix of a linear map does not depend on the bases, which closes a question left open in Section 3.5, and dimension counting, a way to prove that vectors with some property exist without finding them. We end with an application outside the course, error-correcting codes over $\mathbb{F}_2$.

Throughout, V and W are vector spaces over a field F.

4.2.1 Restriction

Restricting a function keeps its rule and shrinks its set of inputs. We have already written $\operatorname{im}(T\restriction_X)$ for the set of values of T on a subset X (Proposition 2.1.16, Proposition 2.6.15). Here is the definition behind the notation.

Definition 4.2.1 (Restriction). Let $g : X \to Y$ be a function and let $A \subseteq X$. The **restriction** of g to A is the function

$$g\restriction_A : A \to Y, \quad a \mapsto g(a).$$

Only the domain changes. The target is still Y, even when every value of $g{\restriction}_A$ happens to lie in some smaller set. The image of the restriction,

$$\operatorname{im}(g{\restriction}_A) = \{g(a) \mid a \in A\},$$

is the set of values of g on A, which Section 2.1.4 wrote as $T(A)$; from now on we always write $\operatorname{im}(g{\restriction}_A)$. (For any function $g : X \to Y$, $\operatorname{im} g = \{g(x) \mid x \in X\}$, as for linear maps.)

Example 4.2.2 (What restriction can change).

(a) Let $P : \mathbb{R}^2 \to \mathbb{R}^2$, $P(x, y) = (x, 0)$, be the projection onto the x-axis, and let $\ell = \operatorname{span}((1, 1))$. Then $P{\restriction}_\ell : \ell \to \mathbb{R}^2$ sends $(t, t) \mapsto (t, 0)$. It is injective, although P is not, and its image is the x-axis, all of $\operatorname{im} P$. Restricted to the y-axis instead, P becomes the zero map.

(b) Let $Q : \mathbb{R}^2 \to \mathbb{R}$, $Q(x, y) = x + y$. It is surjective. Its restriction to the x-axis is still surjective (and now injective as well), but its restriction to the line $y = -x$ is the zero map, which is not surjective.

The lecture records the general behavior.

Proposition 4.2.3 (Properties of restriction).

(a) If $g : X \to Y$ is injective, then $g{\restriction}_A$ is injective for every $A \subseteq X$.

(b) If $f : X \to Y$ and $g : Y \to Z$ are functions, then $\operatorname{im}(g \circ f) = \operatorname{im}\left(g{\restriction}_{\operatorname{im} f}\right)$.

(c) If $T \in \mathrm{L}(V, W)$ and U is a subspace of V, then $T{\restriction}_U \in \mathrm{L}(U, W)$ and $\ker(T{\restriction}_U) = U \cap \ker T$. In particular, $T{\restriction}_U$ is injective if and only if $U \cap \ker T = \{\mathbf{0}\}$.

Proof. ROUTINE

Uses. Definition 4.2.1 (Restriction), Definition 1.4.1 (Subspace), Definition 2.1.1 (Linear map), Definition 2.6.5 (Kernel), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Each part is an unwinding of definitions. For (b), write out both sides as sets of values.
Hint 2. For (c), the operations of U are those of V, so the linearity identities for $T{\restriction}_U$ are instances of those for T. Which $\mathbf{u} \in U$ does $T{\restriction}_U$ send to $\mathbf{0}$?
Sketch. (a) $g(a) = g(a')$ forces $a = a'$ already in X. (b) Both sides are $\{g(f(x)) \mid x \in X\}$. (c) Linearity is inherited, since U is closed under the operations; the kernel is the set of $\mathbf{u} \in U$ with $T(\mathbf{u}) = \mathbf{0}$, which is $U \cap \ker T$; apply Theorem 2.6.11.

(a) If $a, a' \in A$ and $g{\restriction}_A(a) = g{\restriction}_A(a')$, then $g(a) = g(a')$, so $a = a'$ because g is injective.

(b) By definition,

$$\operatorname{im}(g \circ f) = \{g(f(x)) \mid x \in X\} = \{g(y) \mid y \in \operatorname{im} f\} = \operatorname{im}\left(g{\restriction}_{\operatorname{im} f}\right).$$

For linear maps this is Proposition 2.6.15(b), whose proof did not use linearity.

(c) By Definition 1.4.1, U is a vector space with the operations of V. For $\mathbf{u}, \mathbf{u}' \in U$ and $a \in F$, the vectors $\mathbf{u} + \mathbf{u}'$ and $a\mathbf{u}$ lie in U, and

$$T{\restriction}_U(\mathbf{u} + \mathbf{u}') = T(\mathbf{u}) + T(\mathbf{u}'), \quad T{\restriction}_U(a\mathbf{u}) = aT(\mathbf{u})$$

because T is linear. So $T{\restriction}_U$ is linear (Definition 2.1.1). Its kernel is $\{\mathbf{u} \in U \mid T(\mathbf{u}) = \mathbf{0}\} = U \cap \ker T$. The last statement is Theorem 2.6.11 applied to $T{\restriction}_U$. □

Part (a) and Example 4.2.2(b) are the lecture's remarks that restriction preserves injectivity but not surjectivity. Shrinking the domain removes inputs, which can remove coincidences of values but can also remove values. Part (b) says that the composite $g \circ f$ uses g only on the part of Y that f reaches.

Exercise 4.2.4. ROUTINE Let $D : \mathbb{R}[x]_{\leq 3} \to \mathbb{R}[x]_{\leq 3}$ be differentiation.

(a) Let $U = \operatorname{span}(x^2, x^3)$. Find $\ker(D|_U)$ and $\operatorname{im}(D|_U)$.

(b) Let $U' = \operatorname{span}(1, x)$. Find $\ker(D|_{U'})$ and $\operatorname{im}(D|_{U'})$.

(c) Find a subspace U'' such that $D|_{U''}$ is injective and $\operatorname{im}(D|_{U''}) = \operatorname{im} D$.

Solution on page 1296

Restriction is a small idea that we will use constantly, so we name it.

Technique 4.2.5 (Restrict to a subspace). To study a linear map on part of its domain, restrict it. For a subspace U of V, the map $T|_U : U \to W$ is linear, its kernel is $U \cap \ker T$, and its image is the set of values of T on U (Proposition 4.2.3). Everything known about linear maps then applies to $T|_U$, which is often better behaved than T: it can be injective when T is not, and its domain is smaller. The most common entry point is the identity $\operatorname{im}(TS) = \operatorname{im}(T|_{\operatorname{im} S})$, which says that the composite TS sees T only on the subspace $\operatorname{im} S$.

Reach for it when a composite is involved, when a hypothesis concerns only part of V (an image, a kernel, a subspace that T maps into itself), or when T is injective on a subspace you care about. The canonical example is the proof of Theorem 4.2.7, which restricts $T_{G^{-1}}$ to $\operatorname{im} T_{A_r}$. Later the same move bounds the rank of a product (Section 4.3) and turns an operator into a smaller operator on an invariant subspace (Section 7.2).

Used in. Proposition 4.2.6, Theorem 4.2.7, Proposition 4.2.15, Proposition 4.2.26, Exercise 4.2.44, Exercise 4.2.52, Exercise 4.2.53, Proposition 4.3.1, Theorem 4.3.3, Proposition 4.3.5, Proposition 4.6.29, Exercise 4.8.7, Exercise 4.8.10, Exercise 4.8.11, Problem 4.2, Problem 4.18, Corollary 6.6.22, Proposition 7.6.10, Proposition 7.6.12, Theorem 7.6.31, Notation 8.3.10, Exercise 4.6.52.

4.2.2 Injective maps preserve dimension

An injective linear map loses nothing, so it should not lower dimension. The lecture proves this by pushing a basis forward, which we already did in Proposition 2.6.22.

Proposition 4.2.6 (Injective maps preserve dimension). Let $T \in \mathrm{L}(V, W)$.

(a) If V is finite-dimensional and T is injective, then $\operatorname{im} T$ is finite-dimensional and $\dim \operatorname{im} T = \dim V$. More precisely, for every basis $\mathcal{B}$ of V, the list $T(\mathcal{B})$ is a basis of $\operatorname{im} T$.

(b) If U is a finite-dimensional subspace of V with $U \cap \ker T = \{\mathbf{0}\}$, then $\dim \operatorname{im}(T|_U) = \dim U$.

Proof. ROUTINE

Uses. Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Proposition 4.2.3 (Properties of restriction), Definition 4.1.5 (Dimension)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. Take a basis $\mathcal{B}$ of V. What does Proposition 2.6.22(d) say about $T(\mathcal{B})$?
Hint 2. For (b), find an injective map to which (a) applies.
Sketch. (a) For injective T, the list $T(\mathcal{B})$ spans $\operatorname{im} T$ and is independent (Proposition 2.6.22(d)), so it is a basis of $\operatorname{im} T$ with $\dim V$ entries. (b) $T|_U$ is injective by Proposition 4.2.3(c); apply (a) to it.

(a) Let $\mathcal{B}$ be a basis of V, of length $\dim V$. By Proposition 2.6.22(d), $T(\mathcal{B})$ spans $\operatorname{im} T$, and it is a basis of $\operatorname{im} T$ because T is injective. So $\operatorname{im} T$ is finite-dimensional, and $\dim \operatorname{im} T$ is the length of $T(\mathcal{B})$, which is $\dim V$.

(b) By Proposition 4.2.3(c), $T|_U \in \mathrm{L}(U, W)$ is injective. Applying (a) to it gives $\dim \operatorname{im}(T|_U) = \dim U$. □

Without injectivity, $T(\mathcal{B})$ still spans $\operatorname{im} T$, so $\operatorname{im} T$ is finite-dimensional with $\dim \operatorname{im} T \leq \dim V$ (Corollary 4.1.11(a)), and equality holds only if the spanning list $T(\mathcal{B})$ is a basis, that is, only if T is injective.

4.2.3 The image of a matrix

We now compute $\dim \operatorname{im} T_A$. The lecture's strategy has three steps:

- write $A = G^{-1}A_r$, where $A_r = \operatorname{rref}(A)$ and G is invertible;
- show that the pivot columns of A_r form a basis of $\operatorname{im} T_{A_r}$, which is easy because they are standard basis vectors;
- slide this basis along G^{-1}, which is injective, to get a basis of $\operatorname{im} T_A$.

The last step lands on the pivot columns of A itself, because G^{-1} times a column of A_r is the corresponding column of A.

Theorem 4.2.7 (The image of T_A). Let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, let $k = \operatorname{rank} A$, and let $p_1 < \cdots < p_k$ be the pivot columns of A. Then $\mathbf{a}_{p_1}, \dots, \mathbf{a}_{p_k}$ is a basis of $\operatorname{im} T_A$. In particular,

$$\dim \operatorname{im} T_A = \operatorname{rank} A.$$

Proof (through the reduced form, by restriction). STANDARD

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Proposition 3.3.8 (Properties of the output), Proposition 2.6.18 (The column space is spanned by the columns), Corollary 1.3.19 (Equal spans), Theorem 2.4.2 (The matrix of a composite), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Definition 2.4.1 (Matrix product)
Techniques. Restrict to a subspace (Technique 4.2.5), Reduce to RREF (Technique 3.3.22)
Hint 1. First solve the problem for A_r. Which columns of A_r are pivot columns, and in which rows can the other columns have nonzero entries?
Hint 2. Write $A = G^{-1}A_r$, so $T_A = T_{G^{-1}} \circ T_{A_r}$. By Proposition 4.2.3(b), $\operatorname{im} T_A$ is the image of $T_{G^{-1}}$ restricted to $\operatorname{im} T_{A_r}$.
Hint 3. $T_{G^{-1}}$ is injective, so it carries a basis of $\operatorname{im} T_{A_r}$ to a basis of $\operatorname{im} T_A$. Compute $G^{-1}\mathbf{e}_s$ column by column from $A = G^{-1}A_r$.
Sketch. The pivot columns of A_r are $\mathbf{e}_1, \dots, \mathbf{e}_k$, and every column of A_r vanishes below row k, so $\mathbf{e}_1, \dots, \mathbf{e}_k$ is a basis of $\operatorname{im} T_{A_r}$. With $GA = A_r$, $T_A = T_{G^{-1}}T_{A_r}$, so $\operatorname{im} T_A = \operatorname{im}\left(T_{G^{-1}}|_{\operatorname{im} T_{A_r}}\right)$. The injective map $T_{G^{-1}}$ sends the basis $\mathbf{e}_1, \dots, \mathbf{e}_k$ to a basis of $\operatorname{im} T_A$, and $G^{-1}\mathbf{e}_s = G^{-1}$ (column p_s of A_r) $= \mathbf{a}_{p_s}$.

Let $G \in F^{m\times m}$ be invertible with $GA = A_r$ (Theorem 3.3.10), so that $A = G^{-1}A_r$.

The reduced matrix. By Lemma 3.3.16(b), the pivot columns of A_r are $\mathbf{e}_1, \dots, \mathbf{e}_k \in F^m$, which are linearly independent. By Proposition 3.3.8, rows $k+1, \dots, m$ of A_r are zero, so every column of A_r lies in $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$. Since $\mathbf{e}_1, \dots, \mathbf{e}_k$ are themselves columns of A_r, Corollary 1.3.19 and Proposition 2.6.18 give

$$\operatorname{im} T_{A_r} = \operatorname{span}(\text{columns of } A_r) = \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k),$$

and $\mathbf{e}_1, \dots, \mathbf{e}_k$ is a basis of $\operatorname{im} T_{A_r}$.

Sliding by G^{-1}. By Theorem 2.4.2, $T_A = T_{G^{-1}A_r} = T_{G^{-1}} \circ T_{A_r}$, so Proposition 4.2.3(b) gives

$$\operatorname{im} T_A = \operatorname{im}\left(T_{G^{-1}}|_{\operatorname{im} T_{A_r}}\right).$$

The map $T_{G^{-1}}$ is injective, because T_G is a left inverse of it: $T_G \circ T_{G^{-1}} = T_{GG^{-1}} = T_{I_m} = \mathrm{id}$ (Proposition 2.5.4). Its restriction to $\operatorname{im} T_{A_r}$ is injective as well (Proposition 4.2.3(a)), so by Proposition 4.2.6(a) it sends the basis $\mathbf{e}_1, \dots, \mathbf{e}_k$ of $\operatorname{im} T_{A_r}$ to a basis $G^{-1}\mathbf{e}_1, \dots, G^{-1}\mathbf{e}_k$ of $\operatorname{im} T_A$. Finally, column p_s of $A = G^{-1}A_r$ is G^{-1} times column p_s of A_r (Definition 2.4.1), that is,

$$\mathbf{a}_{p_s} = G^{-1}\mathbf{e}_s \quad (s = 1, \dots, k).$$

So $\mathbf{a}_{p_1}, \dots, \mathbf{a}_{p_k}$ is a basis of $\operatorname{im} T_A$, and $\dim \operatorname{im} T_A = k = \operatorname{rank} A$. □

The second proof skips G and uses the description of pivot columns from Section 3.3.4, which never mentions the algorithm.

Proof (by sifting the columns). ROUTINE

Uses. Corollary 3.3.19 (Pivot columns, intrinsically), Corollary 1.5.13 (Independence, one entry at a time), Corollary 1.3.19 (Equal spans), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 1.6.2 (Bases are independent spanning lists)

Hint 1. Corollary 3.3.19 describes the pivot columns of A directly. Which of its two parts gives independence, and which gives spanning?
Hint 2. Independence: $\mathbf{a}_{p_s}$ is not a combination of $\mathbf{a}_1, \dots, \mathbf{a}_{p_s-1}$, so in particular not of the earlier pivot columns. Use Corollary 1.5.13.
Hint 3. Spanning: every free column is a combination of pivot columns (Corollary 3.3.19(b)).
Sketch. By Corollary 3.3.19(a), $\mathbf{a}_{p_s} \notin \operatorname{span}\left(\mathbf{a}_1, \dots, \mathbf{a}_{p_s-1}\right) \supseteq \operatorname{span}\left(\mathbf{a}_{p_1}, \dots, \mathbf{a}_{p_{s-1}}\right)$, so the pivot columns are independent by Corollary 1.5.13. By (b), every free column lies in their span, so their span contains all the columns and equals $\operatorname{im} T_A$.

Let $\mathcal{P} = \mathbf{a}_{p_1}, \dots, \mathbf{a}_{p_k}$.

Independent. Let $1 \leq s \leq k$. By Corollary 3.3.19(a), $\mathbf{a}_{p_s}$ is not a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_{p_s-1}$. Since $p_1, \dots, p_{s-1} < p_s$, it is in particular not a combination of $\mathbf{a}_{p_1}, \dots, \mathbf{a}_{p_{s-1}}$. By Corollary 1.5.13, $\mathcal{P}$ is linearly independent.

Spanning. Each pivot column is an entry of $\mathcal{P}$, and by Corollary 3.3.19(b) each free column is a linear combination of pivot columns. So every column of A lies in $\operatorname{span}(\mathcal{P})$, and every entry of $\mathcal{P}$ is a column of A. By Corollary 1.3.19 and Proposition 2.6.18, $\operatorname{span}(\mathcal{P}) = \operatorname{span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \operatorname{im} T_A$.

So $\mathcal{P}$ is a basis of $\operatorname{im} T_A$ (Proposition 1.6.2), of length $k = \operatorname{rank} A$. □

The first proof does the easy case A_r and transports the answer along an injective map. The second is the sifting procedure of Section 1.6.3 applied to the columns of A: scanning left to right and keeping each column that is not a combination of the earlier ones keeps exactly the pivot columns. Both rest on one fact, that row operations do not change the linear relations among the columns (Corollary 3.2.17). The first proof uses it through G; the second uses it through Corollary 3.3.19, which was proved from it.

The theorem names columns of A, not of A_r, and the distinction matters.

Non-example 4.2.8 (The columns of A_r are the wrong vectors). Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, so $A_r = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$. Then $\operatorname{im} T_{A_r} = \operatorname{span}(\mathbf{e}_1)$, while $\operatorname{im} T_A = \operatorname{span}((1,2))$, a different line. The pivot column $\mathbf{e}_1$ of A_r does not even lie in $\operatorname{im} T_A$. Row operations change the column space. What they preserve is the linear relations among the columns, so A_r tells us *which* columns to take, and we must then take them from A.

For the kernel, the free columns do the same job. Theorem 3.4.9 writes every vector of $\ker T_A$ as a combination of the basic solutions, one for each free column, and Remark 3.4.11 observed that they form a basis. So the reduced form gives bases of both subspaces at once.

Example 4.2.9 (Bases for the kernel and the image). Let

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 & 1 \\ 2 & 4 & 1 & 8 & 1 \\ -1 & -2 & 2 & 1 & -3 \end{pmatrix} \in \mathbb{Q}^{3\times 5}, \quad A_r = \begin{pmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The pivot columns are $1, 3$ and the free columns are $2, 4, 5$, so $\operatorname{rank} A = 2$.

Image. By Theorem 4.2.7, columns 1 and 3 of A form a basis of $\operatorname{im} T_A$:

$$\operatorname{im} T_A = \operatorname{span}((1,2,-1),(0,1,2)),$$

a plane in $\mathbb{Q}^3$. (It does not contain the pivot column $\mathbf{e}_1$ of A_r: $a(1,2,-1) + b(0,1,2) = (1,0,0)$ forces $a = 1$, then $b = -2$, and then the third entry is -5.)

Kernel. Copying the free columns of A_r into Definition 3.4.8, the basic solutions are

$$\mathbf{s}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} -3 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

and they form a basis of $\ker T_A$. Each one records how a free column is built from pivot columns: $A\mathbf{s}_2 = \mathbf{0}$ says $\mathbf{a}_4 = 3\mathbf{a}_1 + 2\mathbf{a}_3$, and $A\mathbf{s}_3 = \mathbf{0}$ says $\mathbf{a}_5 = \mathbf{a}_1 - \mathbf{a}_3$.

So $\dim \operatorname{im} T_A = 2$ and $\dim \ker T_A = 3$, one for each pivot column and one for each free column, and $2 + 3 = 5$ is the number of columns.

$$\begin{pmatrix} 1 & 2 & 0 & 3 & 1 \\ 2 & 4 & 1 & 8 & 1 \\ -1 & -2 & 2 & 1 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The reduction of Example 4.2.9. At the end the summary lists the pivot columns, whose columns *in* A form a basis of $\operatorname{im} T_A$, and the free columns, which index the basic solutions, a basis of $\ker T_A$. Use "Edit" to try your own matrix: the number of pivot columns plus the number of free columns is always the number of columns. (Interactive in the web version.)

Exercise 4.2.10. ROUTINE Let $A = \begin{pmatrix} 1 & -1 & 2 & 0 \\ 2 & -2 & 5 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 4}$.

(a) Find a basis of $\operatorname{im} T_A$ consisting of columns of A, and a basis of $\ker T_A$.

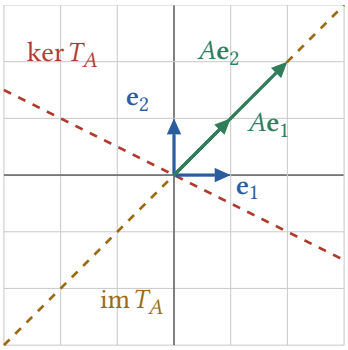
(b) Is $(1, 0, 0)$ in $\operatorname{im} T_A$?

Solution on page 1297

4.2.4 Rank–nullity

Intuition (Lost plus reached). Take $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$. Both columns lie on the line $y = x$, so T_A flattens the plane onto that line: $\operatorname{im} T_A = \operatorname{span}((1, 1))$, of dimension 1. The vectors it sends to $\mathbf{0}$ are those with $x + 2y = 0$, so $\ker T_A = \operatorname{span}((2, -1))$, also of dimension 1. One dimension of $\mathbb{R}^2$ is lost and the other survives, and $1 + 1 = 2$.

There is more to see. Any line ℓ through $\mathbf{0}$ other than $\ker T_A$ meets the kernel only in $\mathbf{0}$, so $T_A|_\ell$ is injective, and it maps ℓ onto the image line. So the plane consists of the directions T_A kills, together with a line that T_A carries faithfully onto $\operatorname{im} T_A$. In general, V is made of $\ker T$ together with further directions on which T is injective; those are carried to independent directions that fill out $\operatorname{im} T$, and counting gives $\dim V = \dim \ker T + \dim \operatorname{im} T$. For a matrix the two kinds of direction have names: free variables and pivot variables.



The map of $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ squashes the whole plane onto its image, the line $y = x$, and sends its kernel, the line $x + 2y = 0$, to $\mathbf{0}$. In the interactive version, drag the columns Ae_1 and Ae_2. While they are independent the image is the whole plane and the kernel is $\{\mathbf{0}\}$; when they line up, the image drops to a line and a line of kernel appears (and if both are $\mathbf{0}$, the image is $\{\mathbf{0}\}$ and the kernel is the plane), so the two dimensions always add up to 2. (Interactive in the web version.)

Theorem 4.2.11 (Rank–nullity).
(a) Let $A \in F^{m\times n}$. Then

$$\dim \ker T_A + \operatorname{rank} A = n.$$

(b) Let V be finite-dimensional and let $T \in \mathrm{L}(V, W)$. Then $\operatorname{im} T$ is finite-dimensional, and

$$\dim \ker T + \dim \operatorname{im} T = \dim V.$$

In (a), n is the number of columns, the dimension of the domain F^n; the number of rows plays no role. In (b), only V is assumed finite-dimensional, and W may be any vector space. The first proof is the lecture's: it counts pivots for matrices and transfers the count to arbitrary maps with coordinates.

Proof (by counting pivots). STANDARD

Uses. Proposition 3.5.2 (First properties of rank), Remark 3.4.11 (A basis of the kernel), Theorem 4.2.7 (The image of T_A), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Definition 1.6.23 (Finite-dimensional), Proposition 2.6.20 (Kernel and image in coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 1.6.16 (The coordinate map is a bijection), Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Reduce to RREF (Technique 3.3.22), Translate between maps and matrices (Technique 2.3.27)
Hint 1. For (a): the free columns index a basis of $\ker T_A$, and the pivot columns of A form a basis of $\operatorname{im} T_A$. How many of each are there?
Hint 2. For (b), choose bases and pass to a matrix of T. A matrix needs a finite-dimensional target. Why may you replace W by $\operatorname{im} T$, and why is $\operatorname{im} T$ finite-dimensional?
Hint 3. The coordinate maps are injective and linear, and by Proposition 2.6.20 they carry $\ker T$ onto $\ker T_A$ and $\operatorname{im} T$ onto $\operatorname{im} T_A$. What do injective linear maps do to dimension?
Sketch. (a) The $n - \operatorname{rank} A$ basic solutions form a basis of $\ker T_A$ (Remark 3.4.11), and $\dim \operatorname{im} T_A = \operatorname{rank} A$ (Theorem 4.2.7). (b) $\operatorname{im} T$ is spanned by $T(\mathcal{B})$ for a basis $\mathcal{B}$ of V, so it is finite-dimensional, say of dimension r. The cases $\dim V = 0$ and $r = 0$ are immediate. Otherwise regard T as a surjective map $T' : V \to \operatorname{im} T$ and let $A = [T']_{\mathcal{C}\mathcal{B}} \in F^{r\times n}$. Then T_A is surjective, so $\operatorname{rank} A = r$, and the coordinate map restricted to $\ker T$ is an injective linear map onto $\ker T_A$, so $\dim \ker T = \dim \ker T_A = n - r$ by (a).

(a) Let $k = \operatorname{rank} A$. By Proposition 3.5.2(b), A has exactly $n - k$ free columns, and by Remark 3.4.11 the basic solutions, one for each free column, form a basis of $\ker T_A$. So $\dim \ker T_A = n - k$, and $\dim \ker T_A + \operatorname{rank} A = (n - k) + k = n$. (By Theorem 4.2.7, the k here is also $\dim \operatorname{im} T_A$: the pivot columns count the image, and the free columns count the kernel.)

(b) Let $n = \dim V$ and let $\mathcal{B}$ be a basis of V. By Proposition 2.1.16(a), $T(\mathcal{B})$ spans $\operatorname{im} T$, so $\operatorname{im} T$ is finite-dimensional (Definition 1.6.23). Let $r = \dim \operatorname{im} T$. If $n = 0$, then $V = \{\mathbf{0}\}$ and both sides are 0; if $r = 0$, then $T = 0$ and $\ker T = V$. So assume $n, r \geq 1$, and let $\mathcal{C}$ be a basis of $\operatorname{im} T$. The function

$$T' : V \to \operatorname{im} T, \quad T'(\mathbf{v}) = T(\mathbf{v}),$$

is T with its target shrunk to $\operatorname{im} T$. It is linear, it is surjective, and $\ker T' = \ker T$. Let $A = [T']_{\mathcal{C}\mathcal{B}} \in F^{r \times n}$.

Since T' is surjective, so is T_A, by Proposition 2.6.20(c). So $\operatorname{im} T_A = F^r$ and $\operatorname{rank} A = r$ by Theorem 4.2.7. By the same proposition, the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ sends $\ker T'$ onto $\ker T_A$. That map is linear and injective (Proposition 1.6.19, Proposition 1.6.16), so its restriction to the subspace $\ker T$ is an injective linear map with image $\ker T_A$ (Proposition 4.2.3). Since $\ker T$ is finite-dimensional (Proposition 4.1.21(a)), Proposition 4.2.6(a) gives $\dim \ker T = \dim \ker T_A$. By (a), $\dim \ker T_A = n - r$. Hence

$$\dim \ker T + \dim \operatorname{im} T = (n - r) + r = n = \dim V. \qquad \square$$

The second proof is the one in LADR. It never chooses coordinates, and it explains the formula rather than verifying it: it finds the directions of the intuition above inside V.

Proof (by extending a basis of the kernel). STANDARD

> *Uses.* Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.6.2 (Bases are independent spanning lists), Definition 2.6.5 (Kernel), Theorem 4.2.7 (The image of T_A)
> *Techniques.* Extend a basis of a subspace (Technique 4.1.18)
> *Hint 1.* Start from a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ of $\ker T$ and extend it to a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_r$ of V. Which vectors should form a basis of $\operatorname{im} T$?
> *Hint 2.* $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)$ span $\operatorname{im} T$: apply T to an arbitrary vector, written in the basis, and watch the $\mathbf{u}_i$ disappear.
> *Hint 3.* For independence: if $\sum_j b_j T(\mathbf{v}_j) = \mathbf{0}$, then $\sum_j b_j \mathbf{v}_j \in \ker T$, so it is a combination of the $\mathbf{u}_i$. Now use the independence of the whole basis of V.
> *Sketch.* $\ker T$ has a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ (Proposition 4.1.21(a)), which extends to a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_r$ of V. Since T kills the $\mathbf{u}_i$, the vectors $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)$ span $\operatorname{im} T$. If $\sum_j b_j T(\mathbf{v}_j) = \mathbf{0}$, then $\sum_j b_j \mathbf{v}_j \in \ker T = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, which gives a linear relation among basis vectors of V, so all $b_j = 0$. Hence $\dim \operatorname{im} T = r = \dim V - \dim \ker T$. Part (a) is (b) for T_A, with Theorem 4.2.7.

We prove (b). Then (a) follows by applying (b) to $T_A : F^n \to F^m$, since $\dim \operatorname{im} T_A = \operatorname{rank} A$ (Theorem 4.2.7).

By Proposition 4.1.21(a), the subspace $\ker T$ of V is finite-dimensional; let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be a basis of it. By Theorem 4.1.16, it extends to a basis

$$\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_r$$

of V, where $k + r = \dim V$. We claim that $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)$ is a basis of $\operatorname{im} T$. Then $\operatorname{im} T$ is finite-dimensional and $\dim \operatorname{im} T = r = \dim V - \dim \ker T$, as desired.

Spanning. Let $\mathbf{w} \in \operatorname{im} T$, say $\mathbf{w} = T(\mathbf{v})$, and write $\mathbf{v} = \sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j$. Since $T(\mathbf{u}_i) = \mathbf{0}$ for every i,

$$\mathbf{w} = \sum_{i=1}^{k} a_i T(\mathbf{u}_i) + \sum_{j=1}^{r} b_j T(\mathbf{v}_j) = \sum_{j=1}^{r} b_j T(\mathbf{v}_j).$$

So the vectors $T(\mathbf{v}_j)$, which lie in $\operatorname{im} T$, span it.

Independent. Suppose $\sum_j b_j T(\mathbf{v}_j) = \mathbf{0}$. Then $T\left(\sum_j b_j \mathbf{v}_j\right) = \mathbf{0}$, so $\sum_j b_j \mathbf{v}_j \in \ker T$, and there are scalars $a_1, \ldots, a_k$ with $\sum_j b_j \mathbf{v}_j = \sum_i a_i \mathbf{u}_i$. Then

$$\sum_{i=1}^{k} (-a_i) \mathbf{u}_i + \sum_{j=1}^{r} b_j \mathbf{v}_j = \mathbf{0}$$

is a linear relation among the vectors of a basis of V, so all its coefficients are 0 (Theorem 1.5.5). In particular $b_1 = \cdots = b_r = 0$. $\square$

In the language of Section 4.2.1, the second proof sets $U = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_r)$. Independence says that $U \cap \ker T = \{\mathbf{0}\}$, so $T|_U$ is injective, and spanning says that $\operatorname{im}(T|_U) = \operatorname{im} T$. So $\operatorname{im} T$ is a faithful copy of U, and Proposition 4.2.6(b) gives $\dim \operatorname{im} T = \dim U = r$.

Remark 4.2.12 (A third proof). The second proof chose the vectors $\mathbf{v}_j$, and different choices give different subspaces U; only $\dim U$ is determined by T. In the plane example of Intuition (Lost plus reached), every line other than $\ker T_A$ is a possible U. Section 4.6 removes the choice: it replaces U by the quotient space $V/\ker T$, built from V and $\ker T$ alone, whose elements are the translates $\mathbf{v} + \ker T$. The map T induces an isomorphism $V/\ker T \to \operatorname{im} T$ (Theorem 4.6.34), and $\dim V/\ker T = \dim V - \dim \ker T$, which is rank–nullity once more. Quotients are not part of the course.

> *Intuition (Comparing the proofs).* The pivot-counting proof is a computation. Row reduction sorts the columns of A into pivot columns and free columns; the free ones index a basis of the kernel, and the pivot ones are a basis of the image. For matrices, the theorem is then the observation that every column is one or the other. The proof produces explicit bases, which is what we want in examples. Its weakness is that it only sees matrices: for an abstract T we had to choose bases, and the real content of (b) moved into the transfer step, which says that coordinates do not change the dimensions of kernels and images.
>
> The basis-extension proof never mentions matrices and works for every T whose domain is finite-dimensional. It shows *why* the formula holds: V is built from $\ker T$ and a subspace U on which T is injective, and T carries U onto $\operatorname{im} T$ without loss. Its price is a choice of U, which the quotient proof removes. When computing, count pivots. When proving something about an abstract map, extend a basis of the kernel.

The bases in the second proof are adapted to T. If W is finite-dimensional, extend the independent list $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)$ to a basis of W. In the bases $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{u}_1, \ldots, \mathbf{u}_k$ of V and this basis of W, the matrix of T is the rank normal form N_r of Theorem 3.5.27. This is Corollary 3.5.29 for abstract maps, and Section 4.5 returns to it.

> **Exercise 4.2.13.** ROUTINE Let $T : F^3 \to F^3$, $T(x_1, x_2, x_3) = (x_1 - x_2, x_2 - x_3, x_3 - x_1)$.
>
> (a) Find $\ker T$.
>
> (b) Show that $\operatorname{im} T = \{\mathbf{y} \in F^3 \mid y_1 + y_2 + y_3 = 0\}$, without solving $T(\mathbf{x}) = \mathbf{y}$.

Solution on page 1297

4.2.5 The rank of a linear map

Rank–nullity suggests measuring a linear map by the dimensions of its kernel and image. Remark 3.5.17 promised an intrinsic definition of rank; here it is.

> **Definition 4.2.14** (Rank and nullity of a linear map). Let $T \in \mathrm{L}(V, W)$. If $\operatorname{im} T$ is finite-dimensional, the **rank** of T is
>
> $$\operatorname{rank} T = \dim \operatorname{im} T.$$
>
> If $\ker T$ is finite-dimensional, the **nullity** of T is $\operatorname{nullity} T = \dim \ker T$.

When V is finite-dimensional both are defined (Theorem 4.2.11, Proposition 4.1.21), and rank–nullity reads

$$\operatorname{rank} T + \operatorname{nullity} T = \dim V,$$

which is where the theorem's name comes from. For a matrix, $\operatorname{rank} T_A = \dim \operatorname{im} T_A = \operatorname{rank} A$ by Theorem 4.2.7, so the two meanings of "rank" agree, and we use $\operatorname{rank} A$ and $\operatorname{rank} T_A$ interchangeably. Likewise the **nullity** of A is $\operatorname{nullity} A = \dim \ker T_A = n - \operatorname{rank} A$.

For a map between finite-dimensional spaces, we can also compute the rank of any of its matrices. The lecture's observation is that the answer does not depend on which matrix we pick.

Proposition 4.2.15 (Rank through any matrix). Let V and W be nonzero finite-dimensional spaces, with bases $\mathcal{B}$ and $\mathcal{C}$, let $T \in \mathrm{L}(V, W)$, and let $A = [T]_{\mathcal{C}\mathcal{B}}$. Then

$$\operatorname{rank} A = \operatorname{rank} T \quad \text{and} \quad \operatorname{nullity} A = \operatorname{nullity} T.$$

In particular, $\operatorname{rank} [T]_{\mathcal{C}\mathcal{B}}$ does not depend on the choice of $\mathcal{B}$ and $\mathcal{C}$.

Proof. ROUTINE

Uses. Proposition 2.6.20 (Kernel and image in coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 1.6.16 (The coordinate map is a bijection), Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Theorem 4.2.7 (The image of T_A)
Techniques. Translate between maps and matrices (Technique 2.3.27), Restrict to a subspace (Technique 4.2.5)
Hint 1. Under $\mathcal{C}$-coordinates, which subspace of F^m corresponds to $\operatorname{im} T$? See Proposition 2.6.20.
Hint 2. The coordinate map of W, restricted to $\operatorname{im} T$, is linear and injective. What does that do to dimension?
Sketch. By Proposition 2.6.20, $\mathbf{w} \mapsto [\mathbf{w}]_{\mathcal{C}}$ maps $\operatorname{im} T$ onto $\operatorname{im} T_A$, and it is the restriction of an injective linear map, so $\dim \operatorname{im} T = \dim \operatorname{im} T_A = \operatorname{rank} A$ (Proposition 4.2.6, Theorem 4.2.7). The same argument with $\mathcal{B}$-coordinates matches $\ker T$ with $\ker T_A$.

By Proposition 2.6.20, the coordinate map $\mathbf{w} \mapsto [\mathbf{w}]_{\mathcal{C}}$ sends $\operatorname{im} T$ onto $\operatorname{im} T_A$. It is linear and injective on all of W (Proposition 1.6.19, Proposition 1.6.16), so its restriction to $\operatorname{im} T$ is an injective linear map with image $\operatorname{im} T_A$ (Proposition 4.2.3). The space $\operatorname{im} T$ is finite-dimensional (Theorem 4.2.11), so Proposition 4.2.6(a) and Theorem 4.2.7 give

$$\operatorname{rank} T = \dim \operatorname{im} T = \dim \operatorname{im} T_A = \operatorname{rank} A.$$

In the same way, $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ restricts to an injective linear map from $\ker T$ onto $\ker T_A$, so $\operatorname{nullity} T = \dim \ker T_A = \operatorname{nullity} A$. □

This settles Remark 3.5.17. Once the change-of-basis formula of Section 4.5 expresses two matrices of T through each other by invertible factors, the proposition also shows that such factors do not change the rank; Section 4.3 proves that directly.

Example 4.2.16 (One map, two matrices). Let $\ell = \operatorname{span}(\mathbf{u})$ with $\mathbf{u} = (c, s) = (\cos\theta, \sin\theta)$, and let Pr_ℓ be the orthogonal projection onto ℓ, so $\mathrm{Pr}_\ell(x, y) = (xc + ys)\mathbf{u}$ (Example 2.6.7(d)). In the standard basis,

$$[\mathrm{Pr}_\ell]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix},$$

whose columns $c\mathbf{u}$ and $s\mathbf{u}$ are multiples of the nonzero vector $\mathbf{u}$, not both $\mathbf{0}$; so its rank is 1. Let $\mathbf{u}^\perp = (-s, c)$. By Example 2.6.7(d), every vector of $\mathbb{R}^2$ is a combination of $\mathbf{u}$ and $\mathbf{u}^\perp$, so $\mathcal{B} = \mathbf{u}, \mathbf{u}^\perp$ is a basis (Corollary 4.1.11(b)). In this basis $\mathrm{Pr}_\ell(\mathbf{u}) = \mathbf{u}$ and $\mathrm{Pr}_\ell(\mathbf{u}^\perp) = \mathbf{0}$, so

$$[\mathrm{Pr}_\ell]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

The two matrices look different, but both have rank $1 = \dim \ell = \operatorname{rank} \mathrm{Pr}_\ell$.

Exercise 4.2.17. ROUTINE Let $T : F[x]_{\leq 2} \to F^3$, $T(p) = (p(0), p(1), p(2))$. Using the matrix of T in the bases $1, x, x^2$ and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, find $\operatorname{rank} T$ and $\operatorname{nullity} T$ for $F = \mathbb{Q}$ and for $F = \mathbb{F}_2$. For $F = \mathbb{F}_2$, find a nonzero $p \in \ker T$.

Solution on page 1297

4.2.6 Consequences

Rank–nullity turns statements about vectors into statements about numbers. This subsection collects the consequences that the course uses.

Injective, surjective, bijective

For a map between finite sets of the same size, injective and surjective are equivalent. Dimension plays the role of size.

Corollary 4.2.18 (Injective if and only if surjective). Let V and W be finite-dimensional with $\dim V = \dim W$, and let $T \in \mathrm{L}(V, W)$. The following are equivalent.

(i) T is injective.
(ii) T is surjective.
(iii) T is bijective.

In particular, for V finite-dimensional, an operator $T \in \mathrm{L}(V)$ is injective if and only if it is surjective.

Proof. ROUTINE

Uses. Theorem 4.2.11 (Rank–nullity), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 4.1.21 (Dimensions of subspaces), Definition 2.6.3 (Bijective)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Translate injectivity and surjectivity into statements about $\dim\ker T$ and $\dim\operatorname{im} T$.
Hint 2. $\operatorname{im} T$ is a subspace of W. When is a subspace of W equal to W?
Sketch. Let $n = \dim V = \dim W$. T is injective iff $\dim\ker T = 0$, iff $\dim\operatorname{im} T = n$ by rank–nullity, iff $\operatorname{im} T = W$ by Proposition 4.1.21(b). Bijective means both.

Let $n = \dim V = \dim W$. By Theorem 2.6.11, T is injective if and only if $\ker T = \{\mathbf{0}\}$, that is, $\dim\ker T = 0$. By Theorem 4.2.11 this holds if and only if $\dim\operatorname{im} T = n$. Since $\operatorname{im} T$ is a subspace of W and $\dim W = n$, Proposition 4.1.21(b) says that $\dim\operatorname{im} T = n$ if and only if $\operatorname{im} T = W$, that is, T is surjective. So (i) $\Leftrightarrow$ (ii), and (iii) is (i) and (ii) together (Definition 2.6.3). □

So for maps between spaces of the same finite dimension, uniqueness gives existence, which extends Technique 3.4.21 from square matrices to linear maps: to show that $T(\mathbf{v}) = \mathbf{w}$ is solvable for every $\mathbf{w}$, it suffices to show that $T(\mathbf{v}) = \mathbf{0}$ only for $\mathbf{v} = \mathbf{0}$.

Example 4.2.19 (Uniqueness gives existence, for polynomials). Let $T : \mathbb{R}[x]_{\le n} \to \mathbb{R}[x]_{\le n}$, $T(p) = p + p'$. If $p \neq \mathbf{0}$ has degree d with leading coefficient a_d, then p' has degree less than d, so the coefficient of x^d in $p + p'$ is $a_d \neq 0$. Hence $\ker T = \{\mathbf{0}\}$, and T is surjective by Corollary 4.2.18: for every $q \in \mathbb{R}[x]_{\le n}$, the equation $p + p' = q$ has a solution $p \in \mathbb{R}[x]_{\le n}$, and only one. For instance, for $q = x^2$, trying $p = x^2 + bx + c$ gives $p + p' = x^2 + (b+2)x + (b+c)$, so $b = -2$, $c = 2$, and $p = x^2 - 2x + 2$. We knew a solution existed before we looked for it.

The finite dimension is essential. On $F[x]$, multiplication by x is injective but not surjective, and in characteristic 0 differentiation is surjective but not injective (Example 2.6.4(e)).

Remark 4.2.20 (One-sided inverses of operators). Let V be finite-dimensional and $S, T \in \mathrm{L}(V)$ with $ST = \mathrm{id}_V$. Then T is injective (Proposition 2.5.4), hence bijective by Corollary 4.2.18, and its inverse is linear (Proposition 2.7.7). So $S = S\big(TT^{-1}\big) = (ST)T^{-1} = T^{-1}$, and $TS = \mathrm{id}_V$ as well. A one-sided inverse of an operator on a finite-dimensional space is two-sided. This is Theorem 3.5.23 for operators, and the same argument works for $T \in \mathrm{L}(V, W)$ and $S \in \mathrm{L}(W, V)$ when $\dim V = \dim W < \infty$.

Dimension obstructions

Corollary 4.2.21 (Dimension obstructions). Let V and W be finite-dimensional, and let $T \in \mathrm{L}(V, W)$.

(a) $\operatorname{rank} T \le \min(\dim V, \dim W)$, and $\dim \ker T \ge \dim V - \dim W$.
(b) If T is injective, then $\dim V \le \dim W$. So if $\dim V > \dim W$, no linear map $V \to W$ is injective.
(c) If T is surjective, then $\dim V \ge \dim W$. So if $\dim V < \dim W$, no linear map $V \to W$ is surjective.

Proof. ROUTINE

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $\operatorname{im} T$ is a subspace of W, and rank–nullity gives $\operatorname{rank} T = \dim V - \dim \ker T$.
Hint 2. For (b) and (c), put $\dim \ker T = 0$, respectively $\operatorname{rank} T = \dim W$, into (a).
Sketch. (a) $\operatorname{rank} T \le \dim W$ since $\operatorname{im} T \subseteq W$, and $\operatorname{rank} T = \dim V - \dim \ker T \le \dim V$; rearranging $\operatorname{rank} T \le \dim W$ gives the bound on $\dim \ker T$. (b) Injective means $\dim \ker T = 0$, so $\dim V = \operatorname{rank} T \le \dim W$. (c) Surjective means $\operatorname{rank} T = \dim W$, so $\dim W \le \dim V$.

(a) Since $\operatorname{im} T$ is a subspace of W, $\operatorname{rank} T \le \dim W$ (Proposition 4.1.21(a)). By Theorem 4.2.11, $\operatorname{rank} T = \dim V - \dim \ker T \le \dim V$. Rearranging $\dim V - \dim \ker T = \operatorname{rank} T \le \dim W$ gives $\dim \ker T \ge \dim V - \dim W$.

(b) If T is injective, then $\dim \ker T = 0$ (Theorem 2.6.11), so $\dim V = \operatorname{rank} T \le \dim W$ by rank–nullity and (a).

(c) If T is surjective, then $\operatorname{rank} T = \dim W$, and $\operatorname{rank} T \le \dim V$ by (a). □

For T_A with $A \in F^{m\times n}$, (b) and (c) are the "moreover" parts of Theorem 3.5.12 and Theorem 3.5.13: an injective T_A has $n \le m$, and a surjective one has $n \ge m$. Now they hold for all linear maps. Part (b) is the precise sense in which a bigger space cannot be fitted into a smaller one. As sets, $\mathbb{R}^2$ and $\mathbb{R}$ are in bijection, but no injective map $\mathbb{R}^2 \to \mathbb{R}$ is linear.

Exercise 4.2.22. ROUTINE For each item, give an example or show that none exists.

(a) An injective linear map $F^{2\times 3} \to F[x]_{\le 4}$.
(b) A surjective linear map $F[x]_{\le 3} \to F^{2\times 2}$.
(c) A linear map $T : \mathbb{R}^5 \to \mathbb{R}^2$ with $\ker T = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$.
(d) A linear map $T : \mathbb{R}^4 \to \mathbb{R}^3$ with $\ker T = \operatorname{span}(\mathbf{e}_1)$.

Solution on page 1298

Linear systems, revisited

Rank–nullity gives a second view of Corollary 3.4.19, and it says how many parameters a description of a solution set needs.

Corollary 4.2.23 (Dimension of solution sets). Let $A \in F^{m\times n}$.
(a) $\dim \ker T_A = n - \operatorname{rank} A \ge n - m$. In particular, if $n > m$, then $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution.
(b) Suppose $A\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x}_0$, and that its solution set is $\mathbf{x}_0 + \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_r)$ for some $\mathbf{u}_1, \dots, \mathbf{u}_r \in F^n$. Then $r \ge n - \operatorname{rank} A$, with equality if and only if $\mathbf{u}_1, \dots, \mathbf{u}_r$ is linearly independent.
(c) Conversely, if $\mathbf{u}_1, \dots, \mathbf{u}_r \in \ker T_A$ are linearly independent and $r = n - \operatorname{rank} A$, then the solution set of $A\mathbf{x} = \mathbf{y}$ is $\mathbf{x}_0 + \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_r)$ for every solution $\mathbf{x}_0$.

Proof. ROUTINE

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 3.5.2 (First properties of rank), Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 4.1.11 (Spanning lists are long), Theorem 4.1.10 (Independent lists are short), Proposition 1.6.2 (Bases are independent spanning lists)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. (a) is rank–nullity for T_A, with $\operatorname{rank} A \le m$.

Hint 2. For (b), the solution set is also $\mathbf{x}_0 + \ker T_A$ (Proposition 3.1.19). Deduce $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r) = \ker T_A$, and compare the length of a spanning list with the dimension.
Sketch. (a) $\dim \ker T_A = n - \operatorname{rank} A$ and $\operatorname{rank} A \leq m$. (b) The solution set is $\mathbf{x}_0 + \ker T_A$, so $\operatorname{span}(\mathbf{u}_i) = \ker T_A$. A spanning list of a space of dimension $n - \operatorname{rank} A$ has at least that many entries, with equality iff it is a basis, iff it is independent. (c) An independent list of length $\dim \ker T_A$ in $\ker T_A$ is a basis of it.

(a) By Theorem 4.2.11(a), $\dim \ker T_A = n - \operatorname{rank} A$, and $\operatorname{rank} A \leq m$ (Proposition 3.5.2(a)). If $n > m$, then $\dim \ker T_A \geq n - m \geq 1$, so $\ker T_A$ contains a nonzero vector.

(b) By Proposition 3.1.19, the solution set is $\mathbf{x}_0 + \ker T_A$. For $\mathbf{v} \in F^n$, the vector $\mathbf{x}_0 + \mathbf{v}$ is a solution if and only if $\mathbf{v} \in \ker T_A$, and also if and only if $\mathbf{v} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$. So $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r) = \ker T_A$, a space of dimension $d = n - \operatorname{rank} A$. A list spanning it has length $r \geq d$ (Corollary 4.1.11(a)). If $r = d$, the list is a basis (Corollary 4.1.11(b)), hence independent. If the list is independent, it is a basis of $\ker T_A$ (Proposition 1.6.2), so $r = d$.

(c) An independent list in $\ker T_A$ of length $\dim \ker T_A$ is a basis of $\ker T_A$ (Theorem 4.1.10(b)), so its span is $\ker T_A$, and the solution set is $\mathbf{x}_0 + \ker T_A$ by Proposition 3.1.19. □

Part (a) is Corollary 3.4.19(a) from the other side. There, more columns than pivots forced a free variable. Here, the image has room for at most m of the n dimensions of F^n, so the kernel must absorb at least $n - m$ of them. Part (b) says that the number of parameters in a description of the solutions is at least $n - \operatorname{rank} A$, however the description was found, and part (c) is the quick completeness test promised in Remark 3.4.12: $n - \operatorname{rank} A$ independent solutions of the homogeneous system always span its solution space. Over a field with q elements, a consistent system has $q^{n - \operatorname{rank} A}$ solutions (Corollary 3.4.14(b)).

Dimension counting

Rank–nullity and the bounds of Section 4.1 prove that certain vectors exist without exhibiting them. That is worth a name.

Technique 4.2.24 (Dimension counting). To show that a subspace is nonzero, that two subspaces are equal, or that a linear map is injective or surjective, compute dimensions instead of vectors. The tools are rank–nullity, which ties $\dim \ker T$ to $\dim \operatorname{im} T$; Proposition 4.1.21(b), which turns one containment and equal dimensions into equality; and Theorem 4.1.10, which makes long lists dependent. A typical argument computes whichever of $\dim \ker T$ and $\dim \operatorname{im} T$ is easier and reads off the other. Often the image is easy because it obviously contains a basis of the target, or the kernel is easy because a structural reason (degrees, a top coefficient) forces it to be small.

Reach for it when a direct computation would mean solving equations you would rather not solve, or when you need a nonzero vector with some properties (in an intersection, in a kernel) and only its existence matters. The canonical example is that two 3-dimensional subspaces of $\mathbb{R}^5$ meet in a nonzero vector (Proposition 4.2.26, Example 4.2.27). Later the same counting gives the rank inequalities (Section 4.3), the dimension formula for sums (Section 4.4), the bounds on eigenspaces (Section 7.4), and the min–max principles (Section 12.3).

Used in. Exercise 4.2.13, Corollary 4.2.18, Corollary 4.2.21, Corollary 4.2.23, Proposition 4.2.26, Exercise 4.2.39, Exercise 4.2.42, Exercise 4.2.46, Exercise 4.2.47, Exercise 4.2.49, Exercise 4.2.51, Exercise 4.2.52, Proposition 4.3.1, Theorem 4.3.3, Proposition 4.3.5, Proposition 4.3.20, Proposition 4.3.23, Exercise 4.3.42, Theorem 4.4.13, Theorem 4.4.19, Proposition 4.4.21, Exercise 4.4.23, Theorem 4.4.27, Exercise 4.4.34, Exercise 4.4.40, Exercise 4.4.45, Exercise 4.4.46, Exercise 4.4.48, Exercise 4.4.49, Exercise 4.4.50, Theorem 4.5.6, Exercise 4.5.41, Theorem 4.6.5, Corollary 4.6.10, Proposition 4.6.11, Theorem 4.6.27, Exercise 4.6.47, Exercise 4.6.48, Theorem 4.7.8, Exercise 4.7.52, Exercise 4.8.4, Exercise 4.8.6, Exercise 4.8.7, Exercise 4.8.8, Exercise 4.8.9, Problem 4.1, Problem 4.5, Problem 4.6, Problem 4.8, Problem 4.9, Problem 4.11, Problem 4.16, Problem 4.17, Corollary 6.4.25, Theorem 6.4.26, Corollary 6.6.22, Problem 6.10, Theorem 7.1.14, Exercise 7.1.57, Theorem 7.4.16, Theorem 7.4.30, Proposition 7.5.8, Proposition 7.6.12, Exercise 7.7.9, Problem 7.2, Problem 7.3, Problem 7.12, Exercise 4.2.40, Exercise 4.2.50, Exercise 4.3.30, Exercise 4.3.31, Exercise 4.6.38, Exercise 4.6.54, Exercise 6.5.45, Exercise 6.6.43, Exercise 7.2.45, Exercise 7.2.46, Exercise 7.2.50.

The simplest instance: one nonzero linear condition cuts the dimension by exactly one.

Example 4.2.25 (Kernels of linear functionals). Let V be finite-dimensional, and let $f : V \to F$ be linear and not the zero map. Then $\operatorname{im} f$ is a nonzero subspace of F, so $\operatorname{im} f = F$ (Proposition 4.1.21(b), as $\dim F = 1$), and by rank–nullity

$$\dim \ker f = \dim V - 1.$$

For instance:

(a) For $\mathbf{a} \in F^n$ with $\mathbf{a} \neq \mathbf{0}$, the solutions of $a_1x_1 + \cdots + a_nx_n = 0$ form a subspace of F^n of dimension $n - 1$.

(b) The trace-zero matrices form a subspace of $F^{n\times n}$ of dimension $n^2 - 1$, since $\operatorname{tr} : F^{n\times n} \to F$ is linear (Proposition 2.4.27) and $\operatorname{tr} U_{11} = 1$.

(c) For $c \in F$, the polynomials $p \in F[x]_{\leq n}$ with $p(c) = 0$ form a subspace of dimension n, since evaluation at c is linear (Example 2.1.9) and sends 1 to 1.

In Section 4.1, (b) and (c) needed an explicit basis or a squeeze (Exercise 4.1.43, Exercise 4.1.24). Here we only had to check that the functional is not zero.

The next proposition is the canonical example of the technique. In $\mathbb{R}^3$, two planes through $\mathbf{0}$ always share a line, since $2 + 2 > 3$. In $\mathbb{R}^4$ they need not: $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) \cap \operatorname{span}(\mathbf{e}_3, \mathbf{e}_4) = \{\mathbf{0}\}$. The count behind this: U and W together supply $\dim U + \dim W$ independent directions, and V has room for only $\dim V$ of them, so the excess must be shared.

Proposition 4.2.26 (Subspaces that must meet). Let V be finite-dimensional, and let U and W be subspaces of V. Then

$$\dim(U \cap W) \geq \dim U + \dim W - \dim V.$$

In particular, if $\dim U + \dim W > \dim V$, then $U \cap W \neq \{\mathbf{0}\}$.

To make the count precise, we build a linear map whose kernel records $U \cap W$. The first proof uses both subspaces symmetrically.

Proof (with a difference map). STANDARD

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.2.11 (Rank–nullity), Corollary 4.2.21 (Dimension obstructions), Theorem 1.5.5 (Independence is decided at zero), Proposition 4.2.6 (Injective maps preserve dimension), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Take bases $\mathbf{u}_1, \ldots, \mathbf{u}_k$ of U and $\mathbf{w}_1, \ldots, \mathbf{w}_l$ of W, and consider $L : F^{k+l} \to V$, $L(\mathbf{a}, \mathbf{b}) = \sum_i a_i\mathbf{u}_i - \sum_j b_j\mathbf{w}_j$. How large is $\ker L$?
Hint 2. An element of $\ker L$ says that $\sum_i a_i\mathbf{u}_i = \sum_j b_j\mathbf{w}_j$, a vector of $U \cap W$. Show that $(\mathbf{a}, \mathbf{b}) \mapsto \sum_i a_i\mathbf{u}_i$ is injective on $\ker L$.
Sketch. With L as in the hint, rank–nullity gives $\dim \ker L = k + l - \operatorname{rank} L \geq k + l - n$. The linear map $\varphi : \ker L \to V$, $(\mathbf{a}, \mathbf{b}) \mapsto \sum_i a_i\mathbf{u}_i$, takes values in $U \cap W$ and is injective: if $\sum_i a_i\mathbf{u}_i = \mathbf{0}$, then $\mathbf{a} = \mathbf{0}$, and then $\sum_j b_j\mathbf{w}_j = \mathbf{0}$ gives $\mathbf{b} = \mathbf{0}$. So $\dim(U \cap W) \geq \dim \ker L \geq k + l - n$.

Let $n = \dim V$. By Proposition 4.1.21(a), U and W are finite-dimensional; let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_l$ be bases. Write elements of F^{k+l} as $(\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in F^k$ and $\mathbf{b} \in F^l$, and define

$$L : F^{k+l} \to V, \quad L(\mathbf{a}, \mathbf{b}) = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k - b_1\mathbf{w}_1 - \cdots - b_l\mathbf{w}_l.$$

It is linear, since each coordinate enters linearly. By Theorem 4.2.11 and Corollary 4.2.21(a),

$$\dim \ker L = (k + l) - \operatorname{rank} L \geq k + l - n.$$

Define $\varphi : \ker L \to V$ by $\varphi(\mathbf{a}, \mathbf{b}) = \sum_i a_i\mathbf{u}_i$; it is linear. If $(\mathbf{a}, \mathbf{b}) \in \ker L$, then $\varphi(\mathbf{a}, \mathbf{b}) = \sum_i a_i\mathbf{u}_i = \sum_j b_j\mathbf{w}_j$ lies in U and in W, so $\operatorname{im}\varphi \subseteq U \cap W$. And φ is injective: if $\sum_i a_i\mathbf{u}_i = \mathbf{0}$, then $\mathbf{a} = \mathbf{0}$ by independence of the $\mathbf{u}_i$ (Theorem 1.5.5), so $\sum_j b_j\mathbf{w}_j = \sum_i a_i\mathbf{u}_i = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$ likewise. By Proposition 4.2.6(a) and Proposition 4.1.21(a),

$$\dim(U \cap W) \geq \dim \operatorname{im} \varphi = \dim \ker L \geq k + l - n,$$

as desired. □

The second proof describes W by linear equations and restricts them to U.

Proof (by restricting a map that kills W). STANDARD

> *Uses.* Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 4.2.3 (Properties of restriction), Theorem 4.2.11 (Rank–nullity), Corollary 4.2.21 (Dimension obstructions)
> *Techniques.* Restrict to a subspace (Technique 4.2.5), Extend a basis of a subspace (Technique 4.1.18), Dimension counting (Technique 4.2.24)
> *Hint 1.* Find a linear map $L : V \to F^{n-l}$, where $l = \dim W$ and $n = \dim V$, with $\ker L = W$. Extending a basis of W helps.
> *Hint 2.* Restrict L to U. What is $\ker(L|_U)$, and how large can $\operatorname{im}(L|_U)$ be?
> *Sketch.* Extend a basis $\mathbf{w}_1, \ldots, \mathbf{w}_l$ of W to a basis $\mathbf{w}_1, \ldots, \mathbf{w}_l, \mathbf{v}_1, \ldots, \mathbf{v}_{n-l}$ of V, and let L send $\mathbf{w}_j \mapsto \mathbf{0}$ and $\mathbf{v}_i \mapsto \mathbf{e}_i \in F^{n-l}$. Then $\ker L = W$, so $\ker(L|_U) = U \cap W$, and rank–nullity for $L|_U$ gives $\dim U = \dim(U \cap W) + \operatorname{rank}(L|_U) \leq \dim(U \cap W) + n - l$.

Let $n = \dim V$ and $l = \dim W$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_l$ be a basis of W (Proposition 4.1.21(a)), and extend it to a basis $\mathbf{w}_1, \ldots, \mathbf{w}_l, \mathbf{v}_1, \ldots, \mathbf{v}_{n-l}$ of V (Theorem 4.1.16). By Theorem 2.2.1 there is a linear map $L : V \to F^{n-l}$ with $L(\mathbf{w}_j) = \mathbf{0}$ for all j and $L(\mathbf{v}_i) = \mathbf{e}_i$ for all i. If $\mathbf{v} = \sum_j b_j \mathbf{w}_j + \sum_i c_i \mathbf{v}_i$, then $L(\mathbf{v}) = (c_1, \ldots, c_{n-l})$, which is $\mathbf{0}$ exactly when all $c_i = 0$, that is, when $\mathbf{v} \in W$. So $\ker L = W$.

By Proposition 4.2.3(c), $L|_U : U \to F^{n-l}$ is linear with kernel $U \cap W$. By Theorem 4.2.11 and Corollary 4.2.21(a),

$$\dim U = \dim(U \cap W) + \operatorname{rank}(L|_U) \leq \dim(U \cap W) + (n - l).$$

Rearranging gives $\dim(U \cap W) \geq \dim U + \dim W - \dim V$. □

The first proof gives more if we look closer. The image of L is $U + W$, and φ maps $\ker L$ onto $U \cap W$, so rank–nullity for L is an equation, $\dim(U + W) + \dim(U \cap W) = \dim U + \dim W$; that is the dimension formula of Section 4.4 (Theorem 4.4.19). The second proof is the one to remember in practice. The map L has $n - l$ coordinates, each a linear condition, and W is the set where all of them vanish. Restricted to U, each condition lowers the dimension by at most one. So a subspace of U cut out by s linear conditions has dimension at least $\dim U - s$.

Example 4.2.27 (Two subspaces of $\mathbb{R}^5$).

(a) Two 3-dimensional subspaces of $\mathbb{R}^5$ meet in at least a line, since $3 + 3 - 5 = 1$. The bound is attained by $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and $\operatorname{span}(\mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5)$, which meet in $\operatorname{span}(\mathbf{e}_3)$, and the intersection can be larger (take $U = W$). In $\mathbb{R}^6$, two 3-dimensional subspaces need not share a nonzero vector: $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \cap \operatorname{span}(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6) = \{\mathbf{0}\}$.

(b) The polynomials $p \in \mathbb{R}[x]_{\leq 4}$ with $p(0) = p'(0) = 0$ form a 3-dimensional subspace W, with basis x^2, x^3, x^4. So by Proposition 4.2.26, every 3-dimensional subspace U of the 5-dimensional space $\mathbb{R}[x]_{\leq 4}$ contains a nonzero p with $p(0) = p'(0) = 0$. For $U = \operatorname{span}\big(1, 1 + x + x^2, x + x^3\big)$, writing $p = a + b\big(1 + x + x^2\big) + c\big(x + x^3\big)$, the conditions are $a + b = 0$ and $b + c = 0$, solved by $(a, b, c) = (1, -1, 1)$:

$$p = 1 - \big(1 + x + x^2\big) + \big(x + x^3\big) = x^3 - x^2.$$

Dimension counting also proves existence statements about polynomials, sometimes with a formula as a by-product.

Example 4.2.28 (Sums of powers). Let $n \geq 1$, and let $\Delta : \mathbb{Q}[x]_{\leq n} \to \mathbb{Q}[x]_{\leq n-1}$ be the linear map with

$$\Delta\big(x^j\big) = (x + 1)^j - x^j = \sum_{i=0}^{j-1} \binom{j}{i} x^i \quad (0 \leq j \leq n)$$

(Theorem 2.2.1). By the binomial theorem in $\mathbb{Q}$, $(\Delta p)(t) = p(t+1) - p(t)$ for every p and every $t \in \mathbb{Q}$.

Kernel. Constants are in $\ker \Delta$. Conversely, if $\Delta p = \mathbf{0}$, then $p(t+1) = p(t)$ for all t, so $p(0) = p(1) = \cdots = p(n)$. Then p and the constant polynomial $p(0)$ have the same values at the $n+1$ distinct points $0, 1, \ldots, n$, so they are equal by the uniqueness in Theorem 3.4.24. Hence $\ker \Delta$ is the constants, and $\dim \ker \Delta = 1$.

Image. By rank–nullity, $\operatorname{rank} \Delta = (n+1) - 1 = n = \dim \mathbb{Q}[x]_{\leq n-1}$, so Δ is surjective (Proposition 4.1.21(b)). So for every $q \in \mathbb{Q}[x]_{\leq n-1}$ there is $p \in \mathbb{Q}[x]_{\leq n}$ with $p(t+1) - p(t) = q(t)$ for all t, and the sum telescopes:

$$q(0) + q(1) + \cdots + q(N-1) = \sum_{t=0}^{N-1} (p(t+1) - p(t)) = p(N) - p(0).$$

So for every nonzero polynomial q, the sum $\sum_{t=0}^{N-1} q(t)$ is a polynomial in N of degree at most $\deg q + 1$.

For $q = x^2$, some $p \in \mathbb{Q}[x]_{\leq 3}$ has $\Delta p = q$, and subtracting its constant term (which lies in $\ker \Delta$) we may take $p = ax^3 + bx^2 + cx$. Since $\Delta\left(x^3\right) = 3x^2 + 3x + 1$, $\Delta\left(x^2\right) = 2x + 1$, and $\Delta(x) = 1$,

$$\Delta p = 3ax^2 + (3a + 2b)x + (a + b + c),$$

and $\Delta p = x^2$ gives $a = 1/3$, $b = -1/2$, $c = 1/6$, that is, $p = x(x-1)(2x-1)/6$. Hence

$$0^2 + 1^2 + \cdots + (N-1)^2 = \frac{N(N-1)(2N-1)}{6}.$$

The dimension count guaranteed in advance that this linear system for a, b, c had a solution.

Linear maps on spaces of matrices

Spaces of matrices are vector spaces of known dimension, so maps between them are natural targets for dimension counting.

> **Technique 4.2.29** (Study a linear map on a space of matrices). Many questions about matrices ask which X satisfy a linear condition, such as $AX = XA$, $AX = XB$, or $X^{\mathsf{T}} = -X$, or which matrices can be written in a given linear form, such as $AX - XA$. Treat $X \mapsto (AX - XA)$, or whatever the expression is, as a linear map on $F^{n\times n}$ (or $F^{m\times n}$), a space of dimension n^2 (or mn) by Example 4.1.7. The first kind of question asks for its kernel and the second for its image, and rank–nullity relates the two. So an easy description of the kernel gives the dimension of an image that is hard to describe, and with one easy containment the dimension often identifies the image completely.
>
> Reach for it when a matrix equation is linear in the unknown matrix. The canonical example is Example 4.2.30. The trace argument that $AB - BA \neq I_n$ when $n \cdot 1 \neq 0$ in F (Technique 2.4.29) is an instance: the image of $X \mapsto AX - XA$ lies in the trace-zero matrices. Later instances are Sylvester equations $AX - XB = C$ (Section 10.3) and transposition as an operator with eigenvalues ± 1 (Section 7.4).
> *Used in.* Exercise 4.2.41, Exercise 4.2.43, Problem 4.4, Exercise 7.3.55, Exercise 7.4.41, Exercise 7.5.45, Exercise 7.7.11, Problem 7.13, Exercise 4.5.48, Exercise 7.5.44.

Example 4.2.30 (Commutators with a diagonal matrix). Let $D = \operatorname{diag}(d_1, \ldots, d_n) \in F^{n\times n}$ with distinct diagonal entries $d_1, \ldots, d_n$, and let $\operatorname{ad}_D : F^{n\times n} \to F^{n\times n}$, $\operatorname{ad}_D(X) = DX - XD$. It is linear (Proposition 2.4.14). Multiplying by D on the left scales row i by d_i, and on the right scales column j by d_j (Example 2.4.10), so

$$(DX - XD)_{ij} = \left(d_i - d_j\right)x_{ij}.$$

Kernel. $\operatorname{ad}_D(X) = 0$ means $\left(d_i - d_j\right)x_{ij} = 0$ for all i, j. For $i \neq j$, $d_i - d_j \neq 0$, so $x_{ij} = 0$. So $\ker \operatorname{ad}_D$ is the space of diagonal matrices, of dimension n, and $\operatorname{rank} \operatorname{ad}_D = n^2 - n$.

Image. Every $DX - XD$ has zero diagonal, since its (i,i) entry is $(d_i - d_i)x_{ii} = 0$. The matrices with zero diagonal form a subspace Z with basis U_{ij}, $i \neq j$, so $\dim Z = n^2 - n$. Since $\operatorname{im} \operatorname{ad}_D \subseteq Z$ and the dimensions agree, $\operatorname{im} \operatorname{ad}_D = Z$ (Proposition 4.1.21(b)): every matrix with zero diagonal is $DX - XD$ for some X.

Here we could also have solved for X directly, $x_{ij} = c_{ij}/(d_i - d_j)$ for $i \neq j$. The count did not need the formula, which is what makes it useful when no formula is in sight.

4.2.7 Application: error-correcting codes BEYOND 21-242

Coding theory is not part of 21-242. This subsection uses only rank–nullity and arithmetic in $\mathbb{F}_2$.

Suppose we send blocks of 4 bits over a channel that occasionally flips a bit. Sending each bit three times and taking a majority vote corrects any single flipped bit, at the cost of 12 bits per block. Hamming found a way to correct any single flipped bit in a block using 7 bits. The construction is a kernel, and rank–nullity is what makes the numbers work.

We work in $\mathbb{F}_2^n$, where $-\mathbf{x} = \mathbf{x}$, so subtraction and addition agree.

Definition 4.2.31 (Linear codes). A **linear code** of length n is a subspace C of $\mathbb{F}_2^n$, and its elements are **codewords**. If $\dim C = k$, we call C an $[n,k]$ code. The **weight** $\operatorname{wt}(\mathbf{x})$ of $\mathbf{x} \in \mathbb{F}_2^n$ is the number of nonzero entries of $\mathbf{x}$, and the **distance** between $\mathbf{x}$ and $\mathbf{y}$ is $d(\mathbf{x},\mathbf{y}) = \operatorname{wt}(\mathbf{x} - \mathbf{y})$, the number of positions in which they differ. The **minimum distance** of $C \neq \{\mathbf{0}\}$ is the smallest weight of a nonzero codeword.

Since $\mathbf{c} - \mathbf{c}' \in C$ for codewords $\mathbf{c}, \mathbf{c}'$, and $\operatorname{wt}(\mathbf{c}) = d(\mathbf{c},\mathbf{0})$, the minimum distance is also the smallest distance between two different codewords. The distance satisfies the triangle inequality $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$: a position in which $\mathbf{x}$ and $\mathbf{z}$ differ is a position in which $\mathbf{x}$ differs from $\mathbf{y}$ or $\mathbf{y}$ differs from $\mathbf{z}$.

A convenient way to specify a code is as a kernel. If $H \in \mathbb{F}_2^{r \times n}$, then $C = \ker T_H$ is a linear code of length n, the codewords are the vectors that pass the r **parity checks** given by the rows of H, and H is a **parity-check matrix** of C. By rank–nullity,

$$\dim C = n - \operatorname{rank} H, \quad |C| = 2^{n - \operatorname{rank} H}$$

(Proposition 4.1.34).

Example 4.2.32 (The Hamming code). Let $H \in \mathbb{F}_2^{3\times 7}$ be the matrix whose jth column $\mathbf{h}_j$ is the binary expansion of j, highest digit on top:

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Columns $4, 2, 1$ are $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, so the columns span $\mathbb{F}_2^3$, and $\operatorname{rank} H = \dim \operatorname{im} T_H = 3$ (Proposition 2.6.18, Theorem 4.2.7). So the **Hamming code** $C = \ker T_H$ is a $[7,4]$ code with $2^4 = 16$ codewords.

Listing the rows of H from the bottom up already gives the reduced form,

$$H_r = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix},$$

with pivot columns $1, 2, 4$ and free columns $3, 5, 6, 7$. By Theorem 3.4.9, a codeword is determined by its free coordinates, which are arbitrary, and the others are

$$\begin{aligned} x_1 &= x_3 + x_5 + x_7, \\ x_2 &= x_3 + x_6 + x_7, \\ x_4 &= x_5 + x_6 + x_7. \end{aligned}$$

So to encode a message (m_1, m_2, m_3, m_4), put it in positions $3, 5, 6, 7$ and fill in the **parity bits** in positions $1, 2, 4$. The message 1011 becomes the codeword 0110011.

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{over } \mathbb{F}_2$$

The parity-check matrix of the Hamming code over $\mathbb{F}_2$. The only row operation is the swap of rows 1 and 3. The summary reports rank 3, pivot columns $1, 2, 4$ (the parity bits), and free columns $3, 5, 6, 7$ (the message bits). (Interactive in the web version.)

The columns of H are nonzero and pairwise distinct, and that is what makes the code correct errors.

Proposition 4.2.33 (Distinct columns give distance three). Let $H \in \mathbb{F}_2^{r \times n}$ have nonzero, pairwise distinct columns $\mathbf{h}_1, \ldots, \mathbf{h}_n$, and let $C = \ker T_H$. Then every nonzero codeword has weight at least 3.

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 4.2.31 (Linear codes)
Hint 1. Over $\mathbb{F}_2$, $H\mathbf{x}$ is the sum of the columns $\mathbf{h}_j$ at the positions where $\mathbf{x}$ has a 1.
Hint 2. Rule out codewords of weight 1 and 2.
Sketch. $H\mathbf{x} = \sum_{x_j=1} \mathbf{h}_j$. A vector of weight 1 has $H\mathbf{x} = \mathbf{h}_j \neq \mathbf{0}$, and one of weight 2 has $H\mathbf{x} = \mathbf{h}_j + \mathbf{h}_k$, which is $\mathbf{0}$ only if $\mathbf{h}_j = \mathbf{h}_k$.

By Definition 2.3.8, $H\mathbf{x} = \sum_j x_j \mathbf{h}_j$, which over $\mathbb{F}_2$ is the sum of the $\mathbf{h}_j$ with $x_j = 1$. If $\mathrm{wt}(\mathbf{x}) = 1$, with its 1 in position j, then $H\mathbf{x} = \mathbf{h}_j \neq \mathbf{0}$. If $\mathrm{wt}(\mathbf{x}) = 2$, with 1s in positions $j \neq k$, then $H\mathbf{x} = \mathbf{h}_j + \mathbf{h}_k$, which is $\mathbf{0}$ only if $\mathbf{h}_j = -\mathbf{h}_k = \mathbf{h}_k$, and the columns are distinct. So no vector of weight 1 or 2 lies in C. □

For the Hamming code the bound is attained: $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = (0,0,1) + (0,1,0) + (0,1,1) = \mathbf{0}$, so 1110000 is a codeword of weight 3, and the minimum distance is 3.

Theorem 4.2.34 (Syndrome decoding). Let H and C be as in Proposition 4.2.33, let $\mathbf{c} \in C$, and let $\mathbf{r} = \mathbf{c} + \mathbf{e}_j$ be $\mathbf{c}$ with its jth bit flipped.

(a) The **syndrome** $H\mathbf{r}$ equals $\mathbf{h}_j$. So $H\mathbf{r}$ determines j, and $\mathbf{c} = \mathbf{r} + \mathbf{e}_j$.
(b) $\mathbf{c}$ is the only codeword at distance at most 1 from $\mathbf{r}$.

For the Hamming code, the syndrome of a word with one flipped bit is the binary expansion of the position of the flip, and the syndrome of a codeword is $\mathbf{0}$.

Proof. ROUTINE

Uses. Proposition 4.2.33 (Distinct columns give distance three), Definition 4.2.31 (Linear codes), Definition 2.6.5 (Kernel)
Hint 1. $H\mathbf{c} = \mathbf{0}$ and $H\mathbf{e}_j = \mathbf{h}_j$.
Hint 2. For (b), use the triangle inequality and Proposition 4.2.33.
Sketch. (a) $H\mathbf{r} = H\mathbf{c} + H\mathbf{e}_j = \mathbf{h}_j$; distinct columns determine j; $\mathbf{r} + \mathbf{e}_j = \mathbf{c} + 2\mathbf{e}_j = \mathbf{c}$. (b) If $\mathbf{c}' \in C$ and $d(\mathbf{r}, \mathbf{c}') \leq 1$, then $d(\mathbf{c}, \mathbf{c}') \leq 2$, so $\mathbf{c} - \mathbf{c}'$ is a codeword of weight at most 2, hence $\mathbf{0}$.

(a) Since $\mathbf{c} \in \ker T_H$, $H\mathbf{r} = H\mathbf{c} + H\mathbf{e}_j = \mathbf{0} + \mathbf{h}_j = \mathbf{h}_j$. The columns are distinct, so $\mathbf{h}_j$ determines j. Over $\mathbb{F}_2$, $\mathbf{e}_j + \mathbf{e}_j = \mathbf{0}$, so $\mathbf{r} + \mathbf{e}_j = \mathbf{c}$.

(b) $d(\mathbf{r}, \mathbf{c}) = \mathrm{wt}(\mathbf{e}_j) = 1$. If $\mathbf{c}' \in C$ and $d(\mathbf{r}, \mathbf{c}') \leq 1$, the triangle inequality gives $d(\mathbf{c}, \mathbf{c}') \leq 2$, so $\mathbf{c} - \mathbf{c}'$ is a codeword of weight at most 2, and it is $\mathbf{0}$ by Proposition 4.2.33.

For the Hamming code, $\mathbf{h}_j$ is the binary expansion of j by construction. □

For example, suppose 0110011 is sent and $\mathbf{r} = 0110111$ is received. Its syndrome is the sum of the columns at the 1s of $\mathbf{r}$, which is $\mathbf{h}_2 + \mathbf{h}_3 + \mathbf{h}_5 + \mathbf{h}_6 + \mathbf{h}_7 = (0,1,0) + (0,1,1) + (1,0,1) + (1,1,0) + (1,1,1) = (1,0,1)$, binary for 5. Flipping bit 5 recovers 0110011, and reading positions $3, 5, 6, 7$ recovers the message 1011.

The numbers fit exactly. The sets $B(\mathbf{c}) = \{\mathbf{x} \mid d(\mathbf{x}, \mathbf{c}) \le 1\}$ for the 16 codewords $\mathbf{c}$ are pairwise disjoint: if $\mathbf{x} \in B(\mathbf{c}) \cap B(\mathbf{c}')$, then $d(\mathbf{c}, \mathbf{c}') \le 2$ by the triangle inequality, so $\mathbf{c} = \mathbf{c}'$ by Proposition 4.2.33. Each $B(\mathbf{c})$ has $1 + 7 = 8$ elements. Together they contain $16 \cdot 8 = 128 = 2^7$ vectors, all of $\mathbb{F}_2^7$: every word of length 7 is within distance 1 of exactly one codeword. Codes with this property are called **perfect**. In terms of the syndrome map $T_H : \mathbb{F}_2^7 \to \mathbb{F}_2^3$, which is surjective, the 8 possible syndromes correspond to "no error" and to the 7 possible error positions, and the count $7 = 4 + 3$ is rank–nullity. The Hamming code transmits 4 bits in 7, against 4 in 12 for the triple repetition code. Dual codes, which reverse the roles of the rows of H and the codewords, return in Section 12.4.

Exercises

Exercise 4.2.35. ROUTINE Let $A = \begin{pmatrix} 1 & i & 0 & 1 \\ i & -1 & 1 & 2i \\ 0 & 0 & 1+i & -1+i \end{pmatrix} \in \mathbb{C}^{3\times 4}$. Find bases of $\ker T_A$ and of $\operatorname{im} T_A$, and check rank–nullity.

Solution on page 1298

Exercise 4.2.36. ROUTINE Let $M = \begin{pmatrix} 0 & 0 & 1 & 1 \\ -1 & -2 & 2 & 0 \\ -2 & 2 & 3 & -1 \end{pmatrix}$, a matrix with integer entries. Find $\operatorname{rank} M$ and a basis of $\ker T_M$ over $\mathbb{Q}$ and over $\mathbb{F}_3$. Explain why the rational kernel vector, read modulo 3, lies in the kernel over $\mathbb{F}_3$, and why the converse fails.

Solution on page 1298

Exercise 4.2.37. ROUTINE Let $A \in \mathbb{C}^{m\times n}$. Since real scalars are complex scalars, $T_A : \mathbb{C}^n \to \mathbb{C}^m$ is also a linear map of vector spaces over $\mathbb{R}$. Show that its rank over $\mathbb{R}$ is $2 \operatorname{rank} A$ and that $\dim_{\mathbb{R}} \ker T_A = 2 \dim_{\mathbb{C}} \ker T_A$.

Solution on page 1299

Exercise 4.2.38. ROUTINE Prove or disprove. All spaces are finite-dimensional.

(a) Every $T \in \mathrm{L}(\mathbb{R}^7, \mathbb{R}^4)$ has $\dim \ker T \ge 3$.

(b) If $T \in \mathrm{L}(V, W)$ is surjective and $\dim V = 5$, then $\dim W \le 5$.

(c) There is an injective linear map $F^{2\times 3} \to F[x]_{\le 4}$.

(d) If $S, T \in \mathrm{L}(V, W)$ have the same kernel, then they have the same image.

(e) If $T \in \mathrm{L}(V)$ satisfies $\ker T = \operatorname{im} T$, then $\dim V$ is even.

Solution on page 1299

Exercise 4.2.39. ROUTINE Let $T : F[x]_{\le 3} \to F^2$, $T(p) = (p(1), p'(1))$, where p' is the formal derivative. Show that T is surjective, find $\dim \ker T$, and show that $(x-1)^2, (x-1)^3$ is a basis of $\ker T$, without solving $T(p) = \mathbf{0}$.

Solution on page 1300

Exercise 4.2.40. STANDARD Let p be a prime and $n \ge 1$, and let D be the formal derivative on $\mathbb{F}_p[x]_{\le n}$. Find nullity D and rank D, and describe $\operatorname{im} D$. Work out the case $p = 3$, $n = 7$.

Solution on page 1300

Exercise 4.2.41. STANDARD Let $n \ge 1$ and $T : F^{n\times n} \to F^{n\times n}$, $T(X) = X - X^{\mathsf{T}}$.

(a) Show that $\ker T$ is the space of symmetric matrices and that $\operatorname{rank} T = n(n-1)/2$.

(b) If $1+1 \neq 0$ in F, show that $\operatorname{im} T$ is the space of skew-symmetric matrices.

(c) If $1+1=0$ in F, show that $\operatorname{im} T$ is the space of symmetric matrices with zero diagonal. In particular, over $\mathbb{F}_2$ the image is a proper subspace of the symmetric matrices, although every value of T is symmetric.

Solution on page 1300

Exercise 4.2.42. STANDARD Let $n \geq 1$, and let $Z_n \subseteq F^{n\times n}$ be the set of matrices all of whose row sums and column sums are 0.

(a) Show that $\dim Z_n = (n-1)^2$. (Consider the map $\Phi : F^{n\times n} \to F^{2n}$ sending X to its n row sums followed by its n column sums.)

(b) Show that the matrices $U_{ij} - U_{in} - U_{nj} + U_{nn}$, for $1 \leq i, j \leq n-1$, form a basis of Z_n.

Solution on page 1301

Exercise 4.2.43. STANDARD Let $A \in F^{m\times n}$ and $p \geq 1$, and let $L_A : F^{n\times p} \to F^{m\times p}$, $L_A(X) = AX$. Show that $\operatorname{rank} L_A = p \cdot \operatorname{rank} A$.

Solution on page 1302

Exercise 4.2.44. STANDARD Let $T \in \mathrm{L}(V, W)$.

(a) Let U be a finite-dimensional subspace of V. Show that $\dim \operatorname{im}(T\restriction_U) = \dim U - \dim(U \cap \ker T)$.

(b) Let V be finite-dimensional. Show that there is a subspace U of V with $U \cap \ker T = \{\mathbf{0}\}$ and $\operatorname{im}(T\restriction_U) = \operatorname{im} T$, and that every such U has $\dim U = \operatorname{rank} T$.

(c) For $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(x, y, z) = (x + y, y + z)$, give two different subspaces U as in (b).

Solution on page 1302

Exercise 4.2.45. STANDARD Find the error in the following argument, show that its claim is false, and repair it.

Claim. Let V be finite-dimensional, $T \in \mathrm{L}(V, W)$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V such that $T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)$ is a basis of $\operatorname{im} T$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of $\ker T$.

Proof. By rank–nullity, $\dim \ker T = n - (n-k) = k$. The list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is linearly independent, as a sublist of a basis, and it has length $k = \dim \ker T$. By Theorem 4.1.10(b), it is a basis of $\ker T$.

Solution on page 1302

Exercise 4.2.46. STANDARD Let $n \geq 1$ and $0 \leq k \leq n$. Show that there is $T \in \mathrm{L}(F^n)$ with $T^2 = 0$ and $\operatorname{rank} T = k$ if and only if $2k \leq n$.

Solution on page 1303

Exercise 4.2.47. STANDARD Let F be any field, and let $T : F[x]_{\leq 3} \to F^4$, $T(p) = (p(0), p'(0), p(1), p'(1))$.

(a) Show that T is bijective. So for any prescribed values and slopes at 0 and 1 there is exactly one polynomial of degree at most 3 with them.

(b) Find the p with $p(0) = p'(0) = p'(1) = 0$ and $p(1) = 1$.

Solution on page 1303

Exercise 4.2.48. STANDARD With Δ as in Example 4.2.28, find $p \in \mathbb{Q}[x]_{\leq 4}$ with $\Delta p = x^3$, and deduce that $1^3 + 2^3 + \cdots + N^3 = (N(N+1)/2)^2$.

Solution on page 1303

Exercise 4.2.49. STANDARD Let V be finite-dimensional with $\dim V = n$.

(a) Let $U_1, \ldots, U_k$ be subspaces of V. Show that

$$\dim(U_1 \cap \cdots \cap U_k) \geq \dim U_1 + \cdots + \dim U_k - (k-1)n.$$

(b) Let $n \geq 2$. Deduce that any $n-1$ subspaces of V of dimension $n-1$ have a nonzero vector in common.

(c) Show that n subspaces of dimension $n-1$ need not have one.

Solution on page 1304

Exercise 4.2.50. STANDARD Let V be finite-dimensional and $T \in \mathrm{L}(V)$. Show that the following are equivalent:

(i) $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$;
(ii) $\ker T^2 = \ker T$;
(iii) $\operatorname{rank} T^2 = \operatorname{rank} T$;
(iv) $\operatorname{im} T^2 = \operatorname{im} T$;
(v) $\ker T + \operatorname{im} T = V$.

Show by an example on an infinite-dimensional space that (i) and (iv) are not equivalent in general.

Solution on page 1304

Exercise 4.2.51. HARD Let V be finite-dimensional with $\dim V = n$, and let $T \in \mathrm{L}(V)$.

(a) Show that $\ker T^k = \ker T^{k+1}$ for some k with $0 \leq k \leq n$.

(b) Show that if $\ker T^k = \ker T^{k+1}$, then $\ker T^m = \ker T^k$ for every $m \geq k$. Deduce that $\ker T^{2n} = \ker T^n$.

(c) Deduce that $\ker T^n \cap \operatorname{im} T^n = \{\mathbf{0}\}$ and $\dim \ker T^n + \dim \operatorname{im} T^n = n$. In particular, for every $A \in F^{n\times n}$, $\ker T_{A^n} \cap \operatorname{im} T_{A^n} = \{\mathbf{0}\}$.

Solution on page 1305

Exercise 4.2.52. HARD Let V be finite-dimensional with $\dim V = n$, let $T \in \mathrm{L}(V)$, and for $k \geq 1$ let $d_k = \dim \ker T^k - \dim \ker T^{k-1}$.

(a) Show that $d_k = \dim\big(\operatorname{im} T^{k-1} \cap \ker T\big)$. (Restrict T^{k-1} to $\ker T^k$.)

(b) Deduce that $d_1 \geq d_2 \geq d_3 \geq \cdots$.

(c) Suppose $T^m = 0$ for some $m \geq 1$. Show that $\dim \ker T \geq n/m$, so that $\operatorname{rank} T \leq (m-1)n/m$.

Solution on page 1305

Exercise 4.2.53. HARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$.

(a) Show that $\mathbf{x} \mapsto B\mathbf{x}$ maps $\ker T_{I_m - AB}$ injectively into $\ker T_{I_n - BA}$.

(b) Deduce that $\dim \ker(I_m - AB) = \dim \ker(I_n - BA)$.

(c) Deduce that $I_m - AB$ is invertible if and only if $I_n - BA$ is.

Solution on page 1306

Exercise 4.2.54. HARD Let S be a subspace of $F^{n\times n}$ in which every nonzero matrix is invertible.

(a) Show that $\dim S \le n$.

(b) Show that $S = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$ is such a subspace of $\mathbb{R}^{2\times 2}$ with $\dim S = 2$.

(c) Find such a subspace of $\mathbb{F}_2^{2\times 2}$ of dimension 2.

Solution on page 1306

Exercise 4.2.55. CHALLENGE For $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$, let $\mathbf{x}\cdot\mathbf{y} = x_1y_1 + \cdots + x_ny_n \in \mathbb{F}_2$.

(a) Let C be a subspace of $\mathbb{F}_2^n$ with $\dim C = k$, and let $C^\perp = \{\mathbf{y} \in \mathbb{F}_2^n \mid \mathbf{c}\cdot\mathbf{y} = 0 \text{ for all } \mathbf{c} \in C\}$. Show that $C^\perp$ is a subspace of dimension $n-k$.

(b) (*Eventown.*) A town has n residents, who form clubs. Every club has an even number of members, every two clubs have an even number of members in common, and no two clubs have exactly the same members. Show that there are at most $2^{\lfloor n/2 \rfloor}$ clubs, and that this many is possible.

Solution on page 1306

Exercise 4.2.56. STANDARD BEYOND 21-242 Let $r \ge 2$, $n = 2^r - 1$, and let $H_r \in \mathbb{F}_2^{r\times n}$ be a matrix whose columns are the n nonzero vectors of $\mathbb{F}_2^r$, in some order. Let $C_r = \ker T_{H_r}$, the Hamming code of length n.

(a) Show that C_r is an $[n, n-r]$ code.

(b) Show that its minimum distance is 3.

(c) Show that every $\mathbf{x} \in \mathbb{F}_2^n$ is within distance 1 of exactly one codeword.

Solution on page 1307

Exercise 4.2.57. ROUTINE BEYOND 21-242 Use the Hamming code of Example 4.2.32.

(a) Encode the message 0110.

(b) Decode the received words 1000110 and 1101011, assuming each has at most one flipped bit.

(c) The codeword 0000000 is sent, and bits 1 and 2 are flipped. What does syndrome decoding return?

Solution on page 1308

4.3 Rank inequalities and the transpose

Rank–nullity measures one matrix at a time. This section compares the ranks of matrices built from one another. Multiplying by an invertible matrix does not change the rank, and multiplying by an arbitrary matrix can only lower it:

$$\operatorname{rank}(AB) \le \min(\operatorname{rank} A, \operatorname{rank} B).$$

We prove this in three ways, by comparing images, by comparing kernels, and by restricting A to the image of B, and the last proof says exactly how much rank is lost. The transpose A^{T} has the same rank as A, although its columns are the rows of A and live in a different space. This equality, row rank equals column rank, is the main theorem of the section. It completes the count of dimensions of the four subspaces attached to a matrix. Finally, matrices of rank one are exactly the products $\mathbf{u}\mathbf{v}^{\mathsf{T}}$, rank is subadditive, and the column–row factorization $A = CR$ writes a matrix of rank r as a sum of r matrices of rank one, which is the fewest possible.

Throughout, F is a field. For $A \in F^{m\times n}$ we write $\operatorname{im} A = \operatorname{im} T_A \subseteq F^m$ and $\ker A = \ker T_A \subseteq F^n$ for the column space and the null space (Definition 2.6.17). So $\operatorname{rank} A = \dim \operatorname{im} A$ (Theorem 4.2.7) and $\dim \ker A = n - \operatorname{rank} A$ (Theorem 4.2.11).

4.3.1 Invertible factors

We already know that $\operatorname{rank}(GA) = \operatorname{rank} A$ for invertible G, because $\operatorname{rref}(GA) = \operatorname{rref}(A)$ (Proposition 3.5.2(c)), and that $\operatorname{rank}(AH) = \operatorname{rank} A$ for invertible H, from the uniqueness of the rank normal form (Exercise 3.5.41(c)). The second proof was roundabout, because row reduction cannot see column operations. Now that rank is a dimension, both statements have short proofs. They are also the reason the rank of a linear map can be read off any of its matrices: changing bases multiplies $[T]_{\mathcal{C}\mathcal{B}}$ on the left and on the right by invertible matrices (Section 4.5), and that cannot change the rank.

Intuition (Outputs and inputs). In GA the matrix G acts after A, on the outputs $A\mathbf{x}$. An invertible G is injective, so it never sends a nonzero output to $\mathbf{0}$: $GA\mathbf{x} = \mathbf{0}$ only when $A\mathbf{x} = \mathbf{0}$, and $\ker(GA) = \ker A$. But G moves the outputs, so $\operatorname{im}(GA)$, the set of vectors $G\mathbf{y}$ with $\mathbf{y} \in \operatorname{im} A$, is in general a different subspace of F^m. It has the same dimension, since an injective map preserves dimension.

In AH the matrix H acts before A, on the inputs. An invertible H is surjective, so AH still sees every input, and $\operatorname{im}(AH) = \operatorname{im} A$. But the kernel moves: $AH\mathbf{x} = \mathbf{0}$ exactly when $H\mathbf{x} \in \ker A$. So each of GA and AH keeps one of the two subspaces and moves the other by an invertible map.

Proposition 4.3.1 (Invertible factors preserve rank). Let $A \in F^{m\times n}$, and let $G \in F^{m\times m}$ and $H \in F^{n\times n}$ be invertible.

(a) $\ker(GA) = \ker A$, and $\operatorname{rank}(GA) = \operatorname{rank} A$.
(b) $\operatorname{im}(AH) = \operatorname{im} A$, and $\operatorname{rank}(AH) = \operatorname{rank} A$.

In particular, $\operatorname{rank}(GAH) = \operatorname{rank} A$.

The first proof is the lecture's. It compares images, using the restriction identity $\operatorname{im}(TS) = \operatorname{im}(T|_{\operatorname{im} S})$.

Proof (by restriction, comparing images). ROUTINE

Uses. Theorem 2.4.2 (The matrix of a composite), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 2.6.15 (Kernels and images of composites), Theorem 4.2.7 (The image of T_A)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. Write $T_{GA} = T_G \circ T_A$. By Proposition 4.2.3(b), $\operatorname{im}(GA)$ is the image of T_G restricted to $\operatorname{im} A$. What do you know about T_G?
Hint 2. T_G is injective, so its restriction to $\operatorname{im} A$ is an injective linear map onto $\operatorname{im}(GA)$. For AH, what is $\operatorname{im} T_H$?
Sketch. $\operatorname{im}(GA) = \operatorname{im}(T_G|_{\operatorname{im} A})$, and T_G is injective because $T_{G^{-1}}$ is a left inverse, so the restriction carries $\operatorname{im} A$ injectively onto $\operatorname{im}(GA)$ and the dimensions agree. On the other side, $\operatorname{im}(AH) = \operatorname{im}(T_A|_{\operatorname{im} H}) = \operatorname{im}(T_A|_{F^n}) = \operatorname{im} A$, because T_H is surjective. The kernel statement in (a) is Proposition 2.6.15(a).

By Theorem 2.4.2, $T_{GA} = T_G \circ T_A$ and $T_{AH} = T_A \circ T_H$. Since $G^{-1}G = I_m$ and $HH^{-1} = I_n$, the map T_G is injective (Proposition 2.5.4) and T_H is surjective (Proposition 2.5.5).

(a) By Proposition 4.2.3(b), $\operatorname{im}(GA) = \operatorname{im}(T_G|_{\operatorname{im} A})$. The restriction $T_G|_{\operatorname{im} A}$ is linear and injective (Proposition 4.2.3(a), (c)), so Proposition 4.2.6(a), applied to it, gives $\dim \operatorname{im}(GA) = \dim \operatorname{im} A$. That is, $\operatorname{rank}(GA) = \operatorname{rank} A$ (Theorem 4.2.7). Since T_G is injective, $\ker(GA) = \ker A$ by Proposition 2.6.15(a).

(b) By Proposition 4.2.3(b), $\operatorname{im}(AH) = \operatorname{im}(T_A|_{\operatorname{im} H})$. Since $\operatorname{im} H = F^n$, this restriction is T_A itself, so $\operatorname{im}(AH) = \operatorname{im} A$, and in particular $\operatorname{rank}(AH) = \operatorname{rank} A$.

Finally, $\operatorname{rank}(GAH) = \operatorname{rank}(G(AH)) = \operatorname{rank}(AH) = \operatorname{rank} A$ by (a) and (b). □

The lecture remarks that the same statements follow from kernels and rank–nullity. Here is that proof.

Proof (by kernels and rank–nullity). ROUTINE

> *Uses.* Proposition 2.6.15 (Kernels and images of composites), Theorem 4.2.11 (Rank–nullity), Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.1.21 (Dimensions of subspaces)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* GA and AH have n columns, as A does. By rank–nullity it suffices to compare $\dim\ker(GA)$ and $\dim\ker(AH)$ with $\dim\ker A$.
> *Hint 2.* G can be cancelled from $GA\mathbf{x}=\mathbf{0}$. For AH, show that $\mathbf{x}\in\ker(AH)$ if and only if $H\mathbf{x}\in\ker A$, and use T_H to move one kernel onto the other.
> *Sketch.* If $GA\mathbf{x}=\mathbf{0}$, then $A\mathbf{x}=G^{-1}GA\mathbf{x}=\mathbf{0}$, so $\ker(GA)=\ker A$, and rank–nullity gives $\operatorname{rank}(GA)=n-\dim\ker A=\operatorname{rank}A$. Next, $\mathbf{x}\in\ker(AH)$ if and only if $H\mathbf{x}\in\ker A$, so T_H restricts to an injective linear map from $\ker(AH)$ onto $\ker A$. The kernels have the same dimension, so $\operatorname{rank}(AH)=\operatorname{rank}A$ by rank–nullity, and then $\operatorname{im}(AH)\subseteq\operatorname{im}A$ with equal dimensions forces $\operatorname{im}(AH)=\operatorname{im}A$.

(a) By Proposition 2.6.15(a), $\ker A\subseteq\ker(GA)$. Conversely, if $GA\mathbf{x}=\mathbf{0}$, then $A\mathbf{x}=G^{-1}(GA\mathbf{x})=\mathbf{0}$. So $\ker(GA)=\ker A$. The matrices GA and A both have n columns, so by Theorem 4.2.11(a),

$$\operatorname{rank}(GA)=n-\dim\ker(GA)=n-\dim\ker A=\operatorname{rank}A.$$

(b) For $\mathbf{x}\in F^n$, we have $AH\mathbf{x}=\mathbf{0}$ if and only if $H\mathbf{x}\in\ker A$. So T_H maps $\ker(AH)$ into $\ker A$. Since $H^{-1}H=I_n$, T_H is injective, so its restriction to $\ker(AH)$ is linear and injective (Proposition 4.2.3). It maps onto $\ker A$: if $\mathbf{y}\in\ker A$, then $\mathbf{x}=H^{-1}\mathbf{y}$ satisfies $AH\mathbf{x}=A\mathbf{y}=\mathbf{0}$ and $H\mathbf{x}=\mathbf{y}$. By Proposition 4.2.6(a), $\dim\ker(AH)=\dim\ker A$, and since AH has n columns, Theorem 4.2.11(a) gives

$$\operatorname{rank}(AH)=n-\dim\ker(AH)=n-\dim\ker A=\operatorname{rank}A.$$

Finally, $\operatorname{im}(AH)\subseteq\operatorname{im}A$ by Proposition 2.6.15(b), and these subspaces of F^m have the same dimension, so they are equal (Proposition 4.1.21(b)).

As in the first proof, $\operatorname{rank}(GAH)=\operatorname{rank}A$ follows from (a) and (b). □

The two proofs split the work in opposite ways. In the first, the left factor needs an argument (an injective map preserves dimension) and the right factor is free (a surjective map on the inputs changes nothing). In the second, the left factor is free (the kernel does not change) and the right factor needs the argument. The same split reappears in Section 4.3.2, where an arbitrary product is compared with its factors.

By Exercise 3.5.41(d), two matrices $A,B\in F^{m\times n}$ satisfy $B=GAH$ for some invertible G and H if and only if they have the same rank. So the rank is the only information about A that row and column operations together preserve. In particular, similar matrices $B=PAP^{-1}$ have the same rank, a fact we return to in Section 4.5.

Exercise 4.3.2. ROUTINE Let $A=\begin{pmatrix}1&2\\2&4\end{pmatrix}$, $G=\begin{pmatrix}1&0\\-2&1\end{pmatrix}$, and $H=\begin{pmatrix}1&-2\\0&1\end{pmatrix}$, over $\mathbb{Q}$. Compute GA and AH, and find the kernel and the image of each of A, GA, and AH. Which subspaces agree, and which do not?

Solution on page 1308

4.3.2 The rank of a product

Intuition (A pipeline is no wider than its narrowest stage). For $\mathbf{x}\in F^p$, $AB\mathbf{x}=A(B\mathbf{x})$: first B, then A. Every output of AB is an output of A, so $\operatorname{im}(AB)\subseteq\operatorname{im}A$; the product cannot reach a direction that A cannot reach. On the other hand, A only ever receives the vectors $B\mathbf{x}$, which fill a space of dimension $\operatorname{rank}B$, and a linear map does not raise dimension. So

$$\operatorname{rank}(AB)\le\operatorname{rank}A\quad\text{and}\quad\operatorname{rank}(AB)\le\operatorname{rank}B.$$

The first bound is about where the outputs land and is immediate. The second is about how many independent directions enter A, and it needs a dimension argument.

Theorem 4.3.3 (Rank of a product).
(a) Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$. Then $\operatorname{im}(AB) \subseteq \operatorname{im} A$ and $\ker B \subseteq \ker(AB)$, and

$$\operatorname{rank}(AB) \leq \min(\operatorname{rank} A, \operatorname{rank} B).$$

(b) More generally, let U and V be finite-dimensional, and let $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$. Then

$$\operatorname{rank}(TS) \leq \min(\operatorname{rank} T, \operatorname{rank} S).$$

Part (a) is part (b) for $S = T_B$ and $T = T_A$, since $T_{AB} = T_A \circ T_B$ (Theorem 2.4.2) and $\operatorname{rank} A = \operatorname{rank} T_A$ (Definition 4.2.14); the two containments in (a) are Proposition 2.6.15. So we prove (b). In both proofs below, the bound by $\operatorname{rank} T$ comes from $\operatorname{im}(TS) \subseteq \operatorname{im} T$. They differ in how they prove the bound by $\operatorname{rank} S$.

Proof (comparing images). ROUTINE

> *Uses.* Proposition 2.6.15 (Kernels and images of composites), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Proposition 4.2.3 (Properties of restriction), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Corollary 4.1.11 (Spanning lists are long)
> *Techniques.* Restrict to a subspace (Technique 4.2.5)
> *Hint 1.* $\operatorname{im}(TS) \subseteq \operatorname{im} T$, which gives one inequality at once.
> *Hint 2.* For the other, $\operatorname{im}(TS) = \operatorname{im}(T{\restriction}_{\operatorname{im} S})$ consists of the values of T on $\operatorname{im} S$. Take a basis of $\operatorname{im} S$ and push it forward with T.
> *Hint 3.* A space spanned by r vectors has dimension at most r.
> *Sketch.* $\operatorname{im}(TS) \subseteq \operatorname{im} T$ gives $\operatorname{rank}(TS) \leq \operatorname{rank} T$. If $\mathbf{s}_1, \dots, \mathbf{s}_r$ is a basis of $\operatorname{im} S$, then $\operatorname{im}(TS) = \operatorname{im}(T{\restriction}_{\operatorname{im} S})$ is spanned by $T(\mathbf{s}_1), \dots, T(\mathbf{s}_r)$, so its dimension is at most $r = \operatorname{rank} S$.

Since U and V are finite-dimensional, $\operatorname{im} S$, $\operatorname{im} T$, and $\operatorname{im}(TS)$ are finite-dimensional (Theorem 4.2.11(b)), so all three ranks are defined.

The bound by $\operatorname{rank} T$. By Proposition 2.6.15(b), $\operatorname{im}(TS) \subseteq \operatorname{im} T$, so $\operatorname{rank}(TS) \leq \operatorname{rank} T$ by Proposition 4.1.21(a).

The bound by $\operatorname{rank} S$. Let $r = \operatorname{rank} S$ and let $\mathbf{s}_1, \dots, \mathbf{s}_r$ be a basis of $\operatorname{im} S$ (the empty list if $r = 0$). By Proposition 4.2.3(b) and Proposition 2.1.16(a),

$$\operatorname{im}(TS) = \operatorname{im}(T{\restriction}_{\operatorname{im} S}) = \operatorname{span}(T(\mathbf{s}_1), \dots, T(\mathbf{s}_r)).$$

So $\operatorname{im}(TS)$ is spanned by a list of length r, and $\dim \operatorname{im}(TS) \leq r$ by Corollary 4.1.11(a). That is, $\operatorname{rank}(TS) \leq \operatorname{rank} S$. □

The second proof moves the comparison to the domain U, where S and TS have a common domain and rank–nullity applies to both.

Proof (comparing kernels, with rank–nullity). ROUTINE

> *Uses.* Proposition 2.6.15 (Kernels and images of composites), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.2.11 (Rank–nullity)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* The bound by $\operatorname{rank} T$ is the image comparison $\operatorname{im}(TS) \subseteq \operatorname{im} T$. For the bound by $\operatorname{rank} S$, compare kernels instead: which of $\ker S$ and $\ker(TS)$ contains the other?
> *Hint 2.* $\ker S \subseteq \ker(TS)$, both subspaces of U. Apply rank–nullity to S and to TS.
> *Sketch.* $\operatorname{rank}(TS) \leq \operatorname{rank} T$ because $\operatorname{im}(TS) \subseteq \operatorname{im} T$. Since $\ker S \subseteq \ker(TS)$, rank–nullity on U gives $\operatorname{rank}(TS) = \dim U - \dim \ker(TS) \leq \dim U - \dim \ker S = \operatorname{rank} S$.

As in the first proof, $\operatorname{im}(TS) \subseteq \operatorname{im} T$ (Proposition 2.6.15(b)), so $\operatorname{rank}(TS) \leq \operatorname{rank} T$ (Proposition 4.1.21(a)).

By Proposition 2.6.15(a), $\ker S \subseteq \ker(TS)$. Both are subspaces of the finite-dimensional space U, so $\dim \ker S \leq \dim \ker(TS)$ (Proposition 4.1.21(a)). By Theorem 4.2.11(b), applied to S and to TS, which are both defined on U,

$$\begin{aligned} \operatorname{rank}(TS) &= \dim U - \dim \ker(TS) \\ &\leq \dim U - \dim \ker S = \operatorname{rank} S, \end{aligned}$$

as desired. □

Both inequalities can be strict, even at the same time.

Example 4.3.4 (A product of rank zero). Let $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, so $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_2 = \mathbf{e}_1$. Then $\operatorname{rank} N = 1$, but $N^2 = 0$ has rank $0 < \min(\operatorname{rank} N, \operatorname{rank} N)$. In NN, the left factor kills exactly what the right factor produces: $\operatorname{im} N = \operatorname{span}(\mathbf{e}_1) = \ker N$. The order of the factors matters too: with $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $MN = N$ has rank 1 but $NM = 0$, so $\operatorname{rank}(AB)$ and $\operatorname{rank}(BA)$ can differ even for square matrices.

What a product loses

The example shows where the rank goes. AB applies A to the vectors of $\operatorname{im} B$, and A kills those that lie in $\ker A$. Restricting A to $\operatorname{im} B$ makes this exact. This is the lecture's restriction view of a composite, and it gives the third proof of the bound by $\operatorname{rank} B$.

Proposition 4.3.5 (Rank of a product, exactly). Let $S \in \mathrm{L}(U, V)$ and $T \in \mathrm{L}(V, W)$, with U finite-dimensional.

(a) $\operatorname{rank}(TS) = \operatorname{rank} S - \dim(\operatorname{im} S \cap \ker T)$.

(b) $\operatorname{rank}(TS) = \operatorname{rank} S$ if and only if $\operatorname{im} S \cap \ker T = \{\mathbf{0}\}$. If V is also finite-dimensional, then $\operatorname{rank}(TS) = \operatorname{rank} T$ if and only if $\operatorname{im} S + \ker T = V$.

For matrices $A \in F^{m\times n}$ and $B \in F^{n\times p}$, (a) reads

$$\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A). \tag{4.3.1}$$

Proof (by restriction). STANDARD

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.2.3 (Properties of restriction), Exercise 4.2.44, Proposition 2.6.15 (Kernels and images of composites), Proposition 4.1.21 (Dimensions of subspaces), Exercise 2.6.39, Theorem 2.4.2 (The matrix of a composite)
Techniques. Restrict to a subspace (Technique 4.2.5), Dimension counting (Technique 4.2.24)
Hint 1. $\operatorname{im}(TS) = \operatorname{im}(T|_{\operatorname{im} S})$. What is the kernel of $T|_{\operatorname{im} S}$?
Hint 2. Apply rank–nullity to $T|_{\operatorname{im} S}$, whose domain is the finite-dimensional space $\operatorname{im} S$.
Hint 3. For the last statement, $\operatorname{im}(TS) \subseteq \operatorname{im} T$; when do two finite-dimensional nested subspaces have the same dimension? Then use Exercise 2.6.39(b).
Sketch. The restriction $R = T|_{\operatorname{im} S}$ has domain $\operatorname{im} S$, of dimension $\operatorname{rank} S$, kernel $\operatorname{im} S \cap \ker T$, and image $\operatorname{im}(TS)$. Rank–nullity for R is (a), and (a) gives the first half of (b). For the second half, $\operatorname{im}(TS) \subseteq \operatorname{im} T$, so the ranks agree exactly when the images agree, which by Exercise 2.6.39(b) means $\operatorname{im} S + \ker T = V$.

(a) Since U is finite-dimensional, so is $\operatorname{im} S$, and $\dim \operatorname{im} S = \operatorname{rank} S$ (Theorem 4.2.11(b)). Let $R = T|_{\operatorname{im} S} : \operatorname{im} S \to W$. By Proposition 4.2.3, R is linear, $\ker R = \operatorname{im} S \cap \ker T$, and $\operatorname{im} R = \operatorname{im}(TS)$. Rank–nullity for R (Theorem 4.2.11(b)) gives

$$\operatorname{rank} S = \dim \operatorname{im} S = \dim(\operatorname{im} S \cap \ker T) + \dim \operatorname{im}(TS),$$

which is (a). (This is Exercise 4.2.44(a) for the subspace $\operatorname{im} S$.)

(b) By (a), $\operatorname{rank}(TS) = \operatorname{rank} S$ if and only if $\dim(\operatorname{im} S \cap \ker T) = 0$, that is, $\operatorname{im} S \cap \ker T = \{\mathbf{0}\}$. Now let V be finite-dimensional, so that $\operatorname{im} T$ is finite-dimensional. Since $\operatorname{im}(TS) \subseteq \operatorname{im} T$ (Proposition 2.6.15(b)), Proposition 4.1.21(b) shows that $\operatorname{rank}(TS) = \operatorname{rank} T$ if and only if $\operatorname{im}(TS) = \operatorname{im} T$, and by Exercise 2.6.39(b) this holds if and only if $\operatorname{im} S + \ker T = V$.

The matrix statement is (a) for $S = T_B$ and $T = T_A$ (Theorem 2.4.2). □

Intuition (Three proofs of $\operatorname{rank}(TS) \le \operatorname{rank} S$*).* The image proof works in the middle space V: it pushes a basis of $\operatorname{im} S$ forward with T and observes that at most $\operatorname{rank} S$ vectors come out. The kernel proof works in the domain U: whatever S kills, TS kills too, and rank–nullity converts a larger kernel into a smaller image. The restriction proof is the sharpest, because it computes the loss instead of bounding it: TS loses exactly the directions that S produces and T destroys, $\operatorname{im} S \cap \ker T$. It answers the question of when the inequality is an equality, and it also gives lower bounds for $\operatorname{rank}(TS)$, which are the subject

of Exercise 4.3.30 and Exercise 4.3.31. When all you need is the inequality, the image proof is the one to remember.

Example 4.3.6 (Powers of a square matrix). Let $A \in F^{n\times n}$. By (4.3.1) with $B = A$,

$$\operatorname{rank}(A^2) = \operatorname{rank} A - \dim(\operatorname{im} A \cap \ker A).$$

So $\operatorname{rank}(A^2) = \operatorname{rank} A$ if and only if $\operatorname{im} A \cap \ker A = \{\mathbf{0}\}$. For the projection $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\operatorname{im} P = \operatorname{span}(\mathbf{e}_1)$ and $\ker P = \operatorname{span}(\mathbf{e}_2)$ meet only in $\mathbf{0}$, and indeed $P^2 = P$. For N of Example 4.3.4, $\operatorname{im} N = \ker N$, and $\operatorname{rank}(N^2) = 1 - 1 = 0$.

With $B = A^k$ instead, the formula gives $\operatorname{rank}(A^{k+1}) = \operatorname{rank}(A^k) - \dim(\operatorname{im} A^k \cap \ker A)$. The subspaces $\operatorname{im} A^k$ shrink as k grows, so the drops $\operatorname{rank}(A^k) - \operatorname{rank}(A^{k+1})$ can only decrease. This is Exercise 4.2.52 in the language of images, and by Exercise 4.2.51 the ranks stop decreasing by $k = n$.

Exercise 4.3.7. ROUTINE Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $m < n$.

(a) Show that $BA \in F^{n\times n}$ is not invertible.

(b) Give an example with $AB = I_m$.

(c) Show that if $AB = I_m$, then $\operatorname{rank}(BA) = m$.

Solution on page 1308

4.3.3 Row rank equals column rank

Intuition (Why this is surprising). Consider

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 2 & 4 & 0 & 6 \\ 1 & 2 & -1 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 4}.$$

Its four columns are vectors in $\mathbb{Q}^3$. They satisfy $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_4 = 3\mathbf{a}_1 + \mathbf{a}_3$, and $\mathbf{a}_1, \mathbf{a}_3$ are independent, so exactly two of the columns are independent. Its three rows are vectors in $\mathbb{Q}^4$. They satisfy one relation, $\text{row}_1 - \text{row}_2 + \text{row}_3 = \mathbf{0}$, and any two of them are independent, so again exactly two of the rows are independent.

Nothing in the definitions suggests that these two numbers should agree. The columns satisfy two independent relations and the rows only one; the columns live in F^3 and the rows in F^4. For a 3×100 matrix, one count is about 100 vectors in F^3 and the other is about 3 vectors in F^{100}. Yet the two counts always agree. Equivalently, the numbers of independent relations always differ by exactly the difference between the numbers of columns and rows: here $2 - 1 = 4 - 3$.

The columns of A^{T} are the rows of A, written as columns (Definition 2.4.21). So $\operatorname{im} A^{\mathsf{T}} \subseteq F^n$ is the span of the rows of A (Proposition 2.6.18), and $\operatorname{rank} A^{\mathsf{T}}$ is its dimension.

Theorem 4.3.8 (Row rank equals column rank). For every $A \in F^{m\times n}$,

$$\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A.$$

That is, the span of the rows of A, viewed as vectors in F^n, has the same dimension as the span of its columns in F^m.

The first proof reduces A to its rank normal form, where the statement is obvious.

Proof (via the rank normal form). ROUTINE

Uses. Theorem 3.5.27 (Rank normal form), Proposition 2.4.23 (Transpose of a product), Proposition 2.5.17 (Algebra of inverses), Proposition 4.3.1 (Invertible factors preserve rank), Proposition 2.6.18 (The column space is spanned by the columns), Theorem 4.2.7 (The image of T_A)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. The statement is obvious for $N_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$. Reduce to it.
Hint 2. Transpose $GAH = N_k$. Which factors of the result are invertible, and what is the rank of N_k^{T}?
Sketch. With $GAH = N_k$ and $k = \operatorname{rank} A$, transposing gives $H^{\mathsf{T}}A^{\mathsf{T}}G^{\mathsf{T}} = N_k^{\mathsf{T}}$, with H^{T} and G^{T} invertible. By Proposition 4.3.1, $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} N_k^{\mathsf{T}}$, and N_k^{T} has columns $\mathbf{e}_1, \dots, \mathbf{e}_k$ followed by zero columns, so its rank is k.

Let $k = \operatorname{rank} A$. By Theorem 3.5.27 there are invertible $G \in F^{m\times m}$ and $H \in F^{n\times n}$ with $GAH = N_k$. Transposing with Proposition 2.4.23,

$$H^{\mathsf{T}}A^{\mathsf{T}}G^{\mathsf{T}} = (GAH)^{\mathsf{T}} = N_k^{\mathsf{T}}.$$

The matrices H^{T} and G^{T} are invertible (Proposition 2.5.17(d)), so $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank}\left(H^{\mathsf{T}}A^{\mathsf{T}}G^{\mathsf{T}}\right) = \operatorname{rank} N_k^{\mathsf{T}}$ by Proposition 4.3.1. The matrix $N_k^{\mathsf{T}} \in F^{n\times m}$ has 1 in positions $(1,1), \dots, (k,k)$ and 0 elsewhere, so its columns are $\mathbf{e}_1, \dots, \mathbf{e}_k \in F^n$ followed by $m-k$ zero columns. Hence $\operatorname{im} N_k^{\mathsf{T}} = \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$ (Proposition 2.6.18), which has the basis $\mathbf{e}_1, \dots, \mathbf{e}_k$. So $\operatorname{rank} N_k^{\mathsf{T}} = k$ (Theorem 4.2.7), and $\operatorname{rank} A^{\mathsf{T}} = k = \operatorname{rank} A$. □

The second proof uses only row operations. They change the column space of A but keep its row space, and the reduced form exposes a basis of the row space.

Proof (row operations keep the row space, and RREF exposes it). STANDARD

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 2.4.23 (Transpose of a product), Proposition 2.5.17 (Algebra of inverses), Proposition 4.3.1 (Invertible factors preserve rank), Proposition 3.3.8 (Properties of the output), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Proposition 2.6.18 (The column space is spanned by the columns), Theorem 1.5.5 (Independence is decided at zero), Theorem 4.2.7 (The image of T_A)
Techniques. Reduce to RREF (Technique 3.3.22), Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. Transpose $GA = A_r$, and recall from Proposition 4.3.1(b) that an invertible factor on the right does not change the image.
Hint 2. So A and A_r have the same row space. Which rows of A_r are nonzero, and why are they linearly independent?
Hint 3. Row s of A_r has a 1 in the pivot column p_s, and every other row has a 0 there.
Sketch. From $GA = A_r$ we get $A_r^{\mathsf{T}} = A^{\mathsf{T}}G^{\mathsf{T}}$ with G^{T} invertible, so $\operatorname{im} A_r^{\mathsf{T}} = \operatorname{im} A^{\mathsf{T}}$ by Proposition 4.3.1(b): A and A_r have the same row space. The nonzero rows of A_r are its first $k = \operatorname{rank} A$ rows. They span its row space, and they are independent because row s is the only one with a nonzero entry in column p_s. So the row space has dimension k.

Let $k = \operatorname{rank} A$, and let $G \in F^{m\times m}$ be invertible with $GA = A_r$ (Theorem 3.3.10). Transposing, $A_r^{\mathsf{T}} = A^{\mathsf{T}}G^{\mathsf{T}}$ (Proposition 2.4.23), and G^{T} is invertible (Proposition 2.5.17(d)). By Proposition 4.3.1(b),

$$\operatorname{im} A_r^{\mathsf{T}} = \operatorname{im} A^{\mathsf{T}}.$$

That is, A and A_r have the same row space.

By Proposition 3.3.8(a), rows $k+1, \dots, m$ of A_r are zero. Let $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_k \in F^n$ be the transposes of rows $1, \dots, k$ of A_r. These are the first k columns of A_r^{T}, and its other columns are zero, so $\operatorname{im} A_r^{\mathsf{T}} = \operatorname{span}(\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_k)$ (Proposition 2.6.18). We show that $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_k$ is linearly independent. Let $p_1 < \cdots < p_k$ be the pivot columns of A_r. Column p_t of A_r is $\mathbf{e}_t$ (Lemma 3.3.16(b)), so entry p_t of $\boldsymbol{\rho}_s$ is 1 if $s = t$ and 0 if $s \neq t$. Hence if $c_1\boldsymbol{\rho}_1 + \cdots + c_k\boldsymbol{\rho}_k = \mathbf{0}$, then entry p_t of the left side is c_t, so $c_t = 0$ for every t (Theorem 1.5.5).

So $\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_k$ is a basis of $\operatorname{im} A_r^{\mathsf{T}} = \operatorname{im} A^{\mathsf{T}}$, and $\operatorname{rank} A^{\mathsf{T}} = \dim \operatorname{im} A^{\mathsf{T}} = k = \operatorname{rank} A$ (Theorem 4.2.7). □

Both proofs reduce A to a form in which the theorem can be seen, but they teach different things. The first uses row and column operations together and lands on N_k, whose transpose is again of the same shape. The second uses only row operations, and it rests on a contrast. Row operations change the column space but keep the relations among the columns (Corollary 3.2.17), hence the number of independent columns. They keep the row space itself. So reduction preserves both numbers, and in A_r both are visibly the number of pivots. The second proof also computes something: the nonzero rows of $\operatorname{rref}(A)$ form a basis of the row

space of A. A third proof, from the column–row factorization, is in Section 4.3.6, and for real matrices Exercise 4.3.33 gives a fourth. The transpose of a matrix is the matrix of the *dual map*, and the coordinate-free form of the theorem is that a linear map and its dual have the same rank (Section 13.1). That point of view is beyond the course.

Remark 4.3.9 (The two halves of the product inequality). With the transpose available, the two bounds in Theorem 4.3.3 are mirror images of each other (Technique 3.2.27). If we only knew the easy bound $\operatorname{rank}(XY) \le \operatorname{rank} X$, which comes from $\operatorname{im}(XY) \subseteq \operatorname{im} X$, then

$$\begin{aligned}\operatorname{rank}(AB) &= \operatorname{rank}\big((AB)^{\mathsf{T}}\big) = \operatorname{rank}\big(B^{\mathsf{T}}A^{\mathsf{T}}\big)\\ &\le \operatorname{rank} B^{\mathsf{T}} = \operatorname{rank} B.\end{aligned}$$

So the bound that needed a dimension argument follows from the bound about images, once rows and columns are known to have the same rank.

The theorem adds a condition about rows to the Second Fundamental Theorem.

Corollary 4.3.10 (Independent rows). Let $A \in F^{m\times n}$. The following are equivalent.

(i) The rows of A are linearly independent.
(ii) $\operatorname{rank} A = m$.
(iii) For every $\mathbf{y} \in F^m$, the equation $A\mathbf{x} = \mathbf{y}$ has a solution.

In particular, a square matrix is invertible if and only if its rows are linearly independent.

Proof. ROUTINE

Uses. Theorem 4.3.8 (Row rank equals column rank), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Corollary 3.5.25 (The invertible matrix theorem)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint. The rows of A are the columns of $A^{\mathsf{T}} \in F^{n\times m}$. Apply the First Fundamental Theorem to A^{T}.
Sketch. The rows of A are independent iff the m columns of A^{T} are, iff $\operatorname{rank} A^{\mathsf{T}} = m$ (Theorem 3.5.12 for A^{T}), iff $\operatorname{rank} A = m$, iff T_A is surjective (Theorem 3.5.13).

The rows of A, written as columns, are the m columns of $A^{\mathsf{T}} \in F^{n\times m}$. By (III) ⇔ (VI) of Theorem 3.5.12 for A^{T}, they are linearly independent if and only if $\operatorname{rank} A^{\mathsf{T}} = m$, that is, $\operatorname{rank} A = m$ by Theorem 4.3.8. This proves (i) ⇔ (ii), and (ii) ⇔ (iii) is (VI) ⇔ (II) of Theorem 3.5.13. For $m = n$, condition (ii) says that A is invertible (Corollary 3.5.25). □

Read as a statement about equations, (i) ⇔ (iii) is natural in one direction. If some combination of the rows vanishes, say $z_1(\text{row } 1) + \cdots + z_m(\text{row } m) = \mathbf{0}$ with not all $z_i = 0$, then every solvable right side satisfies $z_1y_1 + \cdots + z_my_m = 0$, so not every $\mathbf{y}$ is reachable. The corollary says that this is the only obstruction.

Exercise 4.3.11. ROUTINE Over $\mathbb{F}_5$, let $A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 1 & 0 \\ 0 & 4 & 2 & 3 \end{pmatrix}$.

(a) Find $\operatorname{rank} A$, and a basis of the row space of A from $\operatorname{rref}(A)$.

(b) Find a basis of the row space consisting of rows of A, and express the remaining row in terms of it.

Solution on page 1309

4.3.4 The four fundamental subspaces

A matrix $A \in F^{m\times n}$ has a kernel and an image, and so does its transpose. Together they are the four subspaces attached to A.

Definition 4.3.12 (The four fundamental subspaces). Let $A \in F^{m\times n}$. The **four fundamental subspaces** of A are

- the **column space** $\operatorname{im} A \subseteq F^m$, the span of the columns of A;
- the **null space** $\ker A \subseteq F^n$;
- the **row space** $\operatorname{im} A^{\mathsf{T}} \subseteq F^n$, the span of the rows of A written as columns;
- the **left null space** $\ker A^{\mathsf{T}} \subseteq F^m$.

The **column rank** of A is $\dim \operatorname{im} A$, and the **row rank** is $\dim \operatorname{im} A^{\mathsf{T}}$.

So Theorem 4.3.8 says that the row rank equals the column rank, and both equal $\operatorname{rank} A$. Table 4.3.1 summarizes the dimensions and bases found below. The name "left null space" comes from writing the condition on the left: since $\left(A^{\mathsf{T}}\mathbf{y}\right)^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}A$ (Proposition 2.4.23), a vector $\mathbf{y} \in F^m$ lies in $\ker A^{\mathsf{T}}$ exactly when $\mathbf{y}^{\mathsf{T}}A = \mathbf{0}^{\mathsf{T}}$, that is,

$$y_1(\text{row 1 of } A) + \cdots + y_m(\text{row } m \text{ of } A) = \mathbf{0}.$$

So the null space records the linear relations among the columns of A, and the left null space records the linear relations among the rows, or equivalently among the equations of the system $A\mathbf{x} = \mathbf{y}$.

Proposition 4.3.13 (Dimensions of the four subspaces). Let $A \in F^{m\times n}$ have rank r, and let $G \in F^{m\times m}$ be invertible with $GA = A_r$.

(a) $\dim \operatorname{im} A = r$, $\dim \ker A = n - r$, $\dim \operatorname{im} A^{\mathsf{T}} = r$, and $\dim \ker A^{\mathsf{T}} = m - r$.

(b) The pivot columns of A form a basis of $\operatorname{im} A$; the basic solutions form a basis of $\ker A$; the nonzero rows of A_r, written as columns, form a basis of $\operatorname{im} A^{\mathsf{T}}$; and the last $m - r$ rows of G, written as columns, form a basis of $\ker A^{\mathsf{T}}$.

Proof. ROUTINE

Uses. Theorem 4.2.7 (The image of T_A), Theorem 4.2.11 (Rank–nullity), Theorem 4.3.8 (Row rank equals column rank), Remark 3.4.11 (A basis of the kernel), Exercise 3.3.57, Corollary 4.1.11 (Spanning lists are long)

Hint 1. Rank–nullity for A gives two of the dimensions. Rank–nullity for $A^{\mathsf{T}} \in F^{n\times m}$, together with Theorem 4.3.8, gives the other two.

Hint 2. For the basis of $\ker A^{\mathsf{T}}$, Exercise 3.3.57 shows that the last $m - r$ rows of G span the relations among the rows of A. How many of them are there?

Sketch. (a) $\dim \operatorname{im} A = r$ and $\dim \ker A = n - r$ are Theorem 4.2.7 and rank–nullity. For $A^{\mathsf{T}} \in F^{n\times m}$, $\operatorname{rank} A^{\mathsf{T}} = r$ and rank–nullity gives $\dim \ker A^{\mathsf{T}} = m - r$. (b) The first two bases are Theorem 4.2.7 and Remark 3.4.11, and the third is the second proof of Theorem 4.3.8. The last $m - r$ rows of G, transposed, span $\ker A^{\mathsf{T}}$ by Exercise 3.3.57, and $m - r$ vectors spanning a space of dimension $m - r$ form a basis.

(a) By Theorem 4.2.7 and Theorem 4.2.11(a), $\dim \operatorname{im} A = r$ and $\dim \ker A = n - r$. The transpose $A^{\mathsf{T}} \in F^{n\times m}$ has m columns and rank r (Theorem 4.3.8), so $\dim \operatorname{im} A^{\mathsf{T}} = r$ and $\dim \ker A^{\mathsf{T}} = m - r$, by the same two results applied to A^{T}.

(b) The pivot columns of A form a basis of $\operatorname{im} A$ by Theorem 4.2.7, and the basic solutions form a basis of $\ker A$ by Remark 3.4.11. The nonzero rows of A_r, transposed, form a basis of $\operatorname{im} A^{\mathsf{T}}$, as shown in the second proof of Theorem 4.3.8. Finally, $\mathbf{y} \in \ker A^{\mathsf{T}}$ if and only if $\mathbf{y}^{\mathsf{T}}A = \mathbf{0}^{\mathsf{T}}$, that is, $\mathbf{y}^{\mathsf{T}}$ is a relation among the rows of A. By Exercise 3.3.57, these relations are exactly the linear combinations of the last $m - r$ rows of G. So the transposes of these $m - r$ rows span $\ker A^{\mathsf{T}}$, a space of dimension $m - r$ by (a), and they form a basis by Corollary 4.1.11(b). □

Subspace	Dimension	Basis, from $GA = A_r$
$\operatorname{im} A \subseteq F^m$	r	the pivot columns of A
$\ker A \subseteq F^n$	$n - r$	the basic solutions
$\operatorname{im} A^{\mathsf{T}} \subseteq F^n$	r	the nonzero rows of A_r
$\ker A^{\mathsf{T}} \subseteq F^m$	$m - r$	the last $m - r$ rows of G

Table 4.3.1. The four fundamental subspaces of $A \in F^{m\times n}$ of rank r (Proposition 4.3.13). Rows are written as columns. The two subspaces of F^n have dimensions adding to n, and the two subspaces of F^m have dimensions adding to m.

Example 4.3.14 (The four subspaces of a 3×4 matrix). Take the matrix of Intuition (Why this is surprising),

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 2 & 4 & 0 & 6 \\ 1 & 2 & -1 & 2 \end{pmatrix}, \quad A_r = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here $m = 3$, $n = 4$, $r = 2$, and the pivot columns are 1 and 3.

- *Column space.* Columns 1 and 3 of A: $\operatorname{im} A = \operatorname{span}((1,2,1),(1,0,-1))$, of dimension 2.
- *Null space.* The free columns are 2 and 4, with basic solutions $(-2,1,0,0)$ and $(-3,0,-1,1)$, so $\dim\ker A = 2$.
- *Row space.* The nonzero rows of A_r: $\operatorname{im} A^{\mathsf{T}} = \operatorname{span}((1,2,0,3),(0,0,1,1))$, of dimension 2.
- *Left null space.* The relation $\text{row}_1 - \text{row}_2 + \text{row}_3 = \mathbf{0}$ gives $(1,-1,1) \in \ker A^{\mathsf{T}}$, and $\dim\ker A^{\mathsf{T}} = 3 - 2 = 1$, so $\ker A^{\mathsf{T}} = \operatorname{span}((1,-1,1))$.

The dimensions are $2 + 2 = 4 = n$ in F^n and $2 + 1 = 3 = m$ in F^m.

In this example something more is true. Every vector of the row space has dot product 0 with every vector of the null space: for instance $(1,2,0,3)\cdot(-3,0,-1,1) = -3 + 3 = 0$. Likewise $(1,2,1)\cdot(1,-1,1) = 0 = (1,0,-1)\cdot(1,-1,1)$. This always happens. Writing $\mathbf{u}\cdot\mathbf{x} = u_1x_1 + \cdots + u_nx_n$, entry i of $A\mathbf{x}$ is (row i) $\cdot\,\mathbf{x}$, so $\mathbf{x} \in \ker A$ exactly when every vector of the row space has dot product 0 with $\mathbf{x}$ (Exercise 3.3.50(a)); applied to A^{T}, the same holds for the column space and the left null space. Over $\mathbb{R}$, the row space and the null space are then *orthogonal complements* in $\mathbb{R}^n$: they meet only in $\mathbf{0}$, and together they span $\mathbb{R}^n$. So $\mathbb{R}^n$ splits into two perpendicular pieces, of dimensions r and $n - r$, and $\mathbb{R}^m$ into two perpendicular pieces of dimensions r and $m - r$. This picture is behind least squares (Section 6.5), and it is proved in Section 6.6 (Theorem 6.6.21). Over some other fields, such as $\mathbb{F}_2$ and $\mathbb{C}$, the dimensions still add up, but the two pieces can overlap.

Non-example 4.3.15 (Row space and null space can overlap). Over $\mathbb{F}_2$, let $A = \begin{pmatrix} 1 & 1 \end{pmatrix}$. Its row space is $\operatorname{span}((1,1))$, and $(1,1)$ lies in $\ker A$, since $1 + 1 = 0$. So $\operatorname{im} A^{\mathsf{T}} = \ker A$: the two subspaces of $\mathbb{F}_2^2$ coincide, although their dimensions $1 + 1 = 2$ add up correctly.

Over $\mathbb{C}$, let $A = \begin{pmatrix} 1 & i \end{pmatrix}$. Then $(1,i) \in \operatorname{im} A^{\mathsf{T}}$, and $A(1,i) = 1 + i^2 = 0$, so $(1,i) \in \ker A$ as well. Over $\mathbb{C}$ the cure is to replace A^{T} by the conjugate transpose A^* of Definition 6.3.26: the column space of A^* meets $\ker A$ only in $\mathbf{0}$ (Section 6.6). Over $\mathbb{F}_2$ no such repair exists: for every symmetric bilinear pairing on $\mathbb{F}_2^n$ with $n \geq 2$, some nonzero vector pairs to 0 with itself (Section 12.4).

4.3.5 Rank-one matrices and sums

The simplest nonzero matrices have rank one. In

$$\begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \end{pmatrix},$$

every column is a multiple of $(1,2)$ and every row is a multiple of $(2,-1)$, and the matrix is the product of these two vectors.

Proposition 4.3.16 (Rank-one matrices). Let $A \in F^{m\times n}$. Then $\operatorname{rank} A = 1$ if and only if $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ for some nonzero $\mathbf{u} \in F^m$ and $\mathbf{v} \in F^n$. In that case $\operatorname{im} A = \operatorname{span}(\mathbf{u})$ and $\operatorname{im} A^{\mathsf{T}} = \operatorname{span}(\mathbf{v})$, and if also $A = \mathbf{u}'\mathbf{v}'^{\mathsf{T}}$, then $\mathbf{u}' = c\mathbf{u}$ and $\mathbf{v}' = c^{-1}\mathbf{v}$ for some nonzero $c \in F$.

Proof. ROUTINE

Uses. Proposition 2.4.8 (Four views of the product), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 1.2.35 (Zero products), Theorem 4.2.7 (The image of T_A), Proposition 2.4.23 (Transpose of a product)
Hint 1. Column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u}$. What does this say about $\operatorname{im}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right)$?
Hint 2. Conversely, if $\operatorname{rank} A = 1$, then $\operatorname{im} A = \operatorname{span}(\mathbf{u})$ for some $\mathbf{u} \neq \mathbf{0}$. Write each column of A as a multiple of $\mathbf{u}$.
Hint 3. For uniqueness, $\mathbf{u}'$ lies in $\operatorname{im} A = \operatorname{span}(\mathbf{u})$. Then compare $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}'\mathbf{v}'^{\mathsf{T}}$ in a row where $\mathbf{u}$ is nonzero.
Sketch. Column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u}$, so $\operatorname{im}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \operatorname{span}(\mathbf{u})$ when $\mathbf{v} \neq \mathbf{0}$, a line. If $\operatorname{rank} A = 1$, then $\operatorname{im} A = \operatorname{span}(\mathbf{u})$, so column j is $v_j\mathbf{u}$ for some v_j, and $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$. Transposing gives the row space. If $\mathbf{u}\mathbf{v}^{\mathsf{T}} = \mathbf{u}'\mathbf{v}'^{\mathsf{T}}$, then $\operatorname{span}(\mathbf{u}') = \operatorname{span}(\mathbf{u})$, so $\mathbf{u}' = c\mathbf{u}$, and row i with $u_i \neq 0$ gives $\mathbf{v} = c\mathbf{v}'$.

If. Let $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$. By the column view (Proposition 2.4.8(b)), column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u}$. So every column lies in $\operatorname{span}(\mathbf{u})$, and $\operatorname{im}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) \subseteq \operatorname{span}(\mathbf{u})$ by Proposition 2.6.18. Some v_j is nonzero, and then $v_j\mathbf{u} \neq \mathbf{0}$ (Proposition 1.2.35), so $\operatorname{im}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right)$ is a nonzero subspace of the line $\operatorname{span}(\mathbf{u})$, hence equal to it. So $\operatorname{rank}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = 1$ (Theorem 4.2.7) and $\operatorname{im} A = \operatorname{span}(\mathbf{u})$.

Only if. Let $\operatorname{rank} A = 1$. Then $\operatorname{im} A$ is one-dimensional, so $\operatorname{im} A = \operatorname{span}(\mathbf{u})$ for any nonzero $\mathbf{u} \in \operatorname{im} A$. Each column $\mathbf{a}_j$ lies in $\operatorname{im} A$, so $\mathbf{a}_j = v_j\mathbf{u}$ for some $v_j \in F$. With $\mathbf{v} = (v_1, \ldots, v_n)$, column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u} = \mathbf{a}_j$, so $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$. And $\mathbf{v} \neq \mathbf{0}$, since $A \neq 0$.

The row space and uniqueness. By Proposition 2.4.23, $A^{\mathsf{T}} = \mathbf{v}\mathbf{u}^{\mathsf{T}}$, so $\operatorname{im} A^{\mathsf{T}} = \operatorname{span}(\mathbf{v})$ by the first paragraph. If also $A = \mathbf{u}'\mathbf{v}'^{\mathsf{T}}$ with $\mathbf{u}', \mathbf{v}'$ nonzero, then $\operatorname{span}(\mathbf{u}') = \operatorname{im} A = \operatorname{span}(\mathbf{u})$, so $\mathbf{u}' = c\mathbf{u}$ with $c \neq 0$. Then $\mathbf{u}(\mathbf{v} - c\mathbf{v}')^{\mathsf{T}} = \mathbf{u}\mathbf{v}^{\mathsf{T}} - \mathbf{u}'\mathbf{v}'^{\mathsf{T}} = 0$. Choose i with $u_i \neq 0$. By the row view (Proposition 2.4.8(c)), row i of the left side is $u_i(\mathbf{v} - c\mathbf{v}')^{\mathsf{T}}$, which is therefore zero, so $\mathbf{v} = c\mathbf{v}'$, that is, $\mathbf{v}' = c^{-1}\mathbf{v}$. □

The rank normal form gives the same factorization in one line: if $GAH = N_1 = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$, then $A = \left(G^{-1}\mathbf{e}_1\right)\left(\mathbf{e}_1^{\mathsf{T}}H^{-1}\right)$, a column of G^{-1} times a row of H^{-1}. The form $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is worth reaching for whenever rank one appears.

Technique 4.3.17 (Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$). A matrix of rank one is $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ (Proposition 4.3.16). In this form products collapse to scalars, because $\mathbf{v}^{\mathsf{T}}\mathbf{w}$ is a 1×1 matrix and can be moved out of the way:

$$\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right)\mathbf{x} = \left(\mathbf{v}^{\mathsf{T}}\mathbf{x}\right)\mathbf{u},$$
$$\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right)\left(\mathbf{w}\mathbf{z}^{\mathsf{T}}\right) = \left(\mathbf{v}^{\mathsf{T}}\mathbf{w}\right)\mathbf{u}\mathbf{z}^{\mathsf{T}},$$
$$\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right)^2 = \left(\mathbf{v}^{\mathsf{T}}\mathbf{u}\right)\mathbf{u}\mathbf{v}^{\mathsf{T}}.$$

Reach for it when a hypothesis says that a matrix has rank one, when a matrix differs from a known one by a rank-one matrix (as $I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ or $A + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ do), and when a matrix is to be split into simple pieces. The canonical example is Proposition 4.3.16, and Example 4.3.18 shows the collapse at work. Later instances are the matrix determinant lemma for $\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)$ (Theorem 5.5.27) and the singular value decomposition, which writes every matrix as a sum of rank-one pieces in the best possible way (Section 9.1).
Used in. Exercise 4.3.19, Corollary 4.3.24, Exercise 4.3.37, Exercise 4.3.38, Exercise 4.3.40, Exercise 4.3.41, Exercise 4.5.51, Exercise 7.4.40, Problem 7.7, Problem 7.13, Exercise 4.3.39, Exercise 8.1.26.

Example 4.3.18 (Powers of a rank-one matrix). Let $A = \mathbf{u}\mathbf{v}^{\mathsf{T}} \in F^{n\times n}$ have rank one. Its diagonal entries are $u_i v_i$, so $\mathbf{v}^{\mathsf{T}}\mathbf{u} = u_1v_1 + \cdots + u_nv_n = \operatorname{tr} A$. By Technique 4.3.17,

$$A^2 = \mathbf{u}\big(\mathbf{v}^{\mathsf{T}}\mathbf{u}\big)\mathbf{v}^{\mathsf{T}} = (\operatorname{tr} A)A,$$

and by induction $A^k = (\operatorname{tr} A)^{k-1}A$ for $k \geq 1$. So a rank-one matrix is nilpotent if and only if $\operatorname{tr} A = 0$, in which case already $A^2 = 0$; and $A^2 = A$ if and only if $\operatorname{tr} A = 1$, since $A \neq 0$. For instance $\begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \end{pmatrix}$ has trace 0, and its square is 0; while $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \mathbf{e}_1\begin{pmatrix} 1 & 1 \end{pmatrix}$ has trace 1 and equals its square.

Exercise 4.3.19. ROUTINE Let $A = \begin{pmatrix} 3 & 1 & 1 \\ 6 & 2 & 2 \\ -3 & -1 & -1 \end{pmatrix}$.

(a) Over $\mathbb{Q}$, show that $\operatorname{rank} A = 1$ by writing $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$, and compute A^{10}.

(b) Over $\mathbb{F}_2$, compute A^2.

Solution on page 1309

For sums, the image of $A + B$ lies inside $\operatorname{im} A + \operatorname{im} B$, which gives an upper bound but no exact formula.

Proposition 4.3.20 (Rank is subadditive). For $A, B \in F^{m\times n}$,

$$\operatorname{rank}(A + B) \leq \operatorname{rank} A + \operatorname{rank} B.$$

Proof (comparing images). ROUTINE

Uses. Proposition 1.4.24 (The sum of two subspaces), Proposition 4.1.21 (Dimensions of subspaces), Corollary 4.1.11 (Spanning lists are long)
Hint 1. $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$. Which subspace contains $\operatorname{im}(A + B)$?
Hint 2. Concatenate a basis of $\operatorname{im} A$ and a basis of $\operatorname{im} B$.
Sketch. Every $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$ lies in $\operatorname{im} A + \operatorname{im} B$, which is spanned by a basis of $\operatorname{im} A$ followed by a basis of $\operatorname{im} B$, a list of length $\operatorname{rank} A + \operatorname{rank} B$.

For every $\mathbf{x} \in F^n$, $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} \in \operatorname{im} A + \operatorname{im} B$, so $\operatorname{im}(A + B) \subseteq \operatorname{im} A + \operatorname{im} B$, a subspace of F^m (Proposition 1.4.24). Let $\mathbf{a}_1, \ldots, \mathbf{a}_k$ and $\mathbf{b}_1, \ldots, \mathbf{b}_l$ be bases of $\operatorname{im} A$ and $\operatorname{im} B$. Every vector of $\operatorname{im} A + \operatorname{im} B$ is a sum of a combination of the $\mathbf{a}_i$ and a combination of the $\mathbf{b}_j$, so the list $\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{b}_1, \ldots, \mathbf{b}_l$ spans $\operatorname{im} A + \operatorname{im} B$, and $\dim(\operatorname{im} A + \operatorname{im} B) \leq k + l$ (Corollary 4.1.11(a)). Hence

$$\begin{aligned} \operatorname{rank}(A + B) &\leq \dim(\operatorname{im} A + \operatorname{im} B) \\ &\leq k + l = \operatorname{rank} A + \operatorname{rank} B \end{aligned}$$

by Proposition 4.1.21(a). □

Proof (comparing kernels). ROUTINE

Uses. Proposition 4.2.26 (Subspaces that must meet), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $\ker A \cap \ker B \subseteq \ker(A + B)$.
Hint 2. How large must $\ker A \cap \ker B$ be? Use Proposition 4.2.26 in F^n.
Sketch. $\ker A \cap \ker B \subseteq \ker(A + B)$, and by Proposition 4.2.26, $\dim(\ker A \cap \ker B) \geq (n - \operatorname{rank} A) + (n - \operatorname{rank} B) - n$. So $\dim \ker(A + B) \geq n - \operatorname{rank} A - \operatorname{rank} B$, and rank–nullity gives the claim.

If $A\mathbf{x} = B\mathbf{x} = \mathbf{0}$, then $(A + B)\mathbf{x} = \mathbf{0}$, so $\ker A \cap \ker B \subseteq \ker(A + B)$. By Proposition 4.1.21(a), Proposition 4.2.26 in F^n, and Theorem 4.2.11(a),

$$\begin{aligned} \dim \ker(A + B) &\geq \dim(\ker A \cap \ker B) \\ &\geq (n - \operatorname{rank} A) + (n - \operatorname{rank} B) - n \\ &= n - \operatorname{rank} A - \operatorname{rank} B. \end{aligned}$$

By Theorem 4.2.11(a) again, $\operatorname{rank}(A+B) = n - \dim\ker(A+B) \le \operatorname{rank} A + \operatorname{rank} B$. □

As for products, the image proof works in the target and the kernel proof in the domain. On the kernel side, the inequality is the statement that two large subspaces of F^n must overlap (Proposition 4.2.26).

By induction on k, a sum of k matrices of rank one has rank at most k. The next subsection shows that this bound is attained: every matrix of rank r is a sum of r matrices of rank one.

4.3.6 The column–row factorization

The pivot columns of A form a basis of $\operatorname{im} A$ (Theorem 4.2.7), so every column of A is a combination of them, and the coefficients can be read off $\operatorname{rref}(A)$ (Lemma 3.3.16(d), Corollary 3.2.17). Recording all these combinations at once is a matrix product.

Theorem 4.3.21 (Column–row factorization). Let $A \in F^{m\times n}$ have rank $r \ge 1$ and pivot columns $p_1 < \cdots < p_r$. Let $C \in F^{m\times r}$ be the matrix whose columns are the pivot columns $\mathbf{a}_{p_1}, \ldots, \mathbf{a}_{p_r}$ of A, and let $R \in F^{r\times n}$ consist of the first r rows of $\operatorname{rref}(A)$, which are its nonzero rows. Then

$$A = CR.$$

Moreover, the columns of C form a basis of $\operatorname{im} A$, and the rows of R, written as columns, form a basis of $\operatorname{im} A^{\mathsf{T}}$. In particular $\operatorname{rank} C = \operatorname{rank} R = r$.

Proof. ROUTINE

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.8 (Properties of the output), Theorem 2.4.34 (Block multiplication), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.4.1 (Matrix product), Theorem 4.2.7 (The image of T_A), Theorem 4.3.8 (Row rank equals column rank), Theorem 3.3.12 (Matrices in RREF)
Hint 1. Write $A = G^{-1}A_r$, where $A_r = \begin{pmatrix} R \\ 0 \end{pmatrix}$ in block form.
Hint 2. Split G^{-1} into its first r columns and the rest, and multiply in blocks. Then identify column s of G^{-1} by looking at column p_s of $A = G^{-1}A_r$.
Sketch. $GA = A_r = \begin{pmatrix} R \\ 0 \end{pmatrix}$. With $G^{-1} = (C' \;\; C'')$ split after r columns, block multiplication gives $A = C'R$. Column s of C' is $G^{-1}\mathbf{e}_s$, and since column p_s of A_r is $\mathbf{e}_s$, column p_s of A is also $G^{-1}\mathbf{e}_s$; so $C' = C$. The basis statements are Theorem 4.2.7 and the second proof of Theorem 4.3.8.

Let $G \in F^{m\times m}$ be invertible with $GA = A_r$ (Theorem 3.3.10). By Proposition 3.3.8(a), rows $r+1, \ldots, m$ of A_r are zero, so in block form $A_r = \begin{pmatrix} R \\ 0 \end{pmatrix}$, with a zero block of $m-r$ rows. Split $G^{-1} = (C' \;\; C'')$, where $C' \in F^{m\times r}$ consists of the first r columns of G^{-1}. By Theorem 2.4.34,

$$A = G^{-1}A_r = (C' \;\; C'')\begin{pmatrix} R \\ 0 \end{pmatrix} = C'R + C''0 = C'R.$$

Column s of C' is $G^{-1}\mathbf{e}_s$. Column p_s of A_r is $\mathbf{e}_s$ (Lemma 3.3.16(b)), so column p_s of $A = G^{-1}A_r$ is $G^{-1}\mathbf{e}_s$ (Definition 2.4.1). Hence column s of C' is $\mathbf{a}_{p_s}$, that is, $C' = C$, and $A = CR$.

The columns of C form a basis of $\operatorname{im} A$ by Theorem 4.2.7, so $\operatorname{rank} C = r$. The rows of R are the nonzero rows of A_r, and the second proof of Theorem 4.3.8 showed that, written as columns, they form a basis of $\operatorname{im} A^{\mathsf{T}}$. Finally, R satisfies (R1)–(R4) with r pivots, so it is in RREF and $\operatorname{rank} R = r$ (Theorem 3.3.12). □

Read column by column (Proposition 2.4.8(b)), $A = CR$ says that column j of A is the combination of the pivot columns of A whose coefficients are column j of $\operatorname{rref}(A)$:

$$\mathbf{a}_j = R_{1j}\mathbf{a}_{p_1} + \cdots + R_{rj}\mathbf{a}_{p_r}.$$

This is Corollary 3.2.17 once more: the free columns of $\operatorname{rref}(A)$ record how the free columns of A are built from its pivot columns.

Example 4.3.22 (A column–row factorization). For the matrix A of Example 4.3.14, the pivot columns are 1 and 3, so $A = CR$ with

$$C = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Column 4 of R is $(3, 1)$, and indeed $\mathbf{a}_4 = 3\mathbf{a}_1 + \mathbf{a}_3$; column 2 of R is $(2, 0)$, and $\mathbf{a}_2 = 2\mathbf{a}_1$. By the column–row view of the product (Proposition 2.4.8(d)), A is a sum of two matrices of rank one:

$$\begin{aligned} A &= \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 0 & 6 \\ 1 & 2 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{pmatrix}. \end{aligned}$$

$$\begin{pmatrix} 1 & 2 & 1 & 4 \\ 2 & 4 & 0 & 6 \\ 1 & 2 & -1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The reduction of Example 4.3.22. At the end, the pivot columns listed in the summary pick out the columns of A that form C, and the nonzero rows of the final matrix form R. Use "Edit" to try your own matrix: the number of pivots is both the number of columns of C and the number of rows of R. (Interactive in the web version.)

The column–row factorization is one of many factorizations of A through F^r, and the next proposition says that all of them carry the same information.

Proposition 4.3.23 (Factorizations through F^k). Let $A = XY$ with $X \in F^{m\times k}$ and $Y \in F^{k\times n}$.

(a) $\operatorname{rank} A \le k$.

(b) If $k = \operatorname{rank} A$, then the columns of X form a basis of $\operatorname{im} A$, and the rows of Y, written as columns, form a basis of $\operatorname{im} A^{\mathsf{T}}$.

Proof. ROUTINE

Uses. Theorem 4.3.3 (Rank of a product), Proposition 3.5.2 (First properties of rank), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 4.1.21 (Dimensions of subspaces), Corollary 4.1.11 (Spanning lists are long), Proposition 2.4.23 (Transpose of a product), Theorem 4.3.8 (Row rank equals column rank)
Techniques. Dimension counting (Technique 4.2.24), Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. X has only k columns.
Hint 2. For (b), $\operatorname{im} A \subseteq \operatorname{im} X$, and $\operatorname{im} X$ is spanned by k vectors. Compare dimensions. For Y, transpose.
Sketch. (a) $\operatorname{rank} A \le \operatorname{rank} X \le k$. (b) $\operatorname{im} A \subseteq \operatorname{im} X$, and $\operatorname{im} X$ is spanned by the k columns of X, so $\dim \operatorname{im} X \le k = \dim \operatorname{im} A$. Hence $\operatorname{im} A = \operatorname{im} X$, and k vectors spanning a k-dimensional space form a basis. Apply the same argument to $A^{\mathsf{T}} = Y^{\mathsf{T}}X^{\mathsf{T}}$, of rank k by Theorem 4.3.8.

(a) By Theorem 4.3.3 and Proposition 3.5.2(a), $\operatorname{rank} A = \operatorname{rank}(XY) \le \operatorname{rank} X \le k$.

(b) Let $k = \operatorname{rank} A$. By Theorem 4.3.3, $\operatorname{im} A \subseteq \operatorname{im} X$, and $\operatorname{im} X$ is spanned by the k columns of X (Proposition 2.6.18), so $\dim \operatorname{im} X \le k = \dim \operatorname{im} A$ (Corollary 4.1.11(a)). Also $\dim \operatorname{im} A \le \dim \operatorname{im} X$ (Proposition 4.1.21(a)), so the dimensions agree and $\operatorname{im} A = \operatorname{im} X$ by Proposition 4.1.21(b). This is a space of dimension k spanned by the k columns of X; so these columns form a basis of $\operatorname{im} A$ (Corollary 4.1.11(b)). Next, $A^{\mathsf{T}} = Y^{\mathsf{T}}X^{\mathsf{T}}$ (Proposition 2.4.23), with $Y^{\mathsf{T}} \in F^{n\times k}$, and $\operatorname{rank} A^{\mathsf{T}} = k$ (Theorem 4.3.8). The argument just given, applied to A^{T}, shows that the columns of Y^{T}, which are the rows of Y written as columns, form a basis of $\operatorname{im} A^{\mathsf{T}}$. □

So in any factorization $A = XY$ with inner size $\operatorname{rank} A$, the left factor is a basis of the column space and the right factor a basis of the row space. Two such factorizations differ by an invertible $r \times r$ matrix (Exercise 4.3.44), and the column–row factorization is the one in which the right factor is in RREF. Every matrix of rank r is therefore a sum of r rank-one matrices, and Proposition 4.3.20 says that fewer will not do.

Corollary 4.3.24 (Rank as a number of rank-one pieces). Let $A \in F^{m\times n}$. Then $\operatorname{rank} A$ is the smallest $k \geq 0$ such that A is a sum of k matrices of rank one (the empty sum being 0).

Proof. ROUTINE

Uses. Theorem 4.3.21 (Column–row factorization), Proposition 2.4.8 (Four views of the product), Proposition 4.3.16 (Rank-one matrices), Proposition 4.3.20 (Rank is subadditive)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. For a decomposition with rank A terms, write $A = CR$ and use the column–row view of the product.
Hint 2. For the lower bound, use Proposition 4.3.20 and induction on the number of terms.
Sketch. If $r = \operatorname{rank} A \geq 1$, then $A = CR = \sum_{s=1}^{r} \mathbf{c}_s R_{s,\cdot}$, where $\mathbf{c}_s$ is column s of C and $R_{s,\cdot}$ is row s of R; both are nonzero, so each term has rank one. Conversely, a sum of k rank-one matrices has rank at most k by subadditivity.

Let $r = \operatorname{rank} A$. If $r = 0$, then $A = 0$ is the empty sum. If $r \geq 1$, write $A = CR$ as in Theorem 4.3.21. By the column–row view (Proposition 2.4.8(d)),

$$A = \mathbf{c}_1 R_{1,\cdot} + \cdots + \mathbf{c}_r R_{r,\cdot},$$

where $\mathbf{c}_s$ is column s of C and $R_{s,\cdot}$ is row s of R. Each $\mathbf{c}_s$ is nonzero, as a vector of a basis, and each row of R is nonzero, as it contains a pivot. So each term has rank one by Proposition 4.3.16, and A is a sum of r matrices of rank one.

Conversely, if $A = A_1 + \cdots + A_k$ with $\operatorname{rank} A_i = 1$ for each i, then by Proposition 4.3.20 and induction on k, $\operatorname{rank} A \leq \operatorname{rank} A_1 + \cdots + \operatorname{rank} A_k = k$. So no sum of fewer than r rank-one matrices equals A. □

The lower bound can also be seen from Proposition 4.3.23(a): a sum $\sum_{i=1}^{k} \mathbf{u}_i \mathbf{v}_i^{\mathsf{T}}$ is the product of the $m \times k$ matrix with columns $\mathbf{u}_i$ and the $k \times n$ matrix with rows $\mathbf{v}_i^{\mathsf{T}}$, so its rank is at most k. Storing C and R takes $r(m+n)$ numbers instead of the mn entries of A, which is far fewer when r is small. Low-rank approximation (Section 16.3) exploits this to compress data.

The factorization also gives a third proof of Theorem 4.3.8, the one in LADR. It never row reduces, and its idea is to prove an inequality for every matrix and then apply it to the transpose.

Proof of Theorem 4.3.8 (via a column–row factorization). ROUTINE

Uses. Theorem 4.2.7 (The image of T_A), Proposition 2.4.8 (Four views of the product), Proposition 2.4.23 (Transpose of a product), Theorem 4.3.3 (Rank of a product), Proposition 3.5.2 (First properties of rank), Proposition 2.4.22 (Transpose is a linear involution)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. It suffices to prove $\operatorname{rank} A^{\mathsf{T}} \leq \operatorname{rank} A$ for every matrix A. Why?
Hint 2. Write $A = CR$ with C consisting of a basis of $\operatorname{im} A$, so that R has only $r = \operatorname{rank} A$ rows. What does $A^{\mathsf{T}} = R^{\mathsf{T}} C^{\mathsf{T}}$ say about $\operatorname{rank} A^{\mathsf{T}}$?
Sketch. Let $r = \operatorname{rank} A$ and let C have as columns a basis of $\operatorname{im} A$. Every column of A is a combination of these, so $A = CR$ with $R \in F^{r\times n}$. Then $A^{\mathsf{T}} = R^{\mathsf{T}} C^{\mathsf{T}}$, and $\operatorname{rank} A^{\mathsf{T}} \leq \operatorname{rank} R^{\mathsf{T}} \leq r$, since R^{T} has r columns. So $\operatorname{rank} A^{\mathsf{T}} \leq \operatorname{rank} A$ for every A; applied to A^{T}, it gives the reverse inequality.

We first show that $\operatorname{rank} A^{\mathsf{T}} \leq \operatorname{rank} A$ for every $A \in F^{m\times n}$. Let $r = \operatorname{rank} A$. If $r = 0$, then $A = 0$, so $A^{\mathsf{T}} = 0$ and both ranks are 0. Otherwise let $\mathbf{c}_1, \ldots, \mathbf{c}_r$ be a basis of $\operatorname{im} A$ (for instance the pivot columns, Theorem 4.2.7), and let $C \in F^{m\times r}$ have these columns. Each column $\mathbf{a}_j$ of A lies in $\operatorname{im} A$, so $\mathbf{a}_j = x_{1j}\mathbf{c}_1 + \cdots + x_{rj}\mathbf{c}_r$ for some scalars x_{sj}. Let $R = (x_{sj}) \in F^{r\times n}$. By the column view (Proposition 2.4.8(b)), column j of CR is $\sum_s x_{sj}\mathbf{c}_s = \mathbf{a}_j$, so $A = CR$. Transposing, $A^{\mathsf{T}} = R^{\mathsf{T}} C^{\mathsf{T}}$ (Proposition 2.4.23), and by Theorem 4.3.3 and Proposition 3.5.2(a),

$$\operatorname{rank} A^{\mathsf{T}} \leq \operatorname{rank} R^{\mathsf{T}} \leq r = \operatorname{rank} A,$$

since $R^{\mathsf{T}} \in F^{n\times r}$ has only r columns.

Applying this to A^{T}, whose transpose is A (Proposition 2.4.22), gives $\operatorname{rank} A \leq \operatorname{rank} A^{\mathsf{T}}$. Hence $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$. □

Intuition (Comparing the proofs that row rank equals column rank). The normal-form proof reduces A to the simplest matrix of its rank and observes that transposing does not disturb it; the work was done in proving Theorem 3.5.27 and Proposition 4.3.1. The row-operation proof is the computational one: it finds a basis of the row space, and it explains the theorem through what row operations preserve. The factorization proof is the shortest, given Theorem 4.3.3, and it shows the mechanism most directly: every row of $A = CR$ is a combination of the r rows of R, so the rows cannot span more than r dimensions. Its last step, proving an inequality for all matrices and then applying it to the transpose to get the reverse inequality, is a pattern worth remembering.

Exercise 4.3.25. ROUTINE Find the column–row factorization of

$$A = \begin{pmatrix} 1 & i & 0 & 2 \\ i & -1 & 1 & 2i \\ 1+i & -1+i & 1 & 2+2i \end{pmatrix} \in \mathbb{C}^{3\times 4},$$

and write A as a sum of rank A matrices of rank one.

Solution on page 1310

Exercises

Exercise 4.3.26. ROUTINE Over $\mathbb{Q}$, let

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

(a) Find $\ker A$, and for $i = 1, 2$ find $\operatorname{im} B_i \cap \ker A$. Use (4.3.1) to find $\operatorname{rank}(AB_i)$, and check by computing AB_i.

(b) Is there a matrix $B \in \mathbb{Q}^{3\times 2}$ of rank 2 with $AB = 0$?

Solution on page 1310

Exercise 4.3.27. ROUTINE Let $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$.

(a) Over $\mathbb{F}_2$, find bases of the four fundamental subspaces of A, and check the dimensions against Proposition 4.3.13. Show that $\operatorname{im} A^{\mathsf{T}} = \ker A$.

(b) Over $\mathbb{Q}$, do the same, and show that $\operatorname{im} A^{\mathsf{T}} \cap \ker A = \{\mathbf{0}\}$.

Solution on page 1311

Exercise 4.3.28. ROUTINE For each item, find a matrix with the given properties, or show that none exists.

(a) $A \in \mathbb{R}^{3\times 3}$ with $\operatorname{im} A = \operatorname{span}((1,0,1))$ and $\ker A = \operatorname{span}((1,1,0),(0,1,1))$.

(b) $A \in F^{2\times 4}$ with $\ker A = \operatorname{span}(\mathbf{e}_1)$.

(c) $A \in F^{3\times 3}$ whose row space contains $(1,1,0)$ and whose null space contains $(1,0,0)$.

(d) $A \in F^{2\times 2}$ with $\operatorname{im} A = \ker A$.

(e) $A \in F^{3\times 3}$ with $\operatorname{im} A = \ker A$.

Solution on page 1311

Exercise 4.3.29. ROUTINE Let $A, B \in F^{m\times n}$. Prove or disprove.

(a) If A and B have the same four fundamental subspaces, then $A = B$.

(b) If A and B have the same row space, then $\ker A = \ker B$.

(c) If A and B have the same column space, then $\ker A^{\mathsf{T}} = \ker B^{\mathsf{T}}$.

(d) If $\ker A = \ker B$, then $\operatorname{im} A = \operatorname{im} B$.

Solution on page 1312

Exercise 4.3.30. STANDARD (*Sylvester's rank inequality.*) Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$.

(a) Show that $\operatorname{rank}(AB) \geq \operatorname{rank} A + \operatorname{rank} B - n$.

(b) Deduce that if $AB = 0$, then $\operatorname{rank} A + \operatorname{rank} B \leq n$.

(c) Show that for $A_1, \ldots, A_k \in F^{n\times n}$,

$$\operatorname{rank}(A_1 A_2 \cdots A_k) \geq \operatorname{rank} A_1 + \cdots + \operatorname{rank} A_k - (k-1)n.$$

(d) Deduce that if $A \in F^{n\times n}$ has rank $n-1$, then $\operatorname{rank}\left(A^k\right) \geq n-k$ for every $k \geq 1$, and show by an example that equality can hold for every $k \leq n$.

Solution on page 1312

Exercise 4.3.31. HARD (*Frobenius's rank inequality.*) Let $A \in F^{m\times n}$, $B \in F^{n\times p}$, and $C \in F^{p\times q}$.

(a) Show that

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leq \operatorname{rank} B + \operatorname{rank}(ABC).$$

(b) Deduce Sylvester's inequality (Exercise 4.3.30(a)).

(c) Let $A \in F^{n\times n}$ and $d_k = \operatorname{rank}\left(A^k\right) - \operatorname{rank}\left(A^{k+1}\right)$ for $k \geq 0$. Deduce from (a) that $d_0 \geq d_1 \geq d_2 \geq \cdots$ (compare Example 4.3.6).

Solution on page 1313

Exercise 4.3.32. HARD Let $0 \leq a \leq \min(m,n)$ and $0 \leq b \leq \min(n,p)$ be integers. Show that there are $A \in F^{m\times n}$ of rank a and $B \in F^{n\times p}$ of rank b with $\operatorname{rank}(AB) = r$ if and only if

$$\max(0, a+b-n) \leq r \leq \min(a,b).$$

For $m = n = p = 5$, $a = 3$, and $b = 4$, list the possible values of $\operatorname{rank}(AB)$ and give a pair A, B for each.

Solution on page 1313

Exercise 4.3.33. STANDARD Let $A \in F^{m\times n}$.

(a) Show that, over any field, $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A - \dim\left(\operatorname{im} A \cap \ker A^{\mathsf{T}}\right)$.

(b) Let $F = \mathbb{R}$. Show that $\operatorname{im} A \cap \ker A^{\mathsf{T}} = \{\mathbf{0}\}$, and deduce that $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$.

(c) Let $F = \mathbb{R}$. Without using Theorem 4.3.8, deduce from (b) that $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$ and $\operatorname{rank}\left(AA^{\mathsf{T}}\right) = \operatorname{rank} A$.

Solution on page 1314

Exercise 4.3.34. STANDARD Prove or disprove.

(a) Over $\mathbb{F}_2$, $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$ for every matrix A.

(b) Over $\mathbb{C}$, $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$ for every matrix A.

(c) Over $\mathbb{Q}$, $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$ for every matrix A.

(d) Over $\mathbb{F}_2$, if the columns of A are linearly independent, then $A^{\mathsf{T}}A$ is invertible.

Which step of Exercise 4.3.33(b) fails in the false cases?

Solution on page 1314

Exercise 4.3.35. STANDARD Find the error in the following argument. Is its conclusion true? Give a matrix for which the false step fails, and repair the argument.

Claim. For every $A \in F^{m\times n}$, $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$.

Proof. Let $A_r = \operatorname{rref}(A)$, with r pivots. Row operations do not change the column space, so $\operatorname{im} A = \operatorname{im} A_r$. The columns of A_r lie in $\operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_r)$ and include $\mathbf{e}_1, \ldots, \mathbf{e}_r$, so $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_r)$, and $\operatorname{rank} A = r$. Row operations do not change the row space either, and the r nonzero rows of A_r are linearly independent, so $\operatorname{rank} A^{\mathsf{T}} = r$ as well.

Solution on page 1315

Exercise 4.3.36. STANDARD Prove or disprove. All matrices are over a field F.

(a) $\operatorname{rank}(AB) = \operatorname{rank}(BA)$ for all $A, B \in F^{n\times n}$.

(b) If $A \in F^{m\times n}$ and $B \in F^{n\times p}$ both have rank n, then $\operatorname{rank}(AB) = n$.

(c) If $A, B \in F^{m\times n}$ and $\operatorname{im} A \cap \operatorname{im} B = \{\mathbf{0}\}$, then $\operatorname{rank}(A + B) = \operatorname{rank} A + \operatorname{rank} B$.

(d) If $A \in F^{n\times n}$ and $\operatorname{rank}\left(A^2\right) = \operatorname{rank} A$, then $\operatorname{rank}\left(A^k\right) = \operatorname{rank} A$ for every $k \geq 1$.

Solution on page 1315

Exercise 4.3.37. STANDARD Let $\mathbf{u}, \mathbf{v} \in F^n$, and let $s = \mathbf{v}^{\mathsf{T}}\mathbf{u}$.

(a) Show that if $1 + s \neq 0$, then $I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ is invertible, with

$$\left(I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)^{-1} = I_n - \frac{1}{1+s}\mathbf{u}\mathbf{v}^{\mathsf{T}}.$$

(b) Show that if $1 + s = 0$, then $I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ is not invertible, and that its rank is $n - 1$.

(c) Let $J \in F^{n\times n}$ have every entry equal to 1. For which $c \in F$ is $I_n + cJ$ invertible, and what is its inverse? Is $I_4 + J$ invertible over $\mathbb{Q}$? Over $\mathbb{F}_5$?

Solution on page 1316

Exercise 4.3.38. STANDARD Let $A, B \in F^{n\times n}$ and suppose that $\operatorname{rank}(AB - BA) = 1$. Show that $(AB - BA)^2 = 0$. Give an example in which $\operatorname{rank}(AB - BA) = 1$.

Solution on page 1317

Exercise 4.3.39. ROUTINE

(a) Show that $\operatorname{rank}(A+B) \geq |\operatorname{rank} A - \operatorname{rank} B|$ for $A, B \in F^{m\times n}$.

(b) Show that changing one entry of a matrix changes its rank by at most 1.

(c) Let $M_t \in \mathbb{R}^{5\times 5}$ have every diagonal entry equal to 1 and every other entry equal to t. Show that $\operatorname{rank} M_t \geq 4$ for every $t \neq 1$.

Solution on page 1317

Exercise 4.3.40. HARD Let $n \geq 3$, and let $A \in F^{n\times n}$ have entries $a_{ij} = (i-j)^2$, read as elements of F.

(a) Show that $\operatorname{rank} A \leq 3$ over every field, by writing A as a sum of three matrices of rank at most one.

(b) Show that $\operatorname{rank} A = 3$ if $1+1 \neq 0$ in F.

(c) Find $\operatorname{rank} A$ over $\mathbb{F}_2$.

Solution on page 1317

Exercise 4.3.41. STANDARD

(a) Let $A \in F^{m\times n}$. Show that $\operatorname{rank} A \leq 1$ if and only if $a_{ij}a_{kl} = a_{il}a_{kj}$ for all $i,k \in \{1,\ldots,m\}$ and $j,l \in \{1,\ldots,n\}$.

(b) Find all $t \in \mathbb{R}$ for which $\begin{pmatrix} 1 & t & 2 \\ t & 4 & 2t \end{pmatrix}$ has rank 1.

Solution on page 1318

Exercise 4.3.42. HARD Let $A \in F^{n\times n}$. Show that

$$\operatorname{rank} A + \operatorname{rank}(I_n - A) = n$$

if and only if $A^2 = A$.

Solution on page 1319

Exercise 4.3.43. HARD Let $A \in F^{n\times n}$ satisfy $A^2 = A$, and let $r = \operatorname{rank} A \geq 1$.

(a) Let $A = CR$ be a factorization with $C \in F^{n\times r}$ and $R \in F^{r\times n}$, as in Theorem 4.3.21. Show that $RC = I_r$.

(b) Deduce that $\operatorname{tr} A = r \cdot 1$, the sum of r copies of 1 in F.

(c) Deduce that over $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, the rank of an idempotent matrix equals its trace. Show by an example that over $\mathbb{F}_2$ the rank of an idempotent is not determined by its trace.

Solution on page 1319

Exercise 4.3.44. HARD Let $A \in F^{m\times n}$ have rank $r \geq 1$, and suppose $A = CR = C'R'$ with $C, C' \in F^{m\times r}$ and $R, R' \in F^{r\times n}$.

(a) Show that there is exactly one invertible $P \in F^{r\times r}$ with $C' = CP$, and that then $R' = P^{-1}R$.

(b) Deduce that the factorization of Theorem 4.3.21 is the only factorization $A = C'R'$ with inner size r in which R' is in RREF.

Solution on page 1319

Exercise 4.3.45. CHALLENGE Let $A \in F^{n\times n}$.

(a) Suppose $\operatorname{rank}(A^2) = \operatorname{rank} A = r \geq 1$, and let $A = CR$ be a factorization with inner size r. Show that $RC \in F^{r\times r}$ is invertible.

(b) Under the hypothesis of (a), show that $X = C(RC)^{-2}R$ satisfies

$$AXA = A, \quad XAX = X, \quad AX = XA.$$

(c) Conversely, show that if some $X \in F^{n\times n}$ satisfies the three equations in (b), then $\operatorname{rank}(A^2) = \operatorname{rank} A$.

Solution on page 1320

4.4 Sums and direct sums

In Section 1.4.5 we met the sum $U + W$ of two subspaces, the smallest subspace that contains both. This section answers two questions about it: how large it is, and when a vector of $U + W$ can be written as $\mathbf{u} + \mathbf{w}$ in only one way. Two pictures in $\mathbb{R}^3$ already suggest the answers. Every vector of $\mathbb{R}^3$ is, in exactly one way, a vector of the xy-plane plus a vector of the z-axis, and $2 + 1 = 3$. The xy-plane and the xz-plane also add up to $\mathbb{R}^3$, but now $(1, 0, 0) = (1, 0, 0) + \mathbf{0} = \mathbf{0} + (1, 0, 0)$ has two decompositions, and $2 + 2$ overshoots 3 by exactly 1, the dimension of the x-axis, where the two planes overlap. Both questions are answered by looking at the intersection.

Sums in which decompositions are unique, the direct sums, are to subspaces what bases are to vectors. A basis cuts V into lines, and the coordinates of a vector name its pieces; a direct sum cuts V into subspaces of any dimension. Later chapters understand an operator by cutting V into pieces that it maps into themselves. A diagonalizable operator is one for which V is the direct sum of its eigenspaces (Section 7.4), and the Jordan form starts from a direct sum of generalized eigenspaces (Section 10.2). Both need direct sums of more than two subspaces, and that is where the subtleties are.

Throughout, V is a vector space over a field F, and U, W, X, U_1, U_2, ... denote subspaces of V.

4.4.1 Sums of several subspaces

The sum of two subspaces extends to any finite number of them.

Definition 4.4.1 (Sum of subspaces). Let $U_1, \ldots, U_k$ be subspaces of V, where $k \geq 1$. Their **sum** is

$$U_1 + \cdots + U_k = \{\mathbf{u}_1 + \cdots + \mathbf{u}_k \mid \mathbf{u}_i \in U_i \text{ for } i = 1, \ldots, k\},$$

also written $\sum_{i=1}^{k} U_i$. An expression $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k$ with $\mathbf{u}_i \in U_i$ for each i is a **decomposition** of $\mathbf{v}$ along $U_1, \ldots, U_k$, with **components** $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

For $k = 2$ this is the sum $U_1 + U_2$ of Proposition 1.4.24, and for $k = 1$ it is U_1. By convention, a sum of no subspaces is $\{\mathbf{0}\}$, just as an empty sum of vectors is $\mathbf{0}$. The sum of k subspaces is the sum of two, applied repeatedly, so the two-subspace facts carry over.

Proposition 4.4.2 (Sums are the smallest containing subspaces). Let $U_1, \ldots, U_k$ be subspaces of V.

(a) If $k \geq 2$, then $U_1 + \cdots + U_k = (U_1 + \cdots + U_{k-1}) + U_k$.

(b) $U_1 + \cdots + U_k$ is a subspace of V that contains every U_i, and it is contained in every subspace of V that contains every U_i.

(c) If $U_i = \operatorname{span}(\mathcal{L}_i)$ for lists $\mathcal{L}_1, \ldots, \mathcal{L}_k$, then $U_1 + \cdots + U_k = \operatorname{span}(\mathcal{L}_1, \ldots, \mathcal{L}_k)$.

Proof. ROUTINE

Uses. Definition 4.4.1 (Sum of subspaces), Proposition 1.4.24 (The sum of two subspaces), Exercise 1.3.36
Techniques. Double containment (Technique 1.3.14), Induction on the number of summands (Technique 4.4.10)

Hint 1. For (a), an element of the left side is $(\mathbf{u}_1 + \cdots + \mathbf{u}_{k-1}) + \mathbf{u}_k$. Check both containments.
Hint 2. For (b) and (c), induct on k. By (a), the case of k subspaces is a sum of two, $S + U_k$ with $S = U_1 + \cdots + U_{k-1}$, where Proposition 1.4.24 and Exercise 1.3.36 apply.
Sketch. (a) is regrouping. By (a), the sum is $S + U_k$ with $S = U_1 + \cdots + U_{k-1}$, so (b) and (c) follow by induction on k from the two-subspace facts: $S + U_k$ is the smallest subspace containing S and U_k, and $\operatorname{span}(\mathcal{L}) + \operatorname{span}(\mathcal{M}) = \operatorname{span}(\mathcal{L}, \mathcal{M})$.

(a) Let $S = U_1 + \cdots + U_{k-1}$. An element of $U_1 + \cdots + U_k$ has the form $\mathbf{u}_1 + \cdots + \mathbf{u}_k = (\mathbf{u}_1 + \cdots + \mathbf{u}_{k-1}) + \mathbf{u}_k$ with $\mathbf{u}_i \in U_i$, and this lies in $S + U_k$. Conversely, an element of $S + U_k$ is $\mathbf{s} + \mathbf{u}_k$ with $\mathbf{s} = \mathbf{u}_1 + \cdots + \mathbf{u}_{k-1} \in S$ and $\mathbf{u}_k \in U_k$, which is an element of $U_1 + \cdots + U_k$.

(b) We induct on k. For $k = 1$ the sum is U_1, and there is nothing to prove. Let $k \geq 2$, and assume that $S = U_1 + \cdots + U_{k-1}$ is a subspace containing $U_1, \ldots, U_{k-1}$ and contained in every subspace that contains them. By (a) and Proposition 1.4.24(a), $U_1 + \cdots + U_k = S + U_k$ is a subspace containing S and U_k, hence every U_i. If X is a subspace containing every U_i, then $S \subseteq X$ by the inductive hypothesis and $U_k \subseteq X$, so $S + U_k \subseteq X$ by Proposition 1.4.24(b).

(c) Again we induct on k, and $k = 1$ is trivial. For $k \geq 2$, the inductive hypothesis gives $U_1 + \cdots + U_{k-1} = \operatorname{span}(\mathcal{L}_1, \ldots, \mathcal{L}_{k-1})$. So by (a) and Exercise 1.3.36, applied to the lists $\mathcal{L}_1, \ldots, \mathcal{L}_{k-1}$ and $\mathcal{L}_k$,

$$U_1 + \cdots + U_k = \operatorname{span}(\mathcal{L}_1, \ldots, \mathcal{L}_{k-1}) + \operatorname{span}(\mathcal{L}_k) = \operatorname{span}(\mathcal{L}_1, \ldots, \mathcal{L}_k). \qquad \square$$

By (b), the sum depends only on the subspaces involved, not on their order or grouping. For instance $U + W = W + U$ and $(U + W) + X = U + (W + X)$, since each side is the smallest subspace containing U, W (and X).

Example 4.4.3 (Sums).

(a) In F^3, $\operatorname{span}(\mathbf{e}_1) + \operatorname{span}(\mathbf{e}_2) + \operatorname{span}(\mathbf{e}_3) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = F^3$ by Proposition 4.4.2(c). In $\mathbb{R}^2$, the three lines $\operatorname{span}(\mathbf{e}_1)$, $\operatorname{span}(\mathbf{e}_2)$, and $\operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$ also add up to $\mathbb{R}^2$, although the third is not needed.

(b) In $F^{n\times n}$, let Y be the subspace of upper triangular matrices and Z the subspace of lower triangular matrices. Every matrix is the sum of its part on and above the diagonal, which lies in Y, and its part below the diagonal, which lies in Z. So $Y + Z = F^{n\times n}$. The two subspaces overlap in the diagonal matrices, so many matrices have more than one decomposition: $I_n = I_n + 0 = 0 + I_n$.

(c) For $S, T \in \mathrm{L}(V, W)$ we have $(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$, so $\operatorname{im}(S + T) \subseteq \operatorname{im} S + \operatorname{im} T$. Together with the bound $\dim(\operatorname{im} S + \operatorname{im} T) \leq \dim \operatorname{im} S + \dim \operatorname{im} T$ of Theorem 4.4.13 below, this is the reason that rank is subadditive (Proposition 4.3.20). The containment can be strict: for $T \neq 0$, $\operatorname{im}(T + (-T)) = \{\mathbf{0}\}$ is smaller than $\operatorname{im} T + \operatorname{im}(-T) = \operatorname{im} T$.

4.4.2 Direct sums

A basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V writes every vector in exactly one way as $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. Each term $a_i\mathbf{v}_i$ lies in the line $\operatorname{span}(\mathbf{v}_i)$, so the basis exhibits V as a sum of lines in which every vector has exactly one decomposition. Direct sums keep the uniqueness and allow pieces of any dimension.

Definition 4.4.4 (Direct sum). Let $U_1, \ldots, U_k$ be subspaces of V. The sum $U_1 + \cdots + U_k$ is **direct** if every vector of $U_1 + \cdots + U_k$ has exactly one decomposition along $U_1, \ldots, U_k$. In that case we also say that $U_1, \ldots, U_k$ are **independent**, and we write the sum as

$$U_1 \oplus \cdots \oplus U_k \quad \text{or} \quad \bigoplus_{i=1}^{k} U_i.$$

If moreover $U_1 + \cdots + U_k = V$, we write $V = U_1 \oplus \cdots \oplus U_k$ and say that V is the **direct sum** of $U_1, \ldots, U_k$.

So writing $U \oplus W$ is an assertion, not a new operation: it denotes the subspace $U + W$, together with the claim that the sum is direct. (The *external* direct sum, which builds a new space out of two unrelated ones, is the product of Section 4.6.) Some texts define independent subspaces by condition (II) of Theorem 4.4.7 below instead; the theorem shows that the two definitions agree.

Example 4.4.5 (Direct sums).

(a) *Bases.* If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of V, then $V = \operatorname{span}(\mathbf{v}_1) \oplus \cdots \oplus \operatorname{span}(\mathbf{v}_n)$. A decomposition of $\mathbf{v}$ has components $a_i\mathbf{v}_i$ with $a_i \in F$, and the component determines the scalar: if $a\mathbf{v}_i = b\mathbf{v}_i$, then $(a-b)\mathbf{v}_i = \mathbf{0}$, so $a = b$ by Proposition 1.2.35, since $\mathbf{v}_i \neq \mathbf{0}$. So decompositions of $\mathbf{v}$ correspond to representations $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, and there is exactly one. The components of $\mathbf{v}$ are its coordinates times the basis vectors.

(b) *Triangular parts.* Let Y be the upper triangular matrices in $F^{n\times n}$ and Z° the *strictly* lower triangular ones, whose diagonal entries are 0. Then $F^{n\times n} = Y \oplus Z^\circ$: in a decomposition $A = B + C$ with $B \in Y$ and $C \in Z^\circ$, the entries of A on and above the diagonal must be those of B and the entries below must be those of C.

(c) *Even and odd functions.* In $\mathbb{R}^{\mathbb{R}}$, the even functions E and the odd functions O satisfy $E \cap O = \{\mathbf{0}\}$ and $E + O = \mathbb{R}^{\mathbb{R}}$ (Exercise 1.4.35). By Theorem 4.4.7 below, this says $\mathbb{R}^{\mathbb{R}} = E \oplus O$.

(d) *Even and odd polynomials.* In $F[x]$, let E be the polynomials whose coefficients of $x, x^3, x^5, \ldots$ vanish and O those whose coefficients of $1, x^2, x^4, \ldots$ vanish. Every polynomial is the sum of its even-degree terms and its odd-degree terms, and in a decomposition $p = q + r$ with $q \in E$ and $r \in O$, each coefficient of p belongs to exactly one of q and r. So $F[x] = E \oplus O$ over every field, including $\mathbb{F}_2$. Compare Exercise 1.4.35(b), where even and odd *functions* on $\mathbb{F}_2$ coincide.

Non-example 4.4.6 (Sums that are not direct).

(a) The xy-plane and the xz-plane add up to $\mathbb{R}^3$, but $\mathbf{0} = (1,0,0) + (-1,0,0)$ is a second decomposition of $\mathbf{0}$, besides $\mathbf{0} + \mathbf{0}$.

(b) In $\mathbb{R}^2$, the three lines $\ell_1 = \operatorname{span}(\mathbf{e}_1)$, $\ell_2 = \operatorname{span}(\mathbf{e}_2)$, $\ell_3 = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$ meet pairwise only in $\mathbf{0}$. Their sum is still not direct:

$$\mathbf{0} = \mathbf{e}_1 + \mathbf{e}_2 + (-(\mathbf{e}_1 + \mathbf{e}_2))$$

is a decomposition of $\mathbf{0}$ with nonzero components.

Part (b) is the trap of this section. For three or more subspaces, "no two of them overlap" is not the right notion of independence. The analogy with lists explains why: three vectors of $\mathbb{R}^2$, no two of them parallel, are still dependent. Independence of a list compares each vector with the span of the *others* (Corollary 1.5.13), and the right condition for subspaces compares each subspace with the sum of the others.

> *Intuition (What the criteria say).* The theorem below gives three tests for directness. Condition (II) is uniqueness at $\mathbf{0}$; as for lists, uniqueness everywhere follows by subtracting two decompositions. Conditions (III) and (IV) are the subspace versions of "no vector is a combination of the others" and "no vector is a combination of the earlier ones". Condition (IV) is the economical one to check: it asks for $k-1$ intersections to be $\{\mathbf{0}\}$, each with a sum that is built up one subspace at a time.

Theorem 4.4.7 (Criteria for direct sums). Let $U_1, \ldots, U_k$ be subspaces of V. The following are equivalent.

(I) The sum $U_1 + \cdots + U_k$ is direct.

(II) $\mathbf{0}$ has only the trivial decomposition: if $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i \in U_i$ for each i, then $\mathbf{u}_1 = \cdots = \mathbf{u}_k = \mathbf{0}$.

(III) Each U_j meets the sum of the others only in $\mathbf{0}$: $U_j \cap \sum_{i \neq j} U_i = \{\mathbf{0}\}$ for $j = 1, \ldots, k$.

(IV) Each U_j meets the sum of the earlier ones only in $\mathbf{0}$: $U_j \cap \left(U_1 + \cdots + U_{j-1}\right) = \{\mathbf{0}\}$ for $j = 2, \ldots, k$.

In particular, for two subspaces, $U + W$ is direct if and only if $U \cap W = \{\mathbf{0}\}$.

Here $\sum_{i \neq j} U_i$ is the sum of all U_i except U_j. (For $k = 1$ it is the empty sum $\{\mathbf{0}\}$, and all four conditions hold.) The first proof is a cycle of implications.

Technique 4.4.8 (Compare components). In a direct sum, an equation between decompositions splits into one equation per summand. If $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{u}_1' + \cdots + \mathbf{u}_k'$ with $\mathbf{u}_i, \mathbf{u}_i' \in U_i$ and the sum is direct, then $\mathbf{u}_i = \mathbf{u}_i'$ for every i. This is the analogue of comparing coefficients with respect to a basis. To use it, sort every term of an equation into the summand it belongs to, then read off the components one at a time.

Proving that a sum is direct runs the same move backwards. Take two decompositions of one vector, subtract them, and move terms across so that the two sides of the resulting equation lie in different subspaces. Then both sides lie in the intersection, and if the intersection is $\{\mathbf{0}\}$, both sides vanish. For two subspaces this is the whole argument: $\mathbf{u} + \mathbf{w} = \mathbf{u}' + \mathbf{w}'$ gives $\mathbf{u} - \mathbf{u}' = \mathbf{w}' - \mathbf{w} \in U \cap W$.

Reach for it whenever a vector is given as a sum of pieces from known subspaces. The canonical example is Theorem 4.4.7. Later uses include orthogonal decompositions (Section 6.4), the independence of eigenvectors with distinct eigenvalues (Section 7.2), and the generalized eigenspace decomposition (Section 10.2).

Used in. Theorem 4.4.7, Proposition 4.4.9, Exercise 4.4.11, Proposition 4.4.12, Proposition 4.4.26, Theorem 4.4.27, Proposition 4.4.29, Proposition 4.4.31, Exercise 4.4.36, Exercise 4.4.44, Exercise 4.4.51, Problem 4.7.

Proof (by a cycle of implications). STANDARD

Uses. Definition 4.4.4 (Direct sum), Definition 4.4.1 (Sum of subspaces)
Techniques. Proving TFAE (Technique 2.7.1), Compare components (Technique 4.4.8), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Prove (I) ⇔ (II) first, by subtracting two decompositions. Then prove (II) ⇒ (III) ⇒ (IV) ⇒ (II).
Hint 2. For (II) ⇒ (III), a vector $\mathbf{v} \in U_j$ that also equals $\sum_{i \neq j} \mathbf{u}_i$ gives a decomposition of $\mathbf{0}$ whose jth component is $-\mathbf{v}$.
Hint 3. For (IV) ⇒ (II), take a decomposition $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$. If some $\mathbf{u}_i \neq \mathbf{0}$, look at the largest such index j and solve for $\mathbf{u}_j$.
Sketch. (I) ⇔ (II): $\mathbf{0} + \cdots + \mathbf{0}$ is one decomposition of $\mathbf{0}$, and the difference of two decompositions of $\mathbf{v}$ is a decomposition of $\mathbf{0}$. (II) ⇒ (III): if $\mathbf{v} \in U_j$ equals $\sum_{i \neq j} \mathbf{u}_i$, then $\sum_{i \neq j} \mathbf{u}_i + (-\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$. (III) ⇒ (IV): the sum of the earlier subspaces is contained in the sum of the others. (IV) ⇒ (II): if j is the largest index with $\mathbf{u}_j \neq \mathbf{0}$, then $\mathbf{u}_j = -\left(\mathbf{u}_1 + \cdots + \mathbf{u}_{j-1}\right) \in U_j \cap \left(U_1 + \cdots + U_{j-1}\right)$, a contradiction.

We prove (I) ⇔ (II), and then (II) ⇒ (III) ⇒ (IV) ⇒ (II).

(I) ⇒ (II). Since $\mathbf{0} \in U_i$ for each i, $\mathbf{0} = \mathbf{0} + \cdots + \mathbf{0}$ is a decomposition of $\mathbf{0}$. By (I) it is the only one.

(II) ⇒ (I). Suppose $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{u}_1' + \cdots + \mathbf{u}_k'$ with $\mathbf{u}_i, \mathbf{u}_i' \in U_i$. Then

$$\left(\mathbf{u}_1 - \mathbf{u}_1'\right) + \cdots + \left(\mathbf{u}_k - \mathbf{u}_k'\right) = \mathbf{v} - \mathbf{v} = \mathbf{0},$$

and $\mathbf{u}_i - \mathbf{u}_i' \in U_i$, since U_i is a subspace. By (II), $\mathbf{u}_i - \mathbf{u}_i' = \mathbf{0}$ for each i, so the two decompositions agree.

(II) ⇒ (III). Fix j, and let $\mathbf{v} \in U_j \cap \sum_{i \neq j} U_i$, say $\mathbf{v} = \sum_{i \neq j} \mathbf{u}_i$ with $\mathbf{u}_i \in U_i$. Then

$$\mathbf{u}_1 + \cdots + \mathbf{u}_{j-1} + (-\mathbf{v}) + \mathbf{u}_{j+1} + \cdots + \mathbf{u}_k = \mathbf{0}$$

is a decomposition of $\mathbf{0}$, as $-\mathbf{v} \in U_j$. By (II) all its components vanish. In particular $-\mathbf{v} = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$.

(III) ⇒ (IV). Let $2 \le j \le k$. Every element $\mathbf{u}_1 + \cdots + \mathbf{u}_{j-1}$ of $U_1 + \cdots + U_{j-1}$ is also an element of $\sum_{i \ne j} U_i$: take the components $\mathbf{0}$ in U_i for $i > j$. So

$$U_j \cap \big(U_1 + \cdots + U_{j-1}\big) \subseteq U_j \cap \sum_{i \ne j} U_i = \{\mathbf{0}\}.$$

(IV) ⇒ (II). Let $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i \in U_i$. Suppose, for a contradiction, that some $\mathbf{u}_i \ne \mathbf{0}$, and let j be the largest index with $\mathbf{u}_j \ne \mathbf{0}$. Then $\mathbf{u}_i = \mathbf{0}$ for $i > j$, so

$$\mathbf{u}_j = -\big(\mathbf{u}_1 + \cdots + \mathbf{u}_{j-1}\big) = (-\mathbf{u}_1) + \cdots + \big(-\mathbf{u}_{j-1}\big).$$

If $j = 1$, the right side is the empty sum $\mathbf{0}$, which contradicts $\mathbf{u}_1 \ne \mathbf{0}$. If $j \ge 2$, the left side lies in U_j and the right side in $U_1 + \cdots + U_{j-1}$, so $\mathbf{u}_j \in U_j \cap \big(U_1 + \cdots + U_{j-1}\big) = \{\mathbf{0}\}$, again a contradiction. Hence every $\mathbf{u}_i = \mathbf{0}$.

For $k = 2$, condition (III) says $U_1 \cap U_2 = \{\mathbf{0}\}$, which proves the last sentence. □

The step (IV) ⇒ (II) chose an extremal index, the largest j with $\mathbf{u}_j \ne \mathbf{0}$, exactly as the linear dependence lemma does for lists. The other natural route is induction on the number of summands. It rests on a fact that is useful in its own right: independent subspaces can be grouped.

Proposition 4.4.9 (Grouping summands). Let $U_1, \ldots, U_k$ be subspaces of V, let $1 \le m < k$, and write $S = U_1 + \cdots + U_m$ and $S' = U_{m+1} + \cdots + U_k$. Then $U_1 + \cdots + U_k$ is direct if and only if both of the following hold.

(i) The sums $U_1 + \cdots + U_m$ and $U_{m+1} + \cdots + U_k$ are direct.

(ii) $S \cap S' = \{\mathbf{0}\}$.

Moreover, independence of subspaces does not depend on their order, and if $U_1, \ldots, U_k$ are independent, then so is any selection of them.

Proof. ROUTINE

Uses. Theorem 4.4.7 (Criteria for direct sums), Definition 4.4.1 (Sum of subspaces)
Techniques. Compare components (Technique 4.4.8)
Hint 1. Use condition (II) of Theorem 4.4.7 throughout. Only its equivalence with (I) is needed.
Hint 2. For the backward direction, given $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$, set $\mathbf{s} = \mathbf{u}_1 + \cdots + \mathbf{u}_m$ and $\mathbf{s}' = \mathbf{u}_{m+1} + \cdots + \mathbf{u}_k$. Then $\mathbf{s} = -\mathbf{s}'$.
Sketch. (⇒) A decomposition of $\mathbf{0}$ along $U_1, \ldots, U_m$, padded with zero components, is one along all k subspaces, so (i) follows from (II). If $\mathbf{v} \in S \cap S'$, write $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{u}_{m+1} + \cdots + \mathbf{u}_k$; moving everything to one side gives a decomposition of $\mathbf{0}$, so all $\mathbf{u}_i = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$. (≤) If $\sum_i \mathbf{u}_i = \mathbf{0}$, then $\mathbf{s} = -\mathbf{s}' \in S \cap S' = \{\mathbf{0}\}$, and (i) kills each block. The last sentence holds because (II) does not refer to the order, and padding with zeros works for any selection.

We use the equivalence of (I) and (II) in Theorem 4.4.7.

(⇒) Suppose the sum of all k subspaces is direct. If $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$ with $\mathbf{u}_i \in U_i$, then $\mathbf{u}_1 + \cdots + \mathbf{u}_m + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$ is a decomposition of $\mathbf{0}$ along $U_1, \ldots, U_k$, so every $\mathbf{u}_i = \mathbf{0}$. Hence $U_1 + \cdots + U_m$ is direct, and likewise $U_{m+1} + \cdots + U_k$. For (ii), let $\mathbf{v} \in S \cap S'$, say $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{u}_{m+1} + \cdots + \mathbf{u}_k$ with $\mathbf{u}_i \in U_i$. Then

$$\mathbf{u}_1 + \cdots + \mathbf{u}_m + (-\mathbf{u}_{m+1}) + \cdots + (-\mathbf{u}_k) = \mathbf{0}$$

is a decomposition of $\mathbf{0}$, so all components vanish and $\mathbf{v} = \mathbf{0}$.

(≤) Suppose (i) and (ii), and let $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i \in U_i$. Put $\mathbf{s} = \mathbf{u}_1 + \cdots + \mathbf{u}_m \in S$ and $\mathbf{s}' = \mathbf{u}_{m+1} + \cdots + \mathbf{u}_k \in S'$. Then $\mathbf{s} = -\mathbf{s}'$ lies in S and in S', so $\mathbf{s} = \mathbf{0} = \mathbf{s}'$ by (ii). Now $\mathbf{s} = \mathbf{0}$ is a decomposition of $\mathbf{0}$ along $U_1, \ldots, U_m$, so $\mathbf{u}_1 = \cdots = \mathbf{u}_m = \mathbf{0}$ by (i), and likewise $\mathbf{u}_{m+1} = \cdots = \mathbf{u}_k = \mathbf{0}$. So the sum is direct.

Finally, condition (II) does not refer to the order of $U_1, \ldots, U_k$, since a sum of vectors does not depend on the order of its terms. If $U_{i_1}, \ldots, U_{i_r}$ are some of the U_i (with distinct indices), a decomposition of $\mathbf{0}$ along them becomes a decomposition of $\mathbf{0}$ along $U_1, \ldots, U_k$ once we insert the component $\mathbf{0}$ for every other U_i. If $U_1, \ldots, U_k$ are independent, all its components vanish. □

With $m = k-1$, the proposition says: $U_1 + \cdots + U_k$ is direct exactly when $U_1 + \cdots + U_{k-1}$ is direct and $(U_1 + \cdots + U_{k-1}) \cap U_k = \{\mathbf{0}\}$. That is an inductive description of directness, one new subspace at a time.

Technique 4.4.10 (Induction on the number of summands). To prove a statement about k subspaces $U_1, \ldots, U_k$, group the first $k-1$ of them into $S = U_1 + \cdots + U_{k-1}$ and treat $S + U_k$ as a sum of two subspaces. Proposition 4.4.2(a) says that $S + U_k$ is the whole sum, and Proposition 4.4.9 says that it is direct exactly when S is direct and $S \cap U_k = \{\mathbf{0}\}$. So the two-subspace case and the inductive hypothesis for S combine into the case of k subspaces.

Reach for it when the statement for two subspaces is easy, typically because it reduces to $U \cap W = \{\mathbf{0}\}$, and the general case looks messy. Watch the hypothesis in the inductive step: it needs $S \cap U_k = \{\mathbf{0}\}$, an intersection with a sum, and not just $U_i \cap U_k = \{\mathbf{0}\}$ for each $i < k$ (Non-example 4.4.6(b)). The canonical example is the second proof of Theorem 4.4.7. The same induction shows that eigenspaces for distinct eigenvalues are independent (Section 7.2).

Used in. Proposition 4.4.2, Theorem 4.4.7, Exercise 4.4.37, Exercise 4.4.40.

Proof of Theorem 4.4.7 (of (IV) ⇒ (II), by induction on k). STANDARD

Uses. Proposition 4.4.9 (Grouping summands)
Techniques. Induction on the number of summands (Technique 4.4.10)
Hint 1. Induct on k. Condition (IV) for $U_1, \ldots, U_k$ contains condition (IV) for $U_1, \ldots, U_{k-1}$.
Hint 2. With $S = U_1 + \cdots + U_{k-1}$, condition (IV) for $j = k$ says $S \cap U_k = \{\mathbf{0}\}$. Apply Proposition 4.4.9 with $m = k-1$.
Sketch. For $k = 1$, (II) is trivial. If $U_1, \ldots, U_k$ satisfy (IV), so do $U_1, \ldots, U_{k-1}$, which are therefore independent by induction. Condition (IV) for $j = k$ is $(U_1 + \cdots + U_{k-1}) \cap U_k = \{\mathbf{0}\}$, and Proposition 4.4.9 with $m = k-1$ turns these two facts into directness of the whole sum.

This proof uses only (I) ⇔ (II), which was proved directly above, through Proposition 4.4.9, whose proof uses nothing else. We induct on k. For $k = 1$, condition (II) says that $\mathbf{u}_1 = \mathbf{0}$ implies $\mathbf{u}_1 = \mathbf{0}$.

Let $k \geq 2$, assume the implication for $k-1$ subspaces, and suppose $U_1, \ldots, U_k$ satisfy (IV). The conditions of (IV) for $j = 2, \ldots, k-1$ are condition (IV) for $U_1, \ldots, U_{k-1}$, so by the inductive hypothesis these subspaces satisfy (II), and their sum $S = U_1 + \cdots + U_{k-1}$ is direct. The sum U_k of one subspace is trivially direct. Condition (IV) for $j = k$ says $S \cap U_k = \{\mathbf{0}\}$. By Proposition 4.4.9 with $m = k-1$, the sum $U_1 + \cdots + U_k$ is direct, which is (II). □

The second proof explains where (IV) comes from: it is exactly what the induction consumes, and each step asks only that the new subspace avoid everything built so far.

Exercise 4.4.11. ROUTINE Let $U_1 = \operatorname{span}((1,1,0))$, $U_2 = \operatorname{span}((1,0,1))$, and $U_3 = \operatorname{span}((0,1,1))$ in F^3.

(a) Show that $U_i \cap U_j = \{\mathbf{0}\}$ whenever $i \neq j$, over every field F.

(b) Show that $U_1 + U_2 + U_3$ is direct if $1 + 1 \neq 0$ in F, and not direct if $F = \mathbb{F}_2$.

Solution on page 1320

4.4.3 Bases and dimension

For lists, independence is what makes a spanning list a basis. For subspaces it plays the same role one level up: bases of the summands concatenate to a basis of the sum exactly when the summands are independent. The lists version is a recitation exercise.

Proposition 4.4.12 (Concatenating independent lists). Let $\mathcal{L}_1, \ldots, \mathcal{L}_k$ be linearly independent lists in V. The concatenation $\mathcal{L}_1, \ldots, \mathcal{L}_k$ is linearly independent if and only if $\operatorname{span}(\mathcal{L}_1), \ldots, \operatorname{span}(\mathcal{L}_k)$ are independent subspaces. In particular, for linearly independent lists $\mathcal{L}$ and $\mathcal{M}$, the concatenation $\mathcal{L}, \mathcal{M}$ is linearly independent if and only if $\operatorname{span}(\mathcal{L}) \cap \operatorname{span}(\mathcal{M}) = \{\mathbf{0}\}$.

Proof. ROUTINE

Uses. Theorem 4.4.7 (Criteria for direct sums), Theorem 1.5.5 (Independence is decided at zero), Definition 1.3.10 (Span)
Techniques. Compare components (Technique 4.4.8)
Hint 1. A linear combination of the concatenation is a sum of k pieces, the ith of which is a combination of $\mathcal{L}_i$ and so lies in $\operatorname{span}(\mathcal{L}_i)$.
Hint 2. So a linear relation among the concatenation is the same thing as a decomposition of $\mathbf{0}$ along the spans, together with a representation of each component. Compare (II) of Theorem 1.5.5 with (II) of Theorem 4.4.7.
Sketch. Write $\mathcal{L}_i = \mathbf{v}_{i1}, \dots, \mathbf{v}_{in_i}$. A relation $\sum_i \sum_j a_{ij}\mathbf{v}_{ij} = \mathbf{0}$ is a decomposition $\mathbf{0} = \mathbf{u}_1 + \cdots + \mathbf{u}_k$ with $\mathbf{u}_i = \sum_j a_{ij}\mathbf{v}_{ij} \in \operatorname{span}(\mathcal{L}_i)$, and since $\mathcal{L}_i$ is independent, $\mathbf{u}_i = \mathbf{0}$ exactly when all $a_{ij} = 0$. So "only the trivial relation" and "only the trivial decomposition of $\mathbf{0}$" say the same thing.

Write $\mathcal{L}_i = \mathbf{v}_{i1}, \dots, \mathbf{v}_{in_i}$ and $W_i = \operatorname{span}(\mathcal{L}_i)$.

($\leq$) Suppose $W_1, \dots, W_k$ are independent, and let $\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij}\mathbf{v}_{ij} = \mathbf{0}$. Put $\mathbf{u}_i = \sum_j a_{ij}\mathbf{v}_{ij} \in W_i$. Then $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$, so every $\mathbf{u}_i = \mathbf{0}$ by (II) of Theorem 4.4.7. Since $\mathcal{L}_i$ is independent, $a_{i1} = \cdots = a_{in_i} = 0$ (Theorem 1.5.5). So the concatenation has only the trivial relation and is independent.

($\Rightarrow$) Suppose the concatenation is independent, and let $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i \in W_i$, say $\mathbf{u}_i = \sum_j a_{ij}\mathbf{v}_{ij}$. Then $\sum_i \sum_j a_{ij}\mathbf{v}_{ij} = \mathbf{0}$, so all $a_{ij} = 0$ and every $\mathbf{u}_i = \mathbf{0}$. By (II) of Theorem 4.4.7, $W_1, \dots, W_k$ are independent.

The case $k = 2$ follows from the last sentence of Theorem 4.4.7. □

Theorem 4.4.13 (Bases and dimension of a sum). Let $U_1, \dots, U_k$ be finite-dimensional subspaces of V, and for each i let $\mathcal{B}_i$ be a basis of U_i.

(a) The concatenation $\mathcal{B}_1, \dots, \mathcal{B}_k$ spans $U_1 + \cdots + U_k$. Hence $U_1 + \cdots + U_k$ is finite-dimensional, and
$$\dim(U_1 + \cdots + U_k) \leq \dim U_1 + \cdots + \dim U_k.$$
(b) The sum $U_1 + \cdots + U_k$ is direct if and only if $\mathcal{B}_1, \dots, \mathcal{B}_k$ is a basis of $U_1 + \cdots + U_k$.
(c) The sum $U_1 + \cdots + U_k$ is direct if and only if $\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k$.

Proof. ROUTINE

Uses. Proposition 4.4.2 (Sums are the smallest containing subspaces), Proposition 4.4.12 (Concatenating independent lists), Definition 1.6.23 (Finite-dimensional), Corollary 4.1.11 (Spanning lists are long), Theorem 4.1.4 (All bases have the same length)
Techniques. Glue bases of the summands (Technique 4.4.14), Dimension counting (Technique 4.2.24)
Hint 1. (a) is Proposition 4.4.2(c), together with the fact that a spanning list is at least as long as the dimension.
Hint 2. For (b), the concatenation already spans, so it is a basis exactly when it is independent. For (c), when is a spanning list whose length equals the dimension a basis?
Sketch. (a) By Proposition 4.4.2(c), the concatenation, of length $n_1 + \cdots + n_k$ with $n_i = \dim U_i$, spans the sum, so the sum is finite-dimensional with dimension at most $n_1 + \cdots + n_k$. (b) A spanning list is a basis iff it is independent, iff the sum is direct by Proposition 4.4.12. (c) A spanning list is a basis iff its length is the dimension, so (c) reduces to (b).

Let $n_i = \dim U_i$, so that $\mathcal{B}_i$ has length n_i, and let $N = n_1 + \cdots + n_k$ be the length of the concatenation $\mathcal{B} = \mathcal{B}_1, \dots, \mathcal{B}_k$.

(a) By Proposition 4.4.2(c), $\operatorname{span}(\mathcal{B}) = \operatorname{span}(\mathcal{B}_1) + \cdots + \operatorname{span}(\mathcal{B}_k) = U_1 + \cdots + U_k$. So the sum is spanned by a finite list and is finite-dimensional (Definition 1.6.23), and its dimension is at most N by Corollary 4.1.11(a).

(b) By (a), $\mathcal{B}$ spans the sum, so $\mathcal{B}$ is a basis of the sum if and only if it is linearly independent. Each $\mathcal{B}_i$ is independent with span U_i, so by Proposition 4.4.12, $\mathcal{B}$ is independent if and only if $U_1, \dots, U_k$ are independent.

(c) Let $d = \dim(U_1 + \cdots + U_k)$. If $\mathcal{B}$ is a basis of the sum, then $N = d$ by Theorem 4.1.4. Conversely, if $N = d$, then $\mathcal{B}$ is a spanning list of length d, hence a basis by Corollary 4.1.11(b). So $d = N$ if and only if $\mathcal{B}$ is a basis, and (b) finishes the proof. □

In particular, if $V = U_1 \oplus \cdots \oplus U_k$ is finite-dimensional, then $\dim V = \dim U_1 + \cdots + \dim U_k$. Taking $U_i = \operatorname{span}(\mathbf{v}_i)$ with $\mathbf{v}_i \neq \mathbf{0}$, whose basis is the list $\mathbf{v}_i$, part (b) also gives the converse of Example 4.4.5(a): $V = \operatorname{span}(\mathbf{v}_1) \oplus \cdots \oplus \operatorname{span}(\mathbf{v}_n)$ if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis. And part (c) gives a second look at Non-example 4.4.6(b): three lines in $\mathbb{R}^2$ have dimensions adding up to $3 > 2$, so their sum cannot be direct.

Technique 4.4.14 (Glue bases of the summands). When $V = U_1 \oplus \cdots \oplus U_k$, choose a basis of each U_i and concatenate. The result is a basis of V (Theorem 4.4.13) adapted to the decomposition: each basis vector lies in one summand, so the coordinates of a vector split into blocks, one block for each component. In the other direction, to show that subspaces are independent, it suffices to check that bases of them concatenate to an independent list, or just to compare dimensions.

Reach for it to compute the dimension of a direct sum, or to turn a statement about subspaces into one about lists. The canonical example is Theorem 4.4.13. Later, a basis glued from bases of eigenspaces is an eigenbasis (Section 7.4), and Jordan bases are glued from chains (Section 10.3).
Used in. Theorem 4.4.13, Proposition 4.4.16, Theorem 7.4.16, Theorem 7.6.34, Exercise 8.2.36, Exercise 7.6.49, Exercise 8.1.36.

The span explorer shows the case of a plane and a line in $\mathbb{R}^3$.

$$\text{With } U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) \text{ and } W = \operatorname{span}((1,1,1)):$$
$$(2,1,3) = (-1,-2,0) + (3,3,3) \quad \text{with } (-1,-2,0) \in U, (3,3,3) \in W.$$

The first two vectors span a plane U and the third spans a line W. By Proposition 4.4.12, the list is independent exactly when $U \cap W = \{\mathbf{0}\}$, and then the coefficients of the target $\mathbf{w}$ split it as $(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) + a_3\mathbf{v}_3 \in U \oplus W$. Drag $\mathbf{v}_3$ into the plane: the list becomes dependent, and the dependence relation, rearranged, exhibits a nonzero vector of $U \cap W$. (Interactive in the web version.)

4.4.4 Complements

Given a subspace U of V, can V be split as $U \oplus W$? In finite dimensions it always can, by extending a basis.

Definition 4.4.15 (Complement). Let U be a subspace of V. A **complement** of U in V is a subspace W of V with $V = U \oplus W$.

By Theorem 4.4.7, W is a complement of U if and only if $U + W = V$ and $U \cap W = \{\mathbf{0}\}$. The relation is symmetric: W is a complement of U if and only if U is a complement of W.

Proposition 4.4.16 (Complements exist). Let V be finite-dimensional, and let U be a subspace of V.

(a) If $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is a basis of U and $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_r$ is a basis of V, then $\operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_r)$ is a complement of U. In particular, U has a complement.

(b) Every complement of U arises as in (a), and every complement W of U has $\dim W = \dim V - \dim U$.

Proof. ROUTINE

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Proposition 4.4.2 (Sums are the smallest containing subspaces), Proposition 4.4.12 (Concatenating independent lists), Theorem 4.4.13 (Bases and dimension of a sum)
Techniques. Extend a basis of a subspace (Technique 4.1.18), Glue bases of the summands (Technique 4.4.14)
Hint 1. For (a), the extended list spans V and is independent. Translate these two facts into statements about $U + W$ and $U \cap W$.
Hint 2. For (b), concatenate a basis of U with a basis of a complement W and apply Theorem 4.4.13.
Sketch. (a) Let $W = \operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_r)$. The concatenation spans V, so $U + W = V$ (Proposition 4.4.2(c)), and it is independent, so $U \cap W = \{\mathbf{0}\}$ (Proposition 4.4.12). A basis of U exists and extends (Proposition 4.1.21, Theorem 4.1.16). (b) If $V =$

$U \oplus W$, then W is finite-dimensional, and a basis of U followed by a basis of W is a basis of V (Theorem 4.4.13(b)), so W arises as in (a), with $r = \dim V - k$.

(a) Let $W = \operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_r)$. By Proposition 4.4.2(c), $U + W = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_r) = V$. The lists $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_r$ are independent (the second as part of a basis), and their concatenation is independent, so $U \cap W = \{\mathbf{0}\}$ by Proposition 4.4.12. Hence $V = U \oplus W$ by Theorem 4.4.7.

For the existence of a complement: U is finite-dimensional (Proposition 4.1.21(a)), so it has a basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$, which is independent in V and extends to a basis of V (Theorem 4.1.16). Apply the first part.

(b) Let W be a complement of U. It is finite-dimensional (Proposition 4.1.21(a)); let $\mathbf{w}_1, \ldots, \mathbf{w}_r$ be a basis of W, and $\mathbf{u}_1, \ldots, \mathbf{u}_k$ a basis of U. Since $V = U \oplus W$, Theorem 4.4.13(b) says that $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_r$ is a basis of V, and W is the span of the added vectors, as in (a). Counting, $\dim V = k + r$, so $\dim W = r = \dim V - \dim U$. □

Example 4.4.17 (Complements are not unique). In $\mathbb{R}^2$, every line through $\mathbf{0}$ other than the x-axis is a complement of the x-axis: two distinct lines meet only in $\mathbf{0}$, and their sum has dimension $1 + 1 = 2$ by Theorem 4.4.13(c), so it is $\mathbb{R}^2$.

In $\mathbb{Q}^4$, Example 4.1.17 extended $\mathbf{v}_1 = (1, 1, 0, 0)$, $\mathbf{v}_2 = (1, 1, 1, 1)$ by $\mathbf{e}_1, \mathbf{e}_3$ to a basis of $\mathbb{Q}^4$. So $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$ is a complement of $U = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. So is $\operatorname{span}(\mathbf{e}_2, \mathbf{e}_4)$. A vector $a\mathbf{v}_1 + b\mathbf{v}_2 = (a + b, a + b, b, b)$ of U lies in $\operatorname{span}(\mathbf{e}_2, \mathbf{e}_4)$ only if its first and third coordinates vanish, that is, $a + b = 0$ and $b = 0$, so $U \cap \operatorname{span}(\mathbf{e}_2, \mathbf{e}_4) = \{\mathbf{0}\}$. By Proposition 4.4.12, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_2, \mathbf{e}_4$ is independent; having length 4, it is a basis (Theorem 4.1.10(b)), and (a) applies.

What a subspace does determine is the dimension of its complements, by (b). Exercise 4.4.46 shows that a subspace has only one complement exactly when it is $\{\mathbf{0}\}$ or V, and Exercise 4.4.49 counts the complements over a finite field.

Complements also show that nothing distinguishes kernels or images among subspaces.

Proposition 4.4.18 (Every subspace is a kernel and an image). Let V be finite-dimensional, and let U be a subspace of V with a complement W. There are $S, T \in \mathrm{L}(V)$ with

$$\operatorname{im} S = U, \quad \ker S = W, \quad \ker T = U, \quad \operatorname{im} T = W.$$

In particular, every subspace of V is the image of some operator on V and the kernel of another.

Proof. ROUTINE

Uses. Proposition 4.4.16 (Complements exist), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Extend a basis of a subspace (Technique 4.1.18), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Concatenate bases of U and W to get a basis of V, and define S and T on it.
Hint 2. Let S fix the basis vectors of U and kill those of W; let T do the opposite.
Sketch. With bases $\mathbf{u}_1, \ldots, \mathbf{u}_k$ of U and $\mathbf{w}_1, \ldots, \mathbf{w}_r$ of W, which together form a basis of V, let $S(\mathbf{u}_i) = \mathbf{u}_i$, $S(\mathbf{w}_j) = \mathbf{0}$, $T(\mathbf{u}_i) = \mathbf{0}$, and $T(\mathbf{w}_j) = \mathbf{w}_j$. Then $S(\sum a_i\mathbf{u}_i + \sum b_j\mathbf{w}_j) = \sum a_i\mathbf{u}_i$, which is $\mathbf{0}$ iff all $a_i = 0$; read off kernels and images. Complements exist by Proposition 4.4.16.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_r$ be bases of U and W. Since $V = U \oplus W$, their concatenation is a basis of V (Theorem 4.4.13(b)). By Theorem 2.2.1 there are $S, T \in \mathrm{L}(V)$ with

$$S(\mathbf{u}_i) = \mathbf{u}_i, \quad S(\mathbf{w}_j) = \mathbf{0}, \quad T(\mathbf{u}_i) = \mathbf{0}, \quad T(\mathbf{w}_j) = \mathbf{w}_j$$

for all i and j. For $\mathbf{v} = \sum_i a_i\mathbf{u}_i + \sum_j b_j\mathbf{w}_j$, we get $S(\mathbf{v}) = \sum_i a_i\mathbf{u}_i$ and $T(\mathbf{v}) = \sum_j b_j\mathbf{w}_j$.

The image of S is the span of $S(\mathbf{u}_1), \ldots, S(\mathbf{w}_r)$, which is $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{0}, \ldots, \mathbf{0}) = U$ (Proposition 2.1.16). Since $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is independent, $S(\mathbf{v}) = \mathbf{0}$ if and only if all $a_i = 0$ (Theorem 1.5.5), that is, if and only if $\mathbf{v} \in$

W. So $\ker S = W$. Exchanging the roles of U and W gives $\ker T = U$ and $\operatorname{im} T = W$. Every subspace has a complement by Proposition 4.4.16, which proves the last sentence. □

For $V = F^n$ this says that every subspace of F^n is a column space (which is clear: put a spanning list in the columns) and a null space. The second statement is the useful one: every subspace of F^n is the solution set of a homogeneous linear system. The operators S and T will reappear in Section 4.4.6 as the projections attached to $V = U \oplus W$.

ASIDE: *Complements in infinite dimensions* BEYOND 21-242

The proof of Proposition 4.4.16 extends a basis, which needs V to be finite-dimensional. Some subspaces of infinite-dimensional spaces have evident complements; in $F[x]$, the polynomials with $p(0) = 0$ have the constants as a complement. In fact every subspace of every vector space has a complement, but in general the proof needs Zorn's lemma (Section 15.1), and it produces no explicit complement.

4.4.5 The dimension formula

Intuition (Inclusion–exclusion for subspaces). For finite sets, $|A \cup B| = |A| + |B| - |A \cap B|$, because the elements of $A \cap B$ are counted twice. For subspaces, $U + W$ takes the place of the union (it is the smallest subspace containing it), and dimension takes the place of size. A basis of U together with a basis of W spans $U + W$, and we expect it to overcount by $\dim(U \cap W)$ directions. For arbitrary bases this is not visible: no entry of either basis need lie in $U \cap W$ (Exercise 4.4.38). So the first proof chooses bases that are adapted to $U \cap W$, which is the extension technique again.

Theorem 4.4.19 (Dimension of a sum). Let U and W be finite-dimensional subspaces of V. Then $U + W$ and $U \cap W$ are finite-dimensional, and

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Proof (by extending a basis of $U \cap W$). STANDARD

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 1.5.5 (Independence is decided at zero), Theorem 4.1.4 (All bases have the same length)
Techniques. Extend a basis of a subspace (Technique 4.1.18)
Hint 1. Start with a basis $\mathbf{x}_1, \ldots, \mathbf{x}_m$ of $U \cap W$ and extend it twice: to a basis of U and to a basis of W.
Hint 2. Show that the $\mathbf{x}$'s, the new vectors of U, and the new vectors of W together form a basis of $U + W$. Spanning is easy.
Hint 3. For independence, take a relation and put the terms coming from W on one side. That side lies in W, and the other side lies in U.
Sketch. Extend a basis $\mathbf{x}_1, \ldots, \mathbf{x}_m$ of $U \cap W$ to bases $\mathbf{x}, \mathbf{u}_1, \ldots, \mathbf{u}_j$ of U and $\mathbf{x}, \mathbf{w}_1, \ldots, \mathbf{w}_l$ of W. The combined list spans $U + W$. In a relation $\sum a_i\mathbf{x}_i + \sum b_i\mathbf{u}_i + \sum c_i\mathbf{w}_i = \mathbf{0}$, the vector $\sum c_i\mathbf{w}_i$ equals $-(\sum a_i\mathbf{x}_i + \sum b_i\mathbf{u}_i) \in U$, so it lies in $U \cap W$ and is a combination of the $\mathbf{x}$'s; independence of $\mathbf{x}, \mathbf{w}$ forces $c = \mathbf{0}$, and then independence of $\mathbf{x}, \mathbf{u}$ forces $a = b = \mathbf{0}$. Count: $m + j + l = (m + j) + (m + l) - m$.

Since $U \cap W$ is a subspace of the finite-dimensional space U, it is finite-dimensional (Proposition 4.1.21(a)). Let $\mathbf{x}_1, \ldots, \mathbf{x}_m$ be a basis of $U \cap W$. It is an independent list in U, so it extends to a basis $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{u}_1, \ldots, \mathbf{u}_j$ of U (Theorem 4.1.16, applied in U), and likewise to a basis $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{w}_1, \ldots, \mathbf{w}_l$ of W. Then $\dim U = m + j$ and $\dim W = m + l$. We claim that

$$\mathcal{L} = \mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{u}_1, \ldots, \mathbf{u}_j, \mathbf{w}_1, \ldots, \mathbf{w}_l$$

is a basis of $U + W$. Then $U + W$ is finite-dimensional, and

$$\dim(U + W) = m + j + l = (m + j) + (m + l) - m = \dim U + \dim W - \dim(U \cap W).$$

$\mathcal{L}$ *spans* $U+W$. Every entry of $\mathcal{L}$ lies in U or in W, hence in the subspace $U+W$, so $\operatorname{span}(\mathcal{L}) \subseteq U+W$ (Proposition 1.4.14(b)). Conversely, let $\mathbf{u}+\mathbf{w} \in U+W$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Then $\mathbf{u}$ is a combination of the $\mathbf{x}$'s and $\mathbf{u}$'s, and $\mathbf{w}$ is a combination of the $\mathbf{x}$'s and $\mathbf{w}$'s, so $\mathbf{u}+\mathbf{w} \in \operatorname{span}(\mathcal{L})$.

$\mathcal{L}$ *is independent.* Suppose

$$\sum_{i=1}^{m} a_i \mathbf{x}_i + \sum_{i=1}^{j} b_i \mathbf{u}_i + \sum_{i=1}^{l} c_i \mathbf{w}_i = \mathbf{0}.$$

Then $\mathbf{y} = \sum_i c_i \mathbf{w}_i = -\left(\sum_i a_i \mathbf{x}_i + \sum_i b_i \mathbf{u}_i\right)$ lies in W (by the first expression) and in U (by the second), so $\mathbf{y} \in U \cap W$, and $\mathbf{y} = \sum_i d_i \mathbf{x}_i$ for some scalars d_i. Hence

$$\sum_{i=1}^{l} c_i \mathbf{w}_i - \sum_{i=1}^{m} d_i \mathbf{x}_i = \mathbf{0},$$

and since $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{w}_1, \ldots, \mathbf{w}_l$ is independent, all c_i (and d_i) are 0 (Theorem 1.5.5). The original relation becomes $\sum_i a_i \mathbf{x}_i + \sum_i b_i \mathbf{u}_i = \mathbf{0}$, and since $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{u}_1, \ldots, \mathbf{u}_j$ is independent, all a_i and b_i are 0 as well. □

The second proof reduces the formula to the direct case. Inside U, split off the part of U that overlaps W; what remains of U meets W trivially.

Proof (via a complement inside U). STANDARD

> *Uses.* Proposition 4.4.16 (Complements exist), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.4.7 (Criteria for direct sums), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.4.2 (Sums are the smallest containing subspaces)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* Choose a complement X of $U \cap W$ inside the finite-dimensional space U, so that $U = (U \cap W) \oplus X$. Show that $U + W = X \oplus W$.
> *Hint 2.* Since $X \subseteq U$, a vector of $X \cap W$ lies in $U \cap W$. For the sum, $U = (U \cap W) + X \subseteq W + X$.
> *Sketch.* Let X be a complement of $U \cap W$ in U, so $\dim X = \dim U - \dim(U \cap W)$. Since $X \subseteq U$, $X \cap W = X \cap (U \cap W) = \{\mathbf{0}\}$. Since $U = (U \cap W) + X \subseteq X + W$, $U + W = X + W$. So $U + W = X \oplus W$, and Theorem 4.4.13(c) gives $\dim(U+W) = \dim X + \dim W$.

As in the first proof, $U \cap W$ is finite-dimensional. Apply Proposition 4.4.16 in the finite-dimensional space U to its subspace $U \cap W$: there is a subspace $X \subseteq U$ with $U = (U \cap W) \oplus X$, and $\dim X = \dim U - \dim(U \cap W)$.

First, $X \cap W = \{\mathbf{0}\}$: a vector of $X \cap W$ lies in U (as $X \subseteq U$) and in W, so it lies in $X \cap (U \cap W) = \{\mathbf{0}\}$. Second, $X + W = U + W$: since $X \subseteq U$, we have $X + W \subseteq U + W$; and since $U \cap W \subseteq W$, we have $U = (U \cap W) + X \subseteq X + W$, so the subspace $X + W$ contains U and W and hence $U + W$ (Proposition 4.4.2(b)). By Theorem 4.4.7, $U + W = X \oplus W$, and by Theorem 4.4.13, $U + W$ is finite-dimensional with

$$\dim(U+W) = \dim X + \dim W = \dim U - \dim(U \cap W) + \dim W.$$ □

The third proof counts a kernel and an image. It completes the remark after Proposition 4.2.26, whose first proof used the same map to bound $\dim(U \cap W)$.

Proof (by rank–nullity). STANDARD

> *Uses.* Theorem 4.2.11 (Rank–nullity), Proposition 4.2.6 (Injective maps preserve dimension), Theorem 1.5.5 (Independence is decided at zero), Proposition 4.2.26 (Subspaces that must meet), Proposition 4.1.21 (Dimensions of subspaces)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* With bases $\mathbf{u}_1, \ldots, \mathbf{u}_k$ of U and $\mathbf{w}_1, \ldots, \mathbf{w}_l$ of W, consider $L : F^{k+l} \to V$, $L(\mathbf{a}, \mathbf{b}) = \sum_i a_i \mathbf{u}_i - \sum_j b_j \mathbf{w}_j$. What is $\operatorname{im} L$?
> *Hint 2.* An element of $\ker L$ is a pair of coordinate vectors describing one vector of $U \cap W$. Show that $(\mathbf{a}, \mathbf{b}) \mapsto \sum_i a_i \mathbf{u}_i$ is a bijection from $\ker L$ onto $U \cap W$.
> *Sketch.* $\operatorname{im} L = U + W$, since $\mathbf{b}$ and $-\mathbf{b}$ range over the same set. The linear map $\varphi(\mathbf{a}, \mathbf{b}) = \sum_i a_i \mathbf{u}_i$ sends $\ker L$ into $U \cap W$ and is injective, as in Proposition 4.2.26. It is onto $U \cap W$: if $\mathbf{v} = \sum_i a_i \mathbf{u}_i = \sum_j b_j \mathbf{w}_j$, then $(\mathbf{a}, \mathbf{b}) \in \ker L$ and $\varphi(\mathbf{a}, \mathbf{b}) = \mathbf{v}$. So $\dim \ker L = \dim(U \cap W)$, and rank–nullity for L reads $k + l = \dim(U \cap W) + \dim(U + W)$.

Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ and $\mathbf{w}_1, \dots, \mathbf{w}_l$ be bases of U and W. Write elements of F^{k+l} as $(\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in F^k$ and $\mathbf{b} \in F^l$, and define the linear map

$$L : F^{k+l} \to V, \quad L(\mathbf{a}, \mathbf{b}) = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k - b_1\mathbf{w}_1 - \cdots - b_l\mathbf{w}_l.$$

Every value of L is a vector of U plus a vector of W, and every $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} = \sum_i a_i\mathbf{u}_i$ and $\mathbf{w} = \sum_j b_j\mathbf{w}_j$ equals $L(\mathbf{a}, -\mathbf{b})$. So $\operatorname{im} L = U + W$. By rank–nullity (Theorem 4.2.11(b)), $U + W$ is finite-dimensional and

$$k + l = \dim \ker L + \dim(U + W).$$

Define $\varphi : \ker L \to V$ by $\varphi(\mathbf{a}, \mathbf{b}) = \sum_i a_i\mathbf{u}_i$; it is linear. For $(\mathbf{a}, \mathbf{b}) \in \ker L$ we have $\sum_i a_i\mathbf{u}_i = \sum_j b_j\mathbf{w}_j$, so $\varphi(\mathbf{a}, \mathbf{b})$ lies in U and in W. Conversely, every $\mathbf{v} \in U \cap W$ can be written as $\mathbf{v} = \sum_i a_i\mathbf{u}_i = \sum_j b_j\mathbf{w}_j$, and then $(\mathbf{a}, \mathbf{b}) \in \ker L$ with $\varphi(\mathbf{a}, \mathbf{b}) = \mathbf{v}$. So $\operatorname{im} \varphi = U \cap W$. Finally, φ is injective: if $\sum_i a_i\mathbf{u}_i = \mathbf{0}$, then $\mathbf{a} = \mathbf{0}$, so $\sum_j b_j\mathbf{w}_j = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$ (Theorem 1.5.5). By Proposition 4.2.6(a), applied to φ on the finite-dimensional space $\ker L$, $\dim(U \cap W) = \dim \ker L$. Substituting gives $k + l = \dim(U \cap W) + \dim(U + W)$, as desired. □

The three proofs teach different things. The first builds an explicit basis of $U + W$ and is the one to adapt when a basis is wanted. The second shows that the formula is the direct case, $\dim(X \oplus W) = \dim X + \dim W$, in disguise. The third is how one computes: for $U, W \subseteq F^n$, $\ker L$ is the null space of the matrix whose columns are the $\mathbf{u}_i$ and the $-\mathbf{w}_j$, so one row reduction finds $\dim(U + W)$ and a basis of $U \cap W$ at once. With the quotient spaces of Section 4.6, the formula can also be read as $(U + W)/W \cong U/(U \cap W)$.

Example 4.4.20 (Computing a sum and an intersection). In $\mathbb{Q}^4$, let $U = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$ and $W = \operatorname{span}(\mathbf{w}_1, \mathbf{w}_2)$, where

$$\mathbf{u}_1 = (1, 0, 1, 0), \quad \mathbf{u}_2 = (0, 1, 1, 1), \quad \mathbf{w}_1 = (2, 0, -1, 1), \quad \mathbf{w}_2 = (1, 1, -1, 2).$$

Row reduce the matrix whose columns are $\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_1, \mathbf{w}_2$:

$$\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first two columns are pivot columns, so $\mathbf{u}_1, \mathbf{u}_2$ is independent and $\dim U = 2$. The vectors $\mathbf{w}_1, \mathbf{w}_2$ are not parallel (look at the second coordinates), so $\dim W = 2$. By Proposition 4.4.2(c), $U + W$ is the column space, so $\dim(U + W) = 3$, the number of pivots (Theorem 4.2.7). The formula gives

$$\dim(U \cap W) = 2 + 2 - 3 = 1.$$

To find the line $U \cap W$, read the free column: it says $\mathbf{w}_2 = -\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{w}_1$ (Corollary 3.3.19(b)). Sorting the terms by subspace,

$$-\mathbf{u}_1 + \mathbf{u}_2 = -\mathbf{w}_1 + \mathbf{w}_2 = (-1, 1, 0, 1).$$

The left side lies in U and the right side in W, so this nonzero vector spans $U \cap W$. In the language of the third proof, $(\mathbf{a}, \mathbf{b}) = ((-1, 1), (-1, 1))$ spans $\ker L$.

Since $U + W \subseteq V$, the formula recovers Proposition 4.2.26: $\dim(U \cap W) = \dim U + \dim W - \dim(U + W) \geq \dim U + \dim W - \dim V$. It also says by how much that bound can miss, namely by $\dim V - \dim(U + W)$. The most frequent use is a test for complements.

Proposition 4.4.21 (Two out of three). Let V be finite-dimensional, and let U and W be subspaces of V. Consider the conditions

(i) $U \cap W = \{\mathbf{0}\}$;

(ii) $U + W = V$;

(iii) $\dim U + \dim W = \dim V$.

Any two of them imply the third, and then $V = U \oplus W$.

Proof. ROUTINE

Uses. Theorem 4.4.19 (Dimension of a sum), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.4.7 (Criteria for direct sums)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Write the dimension formula as $\dim(U+W) + \dim(U \cap W) = \dim U + \dim W$ and substitute each pair of conditions.
Hint 2. To get (ii) from a dimension count, use that a subspace of V of dimension $\dim V$ is V.
Sketch. By Theorem 4.4.19, $\dim(U+W) + \dim(U \cap W) = \dim U + \dim W$. (i) and (ii) make the left side $\dim V$. (i) and (iii) give $\dim(U+W) = \dim V$, so $U + W = V$. (ii) and (iii) give $\dim(U \cap W) = 0$. Then $V = U \oplus W$ by Theorem 4.4.7.

By Theorem 4.4.19,

$$\dim(U+W) + \dim(U \cap W) = \dim U + \dim W.$$

If (i) and (ii) hold, the left side is $\dim V + 0$, which gives (iii). If (i) and (iii) hold, then $\dim(U+W) = \dim V$, so $U + W = V$ by Proposition 4.1.21(b), which is (ii). If (ii) and (iii) hold, then $\dim V + \dim(U \cap W) = \dim V$, so $\dim(U \cap W) = 0$ and $U \cap W = \{\mathbf{0}\}$, which is (i). In each case (i) and (ii) hold, so $V = U \oplus W$ by Theorem 4.4.7. □

This is the subspace version of the fact that a list of length $\dim V$ is a basis as soon as it is independent or spanning (Theorem 4.1.10, Corollary 4.1.11): with the right dimensions, checking one of "meet trivially" and "add up to V" gives the other for free.

Remark 4.4.22 (No formula for three subspaces). There is no formula of this kind for three subspaces in terms of the dimensions of their intersections. For the three lines ℓ_1, ℓ_2, ℓ_3 of Non-example 4.4.6(b), all pairwise intersections and the triple intersection are $\{\mathbf{0}\}$, so the inclusion–exclusion expression $1 + 1 + 1 - 0 - 0 - 0 + 0$ gives 3, while $\dim(\ell_1 + \ell_2 + \ell_3) = 2$. The two-subspace proofs break down because $(U+W) \cap X$ can be larger than $(U \cap X) + (W \cap X)$; Exercise 4.4.39 shows that this is the only obstruction.

Exercise 4.4.23. ROUTINE Let $U = \{\mathbf{x} \in F^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$.

(a) Show that the complements of U in F^4 are exactly the lines $\operatorname{span}(\mathbf{v})$ with $\mathbf{v} \notin U$.

(b) Is $\operatorname{span}((1,1,1,1))$ a complement of U when $F = \mathbb{Q}$? When $F = \mathbb{F}_2$? When $F = \mathbb{F}_3$?

Solution on page 1321

4.4.6 Projections

A decomposition $V = U \oplus W$ gives every vector a well-defined U-component. Taking that component is a linear map, and the maps that arise in this way are exactly the operators P with $P^2 = P$.

Definition 4.4.24 (Projection along a complement). Suppose $V = U \oplus W$. The **projection onto** U **along** W is the map

$$\operatorname{Pr}_{U,W} : V \to V, \quad \mathbf{u} + \mathbf{w} \mapsto \mathbf{u} \quad (\mathbf{u} \in U, \mathbf{w} \in W).$$

It is well defined because every vector of V has exactly one decomposition $\mathbf{u} + \mathbf{w}$. The order of the subscripts matters: $\operatorname{Pr}_{U,W}$ keeps the U-component and $\operatorname{Pr}_{W,U}$ keeps the W-component. Changing W changes the map, even though its image is still U. The orthogonal projection Pr_W of Section 6.4 is the special case $\operatorname{Pr}_{W,W^\perp}$.

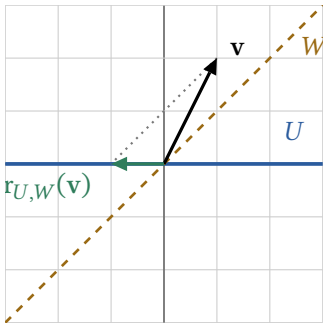
Example 4.4.25 (Projecting along a slanted line). In F^2, let $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \operatorname{span}((1,1))$. Every vector decomposes as

$$(x, y) = (x - y, 0) + (y, y),$$

with $(x - y, 0) \in U$ and $(y, y) \in W$. So $\operatorname{Pr}_{U,W}(x, y) = (x - y, 0)$ and $\operatorname{Pr}_{W,U}(x, y) = (y, y)$, and their matrices with respect to the standard basis are

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Each matrix squares to itself, the two add up to I_2, and their products in either order are 0. Over $\mathbb{R}$, $\operatorname{Pr}_{U,W}$ slides every point parallel to $(1,1)$ until it reaches the x-axis. It is not the perpendicular projection onto the x-axis, which is $(x, y) \mapsto (x, 0)$, the projection onto U along the y-axis.



The projection $\operatorname{Pr}_{U,W}$ of Example 4.4.25, with U the x-axis and $W = \operatorname{span}((1,1))$. The plane collapses onto U, every point moves parallel to W, and the points of W go to $\mathbf{0}$. In the static picture, $\mathbf{v} = (1,2)$ goes to $\operatorname{Pr}_{U,W}(\mathbf{v}) = (-1,0)$, and the dotted segment from $\mathbf{v}$ to its image is parallel to W. In the interactive version you can change the matrix; the result is a projection exactly when $A^2 = A$. (Interactive in the web version.)

Proposition 4.4.26 (Properties of projections). Suppose $V = U \oplus W$, and let $P = \operatorname{Pr}_{U,W}$ and $Q = \operatorname{Pr}_{W,U}$.

(a) P is linear.

(b) $\operatorname{im} P = U$ and $\ker P = W$, and $P(\mathbf{u}) = \mathbf{u}$ for every $\mathbf{u} \in U$.

(c) $P^2 = P$.

(d) $P + Q = \mathrm{id}_V$ and $PQ = QP = 0$.

Proof. ROUTINE

Uses. Definition 4.4.24 (Projection along a complement), Definition 4.4.4 (Direct sum), Definition 2.1.1 (Linear map)
Techniques. Compare components (Technique 4.4.8)
Hint 1. For (a), if $\mathbf{v} = \mathbf{u} + \mathbf{w}$ and $\mathbf{v}' = \mathbf{u}' + \mathbf{w}'$, write down a decomposition of $\mathbf{v} + \mathbf{v}'$ and of $a\mathbf{v}$. Uniqueness does the rest.
Hint 2. For (b), the decomposition of a vector $\mathbf{u} \in U$ is $\mathbf{u} + \mathbf{0}$. For (d), $\mathbf{v} = P(\mathbf{v}) + Q(\mathbf{v})$ by definition.
Sketch. (a) $(\mathbf{u} + \mathbf{u}') + (\mathbf{w} + \mathbf{w}')$ and $a\mathbf{u} + a\mathbf{w}$ are decompositions of $\mathbf{v} + \mathbf{v}'$ and $a\mathbf{v}$, hence the decompositions, so $P(\mathbf{v} + \mathbf{v}') = \mathbf{u} + \mathbf{u}'$ and $P(a\mathbf{v}) = a\mathbf{u}$. (b) $\mathbf{u} = \mathbf{u} + \mathbf{0}$ gives $P(\mathbf{u}) = \mathbf{u}$, so $\operatorname{im} P = U$; and $P(\mathbf{u} + \mathbf{w}) = \mathbf{0}$ iff $\mathbf{u} = \mathbf{0}$. (c) $P(\mathbf{v}) \in U$ is fixed. (d) $\mathbf{v} = \mathbf{u} + \mathbf{w} = P(\mathbf{v}) + Q(\mathbf{v})$, and $P(Q(\mathbf{v})) = P(\mathbf{w}) = \mathbf{0}$ since $\mathbf{w} \in W = \ker P$; symmetrically $QP = 0$.

Throughout, $\mathbf{v} = \mathbf{u} + \mathbf{w}$ and $\mathbf{v}' = \mathbf{u}' + \mathbf{w}'$ with $\mathbf{u}, \mathbf{u}' \in U$ and $\mathbf{w}, \mathbf{w}' \in W$.

(a) Since U and W are subspaces, $\mathbf{v} + \mathbf{v}' = (\mathbf{u} + \mathbf{u}') + (\mathbf{w} + \mathbf{w}')$ and $a\mathbf{v} = a\mathbf{u} + a\mathbf{w}$ are decompositions of $\mathbf{v} + \mathbf{v}'$ and $a\mathbf{v}$ along $U \oplus W$, so they are *the* decompositions. Hence $P(\mathbf{v} + \mathbf{v}') = \mathbf{u} + \mathbf{u}' = P(\mathbf{v}) + P(\mathbf{v}')$ and $P(a\mathbf{v}) = a\mathbf{u} = aP(\mathbf{v})$.

(b) Every value of P lies in U. For $\mathbf{u} \in U$, the decomposition of $\mathbf{u}$ is $\mathbf{u} + \mathbf{0}$, so $P(\mathbf{u}) = \mathbf{u}$; hence $U \subseteq \operatorname{im} P$, and $\operatorname{im} P = U$. Next, $P(\mathbf{v}) = \mathbf{u}$ is $\mathbf{0}$ if and only if $\mathbf{v} = \mathbf{w} \in W$, so $\ker P = W$.

(c) For every $\mathbf{v}$, $P(\mathbf{v}) \in U$, so $P(P(\mathbf{v})) = P(\mathbf{v})$ by (b).

(d) By definition $P(\mathbf{v}) + Q(\mathbf{v}) = \mathbf{u} + \mathbf{w} = \mathbf{v}$, so $P + Q = \mathrm{id}_V$. Since $Q(\mathbf{v}) = \mathbf{w} \in W = \ker P$, we get $PQ(\mathbf{v}) = \mathbf{0}$. Exchanging the roles of U and W (so that $Q = \mathrm{Pr}_{W,U}$ has kernel U) gives $QP = 0$. □

An operator P with $P^2 = P$ is called **idempotent**. The converse of (c) is the main result of this subsection: every idempotent operator is a projection.

> **Theorem 4.4.27** (Idempotents are projections). Let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$. Then
>
> $$V = \operatorname{im} P \oplus \ker P,$$
>
> and $P = \mathrm{Pr}_{\operatorname{im} P, \ker P}$. Moreover, $\operatorname{im} P = \{\mathbf{v} \in V \mid P(\mathbf{v}) = \mathbf{v}\}$.

Together with Proposition 4.4.26, the theorem gives a bijection between the ways of writing $V = U \oplus W$ (with U and W in this order) and the idempotent operators on V: a decomposition gives $\mathrm{Pr}_{U,W}$, and an idempotent P gives the decomposition $V = \operatorname{im} P \oplus \ker P$.

Proof (by writing down the decomposition). ROUTINE

> *Uses.* Theorem 4.4.7 (Criteria for direct sums), Definition 4.4.24 (Projection along a complement), Proposition 2.6.8 (Kernels and images are subspaces)
> *Techniques.* Compare components (Technique 4.4.8)
> *Hint 1.* Every $\mathbf{v}$ can be written as $P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v}))$. Where do the two pieces lie?
> *Hint 2.* If $\mathbf{v} \in \operatorname{im} P \cap \ker P$, write $\mathbf{v} = P(\mathbf{x})$ and apply P.
> *Sketch.* $P(\mathbf{v} - P(\mathbf{v})) = P(\mathbf{v}) - P^2(\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} = P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v})) \in \operatorname{im} P + \ker P$. If $\mathbf{v} = P(\mathbf{x})$ and $P(\mathbf{v}) = \mathbf{0}$, then $\mathbf{v} = P^2(\mathbf{x}) = P(\mathbf{v}) = \mathbf{0}$. So $V = \operatorname{im} P \oplus \ker P$, and the $\operatorname{im} P$-component of $\mathbf{v}$ is $P(\mathbf{v})$. Vectors $P(\mathbf{x})$ of $\operatorname{im} P$ are fixed, as $P(P(\mathbf{x})) = P(\mathbf{x})$.

Let $\mathbf{v} \in V$. Then $P(\mathbf{v}) \in \operatorname{im} P$, and

$$P(\mathbf{v} - P(\mathbf{v})) = P(\mathbf{v}) - P^2(\mathbf{v}) = P(\mathbf{v}) - P(\mathbf{v}) = \mathbf{0},$$

so $\mathbf{v} - P(\mathbf{v}) \in \ker P$. Hence

$$\mathbf{v} = P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v}))$$

is a decomposition of $\mathbf{v}$ along $\operatorname{im} P$ and $\ker P$, and $V = \operatorname{im} P + \ker P$. Next, let $\mathbf{v} \in \operatorname{im} P \cap \ker P$, say $\mathbf{v} = P(\mathbf{x})$. Then $\mathbf{v} = P(\mathbf{x}) = P^2(\mathbf{x}) = P(\mathbf{v}) = \mathbf{0}$. By Theorem 4.4.7, $V = \operatorname{im} P \oplus \ker P$, and by the displayed decomposition, the $\operatorname{im} P$-component of $\mathbf{v}$ is $P(\mathbf{v})$, that is, $P = \mathrm{Pr}_{\operatorname{im} P, \ker P}$.

Finally, if $P(\mathbf{v}) = \mathbf{v}$, then $\mathbf{v} \in \operatorname{im} P$; conversely, if $\mathbf{v} = P(\mathbf{x})$, then $P(\mathbf{v}) = P^2(\mathbf{x}) = P(\mathbf{x}) = \mathbf{v}$. □

In finite dimensions there is a second proof, which never writes down the decomposition $\mathbf{v} = P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v}))$ and lets a dimension count produce the sum instead.

Proof (in finite dimensions, by counting). ROUTINE

> *Uses.* Theorem 4.2.11 (Rank–nullity), Proposition 4.4.21 (Two out of three), Definition 4.4.24 (Projection along a complement)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* Show $\operatorname{im} P \cap \ker P = \{\mathbf{0}\}$ as in the first proof.
> *Hint 2.* Rank–nullity gives $\dim \operatorname{im} P + \dim \ker P = \dim V$. Which result turns this and a trivial intersection into a direct sum?
> *Sketch.* If $\mathbf{v} = P(\mathbf{x})$ and $P(\mathbf{v}) = \mathbf{0}$, then $\mathbf{v} = P^2(\mathbf{x}) = \mathbf{0}$. Rank–nullity gives $\dim \operatorname{im} P + \dim \ker P = \dim V$, so $V = \operatorname{im} P \oplus \ker P$ by Proposition 4.4.21. For $\mathbf{v} = P(\mathbf{x}) + \mathbf{w}$ with $\mathbf{w} \in \ker P$, $P(\mathbf{v}) = P^2(\mathbf{x}) = P(\mathbf{x})$, the $\operatorname{im} P$-component.

Assume that V is finite-dimensional. As in the first proof, if $\mathbf{v} = P(\mathbf{x})$ lies in $\ker P$, then $\mathbf{v} = P^2(\mathbf{x}) = P(\mathbf{v}) = \mathbf{0}$, so $\operatorname{im} P \cap \ker P = \{\mathbf{0}\}$. By Theorem 4.2.11, $\dim \operatorname{im} P + \dim \ker P = \dim V$. Conditions (i) and (iii) of Proposition 4.4.21 hold, so $V = \operatorname{im} P \oplus \ker P$. To identify P, let $\mathbf{v} = P(\mathbf{x}) + \mathbf{w}$ with $\mathbf{w} \in \ker P$. Then $P(\mathbf{v}) = P^2(\mathbf{x}) + P(\mathbf{w}) = P(\mathbf{x})$, which is the $\operatorname{im} P$-component of $\mathbf{v}$. So $P = \mathrm{Pr}_{\operatorname{im} P, \ker P}$. □

The counting proof does not reach infinite-dimensional spaces, and even in finite dimensions the first proof is more informative: it says what the components of $\mathbf{v}$ are.

> **Technique 4.4.28** (Split the space with an idempotent). To decompose V as a direct sum, look for an operator P with $P^2 = P$. Then $V = \operatorname{im} P \oplus \ker P$, every vector splits as $\mathbf{v} = P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v}))$, and $\operatorname{im} P$ is the set of vectors that P fixes (Theorem 4.4.27). In the other direction, every decomposition $V = U \oplus W$ comes with its projections, and statements about decompositions become algebra of operators: $P^2 = P$, $P + Q = \mathrm{id}_V$, $PQ = QP = 0$.
>
> Reach for it when a formula suggests "the part of $\mathbf{v}$ that ...", such as the symmetric part $\left(A + A^{\mathsf{T}}\right)/2$ of a matrix, or more generally the average of $\mathbf{v}$ and $T(\mathbf{v})$ for an operator with $T^2 = \mathrm{id}_V$ (Exercise 4.4.43). Each is $P(\mathbf{v})$ for an idempotent P. The canonical example is Theorem 4.4.27. Later, idempotent matrices over $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ have trace equal to rank (Exercise 4.3.43), and orthogonal projections are the idempotents that are self-adjoint (Section 6.6).
>
> *Used in.* Exercise 4.4.30, Exercise 4.4.43, Exercise 4.4.51, Problem 4.4.

A direct sum of k subspaces comes with k projections, one for each summand. They are the algebraic form of the decomposition, and they are how Chapters 7 and 8 talk about splitting V into eigenspaces.

> **Proposition 4.4.29** (Projections of a direct sum).
>
> (a) Suppose $V = U_1 \oplus \cdots \oplus U_k$, and let $P_i : V \to V$ send each vector to its component in U_i. Then each P_i is linear with $\operatorname{im} P_i = U_i$, and
>
> $$P_1 + \cdots + P_k = \mathrm{id}_V, \quad P_i^2 = P_i, \quad P_i P_j = 0 \quad (i \neq j).$$
>
> (b) Conversely, let $P_1, \ldots, P_k \in \mathrm{L}(V)$ satisfy $P_1 + \cdots + P_k = \mathrm{id}_V$ and $P_i P_j = 0$ for $i \neq j$. Then each P_i is idempotent, $V = \operatorname{im} P_1 \oplus \cdots \oplus \operatorname{im} P_k$, and P_i sends each vector to its component in $\operatorname{im} P_i$.

Proof. STANDARD

> *Uses.* Theorem 4.4.7 (Criteria for direct sums), Proposition 4.4.26 (Properties of projections), Definition 4.4.24 (Projection along a complement), Proposition 4.4.2 (Sums are the smallest containing subspaces)
> *Techniques.* Apply an operator to a relation to kill terms (Technique 4.1.25), Compare components (Technique 4.4.8)
> *Hint 1.* For (a), let W_i be the sum of the summands other than U_i. Show $V = U_i \oplus W_i$ and $P_i = \mathrm{Pr}_{U_i, W_i}$.
> *Hint 2.* For (b), multiply $\mathrm{id}_V = \sum_j P_j$ on the left by P_i to get idempotence.
> *Hint 3.* For directness in (b), take $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i = P_i(\mathbf{x}_i)$ and apply P_j: every term but one is killed.
> *Sketch.* (a) With $W_i = \sum_{j \neq i} U_j$, criterion (III) gives $V = U_i \oplus W_i$, and $P_i = \mathrm{Pr}_{U_i, W_i}$, which is linear and idempotent with image U_i and kernel W_i. For $j \neq i$, $P_j(\mathbf{v}) \in U_j \subseteq W_i$, so $P_i P_j = 0$; and $\mathbf{v} = \sum_i P_i(\mathbf{v})$. (b) $P_i = P_i \sum_j P_j = P_i^2$. Since $\mathbf{v} = \sum_i P_i(\mathbf{v})$, the images add up to V. If $\sum_i P_i(\mathbf{x}_i) = \mathbf{0}$, applying P_j leaves $P_j^2(\mathbf{x}_j) = P_j(\mathbf{x}_j) = \mathbf{0}$, so the sum is direct, and $\mathbf{v} = \sum_i P_i(\mathbf{v})$ is the decomposition of $\mathbf{v}$.

(a) Fix i and let $W_i = \sum_{j \neq i} U_j$. Every $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k$ (with $\mathbf{u}_j \in U_j$) is $\mathbf{u}_i + \sum_{j \neq i} \mathbf{u}_j \in U_i + W_i$, and $U_i \cap W_i = \{\mathbf{0}\}$ by (III) of Theorem 4.4.7. So $V = U_i \oplus W_i$, and the U_i-component of $\mathbf{v}$ in this decomposition is again $\mathbf{u}_i$. That is, $P_i = \mathrm{Pr}_{U_i, W_i}$, which is linear and idempotent with image U_i and kernel W_i by Proposition 4.4.26. For $j \neq i$ and $\mathbf{v} \in V$, $P_j(\mathbf{v}) \in U_j \subseteq W_i = \ker P_i$, so $P_i P_j = 0$. Finally, $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k = P_1(\mathbf{v}) + \cdots + P_k(\mathbf{v})$, so $P_1 + \cdots + P_k = \mathrm{id}_V$.

(b) For each i,

$$P_i = P_i \mathrm{id}_V = P_i(P_1 + \cdots + P_k) = P_i P_1 + \cdots + P_i P_k = P_i^2,$$

since every term $P_i P_j$ with $j \neq i$ is 0. For $\mathbf{v} \in V$, $\mathbf{v} = P_1(\mathbf{v}) + \cdots + P_k(\mathbf{v})$, so $V = \operatorname{im} P_1 + \cdots + \operatorname{im} P_k$. For directness we check (II) of Theorem 4.4.7. Let $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ with $\mathbf{u}_i = P_i(\mathbf{x}_i) \in \operatorname{im} P_i$, and fix j. Applying P_j kills every term with $i \neq j$, because $P_j(\mathbf{u}_i) = P_j P_i(\mathbf{x}_i) = \mathbf{0}$, and leaves $P_j(\mathbf{u}_j) = P_j^2(\mathbf{x}_j) = P_j(\mathbf{x}_j) = \mathbf{u}_j$. So $\mathbf{u}_j = P_j(\mathbf{0}) = \mathbf{0}$. Hence $V = \operatorname{im} P_1 \oplus \cdots \oplus \operatorname{im} P_k$, and since $\mathbf{v} = \sum_i P_i(\mathbf{v})$ with $P_i(\mathbf{v}) \in \operatorname{im} P_i$, the component of $\mathbf{v}$ in $\operatorname{im} P_i$ is $P_i(\mathbf{v})$. □

Exercise 4.4.30. ROUTINE In $\mathbb{Q}^3$, let $U = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 0\}$ and $W = \operatorname{span}((1,1,1))$.

(a) Show that $\mathbb{Q}^3 = U \oplus W$, and find the matrices of $\operatorname{Pr}_{U,W}$ and $\operatorname{Pr}_{W,U}$ with respect to the standard basis.

(b) Show that over $\mathbb{F}_3$, $U + W \neq \mathbb{F}_3^3$. Which step of (a) fails?

Solution on page 1321

4.4.7 Linear maps on direct sums

A linear map is determined by its values on a basis, and those values can be prescribed freely (Theorem 2.2.1). A direct sum decomposition gives the same freedom, one summand at a time.

Proposition 4.4.31 (Defining maps summand by summand). Suppose $V = U_1 \oplus \cdots \oplus U_k$, and let $S_i \in \mathrm{L}(U_i, W)$ for each i. There is exactly one $T \in \mathrm{L}(V, W)$ with $T(\mathbf{u}) = S_i(\mathbf{u})$ for every i and every $\mathbf{u} \in U_i$, namely

$$T(\mathbf{u}_1 + \cdots + \mathbf{u}_k) = S_1(\mathbf{u}_1) + \cdots + S_k(\mathbf{u}_k) \quad (\mathbf{u}_i \in U_i).$$

Proof. ROUTINE

Uses. Definition 4.4.4 (Direct sum), Proposition 4.4.29 (Projections of a direct sum), Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Compare components (Technique 4.4.8)
Hint 1. Uniqueness: a linear T that agrees with each S_i on U_i must satisfy the displayed formula.
Hint 2. Existence: the formula is well defined because decompositions are unique. For linearity, write T in terms of the projections of Proposition 4.4.29.
Sketch. If T is linear and agrees with S_i on U_i, then $T\big(\sum_i \mathbf{u}_i\big) = \sum_i T(\mathbf{u}_i) = \sum_i S_i(\mathbf{u}_i)$, so T is unique. The formula defines a function because each vector has one decomposition, and it equals $\sum_i S_i \circ P_i'$, where $P_i' : V \to U_i$ takes the U_i-component; this is linear. On $\mathbf{u} \in U_i$ the decomposition is $\mathbf{u}$ in slot i and $\mathbf{0}$ elsewhere, so $T(\mathbf{u}) = S_i(\mathbf{u})$.

Uniqueness. If $T \in \mathrm{L}(V,W)$ agrees with S_i on U_i for each i, then for $\mathbf{u}_i \in U_i$,

$$T(\mathbf{u}_1 + \cdots + \mathbf{u}_k) = T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k) = S_1(\mathbf{u}_1) + \cdots + S_k(\mathbf{u}_k).$$

Every vector of V is of this form, so T is determined.

Existence. Since every $\mathbf{v} \in V$ has exactly one decomposition, the displayed formula defines a function $T : V \to W$. Let $P_i' : V \to U_i$ send $\mathbf{v}$ to its component in U_i; it is the projection P_i of Proposition 4.4.29(a) with its target shrunk to U_i, so it is linear. Then $T = S_1 \circ P_1' + \cdots + S_k \circ P_k'$, which is linear by Proposition 2.1.17 and Proposition 2.1.23. Finally, for $\mathbf{u} \in U_i$, the decomposition of $\mathbf{u}$ has component $\mathbf{u}$ in U_i and $\mathbf{0}$ in every other summand, so $T(\mathbf{u}) = S_i(\mathbf{u}) + \sum_{j \neq i} S_j(\mathbf{0}) = S_i(\mathbf{u})$. □

When $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis and $U_i = \operatorname{span}(\mathbf{v}_i)$, a linear map $S_i : U_i \to W$ is determined by $S_i(\mathbf{v}_i)$, which can be any vector of W, and the proposition becomes Theorem 2.2.1.

Remark 4.4.32 (Block diagonal matrices). Suppose $V = U \oplus W$ is finite-dimensional, and $T \in \mathrm{L}(V)$ maps U into U and W into W. Let $T_U \in \mathrm{L}(U)$ and $T_W \in \mathrm{L}(W)$ be the operators $\mathbf{u} \mapsto T(\mathbf{u})$ and $\mathbf{w} \mapsto T(\mathbf{w})$. Glue a basis $\mathcal{B}_U$ of U and a basis $\mathcal{B}_W$ of W into a basis $\mathcal{B}$ of V (Technique 4.4.14). The image of a vector of $\mathcal{B}_U$ is a combination of $\mathcal{B}_U$ alone, and similarly for $\mathcal{B}_W$, so the matrix of T is block diagonal (Definition 2.4.33):

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} [T_U]_{\mathcal{B}_U\mathcal{B}_U} & 0 \\ 0 & [T_W]_{\mathcal{B}_W\mathcal{B}_W} \end{pmatrix}.$$

By Proposition 4.4.31, T is determined by T_U and T_W, so T can be studied through the two smaller operators. This is the strategy of the eigenvalue chapters. An operator is diagonalizable when V is a direct sum of subspaces on each of which T acts as multiplication by a scalar (Section 7.4), and the Jordan form refines such a splitting (Section 10.3). Exercise 4.4.44 computes the kernel and image of such a T.

Exercises

Exercise 4.4.33. ROUTINE In F^4, let $U = \operatorname{span}((1,1,0,0),(0,1,1,0),(0,0,1,1))$ and

$$W = \{\mathbf{x} \in F^4 \mid x_1 + x_2 + x_3 + x_4 = 0 \text{ and } x_1 = x_3\}.$$

(a) For $F = \mathbb{Q}$, find $\dim U$, $\dim W$, a basis of $U \cap W$, and $\dim(U + W)$.

(b) Repeat (a) for $F = \mathbb{F}_2$.

Solution on page 1322

Exercise 4.4.34. ROUTINE Let U, W, X be subspaces of a finite-dimensional space V. Prove or disprove:

(a) If $U \cap W = U \cap X = W \cap X = \{\mathbf{0}\}$, then $U + W + X$ is direct.

(b) If $U + W + X$ is direct, then $U + W$ is direct.

(c) If $V = U \oplus W$ and $V = U \oplus X$, then $W = X$.

(d) If $V = U \oplus W$ and $V = U \oplus X$, then $\dim W = \dim X$.

(e) If $V = U \oplus W$, then $X = (X \cap U) \oplus (X \cap W)$.

(f) If $U + W = V$ and $\dim U + \dim W = \dim V$, then $V = U \oplus W$.

Solution on page 1322

Exercise 4.4.35. ROUTINE Let $T \in \mathrm{L}(V, W)$, and let $U_1, \ldots, U_k$ be subspaces of V.

(a) Show that $\operatorname{im}\left(T\restriction_{U_1+\cdots+U_k}\right) = \operatorname{im}\left(T\restriction_{U_1}\right) + \cdots + \operatorname{im}\left(T\restriction_{U_k}\right)$.

(b) Show that if T is injective and $U_1, \ldots, U_k$ are independent, then $\operatorname{im}\left(T\restriction_{U_1}\right), \ldots, \operatorname{im}\left(T\restriction_{U_k}\right)$ are independent.

(c) Show by an example that (b) can fail if T is not injective.

Solution on page 1323

Exercise 4.4.36. ROUTINE In $\mathbb{C}^2$, let $U = \operatorname{span}((1, i))$ and $W = \operatorname{span}((1, -i))$.

(a) Show that $\mathbb{C}^2 = U \oplus W$. Find the matrix P of $\mathrm{Pr}_{U,W}$ with respect to the standard basis, and check that $P^2 = P$.

(b) Let $R(x, y) = (-y, x)$, the rotation by a quarter turn. Find scalars λ and μ with $R(\mathbf{u}) = \lambda\mathbf{u}$ for all $\mathbf{u} \in U$ and $R(\mathbf{w}) = \mu\mathbf{w}$ for all $\mathbf{w} \in W$.

Solution on page 1323

Exercise 4.4.37. STANDARD Find the error in the following argument.

Claim. Let U_1, U_2, U_3 be finite-dimensional subspaces of V with $U_i \cap U_j = \{\mathbf{0}\}$ whenever $i \neq j$. Then $\dim(U_1 + U_2 + U_3) = \dim U_1 + \dim U_2 + \dim U_3$.

Proof. Let $\mathcal{B}_i$ be a basis of U_i, and let $\mathcal{B}$ be the concatenation $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$, which spans $U_1 + U_2 + U_3$. Suppose $\mathcal{B}$ is dependent. By the linear dependence lemma, some entry $\mathbf{b}$ of $\mathcal{B}$ is a combination of the entries before it. Say $\mathbf{b}$ is in $\mathcal{B}_3$; the other cases are similar. Move the terms that come from $\mathcal{B}_3$ to the left side. The left side is now $\mathbf{b}$ minus a combination of earlier entries of $\mathcal{B}_3$, so it lies in U_3. The right side is a combination of entries of $\mathcal{B}_1$ and $\mathcal{B}_2$, so it lies in U_1 or in U_2. Hence both sides lie in $U_3 \cap U_1$ or in $U_3 \cap U_2$, which are $\{\mathbf{0}\}$. So $\mathbf{b}$ is a combination of earlier entries of $\mathcal{B}_3$, contradicting the independence of $\mathcal{B}_3$. Therefore $\mathcal{B}$ is a basis, and the claim follows.

Solution on page 1324

Exercise 4.4.38. STANDARD Find the error in the following proof of Theorem 4.4.19.

Proof. Let $\mathcal{B}$ be a basis of U and $\mathcal{C}$ a basis of W. The concatenation $\mathcal{B}, \mathcal{C}$ spans $U + W$. The vectors that occur in both $\mathcal{B}$ and $\mathcal{C}$ lie in $U \cap W$, and they form a basis of $U \cap W$. Deleting one copy of each of them from $\mathcal{B}, \mathcal{C}$ leaves a basis of $U + W$ with $\dim U + \dim W - \dim(U \cap W)$ entries.

Solution on page 1324

Exercise 4.4.39. STANDARD For finite-dimensional subspaces U, W, X of V, let

$$e(U, W, X) = \dim U + \dim W + \dim X - \dim(U \cap W) - \dim(U \cap X) - \dim(W \cap X) + \dim(U \cap W \cap X).$$

(a) Find subspaces of $\mathbb{R}^2$ with $\dim(U + W + X) \neq e(U, W, X)$.

(b) Show that $(U \cap X) + (W \cap X) \subseteq (U + W) \cap X$ always.

(c) Show that $\dim(U + W + X) = e(U, W, X)$ whenever $(U + W) \cap X = (U \cap X) + (W \cap X)$.

Solution on page 1324

Exercise 4.4.40. STANDARD Let V have dimension n, and let $U_1, \ldots, U_k$ be subspaces of V.

(a) Show that $\dim(U_1 \cap \cdots \cap U_k) \geq \dim U_1 + \cdots + \dim U_k - (k-1)n$.

(b) Let $k \leq n$. Show that the bound in (a) is attained by the subspaces $U_i = \{\mathbf{x} \in F^n \mid x_i = 0\}$, $i = 1, \ldots, k$.

(c) Deduce that if $n \geq 2$, any $n - 1$ subspaces of V of dimension $n - 1$ have a nonzero vector in common.

Solution on page 1325

Exercise 4.4.41. STANDARD Let $A, B \in F^{m \times n}$, and let $\begin{pmatrix} A & B \end{pmatrix} \in F^{m \times 2n}$ be the matrix whose columns are those of A followed by those of B.

(a) Show that $\operatorname{rank}\begin{pmatrix} A & B \end{pmatrix} = \operatorname{rank} A + \operatorname{rank} B - \dim(\operatorname{im} A \cap \operatorname{im} B)$.

(b) Deduce that $\operatorname{rank}(A + B) \leq \operatorname{rank} A + \operatorname{rank} B - \dim(\operatorname{im} A \cap \operatorname{im} B)$. This sharpens Proposition 4.3.20.

(c) Find $A, B \in F^{2 \times 2}$ with $\operatorname{im} A \cap \operatorname{im} B = \{\mathbf{0}\}$ but $\operatorname{rank}(A + B) < \operatorname{rank} A + \operatorname{rank} B$.

Solution on page 1325

Exercise 4.4.42. STANDARD Let V be finite-dimensional, and let U and W be subspaces of V.

(a) Show that there is $T \in \mathrm{L}(V)$ with $\ker T = U$ and $\operatorname{im} T = W$ if and only if $\dim U + \dim W = \dim V$.

(b) Show that T in (a) can be chosen with $T^2 = T$ if and only if $V = U \oplus W$.

(c) In $\mathbb{R}^2$, let $U = W = \operatorname{span}(\mathbf{e}_1)$. Find T as in (a), and explain directly why no such T satisfies $T^2 = T$.

Solution on page 1326

Exercise 4.4.43. STANDARD Suppose $1 + 1 \neq 0$ in F, and let $T \in \mathrm{L}(V)$ satisfy $T^2 = \mathrm{id}_V$. Let

$$V_+ = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{v}\}, \quad V_- = \{\mathbf{v} \in V \mid T(\mathbf{v}) = -\mathbf{v}\}.$$

(a) Show that $V = V_+ \oplus V_-$.

(b) Deduce that $F^{n \times n}$ is the direct sum of the subspace of symmetric matrices ($A^{\mathsf{T}} = A$) and the subspace of skew-symmetric matrices ($A^{\mathsf{T}} = -A$), and find the dimensions of both.

(c) Over $\mathbb{F}_2$, find $T \in \mathrm{L}\left(\mathbb{F}_2^2\right)$ with $T^2 = \mathrm{id}$ but $V_+ + V_- \neq \mathbb{F}_2^2$.

Solution on page 1326

Exercise 4.4.44. STANDARD Suppose $V = U \oplus W$, and let $T \in \mathrm{L}(V)$ satisfy $T(\mathbf{u}) \in U$ for all $\mathbf{u} \in U$ and $T(\mathbf{w}) \in W$ for all $\mathbf{w} \in W$.

(a) Show that $\ker T = (U \cap \ker T) \oplus (W \cap \ker T)$ and $\operatorname{im} T = \operatorname{im}(T{\restriction_U}) \oplus \operatorname{im}(T{\restriction_W})$.

(b) Deduce that if V is finite-dimensional, then $\operatorname{rank} T = \operatorname{rank}(T{\restriction_U}) + \operatorname{rank}(T{\restriction_W})$.

(c) Show that (a) can fail without the hypothesis on T: find $T \in \mathrm{L}\big(\mathbb{R}^2\big)$ with $\ker T \neq (U \cap \ker T) + (W \cap \ker T)$, where $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \operatorname{span}(\mathbf{e}_2)$.

Solution on page 1327

Exercise 4.4.45. STANDARD Let $A \in F^{m\times n}$. Prove that $F^n = \ker A \oplus \operatorname{im} A^{\mathsf{T}}$ when $F = \mathbb{R}$, and show by examples that this can fail when $F = \mathbb{C}$ and when $F = \mathbb{F}_2$.

Solution on page 1327

Exercise 4.4.46. STANDARD Let V be finite-dimensional, and let U be a subspace of V. Show that U has exactly one complement if and only if $U = \{\mathbf{0}\}$ or $U = V$.

Solution on page 1328

Exercise 4.4.47. STANDARD Let $0 \leq k \leq n$, let $c_0, \ldots, c_k \in F$ be distinct, and let

$$Z = \big\{p \in F[x]_{\leq n} \;\big|\; p(c_0) = p(c_1) = \cdots = p(c_k) = 0\big\}.$$

Show that $F[x]_{\leq n} = Z \oplus F[x]_{\leq k}$, and compute $\dim Z$ in two ways.

Solution on page 1328

Exercise 4.4.48. STANDARD Let $A, B \in F^{n\times n}$ satisfy $AB = 0$, and suppose that $A + B$ is invertible.

(a) Show that $\operatorname{rank} A + \operatorname{rank} B = n$ and $F^n = \operatorname{im} A \oplus \operatorname{im} B$.

(b) Show that $\operatorname{im} B = \ker A$.

(c) Show that if also $BA = 0$, then $\operatorname{im} A = \ker B$. Show by an example that $\operatorname{im} A = \ker B$ can fail without the hypothesis $BA = 0$.

Solution on page 1329

Exercise 4.4.49. HARD Let V be finite-dimensional with $V = U \oplus W_0$.

(a) For $S \in \mathrm{L}(W_0, U)$, let $\Gamma_S = \{\mathbf{w} + S(\mathbf{w}) \mid \mathbf{w} \in W_0\}$. Show that Γ_S is a complement of U.

(b) Show that every complement of U equals Γ_S for exactly one $S \in \mathrm{L}(W_0, U)$.

(c) Deduce that if F has q elements, $\dim V = n$, and $\dim U = k$, then U has exactly $q^{k(n-k)}$ complements in V.

(d) List the complements of $U = \operatorname{span}((1,1,0))$ in $\mathbb{F}_2^3$.

Solution on page 1329

Exercise 4.4.50. HARD Let V be finite-dimensional.

(a) Let U and W be subspaces of V with $\dim U = \dim W$. Show that U and W have a common complement, that is, a subspace X with $V = U \oplus X = W \oplus X$.

(b) Show that the three lines through $\mathbf{0}$ in $\mathbb{F}_2^2$ have no common complement.

Solution on page 1330

Exercise 4.4.51. HARD Let F be $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, and let $P_1, \ldots, P_k \in F^{n\times n}$ satisfy $P_i^2 = P_i$ for each i and $P_1 + \cdots + P_k = I_n$.

(a) Show that $F^n = \operatorname{im} P_1 \oplus \cdots \oplus \operatorname{im} P_k$.

(b) Deduce that $P_i P_j = 0$ whenever $i \neq j$.

(c) Show that the conclusion of (b) can fail over $\mathbb{F}_2$.

Solution on page 1331

Exercise 4.4.52. CHALLENGE Let F be a field with q elements. Call subspaces $U_1, \ldots, U_m$ of F^4 a **spread** if each has dimension 2 and any two of them are complements of each other.

(a) Show that a spread has at most $q^2 + 1$ members.

(b) For $F = \mathbb{F}_2$, construct a spread with 5 members. (For $M \in \mathbb{F}_2^{2\times 2}$, consider the subspace of vectors (x_1, x_2, y_1, y_2) with $(y_1, y_2) = M(x_1, x_2)$.)

Solution on page 1331

4.5 Isomorphisms and change of basis

A basis of length n turns V into F^n: every vector gets a column of coordinates, and every linear map gets a matrix. But a vector space has many bases, and nothing in V prefers one of them. This section studies that freedom. First we make precise when two vector spaces differ only in the names of their vectors; for finite-dimensional spaces this happens exactly when their dimensions agree. Then we work out how coordinates and matrices change with the basis. This takes one formula, and it leads to similarity: two square matrices are similar exactly when they are matrices of one operator in two bases. Whatever all matrices of an operator have in common, such as their rank and their trace, is then a property of the operator itself, and the freedom to pick the basis becomes a proof technique.

Throughout, F is a field.

4.5.1 Isomorphisms

The coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ of a basis $\mathcal{B}$ of length n is a linear bijection $V \to F^n$ (Proposition 1.6.19, Proposition 1.6.16). Through it, V and F^n have the same arithmetic: adding two vectors adds their coordinate columns, and scaling a vector scales its column. So V is F^n with different names for the vectors. Linear bijections are exactly the maps that rename vectors in this way.

Definition 4.5.1 (Isomorphism). An **isomorphism** is an invertible linear map $T \in \mathrm{L}(V, W)$. If there is an isomorphism from V to W, then V and W are **isomorphic**, and we write $V \cong W$.

By Proposition 2.7.7, a linear map is an isomorphism if and only if it is bijective, since the inverse function of a bijective linear map is automatically linear. So to show that T is an isomorphism, we may either show that it is injective and surjective, or exhibit a linear two-sided inverse.

Example 4.5.2 (Isomorphisms).

(a) For every basis $\mathcal{B}$ of V of length $n \geq 1$, the coordinate map $V \to F^n$ is an isomorphism, with inverse $(a_1, \ldots, a_n) \mapsto a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n$ (Example 2.5.14(c)). For $V = F[x]_{\leq n}$ and the basis $1, x, \ldots, x^n$, it is the map $a_0 + a_1x + \cdots + a_nx^n \mapsto (a_0, \ldots, a_n)$, so $F[x]_{\leq n} \cong F^{n+1}$.

(b) Stacking the columns of a matrix, $A \mapsto (a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{mn})$, is a linear bijection $F^{m\times n} \to F^{mn}$: both spaces use entrywise operations, and unstacking is an inverse. So $F^{2\times 3} \cong F^6$.

(c) The map $(x, y) \mapsto (x, y, 0)$ is an isomorphism from F^2 onto the plane $P = \{(x, y, 0) \mid x, y \in F\}$ of F^3, with inverse $(x, y, 0) \mapsto (x, y)$. This is the sense in which P is "a copy of F^2" (Exercise 4.1.41(f)), even though F^2 is not a subset of F^3.

(d) The map $\mathrm{L}(F, V) \to V$, $T \mapsto T(1)$, is an isomorphism. It is linear because the operations on $\mathrm{L}(F, V)$ are pointwise. Its inverse sends $\mathbf{v}$ to the map $a \mapsto a\mathbf{v}$, since every $T \in \mathrm{L}(F, V)$ satisfies $T(a) = aT(1)$.

Non-example 4.5.3 (Bijections that are not isomorphisms).

(a) Complex conjugation $z \mapsto \overline{z}$ is a bijection $\mathbb{C} \to \mathbb{C}$ that respects addition and multiplication by real numbers, so it is an isomorphism of vector spaces over $\mathbb{R}$. It is not linear over $\mathbb{C}$: $\overline{i \cdot 1} = -i$, while $i \cdot \overline{1} = i$. Whether a map is an isomorphism depends on the field, as dimension does (Remark 4.1.6).

(b) The bijection $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$, is not additive, since $(1 + 1)^3 = 8 \neq 2 = 1^3 + 1^3$. A bijection between vector spaces renames the vectors, but unless it respects the operations, the arithmetic does not survive the renaming.

Proposition 4.5.4 (Isomorphism is an equivalence relation). For all vector spaces U, V, and W over F:

(a) $V \cong V$;
(b) if $V \cong W$, then $W \cong V$;
(c) if $U \cong V$ and $V \cong W$, then $U \cong W$.

Proof. ROUTINE

Uses. Definition 4.5.1 (Isomorphism), Proposition 2.5.17 (Algebra of inverses), Example 2.5.14 (Invertible maps and matrices)
Hint 1. Each part asks for one isomorphism. Which maps are available?
Hint 2. Use id_V, the inverse of a given isomorphism, and the composite of two given isomorphisms, and quote Proposition 2.5.17.
Sketch. id_V is invertible. If $T : V \to W$ is invertible, so is $T^{-1} : W \to V$. If $S : U \to V$ and $T : V \to W$ are invertible, so is $TS : U \to W$.

(a) id_V is its own inverse (Example 2.5.14(a)). **(b)** If $T \in \mathrm{L}(V, W)$ is an isomorphism, then $T^{-1} \in \mathrm{L}(W, V)$ is invertible by Proposition 2.5.17(a). **(c)** If $S \in \mathrm{L}(U, V)$ and $T \in \mathrm{L}(V, W)$ are isomorphisms, then $TS \in \mathrm{L}(U, W)$ is invertible by Proposition 2.5.17(b), with inverse $S^{-1}T^{-1}$. □

So the vector spaces over F fall into isomorphism classes (Section A.3). Within a class, every statement that is phrased in terms of the vector space operations transfers from one space to another. Here are the statements we use most.

Proposition 4.5.5 (Isomorphisms preserve linear structure). Let $T \in \mathrm{L}(V, W)$ be an isomorphism.

(a) A list $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ in V is linearly independent if and only if $T(\mathcal{L}) = T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$ is. It spans V if and only if $T(\mathcal{L})$ spans W, and it is a basis of V if and only if $T(\mathcal{L})$ is a basis of W.

(b) For every subspace U of V, $\mathrm{im}(T{\restriction}_U)$ is a subspace of W, and $T{\restriction}_U$ is an isomorphism from U onto $\mathrm{im}(T{\restriction}_U)$.

(c) V is finite-dimensional if and only if W is, and in that case $\dim V = \dim W$. More generally, $\dim \mathrm{im}(T{\restriction}_U) = \dim U$ for every finite-dimensional subspace U of V.

Proof. ROUTINE

Uses. Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 2.5.17 (Algebra of inverses), Proposition 4.2.3 (Properties of restriction), Proposition 2.6.8 (Kernels and images are subspaces), Proposition 2.7.7 (Bijective linear maps are invertible), Definition 4.1.5 (Dimension)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Linearity gives $c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) = T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k)$. For the "only if" directions use that T is injective or surjective; for the "if" directions apply the same argument to T^{-1}.
Hint 2. For (b), a restriction of an injective map is injective, and every map is onto its own image. For (c), apply (a) to a basis.
Sketch. (a) A relation $\sum c_iT(\mathbf{v}_i) = \mathbf{0}$ says $T(\sum c_i\mathbf{v}_i) = \mathbf{0}$, so $\sum c_i\mathbf{v}_i = \mathbf{0}$ by injectivity. If $\mathcal{L}$ spans, every $\mathbf{w} = T\big(T^{-1}\mathbf{w}\big)$ is T of a combination of $\mathcal{L}$. The converses are the same statements for T^{-1}. (b) $T{\restriction_U}$ is linear and injective, and bijective onto its image. (c) A finite spanning list or basis of V is carried to one of W of the same length, and back by T^{-1}; for U, apply this to the isomorphism of (b).

Since T is linear, for all scalars $c_1, \ldots, c_k$,

$$c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) = T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k). \tag{4.5.1}$$

(a) Suppose $\mathcal{L}$ is independent, and $\sum_i c_iT(\mathbf{v}_i) = \mathbf{0}$. By (4.5.1), $\sum_i c_i\mathbf{v}_i \in \ker T$, and $\ker T = \{\mathbf{0}\}$ because T is injective (Theorem 2.6.11). So $\sum_i c_i\mathbf{v}_i = \mathbf{0}$, and all $c_i = 0$. Suppose $\mathcal{L}$ spans V, and let $\mathbf{w} \in W$. Then $T^{-1}(\mathbf{w}) = \sum_i c_i\mathbf{v}_i$ for some scalars, and by (4.5.1), $\mathbf{w} = T\big(T^{-1}(\mathbf{w})\big) = \sum_i c_iT(\mathbf{v}_i)$. So $T(\mathcal{L})$ spans W.

For the converses, recall that $T^{-1} \in \mathrm{L}(W, V)$ is an isomorphism (Proposition 2.5.17(a)) and that $T^{-1}(T(\mathcal{L})) = \mathcal{L}$. Applying what we just proved to T^{-1} and the list $T(\mathcal{L})$ shows that if $T(\mathcal{L})$ is independent or spanning, then so is $\mathcal{L}$. The statement about bases combines the two.

(b) By Proposition 4.2.3(c), $T{\restriction_U} \in \mathrm{L}(U, W)$, so its image is a subspace of W (Proposition 2.6.8). It is injective by Proposition 4.2.3(a), and it maps U onto its own image. So it is a linear bijection $U \to \operatorname{im}(T{\restriction_U})$, hence an isomorphism (Proposition 2.7.7).

(c) If $\mathcal{L}$ is a finite list spanning V, then $T(\mathcal{L})$ is a finite list spanning W by (a); applying this to T^{-1} gives the converse. If V is finite-dimensional with basis $\mathcal{B}$, then $T(\mathcal{B})$ is a basis of W of the same length by (a), so $\dim W = \dim V$. Finally, if U is a finite-dimensional subspace of V, apply this to the isomorphism $T{\restriction_U} : U \to \operatorname{im}(T{\restriction_U})$ of (b). □

Now the main theorem. By (c), isomorphic spaces have equal dimensions. The converse is the surprise: for finite-dimensional spaces, dimension is the only thing that can tell two vector spaces apart.

Theorem 4.5.6 (Classification by dimension). Let V and W be finite-dimensional vector spaces over F. Then $V \cong W$ if and only if $\dim V = \dim W$. In particular, for $n \geq 1$, a vector space is isomorphic to F^n if and only if it is finite-dimensional of dimension n.

We give two proofs. The first works with bases directly; the second assembles the isomorphism from coordinate maps and gets the dimension count from rank–nullity.

Proof (with bases). ROUTINE

Uses. Proposition 4.5.5 (Isomorphisms preserve linear structure), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Proposition 2.7.7 (Bijective linear maps are invertible), Definition 4.1.5 (Dimension)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. For "only if", what does an isomorphism do to a basis?
Hint 2. For "if", choose bases $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of W. Which linear map is the obvious candidate?
Hint 3. The map with $T(\mathbf{v}_i) = \mathbf{w}_i$ sends a basis to a basis. Use Proposition 2.6.22(c).
Sketch. An isomorphism carries a basis to a basis of the same length (Proposition 4.5.5). Conversely, if both spaces have bases of length n, the linear map sending the i-th vector of one to the i-th vector of the other carries a basis to a basis, so it is bijective. The last claim follows since $\dim F^n = n$.

If $V \cong W$, then $\dim V = \dim W$ by Proposition 4.5.5(c).

Conversely, suppose $\dim V = \dim W = n$. If $n = 0$, then $V = \{\mathbf{0}\}$ and $W = \{\mathbf{0}\}$, and the zero map is a bijection between them. If $n \geq 1$, choose bases $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of W. By Theorem 2.2.1 there is a linear map $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i. It carries a basis of V to a basis of W, so it is bijective by Proposition 2.6.22(c), and an isomorphism by Proposition 2.7.7.

For the last statement: $\dim F^n = n$, so a finite-dimensional space of dimension n is isomorphic to F^n by what we just proved. Conversely, if $V \cong F^n$, then V is finite-dimensional with $\dim V = n$ by Proposition 4.5.5(c).□

Proof (with coordinate maps and rank–nullity). ROUTINE

> *Uses.* Example 2.5.14 (Invertible maps and matrices), Proposition 4.5.4 (Isomorphism is an equivalence relation), Theorem 4.2.11 (Rank–nullity), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 4.5.5 (Isomorphisms preserve linear structure)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* For "if", both V and W are isomorphic to the same space. Which one?
> *Hint 2.* The coordinate maps give $V \cong F^n$ and $W \cong F^n$; use Proposition 4.5.4. For "only if", what are $\dim \ker T$ and $\dim \operatorname{im} T$ for an isomorphism T?
> *Sketch.* If $\dim V = \dim W = n \geq 1$, the coordinate maps give $V \cong F^n \cong W$, and the composite $C_{\mathcal{C}}^{-1} C_{\mathcal{B}}$ is an isomorphism. If $T : V \to W$ is an isomorphism, rank–nullity gives $\dim V = \dim \ker T + \dim \operatorname{im} T = 0 + \dim W$.

Suppose $\dim V = \dim W = n$. For $n = 0$ both spaces are $\{\mathbf{0}\}$, as in the first proof. For $n \geq 1$, choose bases $\mathcal{B}$ of V and $\mathcal{C}$ of W. The coordinate maps $C_{\mathcal{B}} : V \to F^n$ and $C_{\mathcal{C}} : W \to F^n$ are isomorphisms (Example 2.5.14(c)), so by Proposition 4.5.4, $C_{\mathcal{C}}^{-1} C_{\mathcal{B}} : V \to W$ is an isomorphism. Explicitly, it sends $\sum_i a_i \mathbf{v}_i$ to $\sum_i a_i \mathbf{w}_i$: it is the map of the first proof.

Conversely, let $T : V \to W$ be an isomorphism. Then $\ker T = \{\mathbf{0}\}$ (Theorem 2.6.11) and $\operatorname{im} T = W$, so by Theorem 4.2.11(b),

$$\dim V = \dim \ker T + \dim \operatorname{im} T = 0 + \dim W.$$

The last statement follows as in the first proof. □

> *Intuition (Choosing an isomorphism is choosing a basis).* The theorem says that up to renaming, the only n-dimensional vector space is F^n. It does not say that there is a preferred renaming. In fact the isomorphisms $V \to F^n$ are exactly the coordinate maps of the bases of V (Exercise 1.6.37): an isomorphism φ is the coordinate map of the basis $\varphi^{-1}(\mathbf{e}_1), \ldots, \varphi^{-1}(\mathbf{e}_n)$. So there are as many ways to identify V with F^n as there are bases, and in a space like $F[x]_{\leq 2}$, or a plane in $\mathbb{R}^3$, none of them is singled out. The rest of this section is about how the identifications compare.

Technique 4.5.7 (Transport along an isomorphism). An isomorphism $T : V \to W$ carries every statement about V that is phrased in terms of the vector space operations (independence, span, bases, subspaces and their dimensions, kernels and images of linear maps) to the corresponding statement about W, and back. So to answer such a question about an unfamiliar space, transport it along an isomorphism to a space where it is easier, answer it there, and transport the answer back.

Reach for it when the space is made of polynomials, matrices, functions, or linear maps, and the question only involves linear structure. The usual target is F^n, through coordinates: Proposition 1.6.20 was an early instance, and Theorem 4.5.6 says that every n-dimensional space can be transported to F^n. The canonical example is the computation of $\dim \mathrm{L}(V, W)$ below, which transports the question to matrices. Be careful to transport only statements made of the linear operations: a property such as "every entry is positive" or "the polynomial has a root" is not preserved by isomorphisms.
Used in. Theorem 4.5.8, Exercise 4.8.12, Exercise 4.5.41, Exercise 4.5.42.

The dimension of $\mathrm{L}(V, W)$

In Section 2.4 we saw that $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is a linear bijection $\mathrm{L}(V, W) \to F^{m \times n}$ (Proposition 2.4.12). In the language of this section, it is an isomorphism, and Technique 4.5.7 gives the dimension of $\mathrm{L}(V, W)$.

Theorem 4.5.8 (Dimension of $\mathrm{L}(V,W)$). Let V and W be finite-dimensional with $\dim V = n$ and $\dim W = m$. Then $\mathrm{L}(V,W)$ is finite-dimensional, and

$$\dim \mathrm{L}(V,W) = mn.$$

If $n, m \geq 1$ and $\mathcal{B}$, $\mathcal{C}$ are bases of V, W, then $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is an isomorphism $\mathrm{L}(V,W) \to F^{m\times n}$.

Proof (by transport to matrices). ROUTINE

> *Uses.* Proposition 2.4.12 (Matrices of sums and scalar multiples), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 4.5.5 (Isomorphisms preserve linear structure), Example 4.1.7 (Dimensions of the standard spaces), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
> *Techniques.* Transport along an isomorphism (Technique 4.5.7)
> *Hint 1.* Dispose of $n = 0$ or $m = 0$ first: what are the linear maps out of $\{\mathbf{0}\}$, or into $\{\mathbf{0}\}$?
> *Hint 2.* For $n, m \geq 1$, Proposition 2.4.12 gives a linear bijection to a space whose dimension you know.
> *Sketch.* If V or W is zero, the only linear map $V \to W$ is 0, so $\dim \mathrm{L}(V,W) = 0 = mn$. Otherwise $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is a linear bijection onto $F^{m\times n}$, which has dimension mn, and isomorphisms preserve dimension.

If $n = 0$, then $V = \{\mathbf{0}\}$, and every linear map $T : V \to W$ is zero since $T(\mathbf{0}) = \mathbf{0}$ (Proposition 2.1.4). If $m = 0$, then every map into $W = \{\mathbf{0}\}$ is zero. In both cases $\mathrm{L}(V,W) = \{0\}$ has dimension $0 = mn$.

Now let $n, m \geq 1$. By Proposition 2.4.12 and Theorem 2.3.20, $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ is a linear bijection from $\mathrm{L}(V,W)$ onto $F^{m\times n}$, that is, an isomorphism. Since $\dim F^{m\times n} = mn$ (Example 4.1.7(a)), Proposition 4.5.5(c) shows that $\mathrm{L}(V,W)$ is finite-dimensional of dimension mn. □

The second proof writes down a basis of $\mathrm{L}(V,W)$. It is longer, but it shows what the isomorphism does: it matches a natural basis of $\mathrm{L}(V,W)$ with the standard basis of $F^{m\times n}$.

Proof (with an explicit basis). STANDARD

> *Uses.* Theorem 2.2.1 (Linear maps are determined by a basis), Definition 2.3.1 (Matrix of a linear map), Definition 1.6.1 (Basis), Definition 1.5.1 (Linear independence), Definition 2.1.21 (Pointwise operations)
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
> *Hint 1.* For bases $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \dots, \mathbf{w}_m$ of W, define mn maps by saying what each does on the basis of V. Which maps correspond to the matrices U_{ij}?
> *Hint 2.* Let φ_{ij} send $\mathbf{v}_j$ to $\mathbf{w}_i$ and every other $\mathbf{v}_k$ to $\mathbf{0}$. To prove that they span, compare T with $\sum a_{ij}\varphi_{ij}$ on the basis. To prove independence, evaluate a relation at $\mathbf{v}_k$.
> *Sketch.* Let φ_{ij} be the linear map with $\varphi_{ij}(\mathbf{v}_k) = \mathbf{w}_i$ if $k = j$ and $\mathbf{0}$ otherwise. If $T(\mathbf{v}_k) = \sum_i a_{ik}\mathbf{w}_i$, then $\sum_{i,j} a_{ij}\varphi_{ij}$ agrees with T on each $\mathbf{v}_k$, so they are equal. If $\sum c_{ij}\varphi_{ij} = 0$, evaluating at $\mathbf{v}_k$ gives $\sum_i c_{ik}\mathbf{w}_i = \mathbf{0}$, so all $c_{ik} = 0$. So the mn maps form a basis.

The zero cases are as in the first proof, so let $n, m \geq 1$, and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ and $\mathbf{w}_1, \dots, \mathbf{w}_m$ be bases of V and W. For $1 \leq i \leq m$ and $1 \leq j \leq n$, let $\varphi_{ij} \in \mathrm{L}(V,W)$ be the linear map with

$$\varphi_{ij}(\mathbf{v}_k) = \begin{cases} \mathbf{w}_i & \text{if } k = j \\ \mathbf{0} & \text{if } k \neq j \end{cases}$$

(Theorem 2.2.1). We show that the mn maps φ_{ij}, listed in any order, form a basis of $\mathrm{L}(V,W)$.

They span. Let $T \in \mathrm{L}(V,W)$, and write $T(\mathbf{v}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{w}_i$ for each k. Let $S = \sum_{i,j} a_{ij}\varphi_{ij}$. Since the operations on $\mathrm{L}(V,W)$ are pointwise,

$$S(\mathbf{v}_k) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\varphi_{ij}(\mathbf{v}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{w}_i = T(\mathbf{v}_k)$$

for each k, because only the terms with $j = k$ survive. Two linear maps that agree on a basis are equal (Theorem 2.2.1), so $T = S$.

They are independent. Suppose $\sum_{i,j} c_{ij}\varphi_{ij} = 0$. Evaluating at $\mathbf{v}_k$ as above gives $\sum_{i=1}^{m} c_{ik}\mathbf{w}_i = \mathbf{0}$, so $c_{1k} = \cdots = c_{mk} = 0$ because $\mathbf{w}_1, \dots, \mathbf{w}_m$ is independent. This holds for every k.

So $\mathrm{L}(V,W)$ has a basis of mn maps, and $\dim \mathrm{L}(V,W) = mn$. □

The coefficients a_{ij} in the second proof are the entries of $[T]_{\mathcal{CB}}$, and $\left[\varphi_{ij}\right]_{\mathcal{CB}} = U_{ij}$, since φ_{ij} sends the j-th basis vector to $\mathbf{w}_i$ and the others to $\mathbf{0}$. So the isomorphism of the first proof carries the basis φ_{ij} to the standard basis U_{ij} of $F^{m\times n}$.

Two special cases come up constantly. For an operator space, $\dim \mathrm{L}(V) = n^2$; so for every $T \in \mathrm{L}(V)$, the $n^2 + 1$ operators $\mathrm{id}_V, T, \ldots, T^{n^2}$ are dependent, which is Proposition 4.1.31(b) again. For $W = F$, $\dim \mathrm{L}(V,F) = n$. The linear maps $V \to F$ are the linear functionals, and $\mathrm{L}(V,F)$ is the dual space of V, the subject of Section 13.1.

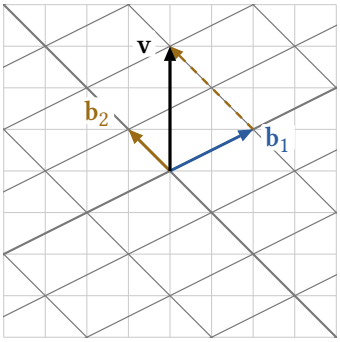
Exercise 4.5.9. ROUTINE Sort the following vector spaces over F into isomorphism classes.

(a) $F^{2\times 3}$.
(b) $F[x]_{\leq 5}$.
(c) $\mathrm{L}\left(F^2, F^3\right)$.
(d) The upper triangular matrices in $F^{3\times 3}$.
(e) $\left\{p \in F[x]_{\leq 6} \;\middle|\; p(0) = 0\right\}$.
(f) $\mathrm{L}\left(F^3, F\right)$.
(g) $\left\{A \in F^{2\times 2} \;\middle|\; \operatorname{tr} A = 0\right\}$.

Solution on page 1332

4.5.2 Change of coordinates

Intuition (One vector, two grids). In $\mathbb{R}^2$, let $\mathbf{b}_1 = (2,1)$ and $\mathbf{b}_2 = (-1,1)$. The vector $\mathbf{v} = (0,3)$ has standard coordinates $(0,3)$. It is also $\mathbf{b}_1 + 2\mathbf{b}_2$, so in the basis $\mathcal{B} = \mathbf{b}_1, \mathbf{b}_2$ its coordinates are $(1,2)$. Draw the grid of integer combinations of $\mathbf{b}_1$ and $\mathbf{b}_2$ on top of the standard grid: the arrow $\mathbf{v}$ is the same arrow, and each grid reads off its own pair of numbers. Changing the basis does not move any vector. It changes the grid on which we read the vectors, and the question is how the two readings are related.



The vector $\mathbf{v} = (0,3)$ on the standard grid and on the grid of $\mathcal{B} = (2,1), (-1,1)$. Read on the first grid, $[\mathbf{v}]_{\mathcal{E}} = (0,3)$; read on the second, $\mathbf{v} = \mathbf{b}_1 + 2\mathbf{b}_2$, so $[\mathbf{v}]_{\mathcal{B}} = (1,2)$. In the interactive version, drag $\mathbf{b}_1$, $\mathbf{b}_2$, or $\mathbf{v}$ (or focus one and use the arrow keys); the panel shows $[\mathrm{id}]_{\mathcal{EB}}$ and the conversion $[\mathbf{v}]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{BE}}[\mathbf{v}]_{\mathcal{E}}$. When $\mathbf{b}_1$ and $\mathbf{b}_2$ become parallel, the second grid collapses and the coordinates are no longer defined. (Interactive in the web version.)

The conversion between the two readings is the matrix of the identity map with different bases on the two sides, which we met in Section 2.3.5. We give it a name.

Definition 4.5.10 (Change-of-coordinates matrix). Let V be a vector space with $\dim V = n \geq 1$, and let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mathring{\mathcal{B}} = \mathring{\mathbf{b}}_1, \ldots, \mathring{\mathbf{b}}_n$ be bases of V. The **change-of-coordinates matrix** from $\mathcal{B}$ to $\mathring{\mathcal{B}}$ is

$$[\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}} = \left([\mathbf{b}_1]_{\mathring{\mathcal{B}}}\ [\mathbf{b}_2]_{\mathring{\mathcal{B}}}\ \cdots\ [\mathbf{b}_n]_{\mathring{\mathcal{B}}}\right) \in F^{n\times n},$$

whose j-th column holds the $\mathring{\mathcal{B}}$-coordinates of the j-th vector of $\mathcal{B}$. By Theorem 2.3.9,

$$[\mathbf{v}]_{\mathring{\mathcal{B}}} = [\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \quad \text{for every } \mathbf{v} \in V.$$

As with every matrix of a map, the subscripts are target first: $[\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}}$ eats $\mathcal{B}$-coordinates and returns $\mathring{\mathcal{B}}$-coordinates. Books disagree on which of the two matrices to call "the change-of-basis matrix from $\mathcal{B}$ to $\mathring{\mathcal{B}}$", so it is safest to trust the subscripts and not the words. We drop the subscript V from id_V when the space is clear.

Example 4.5.11 (Two grids in the plane). Let F be a field in which $3 \neq 0$, let $\mathcal{E}$ be the standard basis of F^2, and let $\mathcal{B} = (2,1), (-1,1)$. The columns of $[\mathrm{id}]_{\mathcal{E}\mathcal{B}}$ are the standard coordinates of the vectors of $\mathcal{B}$, which are those vectors themselves. The columns of $[\mathrm{id}]_{\mathcal{B}\mathcal{E}}$ are the $\mathcal{B}$-coordinates of $\mathbf{e}_1$ and $\mathbf{e}_2$, which we find by solving: $\mathbf{e}_1 = \frac{1}{3}(2,1) - \frac{1}{3}(-1,1)$ and $\mathbf{e}_2 = \frac{1}{3}(2,1) + \frac{2}{3}(-1,1)$. So

$$[\mathrm{id}]_{\mathcal{E}\mathcal{B}} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \quad [\mathrm{id}]_{\mathcal{B}\mathcal{E}} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

For $\mathbf{v} = (0,3)$, we get $[\mathbf{v}]_{\mathcal{B}} = \frac{1}{3}(0+3, 0+6) = (1,2)$, as in the picture. (If $3 = 0$ in F, then $(-1,1) = (2,1)$ and $\mathcal{B}$ is not a basis.)

The general rules are consequences of the composition theorem.

Proposition 4.5.12 (Calculus of change of coordinates). Let $\mathcal{B}$, $\mathcal{B}'$, and $\mathcal{B}''$ be bases of V, where $\dim V = n \geq 1$.

(a) $[\mathrm{id}]_{\mathcal{B}''\mathcal{B}'}[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}''\mathcal{B}}$, and $[\mathrm{id}]_{\mathcal{B}\mathcal{B}} = I_n$.
(b) $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ is invertible, with inverse $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'}$.
(c) $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ is the only matrix $P \in F^{n\times n}$ with $P[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}'}$ for every $\mathbf{v} \in V$.

Proof. ROUTINE

Uses. Theorem 2.4.2 (The matrix of a composite), Example 2.3.6 (The zero map and the identity map), Definition 2.5.12 (Invertible), Theorem 2.3.9 (The matrix of a map computes the map), Definition 2.3.1 (Matrix of a linear map)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. For (a), $\mathrm{id} \circ \mathrm{id} = \mathrm{id}$. Which theorem turns a composite into a product of matrices?
Hint 2. (b) follows from (a) with $\mathcal{B}'' = \mathcal{B}$. For (c), apply P to the coordinates of a basis vector $\mathbf{b}_j$.
Sketch. (a) is Theorem 2.4.2 for $\mathrm{id} \circ \mathrm{id} = \mathrm{id}$, with the middle basis $\mathcal{B}'$. (b) Taking $\mathcal{B}'' = \mathcal{B}$ in (a), and then exchanging the roles of $\mathcal{B}$ and $\mathcal{B}'$, both products are I_n. (c) Existence is Theorem 2.3.9; if P converts coordinates, then $P\mathbf{e}_j = P[\mathbf{b}_j]_{\mathcal{B}} = [\mathbf{b}_j]_{\mathcal{B}'}$ is the j-th column of $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$.

(a) By Theorem 2.4.2 applied to $\mathrm{id}_V \circ \mathrm{id}_V = \mathrm{id}_V$, with $\mathcal{B}$ on the domain, $\mathcal{B}'$ in the middle, and $\mathcal{B}''$ on the target, $[\mathrm{id}]_{\mathcal{B}''\mathcal{B}'}[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}''\mathcal{B}}$. And $[\mathrm{id}]_{\mathcal{B}\mathcal{B}} = I_n$ by Example 2.3.6.

(b) By (a) with $\mathcal{B}'' = \mathcal{B}$, $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'}[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}\mathcal{B}} = I_n$. Exchanging the roles of $\mathcal{B}$ and $\mathcal{B}'$ gives $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}[\mathrm{id}]_{\mathcal{B}\mathcal{B}'} = I_n$. So $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'}$ is a two-sided inverse of $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ (Definition 2.5.12). (This is also Example 2.5.21.)

(c) $P = [\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ converts coordinates by Theorem 2.3.9. Conversely, suppose $P[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}'}$ for every $\mathbf{v}$. Taking $\mathbf{v} = \mathbf{b}_j$, the j-th vector of $\mathcal{B}$, whose $\mathcal{B}$-coordinates are $\mathbf{e}_j$, gives $P\mathbf{e}_j = [\mathbf{b}_j]_{\mathcal{B}'}$. The left side is the j-th column of P, and the right side is the j-th column of $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ (Definition 2.3.1). □

In F^n there is a standard way to compute. For a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of F^n, $[\mathrm{id}]_{\mathcal{E}\mathcal{B}}$ is the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_n$, since $[\mathbf{b}_j]_{\mathcal{E}} = \mathbf{b}_j$. For two bases $\mathcal{B}$ and $\mathring{\mathcal{B}}$ of F^n, route through the standard basis: by Proposition 4.5.12,

$$[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = [\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{E}}[\mathrm{id}]_{\mathcal{E}\mathcal{B}} = \left([\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}}\right)^{-1}[\mathrm{id}]_{\mathcal{E}\mathcal{B}}.$$

To compute $M^{-1}N$ without computing M^{-1} first, row reduce $(M \mid N)$: if $GM = I_n$, then $G = M^{-1}$, and the reduction ends at $\left(I_n \mid M^{-1}N\right)$ (Theorem 3.3.10).

Exercise 4.5.13. ROUTINE Over $\mathbb{F}_5$, let $\mathcal{B} = (1,2),(3,4)$ and $\mathring{\mathcal{B}} = (1,1),(1,4)$.

(a) Show that both lists are bases of $\mathbb{F}_5^2$, and compute $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}}$.
(b) The vector $\mathbf{v}$ has $[\mathbf{v}]_{\mathcal{B}} = (1,1)$. Find $\mathbf{v}$ and $[\mathbf{v}]_{\mathring{\mathcal{B}}}$, and check your answer.

Solution on page 1332

The next example changes coordinates in a polynomial space, where the new coordinates have a meaning of their own.

Example 4.5.14 (Expanding around a point). Let $V = F[x]_{\leq 2}$, with $\mathcal{B} = 1, x, x^2$ and $\mathring{\mathcal{B}} = 1, x-1, (x-1)^2$. Since

$$x = 1 + (x-1), \quad x^2 = 1 + 2(x-1) + (x-1)^2,$$

the three vectors of $\mathring{\mathcal{B}}$ span V (their span contains 1, x, and x^2), so they form a basis (Corollary 4.1.11(b)). The displayed identities give the columns of the change-of-coordinates matrix, and expanding $(x-1)$ and $(x-1)^2$ in powers of x gives the columns of its inverse:

$$[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad [\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

(Multiply them to check that the product is I_3.) So for $p = a + bx + cx^2$,

$$[p]_{\mathring{\mathcal{B}}} = (a+b+c, b+2c, c), \quad \text{that is,} \quad p = (a+b+c) + (b+2c)(x-1) + c(x-1)^2.$$

Over $\mathbb{R}$, the new coordinates are $p(1)$, $p'(1)$, and $p''(1)/2$: changing to the basis $\mathring{\mathcal{B}}$ is Taylor expansion at 1.

The matrix $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}}$ also appeared in Example 2.4.13, as the matrix in the basis $\mathcal{B}$ of the shift $p(x) \mapsto p(x+1)$. The next remark explains why.

Remark 4.5.15 (Moving the grid or moving the vectors). A change-of-coordinates matrix is also the matrix of an operator in a single basis. Given bases $\mathcal{B}$ and $\mathring{\mathcal{B}}$ of V, let $R \in \mathrm{L}(V)$ be the operator with $R(\mathbf{b}_j) = \mathring{\mathbf{b}}_j$ for each j (Theorem 2.2.1). The j-th column of $[R]_{\mathcal{B}\mathcal{B}}$ is $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}}$, which is also the j-th column of $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$. So

$$[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = [R]_{\mathcal{B}\mathcal{B}}, \quad \text{and hence} \quad [\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = ([R]_{\mathcal{B}\mathcal{B}})^{-1} = \left[R^{-1}\right]_{\mathcal{B}\mathcal{B}}$$

by Proposition 4.5.12(b) and Proposition 2.5.20(b) (R is invertible because it carries a basis to a basis). In Example 4.5.14, $R(x^j) = (x-1)^j$, so R is the substitution $p(x) \mapsto p(x-1)$ and R^{-1} is $p(x) \mapsto p(x+1)$.

So there are two ways to read the same matrix. In the passive reading, the vectors stay put and the grid changes from $\mathcal{B}$ to $\mathring{\mathcal{B}}$. In the active reading, the grid stays put and the vectors move by R^{-1}. Moving the grid forward by R looks, in coordinates, exactly like moving the vectors back by R^{-1}.

So every change-of-coordinates matrix is invertible. Conversely, every invertible matrix changes coordinates, from any given basis to a suitable new one. This is what makes the change-of-basis formula below reversible.

Proposition 4.5.16 (Every invertible matrix changes coordinates). Let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V, and let $S = (s_{ij}) \in F^{n\times n}$ be invertible.

(a) There is exactly one basis $\mathring{\mathcal{B}}$ of V with $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$, namely

$$\mathring{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + s_{2j}\mathbf{b}_2 + \cdots + s_{nj}\mathbf{b}_n \quad (1 \le j \le n).$$

(b) There is exactly one basis $\mathcal{B}'$ of V with $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = S$.

Proof (through coordinates). ROUTINE

Uses. Definition 4.5.10 (Change-of-coordinates matrix), Definition 1.6.14 (Coordinates), Corollary 3.5.25 (The invertible matrix theorem), Theorem 4.1.10 (Independent lists are short), Proposition 1.6.20 (Coordinates translate questions about V into F^n), Proposition 4.5.12 (Calculus of change of coordinates), Proposition 2.5.17 (Algebra of inverses)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. The j-th column of $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$ is $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}}$. So if $\mathring{\mathcal{B}}$ exists, its vectors are forced. Which vectors?
Hint 2. With $\mathring{\mathbf{b}}_j$ defined as in the statement, $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}}$ is the j-th column of S. Why do the columns of an invertible matrix form a basis of F^n, and how does that transfer to V?
Hint 3. For (b), apply (a) to S^{-1} and use Proposition 4.5.12(b).
Sketch. (a) The equation $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$ says $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}} = \mathbf{s}_j$ (the j-th column of S) for each j, which determines $\mathring{\mathbf{b}}_j$. These vectors form a basis because their coordinate vectors, the columns of the invertible S, form a basis of F^n (Proposition 1.6.20(c)). (b) $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = S$ if and only if $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'} = S^{-1}$; apply (a) to S^{-1}.

(a) Let $\mathbf{s}_j$ be the j-th column of S. For any basis $\mathring{\mathcal{B}}$, the j-th column of $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$ is $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}}$ (Definition 4.5.10). So $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$ holds exactly when $\left[\mathring{\mathbf{b}}_j\right]_{\mathcal{B}} = \mathbf{s}_j$ for every j, that is, when $\mathring{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots + s_{nj}\mathbf{b}_n$ (Definition 1.6.14). This proves uniqueness, and it remains to show that these vectors form a basis. Since S is invertible, its columns are independent (Corollary 3.5.25), and n independent vectors in F^n form a basis (Theorem 4.1.10(b)). The coordinate vectors of $\mathring{\mathbf{b}}_1, \ldots, \mathring{\mathbf{b}}_n$ are $\mathbf{s}_1, \ldots, \mathbf{s}_n$, so $\mathring{\mathcal{B}}$ is a basis of V by Proposition 1.6.20(c).

(b) By Proposition 4.5.12(b), for a basis $\mathcal{B}'$ we have $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}} = S$ if and only if $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'} = S^{-1}$. The matrix S^{-1} is invertible (Proposition 2.5.17(a)), so by (a) there is exactly one such basis. □

Proof (by moving the basis). ROUTINE

Uses. Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.5.20 (Inverses and matrices), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Remark 4.5.15 (Moving the grid or moving the vectors)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Use the active reading of Remark 4.5.15: look for an operator R with $R(\mathbf{b}_j) = \mathring{\mathbf{b}}_j$.
Hint 2. Let $R \in \mathrm{L}(V)$ have $[R]_{\mathcal{B}\mathcal{B}} = S$. Why is R invertible, and what is $R(\mathcal{B})$?
Sketch. Let $R \in \mathrm{L}(V)$ be the operator with $[R]_{\mathcal{B}\mathcal{B}} = S$. It is invertible because S is, so $\mathring{\mathcal{B}} = R(\mathcal{B})$ is a basis, and $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = [R]_{\mathcal{B}\mathcal{B}} = S$ by Remark 4.5.15. Uniqueness and (b) are as in the first proof.

(a) By Theorem 2.3.20(b) there is $R \in \mathrm{L}(V)$ with $[R]_{\mathcal{B}\mathcal{B}} = S$, and R is invertible because S is (Proposition 2.5.20(b)). So $\mathring{\mathcal{B}} = R(\mathbf{b}_1), \ldots, R(\mathbf{b}_n)$ is a basis of V (Proposition 2.6.22(c)), and by Remark 4.5.15, $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = [R]_{\mathcal{B}\mathcal{B}} = S$. The j-th vector of $\mathring{\mathcal{B}}$ is $R(\mathbf{b}_j) = \sum_i s_{ij}\mathbf{b}_i$, because the j-th column of $[R]_{\mathcal{B}\mathcal{B}}$ is $\mathbf{s}_j$. Uniqueness is proved as in the first proof: the columns of $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$ determine the vectors of $\mathring{\mathcal{B}}$.

(b) As in the first proof, apply (a) to S^{-1}. □

In particular, for a fixed basis $\mathcal{B}$, the map $\mathring{\mathcal{B}} \mapsto [\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$ is a bijection from the set of bases of V onto the set of invertible $n \times n$ matrices. Over a field with q elements, the number of invertible $n \times n$ matrices is therefore the number of bases of F^n, which Corollary 4.1.35 computed: $(q^n - 1)(q^n - q)\cdots(q^n - q^{n-1})$.

4.5.3 The change-of-basis formula

Now let $T \in \mathrm{L}(V, W)$, where $\dim V = n \ge 1$ and $\dim W = m \ge 1$. If we replace the basis $\mathcal{B}$ of V by $\mathring{\mathcal{B}}$ and the basis $\mathcal{C}$ of W by $\mathring{\mathcal{C}}$, how is the new matrix $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}$ related to the old one?

Theorem 4.5.17 (Change of basis). Let $\mathcal{B}$ and $\mathring{\mathcal{B}}$ be bases of V, let $\mathcal{C}$ and $\mathring{\mathcal{C}}$ be bases of W, and let $T \in \mathrm{L}(V, W)$. Then

$$[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}} = [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}[T]_{\mathcal{C}\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}.$$

In particular, if $T \in \mathrm{L}(V)$ and $P = [\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}$, then

$$[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = P^{-1}[T]_{\mathcal{B}\mathcal{B}}P.$$

Intuition (Reading the formula). Read the right side from right to left, as the composite it is. Start with the $\mathring{\mathcal{B}}$-coordinates of a vector $\mathbf{v}$. The factor $[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}$ converts them into $\mathcal{B}$-coordinates, in which we know how to apply T; the factor $[T]_{\mathcal{C}\mathcal{B}}$ does so and produces the $\mathcal{C}$-coordinates of $T(\mathbf{v})$; and $[\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}$ converts those into $\mathring{\mathcal{C}}$-coordinates. The result is what $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}$ does in one step.

There is also a bookkeeping rule. In a product of matrices of maps, adjacent subscripts must match, like dominoes: $\mathring{\mathcal{C}}\mathcal{C} \cdot \mathcal{C}\mathcal{B} \cdot \mathcal{B}\mathring{\mathcal{B}}$. The two outer subscripts, $\mathring{\mathcal{C}}$ and $\mathring{\mathcal{B}}$, are the bases of the result. A formula whose subscripts do not chain this way is wrong, whatever its other merits.

The first proof is the lecture's: the formula is the composition theorem applied to $T = \mathrm{id}_W \circ T \circ \mathrm{id}_V$, with each copy of a space carrying a different basis.

Proof (by composition). ROUTINE

Uses. Theorem 2.4.2 (The matrix of a composite), Proposition 2.1.24 (Algebra of composition), Proposition 4.5.12 (Calculus of change of coordinates)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Write T as a composite of three maps in which identity maps change the basis.
Hint 2. $T = \mathrm{id}_W \circ T \circ \mathrm{id}_V$. Apply Theorem 2.4.2 twice, choosing the bases $\mathring{\mathcal{B}}$, $\mathcal{B}$, $\mathcal{C}$, $\mathring{\mathcal{C}}$ on the four copies of V and W.
Sketch. $T = \mathrm{id}_W \circ T \circ \mathrm{id}_V$ (Proposition 2.1.24(b)). By Theorem 2.4.2, with $\mathring{\mathcal{B}}$ on the first copy of V, $\mathcal{B}$ on the second, $\mathcal{C}$ on the first copy of W, and $\mathring{\mathcal{C}}$ on the second, the matrix of the composite is the product of the three matrices. For operators, $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = P^{-1}$ by Proposition 4.5.12(b).

By Proposition 2.1.24(b), $T = \mathrm{id}_W \circ (T \circ \mathrm{id}_V)$. Apply Theorem 2.4.2 to $T \circ \mathrm{id}_V$, with $\mathring{\mathcal{B}}$ on the domain, $\mathcal{B}$ on the middle copy of V, and $\mathcal{C}$ on W:

$$[T]_{\mathcal{C}\mathring{\mathcal{B}}} = [T \circ \mathrm{id}_V]_{\mathcal{C}\mathring{\mathcal{B}}} = [T]_{\mathcal{C}\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}.$$

Apply it again to $\mathrm{id}_W \circ T$, with $\mathring{\mathcal{B}}$ on V, $\mathcal{C}$ on the middle copy of W, and $\mathring{\mathcal{C}}$ on the target:

$$[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}} = [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}[T]_{\mathcal{C}\mathring{\mathcal{B}}} = [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}[T]_{\mathcal{C}\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}.$$

For $T \in \mathrm{L}(V)$, take $\mathcal{C} = \mathcal{B}$ and $\mathring{\mathcal{C}} = \mathring{\mathcal{B}}$. Then $[\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}} = P^{-1}$ by Proposition 4.5.12(b). □

The second proof chases a single vector through the diagram of Figure 4.5.1. It is the lecture's proof of $[\mathbf{v}]_{\mathring{\mathcal{B}}} = [\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}$, extended by two squares.

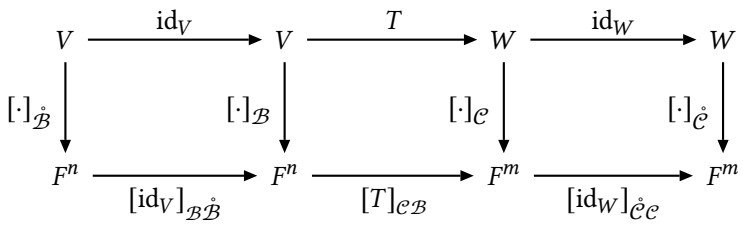


Figure 4.5.1. Three commuting squares. The top row composes to T, and the bottom row multiplies three matrices. Each square says that two routes from its top left corner to its bottom right corner agree (Theorem 2.3.9), so a vector can be pushed along the bottom row instead of the top. The outer rectangle is the change-of-basis formula.

Proof (by chasing a vector). ROUTINE

Uses. Theorem 2.3.9 (The matrix of a map computes the map), Proposition 1.6.16 (The coordinate map is a bijection), Definition 2.3.8 (Matrix–vector product)
Techniques. Translate between maps and matrices (Technique 2.3.27), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Two $m \times n$ matrices are equal if they give the same result on every $\mathbf{x} \in F^n$: compare them on $\mathbf{e}_j$.
Hint 2. Every $\mathbf{x} \in F^n$ is $[\mathbf{v}]_{\mathring{\mathcal{B}}}$ for some $\mathbf{v} \in V$. Multiply $[\mathbf{v}]_{\mathring{\mathcal{B}}}$ by the right side one factor at a time, using Theorem 2.3.9 at each step.
Sketch. For $\mathbf{v} \in V$, multiplying $[\mathbf{v}]_{\mathring{\mathcal{B}}}$ by the three factors in turn gives $[\mathbf{v}]_{\mathcal{B}}$, then $[T(\mathbf{v})]_{\mathcal{C}}$, then $[T(\mathbf{v})]_{\mathring{\mathcal{C}}}$, which is also $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}[\mathbf{v}]_{\mathring{\mathcal{B}}}$. Every column in F^n is a $\mathring{\mathcal{B}}$-coordinate vector, in particular each $\mathbf{e}_j$, so the two matrices have the same columns.

Let M be the right side of the formula, and let $\mathbf{v} \in V$. Applying Theorem 2.3.9 three times (to id_V, to T, and to id_W),

$$\begin{aligned} M[\mathbf{v}]_{\mathring{\mathcal{B}}} &= [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}[T]_{\mathcal{C}\mathcal{B}}\big([\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}[\mathbf{v}]_{\mathring{\mathcal{B}}}\big) \\ &= [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}\big([T]_{\mathcal{C}\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}\big) \\ &= [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}[T(\mathbf{v})]_{\mathcal{C}} = [T(\mathbf{v})]_{\mathring{\mathcal{C}}}, \end{aligned}$$

and $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}[\mathbf{v}]_{\mathring{\mathcal{B}}} = [T(\mathbf{v})]_{\mathring{\mathcal{C}}}$ as well. The coordinate map of $\mathring{\mathcal{B}}$ is onto F^n (Proposition 1.6.16), so for each j we may choose $\mathbf{v}$ with $[\mathbf{v}]_{\mathring{\mathcal{B}}} = \mathbf{e}_j$. Then the j-th columns $M\mathbf{e}_j$ and $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}\mathbf{e}_j$ agree (Definition 2.3.8), and the two matrices are equal. The statement for operators follows as in the first proof. □

The first proof is shorter, because Theorem 2.4.2 already did the vector chase once and for all. The second is the one to remember when the subscripts get confusing: follow one vector around the diagram, and the order of the factors takes care of itself.

Example 4.5.18 (An oblique projection). Let $\ell_1 = \mathrm{span}((1,1))$ and $\ell_2 = \mathrm{span}((1,2))$ in F^2. The vectors $\mathbf{u}_1 = (1,1)$ and $\mathbf{u}_2 = (1,2)$ form a basis $\mathcal{U}$ of F^2 over every field, since $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ has $1 \cdot 2 - 1 \cdot 1 = 1 \neq 0$. Let $P \in \mathrm{L}(F^2)$ be the projection onto ℓ_1 along ℓ_2: $P(\mathbf{u}_1) = \mathbf{u}_1$ and $P(\mathbf{u}_2) = \mathbf{0}$. Its matrix in $\mathcal{U}$ is obvious, and the theorem converts it into the standard matrix. With $Q = [\mathrm{id}]_{\mathcal{E}\mathcal{U}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, whose inverse is $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$,

$$[P]_{\mathcal{E}\mathcal{E}} = Q[P]_{\mathcal{U}\mathcal{U}}Q^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}.$$

Check: $\begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Here we used the theorem with $\mathcal{E}$ as the new basis: $[P]_{\mathcal{E}\mathcal{E}} = [\mathrm{id}]_{\mathcal{E}\mathcal{U}}[P]_{\mathcal{U}\mathcal{U}}[\mathrm{id}]_{\mathcal{U}\mathcal{E}}$.

Example 4.5.19 (Differentiation, revisited). In Example 2.3.22(c), with $2 \neq 0$ in F, we found the matrix of differentiation D on $F[x]_{\leq 2}$ in the basis $\mathring{\mathcal{B}} = 1, x, \frac{1}{2}x^2$ by applying D to each basis vector. The theorem gives it from $[D]_{\mathcal{B}\mathcal{B}}$, where $\mathcal{B} = 1, x, x^2$. Here $P = [\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = \mathrm{diag}\big(1, 1, \frac{1}{2}\big)$, whose inverse is $\mathrm{diag}(1,1,2)$, and

$$P^{-1}[D]_{\mathcal{B}\mathcal{B}}P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

as found there. Scaling the third basis vector by $\frac{1}{2}$ divides the third column by 2 and multiplies the third row by 2.

Exercise 4.5.20. ROUTINE Let $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ and $\mathring{\mathcal{B}} = (1,1), (2,1)$, a basis of F^2. Compute $[T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$ with Theorem 4.5.17, and check the result by computing $A(1,1)$ and $A(2,1)$ directly.

Solution on page 1333

Matrices of one map

Every pair of bases gives a matrix of T, and by the theorem they are all of the form GAH with G and H invertible. The converse also holds, because every invertible matrix is a change-of-coordinates matrix.

Corollary 4.5.21 (Matrices of one linear map). Let $T \in \mathrm{L}(V, W)$ with $\dim V = n \geq 1$ and $\dim W = m \geq 1$, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V and W, and let $A = [T]_{\mathcal{C}\mathcal{B}}$. For $A' \in F^{m\times n}$, the following are equivalent.

(a) $A' = [T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}}$ for some bases $\mathring{\mathcal{B}}$ of V and $\mathring{\mathcal{C}}$ of W.
(b) $A' = GAH$ for some invertible $G \in F^{m\times m}$ and $H \in F^{n\times n}$.
(c) $\operatorname{rank} A' = \operatorname{rank} A$.

Proof. ROUTINE

Uses. Theorem 4.5.17 (Change of basis), Proposition 4.5.12 (Calculus of change of coordinates), Proposition 4.5.16 (Every invertible matrix changes coordinates), Proposition 4.3.1 (Invertible factors preserve rank), Theorem 3.5.27 (Rank normal form), Proposition 2.5.17 (Algebra of inverses)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. (a) ⇒ (b) is the theorem. For (b) ⇒ (a), realize H and G as change-of-coordinates matrices with Proposition 4.5.16.
Hint 2. For (c) ⇒ (b), bring both A and A' to the same rank normal form.
Sketch. (a) ⇒ (b): Theorem 4.5.17 with $G = [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}}$ and $H = [\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}$. (b) ⇒ (a): choose $\mathring{\mathcal{B}}$ with $[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}} = H$ and $\mathring{\mathcal{C}}$ with $[\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}} = G$, and apply the theorem. (b) ⇒ (c) is Proposition 4.3.1. (c) ⇒ (b): $G_1 A H_1 = N_r = G_2 A' H_2$, so $A' = G_2^{-1} G_1 A H_1 H_2^{-1}$.

We prove (a) ⇔ (b) and (b) ⇔ (c).

(a) ⇒ (b). By Theorem 4.5.17, $A' = [\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}} A [\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}$, and both outer factors are invertible (Proposition 4.5.12(b)).

(b) ⇒ (a). By Proposition 4.5.16(a) there is a basis $\mathring{\mathcal{B}}$ of V with $[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}} = H$, and by part (b) there is a basis $\mathring{\mathcal{C}}$ of W with $[\mathrm{id}_W]_{\mathring{\mathcal{C}}\mathcal{C}} = G$. By Theorem 4.5.17, $[T]_{\mathring{\mathcal{C}}\mathring{\mathcal{B}}} = GAH = A'$.

(b) ⇒ (c). This is Proposition 4.3.1.

(c) ⇒ (b). Let $r = \operatorname{rank} A = \operatorname{rank} A'$. By Theorem 3.5.27 there are invertible $G_1, G_2 \in F^{m\times m}$ and $H_1, H_2 \in F^{n\times n}$ with $G_1 A H_1 = N_r = G_2 A' H_2$. Then $A' = \left(G_2^{-1} G_1\right) A \left(H_1 H_2^{-1}\right)$, and both factors are invertible by Proposition 2.5.17. □

So when the domain and the target may carry unrelated bases, the matrix of T can be any matrix of rank $\operatorname{rank} T$, and nothing else about T is visible. Choosing the bases well, the matrix becomes as simple as possible.

Corollary 4.5.22 (Rank normal form for linear maps). Let $T \in \mathrm{L}(V, W)$, where V and W are nonzero and finite-dimensional, and let $r = \operatorname{rank} T$. There are bases $\mathcal{B}$ of V and $\mathcal{C}$ of W with

$$[T]_{\mathcal{C}\mathcal{B}} = N_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Proof (from the matrix version). ROUTINE

Uses. Corollary 4.5.21 (Matrices of one linear map), Proposition 4.2.15 (Rank through any matrix)
Hint 1. Start from any bases. What is the rank of the resulting matrix?
Hint 2. N_r has rank r; use (c) ⇒ (a) of Corollary 4.5.21.
Sketch. For any bases, $\operatorname{rank} [T]_{\mathcal{C}_0\mathcal{B}_0} = \operatorname{rank} T = r = \operatorname{rank} N_r$ (Proposition 4.2.15), so by Corollary 4.5.21, N_r is a matrix of T.

Choose any bases $\mathcal{B}_0$ and $\mathcal{C}_0$. By Proposition 4.2.15, $\operatorname{rank} [T]_{\mathcal{C}_0\mathcal{B}_0} = \operatorname{rank} T = r$, and $\operatorname{rank} N_r = r$. By Corollary 4.5.21, (c) ⇒ (a), there are bases $\mathcal{B}$ and $\mathcal{C}$ with $[T]_{\mathcal{C}\mathcal{B}} = N_r$. □

Proof (by building the bases). STANDARD

Uses. Theorem 4.1.16 (Extending to a basis), Theorem 4.2.11 (Rank–nullity), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Definition 2.3.1 (Matrix of a linear map)
Techniques. Extend a basis of a subspace (Technique 4.1.18)
Hint 1. The last $n-r$ vectors of $\mathcal{B}$ must go to $\mathbf{0}$, and the first r must go to the first r vectors of $\mathcal{C}$. Where should the last $n-r$ vectors come from?
Hint 2. Extend a basis of $\ker T$ to a basis of V, putting the new vectors first. Their images form a basis of $\operatorname{im} T$; extend that to a basis of W.
Sketch. Extend a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of $\ker T$ to a basis $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{u}_1, \dots, \mathbf{u}_k$ of V. The images $T(\mathbf{v}_1), \dots, T(\mathbf{v}_r)$ span $\operatorname{im} T$, and $r = \dim \operatorname{im} T$ by rank–nullity, so they form a basis of $\operatorname{im} T$; extend them to a basis of W. Then $T(\mathbf{v}_j) = \mathbf{c}_j$ and $T(\mathbf{u}_i) = \mathbf{0}$, which is N_r.

Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be a basis of $\ker T$, and extend it to a basis of V (Theorem 4.1.16); list the new vectors first, so that $\mathcal{B} = \mathbf{v}_1, \dots, \mathbf{v}_s, \mathbf{u}_1, \dots, \mathbf{u}_k$. Since $T(\mathbf{u}_i) = \mathbf{0}$, the list $T(\mathcal{B})$, which spans $\operatorname{im} T$ (Proposition 2.6.22(d)), still spans $\operatorname{im} T$ when the zeros are deleted. So $T(\mathbf{v}_1), \dots, T(\mathbf{v}_s)$ spans $\operatorname{im} T$. By Theorem 4.2.11(b), $s = \dim V - k = \dim \operatorname{im} T = r$, so these r vectors span the r-dimensional space $\operatorname{im} T$, and they form a basis of it (Corollary 4.1.11(b)). Write $\mathbf{c}_j = T(\mathbf{v}_j)$ and extend $\mathbf{c}_1, \dots, \mathbf{c}_r$ to a basis $\mathcal{C}$ of W. Then the j-th column of $[T]_{\mathcal{C}\mathcal{B}}$ is $[\mathbf{c}_j]_{\mathcal{C}} = \mathbf{e}_j$ for $j \le r$, and the remaining columns are $[\mathbf{0}]_{\mathcal{C}} = \mathbf{0}$. So $[T]_{\mathcal{C}\mathcal{B}} = N_r$. □

The second proof is the one sketched after the second proof of rank–nullity (Section 4.2), and the first is Corollary 3.5.29 transported to abstract spaces. For operators, where the same basis must be used on both sides, much more survives. That is the subject of the next subsection.

4.5.4 Similarity

For $T \in \mathrm{L}(V)$ we compare $T(\mathbf{v})$ with $\mathbf{v}$, so we use one basis for both. By Theorem 4.5.17, the matrices of T in two bases are related by $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = P^{-1}[T]_{\mathcal{B}\mathcal{B}}P$, with the same invertible P on both sides.

Definition 4.5.23 (Similar matrices). Matrices $A, B \in F^{n\times n}$ are **similar** if

$$B = S^{-1}AS$$

for some invertible $S \in F^{n\times n}$. We then write $A \sim B$.

The lecture writes the same condition as $B = QAQ^{-1}$; take $Q = S^{-1}$. To check a proposed S, it is easier to verify $AS = SB$ (and that S is invertible) than to compute S^{-1}.

Proposition 4.5.24 (Similarity is an equivalence relation). For all $A, B, C \in F^{n\times n}$: $A \sim A$; if $A \sim B$ then $B \sim A$; and if $A \sim B$ and $B \sim C$ then $A \sim C$.

Proof. ROUTINE

Uses. Definition 4.5.23 (Similar matrices), Proposition 2.5.17 (Algebra of inverses), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Each claim asks for an invertible matrix. Try I_n, then S^{-1}, then a product.
Hint 2. If $B = S^{-1}AS$ and $C = R^{-1}BR$, substitute and regroup.
Sketch. $A = I^{-1}AI$. If $B = S^{-1}AS$, then $A = SBS^{-1} = (S^{-1})^{-1}BS^{-1}$. If also $C = R^{-1}BR$, then $C = R^{-1}S^{-1}ASR = (SR)^{-1}A(SR)$.

First, $A = I_n^{-1}AI_n$. If $B = S^{-1}AS$, multiplying by S on the left and S^{-1} on the right gives $A = SBS^{-1} = (S^{-1})^{-1}BS^{-1}$, and S^{-1} is invertible (Proposition 2.5.17(a)). If also $C = R^{-1}BR$, then by associativity (Proposition 2.4.14) and Proposition 2.5.17(b),

$$C = R^{-1}S^{-1}ASR = (SR)^{-1}A(SR),$$

and SR is invertible. □

The definition is a statement about matrices, but its meaning is a statement about operators.

Theorem 4.5.25 (Similar matrices are one operator in two bases). Let $A, B \in F^{n\times n}$ with $n \geq 1$. The following are equivalent.

(a) $A \sim B$.
(b) Whenever $\dim V = n$, $T \in \mathrm{L}(V)$, and $\mathcal{B}$ is a basis of V with $[T]_{\mathcal{B}\mathcal{B}} = A$, there is a basis $\mathring{\mathcal{B}}$ of V with $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = B$.
(c) There are a vector space V with $\dim V = n$, an operator $T \in \mathrm{L}(V)$, and bases $\mathcal{B}$, $\mathring{\mathcal{B}}$ of V with $[T]_{\mathcal{B}\mathcal{B}} = A$ and $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = B$.

Moreover, if $B = S^{-1}AS$, then $B = [T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$, where $\mathring{\mathcal{B}}$ is the list of columns of S.

Proof. ROUTINE

Uses. Theorem 4.5.17 (Change of basis), Proposition 4.5.16 (Every invertible matrix changes coordinates), Proposition 4.5.12 (Calculus of change of coordinates), Proposition 2.3.16 (The matrix of T_A)
Techniques. Proving TFAE (Technique 2.7.1), Translate between maps and matrices (Technique 2.3.27)
Hint 1. Prove (a) ⇒ (b) ⇒ (c) ⇒ (a). Two of the three steps are Theorem 4.5.17 read in the two directions.
Hint 2. For (a) ⇒ (b), given $B = S^{-1}AS$, use Proposition 4.5.16 to find $\mathring{\mathcal{B}}$ with $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$. For (b) ⇒ (c), you need one example of V, T, $\mathcal{B}$: take the obvious one.
Sketch. (a) ⇒ (b): if $B = S^{-1}AS$, choose $\mathring{\mathcal{B}}$ with $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$; then $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = S^{-1}AS = B$. (b) ⇒ (c): apply (b) to T_A on F^n with the standard basis. (c) ⇒ (a): Theorem 4.5.17 with $P = [\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$. For the last claim, the basis of F^n with $[\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}} = S$ is the list of columns of S.

We prove (a) ⇒ (b) ⇒ (c) ⇒ (a).

(a) ⇒ (b). Let $B = S^{-1}AS$, and let V, T, and $\mathcal{B}$ be as in (b). By Proposition 4.5.16(a), there is a basis $\mathring{\mathcal{B}}$ of V with $[\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}} = S$. By Theorem 4.5.17, $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = S^{-1}[T]_{\mathcal{B}\mathcal{B}}S = S^{-1}AS = B$.

(b) ⇒ (c). Take $V = F^n$, $T = T_A$, and the standard basis $\mathcal{E}$; then $[T_A]_{\mathcal{E}\mathcal{E}} = A$ (Proposition 2.3.16), and (b) provides $\mathring{\mathcal{B}}$.

(c) ⇒ (a). Let $P = [\mathrm{id}]_{\mathcal{B}\mathring{\mathcal{B}}}$, which is invertible (Proposition 4.5.12(b)). By Theorem 4.5.17, $B = [T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = P^{-1}AP$.

For the last claim, apply the proof of (a) ⇒ (b) to T_A and $\mathcal{B} = \mathcal{E}$. By Proposition 4.5.16(a), the j-th vector of $\mathring{\mathcal{B}}$ is $\sum_i s_{ij}\mathbf{e}_i$, the j-th column of S. □

Intuition (Similarity is a change of viewpoint). Two similar matrices describe one operator, recorded in two coordinate systems, and S is the dictionary between them: its columns are the new basis, written in the old coordinates. So a question about similarity is a question about whether two matrices can be the same operator. Any property of the operator itself, defined without reference to a basis, must be shared by all of its matrices. That is how we will tell matrices apart.

Example 4.5.26 (A similarity and a non-similarity).

(a) Let F be a field with $2 \neq 0$, and let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Then $A(1,1) = (2,2) = 2(1,1)$ and $A(1,-1) = (0,0)$. So in the basis $\mathring{\mathcal{B}} = (1,1), (1,-1)$, the operator T_A doubles the first basis vector and kills the second, and with $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, the matrix whose columns are these vectors,

$$S^{-1}AS = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

(Check: $AS = \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix} = S\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, and S is invertible since $1\cdot(-1) - 1\cdot 1 = -2 \neq 0$.) Similar matrices need not share a single entry.

(a) Over $\mathbb{F}_2$, the same matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is not similar to $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, which is now the zero matrix: $S^{-1}0S = 0$ for every S, so the only matrix similar to 0 is 0. Whether two matrices are similar can depend on the field.

Similarity invariants

A **similarity invariant** is a quantity or property of square matrices that is the same for any two similar matrices. Here are the ones available so far.

Proposition 4.5.27 (Similarity invariants). Let $A, B \in F^{n\times n}$ with $A \sim B$. Then:

(a) $\operatorname{rank} A = \operatorname{rank} B$;
(b) $\operatorname{tr} A = \operatorname{tr} B$;
(c) $A^k \sim B^k$ for every $k \geq 0$; in particular $\operatorname{rank} A^k = \operatorname{rank} B^k$;
(d) $A - cI_n \sim B - cI_n$ for every $c \in F$; in particular $\operatorname{rank}(A - cI_n) = \operatorname{rank}(B - cI_n)$;
(e) A is invertible if and only if B is, and then $A^{-1} \sim B^{-1}$;
(f) $A^{\mathsf{T}} \sim B^{\mathsf{T}}$.

Proof (by matrix algebra). ROUTINE

Uses. Definition 4.5.23 (Similar matrices), Proposition 4.3.1 (Invertible factors preserve rank), Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.14 (Algebra of matrices), Proposition 2.5.17 (Algebra of inverses), Proposition 2.4.23 (Transpose of a product)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Write $B = S^{-1}AS$. For (a), invertible factors do not change the rank. For (b), move a factor around the trace.
Hint 2. For (c), the inner factors telescope: $(S^{-1}AS)(S^{-1}AS) = S^{-1}A^2S$. For (d), $S^{-1}(cI)S = cI$. For (f), transpose $B = S^{-1}AS$.
Sketch. With $B = S^{-1}AS$: (a) Proposition 4.3.1. (b) $\operatorname{tr}\big(S^{-1}(AS)\big) = \operatorname{tr}\big((AS)S^{-1}\big) = \operatorname{tr} A$. (c) $B^k = S^{-1}A^kS$ by induction. (d) $S^{-1}(A - cI)S = B - cI$. (e) $B^{-1} = S^{-1}A^{-1}S$. (f) $B^{\mathsf{T}} = S^{\mathsf{T}}A^{\mathsf{T}}\big(S^{\mathsf{T}}\big)^{-1}$.

Write $B = S^{-1}AS$ with S invertible.

(a) By Proposition 4.3.1, $\operatorname{rank}\big(S^{-1}AS\big) = \operatorname{rank} A$.

(b) By Proposition 2.4.28 with the factors S^{-1} and AS,

$$\operatorname{tr} B = \operatorname{tr}\big(S^{-1}(AS)\big) = \operatorname{tr}\big((AS)S^{-1}\big) = \operatorname{tr} A.$$

(c) We show $B^k = S^{-1}A^kS$ by induction on k. For $k = 0$ both sides are I_n. If it holds for k, then

$$B^{k+1} = B^kB = \big(S^{-1}A^kS\big)\big(S^{-1}AS\big) = S^{-1}A^k\big(SS^{-1}\big)AS = S^{-1}A^{k+1}S.$$

So $A^k \sim B^k$, and their ranks agree by (a).

(d) $S^{-1}(A - cI_n)S = S^{-1}AS - cS^{-1}S = B - cI_n$ (Proposition 2.4.14), and the ranks agree by (a).

(e) If A is invertible, then B is a product of invertible matrices, hence invertible, with $B^{-1} = S^{-1}A^{-1}S$ (Proposition 2.5.17(b)). So $A^{-1} \sim B^{-1}$. The converse follows by symmetry, since $A = SBS^{-1}$.

(f) By Proposition 2.4.23 and Proposition 2.5.17(d), $B^{\mathsf{T}} = S^{\mathsf{T}}A^{\mathsf{T}}\big(S^{-1}\big)^{\mathsf{T}} = S^{\mathsf{T}}A^{\mathsf{T}}\big(S^{\mathsf{T}}\big)^{-1}$. With $R = \big(S^{\mathsf{T}}\big)^{-1}$, this is $B^{\mathsf{T}} = R^{-1}A^{\mathsf{T}}R$. □

The second proof uses Theorem 4.5.25: since A and B are matrices of one operator, anything defined from the operator alone is the same for both. It reaches (a), (c), (d), and (e), but not (b) or (f), because the trace and the transpose are defined through matrices.

Proof (through the operator). ROUTINE

Uses. Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.2.15 (Rank through any matrix), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.12 (Matrices of sums and scalar multiples), Example 2.3.6 (The zero map and the identity map), Proposition 2.5.20 (Inverses and matrices)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. By Theorem 4.5.25, $A = [T]_{\mathcal{B}\mathcal{B}}$ and $B = [T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$ for a single operator T. What is the rank of a matrix of T?
Hint 2. Which operators have matrices A^k, $A - cI_n$, and A^{-1} in the basis $\mathcal{B}$? Apply Theorem 4.5.25 to them in the other direction.

Sketch. Let $A = [T]_{\mathcal{B}\mathcal{B}}$ and $B = [T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$. Then $\operatorname{rank} A = \operatorname{rank} T = \operatorname{rank} B$. Also A^k and B^k are the matrices of T^k in the two bases, $A - cI_n$ and $B - cI_n$ those of $T - c\mathrm{id}_V$, and A^{-1}, B^{-1} those of T^{-1} (which exists iff either matrix is invertible); so each pair is similar by the theorem.

By Theorem 4.5.25, (a) ⇒ (c), there are V, $T \in \mathrm{L}(V)$, and bases $\mathcal{B}$, $\mathring{\mathcal{B}}$ with $A = [T]_{\mathcal{B}\mathcal{B}}$ and $B = [T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$.

(a) By Proposition 4.2.15, $\operatorname{rank} A = \operatorname{rank} T = \operatorname{rank} B$.

(c) By Theorem 2.4.2 and induction on k, $\left[T^k\right]_{\mathcal{B}\mathcal{B}} = A^k$ and $\left[T^k\right]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = B^k$ (with $T^0 = \mathrm{id}_V$ and Example 2.3.6 for $k = 0$). So A^k and B^k are matrices of the single operator T^k, and they are similar by Theorem 4.5.25, (c) ⇒ (a).

(d) By Proposition 2.4.12 and Example 2.3.6, $A - cI_n$ and $B - cI_n$ are the matrices of $T - c\mathrm{id}_V$ in the two bases.

(e) By Proposition 2.5.20(b), A is invertible if and only if T is, if and only if B is; and then A^{-1} and B^{-1} are the matrices of T^{-1} in the two bases. □

Remark 4.5.28 (What the invariants cannot do). Equal invariants do not make matrices similar. The matrices

$$N = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

both have trace 0 and rank 2. But $N^3 = 0$, since N shifts $\mathbf{e}_1 \mapsto \mathbf{e}_2 \mapsto \mathbf{e}_3 \mapsto \mathbf{0}$, while $D^3 = D$ has rank 2. By (c), $\operatorname{rank} N^3 \neq \operatorname{rank} D^3$ rules out $N \sim D$. Similarly, I_2 and $J = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ have the same rank and trace, and J^k has rank 2 like I_2^k; but $\operatorname{rank}(J - I_2) = 1 \neq 0 = \operatorname{rank}(I_2 - I_2)$, so they are not similar. (More simply, the only matrix similar to I_n is I_n, since $S^{-1} I_n S = I_n$.)

The determinant is one more invariant (Proposition 5.4.17), and the characteristic polynomial, which contains both the trace and the determinant, is a stronger one (Section 7.3). None of these decides similarity on its own. Over $\mathbb{C}$, the ranks of $(A - cI_n)^k$ for all c and k do decide it; this is the content of the Jordan form (Section 10.3).

Technique 4.5.29 (Rule out similarity with an invariant). To show that A and B are not similar, find a similarity invariant on which they differ: trace, rank, rank of a power A^k, rank of $A - cI_n$ for a well-chosen c, and later the determinant and the characteristic polynomial. To show that they are similar, invariants are not enough: produce an S with $AS = SB$, usually by finding a basis in which the operator T_A has matrix B (Theorem 4.5.25).

Reach for the cheapest invariants first. Trace and rank take seconds. If they agree, look for a structural difference: is one matrix nilpotent, or a scalar multiple of I_n, or does $A - cI_n$ drop rank for some c where $B - cI_n$ does not? The canonical example is Remark 4.5.28.

Used in. Exercise 4.5.32, Exercise 4.5.44, Exercise 4.5.45, Exercise 4.8.4, Problem 4.13, Exercise 7.3.38, Exercise 7.4.43, Exercise 4.5.46, Exercise 7.6.43.

Invariants of an operator

Because the trace is a similarity invariant, it can be assigned to an operator, just as the rank was in Definition 4.2.14.

Definition 4.5.30 (Trace of an operator). Let V be finite-dimensional with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. The **trace** of T is

$$\operatorname{tr} T = \operatorname{tr} [T]_{\mathcal{B}\mathcal{B}},$$

where $\mathcal{B}$ is any basis of V.

The definition does not depend on $\mathcal{B}$: by Theorem 4.5.17, the matrices of T in two bases are similar, and similar matrices have the same trace (Proposition 4.5.27(b)). In the same way, $\operatorname{rank}[T]_{\mathcal{B}\mathcal{B}} = \operatorname{rank} T$ for every basis (Proposition 4.2.15), and in Definition 5.4.18 the determinant joins them. For $V = F^n$ and the standard basis, $\operatorname{tr} T_A = \operatorname{tr} A$.

Example 4.5.31 (The trace of transposition). Let $T \in \mathrm{L}(F^{n\times n})$ be $T(X) = X^{\mathsf{T}}$. In the basis of matrix units U_{ij} (Proposition 1.6.7(b)), $T(U_{ij}) = U_{ji}$. The diagonal entry of $[T]$ in the position belonging to U_{ij} is the U_{ij}-coordinate of $T(U_{ij}) = U_{ji}$, which is 1 if $i = j$ and 0 otherwise. So

$$\operatorname{tr} T = n \quad (\text{that is}, n \cdot 1 \in F).$$

Over $\mathbb{R}$ there is a second basis to check this against. T fixes the symmetric matrices, a subspace of dimension $n(n+1)/2$, and negates the skew-symmetric ones, of dimension $n(n-1)/2$. In a basis made of a basis of each, $[T]$ is diagonal with $n(n+1)/2$ entries 1 and $n(n-1)/2$ entries -1, and its trace is again $n(n+1)/2 - n(n-1)/2 = n$.

Exercise 4.5.32. STANDARD Let $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$. Decide which of the following matrices are similar to J. For each one that is, find S with $S^{-1}MS = J$.

(a) $M = I_2$.
(b) $M = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$.
(c) $M = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$.

Solution on page 1333

4.5.5 Choosing a convenient basis

The change-of-basis formula also gives a proof technique: if a statement does not depend on the basis, we may prove it in whichever basis makes it easiest.

Technique 4.5.33 (Choose a convenient basis). Many statements about a linear map T (or a matrix A) are invariant under change of basis: they hold for one matrix of T if and only if they hold for all of them, or for A if and only if they hold for GAH (or for $S^{-1}AS$). To prove such a statement, first check this invariance, then choose bases in which the matrix is as simple as possible, prove the statement there, and transport the conclusion back.

For maps between two spaces, the best possible matrix is the rank normal form N_r (Corollary 4.5.22), and in matrix language this says: to prove something about A that survives $A \mapsto GAH$, assume $A = N_r$. For operators, the bases must agree on both sides, and the convenient bases come from the operator itself: bases adapted to a subspace that T preserves (below), bases of the form $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ (Exercise 4.5.50, Exercise 4.5.43), and later bases of eigenvectors (Section 7.4) and triangularizing bases (Section 7.5).
Used in. Exercise 4.5.41, Exercise 4.5.42, Exercise 4.5.43, Exercise 4.5.49, Exercise 4.5.50, Exercise 4.5.51, Exercise 4.5.52, Exercise 4.5.53, Exercise 4.8.12, Theorem 7.5.22, Problem 7.6, Problem 7.10, Problem 7.11, Proposition 8.2.2, Corollary 8.2.28, Exercise 8.3.43, Exercise 4.5.55.

Example 4.5.34 (A matrix with $AXA = A$). We show that for every $A \in F^{m\times n}$ there is $X \in F^{n\times m}$ with $AXA = A$. (Such an X is a generalized inverse of A; if A is invertible, $X = A^{-1}$ is the only choice.)

For $A = N_r$ the claim is easy: $X = N_r^{\mathsf{T}} \in F^{n\times m}$ works. By block multiplication (Theorem 2.4.34), $N_r N_r^{\mathsf{T}} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in F^{m\times m}$, and then $N_r N_r^{\mathsf{T}} N_r = N_r$.

The statement survives $A \mapsto GAH$: if $N_r X N_r = N_r$ and $A = GN_rH$ with G and H invertible, then $X' = H^{-1}XG^{-1}$ satisfies

$$AX'A = GN_rHH^{-1}XG^{-1}GN_rH = G(N_rXN_r)H = GN_rH = A.$$

By Theorem 3.5.27, every A is GN_rH for some invertible G, H (namely, invert the factors in $G'AH' = N_r$). So every A has a generalized inverse, $X' = H^{-1}N_r^{\mathsf{T}}G^{-1}$.

In terms of maps: if $[T]_{\mathcal{C}\mathcal{B}} = N_r$ with $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_m$, let $S \in \mathrm{L}(W, V)$ send $\mathbf{c}_j \mapsto \mathbf{b}_j$ for $j \le r$ and $\mathbf{c}_j \mapsto \mathbf{0}$ for $j > r$. Then $TST = T$, as one checks on the basis $\mathcal{B}$. The matrix computation is this argument in coordinates.

Bases adapted to a subspace

For an operator, the convenient basis must work on both sides at once. The first source of such bases is a subspace that T maps into itself.

Proposition 4.5.35 (Preserved subspaces give block triangular matrices). Let $\dim V = n \ge 1$, $T \in \mathrm{L}(V)$, and $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ a basis of V. Let $1 \le k \le n$ and $U = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$. Then $\operatorname{im}(T{\restriction_U}) \subseteq U$ if and only if

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

for some $A \in F^{k\times k}$, $B \in F^{k\times(n-k)}$, and $D \in F^{(n-k)\times(n-k)}$, where 0 is the zero block of size $(n-k)\times k$. In that case, $A = [T{\restriction_U}]_{\mathcal{B}_U\mathcal{B}_U}$, where $\mathcal{B}_U = \mathbf{b}_1, \ldots, \mathbf{b}_k$ and $T{\restriction_U}$ is regarded as an operator on U.

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.14 (Coordinates), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Definition 4.2.1 (Restriction)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. The j-th column of $[T]_{\mathcal{B}\mathcal{B}}$ holds the coefficients of $T(\mathbf{b}_j)$. When does a vector lie in U, in terms of its $\mathcal{B}$-coordinates?
Hint 2. $\operatorname{im}(T{\restriction_U})$ is spanned by $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_k)$, so it lies in U exactly when each $T(\mathbf{b}_j)$ with $j \le k$ does.
Sketch. A vector lies in U exactly when its last $n-k$ coordinates are 0. The image of $T{\restriction_U}$ is spanned by $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_k)$, so it lies in U exactly when the first k columns of $[T]_{\mathcal{B}\mathcal{B}}$ vanish below row k. Their top parts are then the coordinates of $T(\mathbf{b}_j)$ in $\mathcal{B}_U$, the columns of $[T{\restriction_U}]$.

A vector $\mathbf{v}$ lies in U if and only if its $\mathcal{B}$-coordinates $(c_1, \ldots, c_n)$ satisfy $c_{k+1} = \cdots = c_n = 0$: if they do, $\mathbf{v} = \sum_{i\le k} c_i\mathbf{b}_i \in U$; conversely, a vector of U is $\sum_{i\le k} c_i\mathbf{b}_i$, and by uniqueness of coordinates (Definition 1.6.14) these are its coordinates, with zeros after position k.

The list $\mathbf{b}_1, \ldots, \mathbf{b}_k$ is a basis of U, so $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_k)$ spans $\operatorname{im}(T{\restriction_U})$ (Proposition 2.6.22(d), applied to $T{\restriction_U}$). Since U is a subspace, $\operatorname{im}(T{\restriction_U}) \subseteq U$ if and only if $T(\mathbf{b}_j) \in U$ for each $j \le k$, that is, if and only if the j-th column $[T(\mathbf{b}_j)]_{\mathcal{B}}$ of $[T]_{\mathcal{B}\mathcal{B}}$ has zeros in rows $k+1, \ldots, n$ for each $j \le k$. This is the block form.

In that case $T(\mathbf{b}_j) = \sum_{i\le k} a_{ij}\mathbf{b}_i$ for $j \le k$, where (a_{ij}) is the top left block. So the $\mathcal{B}_U$-coordinates of $T{\restriction_U}(\mathbf{b}_j)$ are the j-th column of A, and $A = [T{\restriction_U}]_{\mathcal{B}_U\mathcal{B}_U}$ (Definition 2.3.1). □

If also the span of the last $n-k$ basis vectors is mapped into itself, the same argument shows that the block B is zero, and $[T]_{\mathcal{B}\mathcal{B}}$ is block diagonal. Subspaces with $\operatorname{im}(T{\restriction_U}) \subseteq U$ are called invariant subspaces of T; they are studied in Section 7.2.

Technique 4.5.36 (Choose a basis adapted to a preserved subspace). If T maps a subspace U into itself, take a basis of U and extend it to a basis of V (Theorem 4.1.16). In this basis the matrix of T is block upper triangular (Proposition 4.5.35), the top left block is the matrix of T on U, and a question about T splits into a question about T on U and one about the bottom right block. If V is assembled from several preserved subspaces, list bases of each in turn, and the matrix becomes block diagonal.

Reach for it when you can see a subspace that T preserves: a kernel or image of something commuting with T, the span of a vector fixed or scaled by T, or the span of $\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}$. It is the first step of triangularization (Section 7.5) and of the Jordan form (Section 10.3).
Used in. Problem 4.4, Problem 4.12, Problem 4.13, Corollary 7.3.16, Exercise 7.3.41, Theorem 7.4.15, Lemma 7.5.11, Exercise 4.5.42.

Example 4.5.37 (The cyclic permutation of coordinates). Let $Z \in F^{3\times 3}$ be the matrix with $Z(x,y,z) = (z,x,y)$, so that $Z^3 = I_3$. Two subspaces are visibly preserved by T_Z: the line $U = \operatorname{span}(\mathbf{u}_1)$ with $\mathbf{u}_1 = (1,1,1)$, which Z fixes, and the plane $U' = \{(x,y,z) \mid x+y+z = 0\}$, since permuting the coordinates does not change their sum. Let $\mathbf{u}_2 = (1,-1,0)$ and $\mathbf{u}_3 = (0,1,-1)$, both in U'.

Suppose $3 \neq 0$ in F. Then $\mathcal{U} = \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is a basis of F^3: if $a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 = (a+b, a-b+c, a-c) = \mathbf{0}$, then $b = -a$ and $c = a$, so the middle entry is $3a = 0$, and $a = b = c = 0$. Now

$$Z\mathbf{u}_1 = \mathbf{u}_1, \quad Z\mathbf{u}_2 = (0,1,-1) = \mathbf{u}_3, \quad Z\mathbf{u}_3 = (-1,0,1) = -\mathbf{u}_2 - \mathbf{u}_3,$$

so

$$[T_Z]_{\mathcal{U}\mathcal{U}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}, \quad R = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

The matrix is block diagonal because both U and $\operatorname{span}(\mathbf{u}_2, \mathbf{u}_3) = U'$ are preserved. The trace is $0 = 1 + \operatorname{tr} R$, as it must be, and $Z^3 = I_3$ becomes $R^3 = I_2$, which is easy to check by hand. Over $\mathbb{R}$, R turns out to be similar to the rotation by $2\pi/3$ (Exercise 4.5.43), which matches the picture of Z as a rotation of $\mathbb{R}^3$ about the line U.

If $3 = 0$ in F, then $\mathbf{u}_1 \in U'$, so U and U' no longer fill F^3; the list $\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1$ still gives a block upper triangular matrix, but not a block diagonal one.

The normal forms ahead are all answers to one question: in the best basis, how simple can the matrix of an operator be? For maps between two spaces, Corollary 4.5.22 gave the complete answer. For operators the answer is diagonal when there is a basis of eigenvectors (Section 7.4), upper triangular over $\mathbb{C}$ (Section 7.5), and the Jordan form in general over $\mathbb{C}$ (Section 10.3).

Exercises

Exercise 4.5.38. ROUTINE Over $\mathbb{F}_3$, let $\mathcal{B} = (1,1,0), (0,1,1), (1,0,1)$ and $\mathring{\mathcal{B}} = (1,0,0), (1,1,0), (1,1,1)$.

(a) Show that $\mathcal{B}$ is a basis of $\mathbb{F}_3^3$, and compute $[\mathrm{id}]_{\mathcal{B}\mathcal{E}}$.
(b) Find $[(1,2,0)]_{\mathcal{B}}$.
(c) Compute $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}}$.
(d) Is the same list $\mathcal{B}$ a basis of $\mathbb{F}_2^3$?

Solution on page 1333

Exercise 4.5.39. ROUTINE Prove or disprove.

(a) Every change-of-coordinates matrix is invertible.
(b) If $A \sim B$, then A and B have the same diagonal entries, possibly in a different order.
(c) If $T \in \mathrm{L}(V,W)$ is injective and $\dim V = \dim W < \infty$, then T is an isomorphism.
(d) Two matrices in $F^{n\times n}$ with the same rank and the same trace are similar.
(e) If $\dim V = 2$ and $\dim \mathrm{L}(V,W) = 6$, then $W \cong F^3$.
(f) The only matrix similar to cI_n is cI_n.

Solution on page 1334

Exercise 4.5.40. STANDARD Let $a \in F$ and $n \geq 1$. In $F[x]_{\leq n}$, let $\mathcal{B} = 1, x, \ldots, x^n$ and $\mathcal{B}_a = 1, x-a, (x-a)^2, \ldots, (x-a)^n$. Index rows and columns from 0 to n.

(a) Show that $\mathcal{B}_a$ is a basis, and that the (i,j) entry of $[\mathrm{id}]_{\mathcal{B}_a\mathcal{B}}$ is $\binom{j}{i}a^{j-i}$ for $i \leq j$ and 0 for $i > j$.
(b) Find the entries of $[\mathrm{id}]_{\mathcal{B}\mathcal{B}_a}$ in the same way.
(c) Taking $F = \mathbb{Q}$ and $a = 1$, deduce that $\sum_{k=i}^{j}(-1)^{k-i}\binom{j}{k}\binom{k}{i}$ is 1 if $i = j$ and 0 if $i < j$.

Solution on page 1335

Exercise 4.5.41. STANDARD Let V and W be finite-dimensional with $\dim V = n \geq 1$ and $\dim W = m$, and let $\mathbf{v}_0 \in V$ be nonzero. Show that

$$K = \{T \in \mathrm{L}(V,W) \mid T(\mathbf{v}_0) = \mathbf{0}\}$$

is a subspace of $\mathrm{L}(V,W)$ of dimension $(n-1)m$. Give two proofs: one by choosing a basis of V that contains $\mathbf{v}_0$, and one by rank–nullity.

Solution on page 1335

Solution on page 1335

Exercise 4.5.42. HARD Let V and W be finite-dimensional with $\dim V = n \geq 1$ and $\dim W = m \geq 1$, and let $U \subseteq V$ and $U' \subseteq W$ be subspaces with $\dim U = k$ and $\dim U' = k'$. Show that

$$\{T \in \mathrm{L}(V,W) \mid \mathrm{im}(T{\restriction_U}) \subseteq U'\}$$

is a subspace of $\mathrm{L}(V,W)$ of dimension $mn - k(m-k')$. In particular, the operators $T \in \mathrm{L}(V)$ with $\mathrm{im}(T{\restriction_U}) \subseteq U$ form a subspace of $\mathrm{L}(V)$ of dimension $n^2 - k(n-k)$.

Solution on page 1336

Exercise 4.5.43. HARD For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, let $\delta(A) = ad - bc$.

(a) Show that if $A\mathbf{v} \in \mathrm{span}(\mathbf{v})$ for every $\mathbf{v} \in F^2$, then $A = cI_2$ for some $c \in F$.
(b) Suppose A is not a scalar multiple of I_2. Show that there is $\mathbf{v} \in F^2$ such that $\mathbf{v}, A\mathbf{v}$ is a basis of F^2, and that in this basis T_A has the matrix

$$C_A = \begin{pmatrix} 0 & -\delta(A) \\ 1 & \mathrm{tr}\, A \end{pmatrix}.$$

(c) Deduce that two matrices in $F^{2\times 2}$, neither a scalar multiple of I_2, with the same trace and the same value of δ, are similar.
(d) Show that $A \sim A^{\mathsf{T}}$ for every $A \in F^{2\times 2}$.
(e) Show that the matrix $R = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ of Example 4.5.37 is similar over $\mathbb{R}$ to the rotation matrix $\begin{pmatrix} \cos(2\pi/3) & -\sin(2\pi/3) \\ \sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix}$.

Solution on page 1336

Exercise 4.5.44. HARD Keep the notation $\delta(A) = ad - bc$ of Exercise 4.5.43.

(a) Show that $\delta(XY) = \delta(X)\delta(Y)$ for all $X, Y \in F^{2\times 2}$, and deduce that δ is a similarity invariant.
(b) Show that every $A \in F^{2\times 2}$ is similar to exactly one matrix in the list consisting of cI_2 for $c \in F$ and $\begin{pmatrix} 0 & -d \\ 1 & t \end{pmatrix}$ for $t, d \in F$.
(c) Deduce that over $\mathbb{F}_q$ there are exactly $q^2 + q$ similarity classes of 2×2 matrices.

Solution on page 1337

Exercise 4.5.45. ROUTINE Find the error in the following argument, and show that its conclusion is false.

Claim. Any two matrices $A, B \in F^{n\times n}$ of the same rank r are similar.

Proof. By Theorem 3.5.27 there are invertible G, H, G', H' with $GAH = N_r = G'BH'$. Then $B = G'^{-1}GAHH'^{-1}$. Let $S = HH'^{-1}$, which is invertible. Then $B = \big(G'^{-1}G\big)AS$, a product of A with invertible matrices on both sides, so B is similar to A.

Solution on page 1337

Exercise 4.5.46. STANDARD Let $A, B, C, D \in F^{n\times n}$, $n \geq 2$. Prove or disprove.

(a) If $A \sim B$ and $C \sim D$, then $A + C \sim B + D$.
(b) If $A \sim B$, then $AC \sim BC$.
(c) If $A \sim B$, then $A + cI_n \sim B + cI_n$ for every $c \in F$.
(d) If $A^2 \sim B^2$, then $A \sim B$.

Solution on page 1338

Exercise 4.5.47. HARD Let $f : F^{n\times n} \to F$ be a linear map such that $f\big(S^{-1}AS\big) = f(A)$ for every $A \in F^{n\times n}$ and every invertible $S \in F^{n\times n}$. Show that $f = c\,\mathrm{tr}$ for some $c \in F$. That is, up to a scalar factor, the trace is the only linear similarity invariant.

Solution on page 1338

Exercise 4.5.48. STANDARD For $A, B \in F^{n\times n}$, let $T \in \mathrm{L}(F^{n\times n})$ be $T(X) = AXB$. Show that $\mathrm{tr}\,T = (\mathrm{tr}\,A)(\mathrm{tr}\,B)$. Deduce the traces of the operators $X \mapsto AX$, $X \mapsto XB$, and $X \mapsto AX - XA$.

Solution on page 1339

Exercise 4.5.49. STANDARD Let $\dim V = n \geq 1$, and let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$. Let $r = \mathrm{rank}\,P$.

(a) Show that P fixes every vector of $\mathrm{im}\,P$, that $\ker P \cap \mathrm{im}\,P = \{\mathbf{0}\}$, and that a basis of $\mathrm{im}\,P$ followed by a basis of $\ker P$ is a basis of V.
(b) Deduce that in this basis $[P] = N_r = \mathrm{diag}(1, \ldots, 1, 0, \ldots, 0)$, with r ones, and that $\mathrm{tr}\,P = r \cdot 1$.
(c) Show that two matrices $A, B \in F^{n\times n}$ with $A^2 = A$ and $B^2 = B$ are similar if and only if $\mathrm{rank}\,A = \mathrm{rank}\,B$.

Solution on page 1339

Exercise 4.5.50. STANDARD Let $\dim V = n \geq 1$, and let $N \in \mathrm{L}(V)$ satisfy $N^n = 0$ and $N^{n-1} \neq 0$. Let $J_n \in F^{n\times n}$ be the matrix with 1 in positions $(2,1), (3,2), \ldots, (n, n-1)$ and 0 elsewhere.

(a) Show that $[N]_{\mathcal{B}\mathcal{B}} = J_n$ for some basis $\mathcal{B}$ of V.
(b) Deduce that any two matrices $A, B \in F^{n\times n}$ with $A^n = B^n = 0$, $A^{n-1} \neq 0$, and $B^{n-1} \neq 0$ are similar.
(c) Show that $J_n \sim J_n^{\mathsf{T}}$.

Solution on page 1339

Exercise 4.5.51. HARD Let $n \geq 2$, let $A \in F^{n\times n}$ have rank 1, and let $t = \mathrm{tr}\,A$. Show that $A \sim tU_{11}$ if $t \neq 0$, and $A \sim U_{21}$ if $t = 0$. Conclude that two matrices of rank 1 are similar if and only if they have the same trace.

Solution on page 1340

Exercise 4.5.52. STANDARD Let F be a field with at least three elements. Show that every $A \in F^{n\times n}$ is the sum of two invertible matrices. Show that this fails for the matrix $(1) \in \mathbb{F}_2^{1\times 1}$.

Solution on page 1340

Exercise 4.5.53. CHALLENGE Show that for $n \geq 2$, every $A \in \mathbb{F}_2^{n\times n}$ is the sum of two invertible matrices.

Solution on page 1341

Exercise 4.5.54. STANDARD This exercise shows what goes wrong in infinite dimensions.

(a) Let $xF[x]$ be the subspace of polynomials in $F[x]$ whose constant coefficient is 0. Show that $S : F[x] \to xF[x]$, $S(p) = xp$, is an isomorphism. So $F[x]$ is isomorphic to a proper subspace of itself, and S, regarded as an operator on $F[x]$, is injective but not surjective.
(b) Explain why a finite-dimensional space is never isomorphic to a proper subspace of itself.
(c) Show that differentiation $D \in \mathrm{L}(\mathbb{R}[x])$ is surjective but not injective.
(d) Show that if $V \cong W$ and V is infinite-dimensional, then so is W.

Solution on page 1341

Exercise 4.5.55. STANDARD Let $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the rotation by $\pi/2$.

(a) Show that R is similar over $\mathbb{C}$ to $\operatorname{diag}(i, -i)$, by finding $S \in \mathbb{C}^{2\times 2}$.
(b) Show that R is not similar over $\mathbb{R}$ to any diagonal matrix in $\mathbb{R}^{2\times 2}$.
(c) Show that R is similar over $\mathbb{R}$ to $-R$, and interpret the similarity geometrically.

Solution on page 1342

Exercise 4.5.56. ROUTINE Let $T \in \mathrm{L}(F^2)$ satisfy $T(1, 2) = (3, 6)$ and $T(1, 1) = (-1, -1)$. Use Theorem 4.5.17 to find $[T]_{\mathcal{E}\mathcal{E}}$. Check your answer, and compute $\operatorname{tr} T$ in two ways.

Solution on page 1342

4.6 Products and quotients BEYOND 21-242

A vector space can be built out of others in two basic ways. A *product* places several spaces side by side: a vector of $V \times W$ is a pair, one vector from each. A *quotient* goes the other way. It takes a subspace U of V and collapses it to a single point, so that two vectors of V become equal when they differ by an element of U. Recall from Corollary 2.6.31 that a linear map T slices its domain into parallel copies of $\ker T$ and sends each slice to one vector of $\operatorname{im} T$. The quotient $V/\ker T$ turns the set of slices into a vector space on which T becomes injective, and this is the first isomorphism theorem (Theorem 4.6.34). It gives a third proof of rank–nullity. Products, for their part, give the cleanest way to say when a sum of subspaces is direct.

Neither construction appears in the 21-242 lectures (they are in LADR, Section 3E). Quotients come back in Section 7.5, where they give a second route to triangularization, and in Section 13.1.

Throughout, V and W are vector spaces over F.

4.6.1 Product spaces

Definition 4.6.1 (Product space). Let $V_1, \ldots, V_m$ be vector spaces over F. Their **product** is the set of m-tuples

$$V_1 \times \cdots \times V_m = \{(\mathbf{v}_1, \ldots, \mathbf{v}_m) \mid \mathbf{v}_j \in V_j \text{ for } j = 1, \ldots, m\},$$

with the componentwise operations

$$(\mathbf{v}_1, \ldots, \mathbf{v}_m) + (\mathbf{w}_1, \ldots, \mathbf{w}_m) = (\mathbf{v}_1 + \mathbf{w}_1, \ldots, \mathbf{v}_m + \mathbf{w}_m),$$

$$a(\mathbf{v}_1, \ldots, \mathbf{v}_m) = (a\mathbf{v}_1, \ldots, a\mathbf{v}_m).$$

For a single space V we write $V^m = V \times \cdots \times V$ (m factors).

Proposition 4.6.2 (Products are vector spaces). $V_1 \times \cdots \times V_m$ is a vector space over F. Its zero vector is $(\mathbf{0}, \ldots, \mathbf{0})$, and $-(\mathbf{v}_1, \ldots, \mathbf{v}_m) = (-\mathbf{v}_1, \ldots, -\mathbf{v}_m)$.

Proof. ROUTINE

Uses. Definition 1.2.1 (Vector space), Definition 4.6.1 (Product space)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Two tuples are equal exactly when they are equal in every component.
Hint 2. Write out the jth component of both sides of each axiom. It is the same axiom in V_j.
Sketch. Every axiom is an identity between tuples, and its jth component is that axiom in V_j. The zero and the negatives are found the same way.

Sums and scalar multiples of tuples are tuples of the right kind, since each V_j is closed under its operations. Each vector space axiom is an identity between elements of the product, and two tuples are equal when they agree in every component. In the jth component, each axiom becomes the same axiom in V_j. For example, for $\mathbf{x} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ and $\mathbf{y} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$, the jth components of $a(\mathbf{x} + \mathbf{y})$ and $a\mathbf{x} + a\mathbf{y}$ are $a(\mathbf{v}_j + \mathbf{w}_j)$ and $a\mathbf{v}_j + a\mathbf{w}_j$, which are equal by distributivity in V_j. The other identities (commutativity, associativity, the second distributive law, $(ab)\mathbf{x} = a(b\mathbf{x})$, and $1\mathbf{x} = \mathbf{x}$) are checked identically. Finally, $(\mathbf{v}_1, \ldots, \mathbf{v}_m) + (\mathbf{0}, \ldots, \mathbf{0}) = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ and $(\mathbf{v}_1, \ldots, \mathbf{v}_m) + (-\mathbf{v}_1, \ldots, -\mathbf{v}_m) = (\mathbf{0}, \ldots, \mathbf{0})$, componentwise. □

Example 4.6.3 (Products).

(a) F^m is $F \times \cdots \times F$ (m factors); Definition 1.2.4 is the case $V_1 = \cdots = V_m = F$, with the tuples written as columns.

(b) An element of $F^2 \times F^3$ is a pair $((a, b), (c, d, e))$, which is not literally an element of F^5. But removing the inner parentheses, $((a, b), (c, d, e)) \mapsto (a, b, c, d, e)$, is a bijective linear map, so $F^2 \times F^3 \cong F^5$ (Definition 4.5.1). In the same way $F^n \times F^p \cong F^{n+p}$, and we identify the two whenever convenient.

(c) In $V \times W$, the subsets $V \times \{\mathbf{0}\}$ and $\{\mathbf{0}\} \times W$ are subspaces, and $V \times W = (V \times \{\mathbf{0}\}) \oplus (\{\mathbf{0}\} \times W)$: every $(\mathbf{v}, \mathbf{w})$ is $(\mathbf{v}, \mathbf{0}) + (\mathbf{0}, \mathbf{w})$, and this is the only way to write it as such a sum, since the components are forced.

Non-example 4.6.4 (Scaling only one component). Let $W \neq \{\mathbf{0}\}$. On the set $V \times W$ keep componentwise addition, but scale only the first component: $a \odot (\mathbf{v}, \mathbf{w}) = (a\mathbf{v}, \mathbf{w})$. This is not a vector space. In any vector space $0\mathbf{x} = \mathbf{0}$ (Proposition 1.2.32), but here $0 \odot (\mathbf{v}, \mathbf{w}) = (\mathbf{0}, \mathbf{w})$, which is not the zero vector $(\mathbf{0}, \mathbf{0})$ when $\mathbf{w} \neq \mathbf{0}$. (The axiom that fails is distributivity over scalar addition: $(1 + 1) \odot (\mathbf{v}, \mathbf{w})$ has second component $\mathbf{w}$, while $1 \odot (\mathbf{v}, \mathbf{w}) + 1 \odot (\mathbf{v}, \mathbf{w})$ has second component $\mathbf{w} + \mathbf{w}$.)

For each j there are two linear maps connecting V_j with the product. The **inclusion** $\iota_j : V_j \to V_1 \times \cdots \times V_m$ puts a vector in the jth slot and $\mathbf{0}$ in the others, and the **projection** $\mathrm{Pr}_j : V_1 \times \cdots \times V_m \to V_j$ reads off the jth component. Both are linear because the operations are componentwise, $\mathrm{Pr}_j \iota_j = \mathrm{id}_{V_j}$, and every vector of the product splits as

$$(\mathbf{v}_1, \ldots, \mathbf{v}_m) = \iota_1(\mathbf{v}_1) + \cdots + \iota_m(\mathbf{v}_m). \tag{4.6.1}$$

Theorem 4.6.5 (Dimension of a product). Suppose $V_1, \ldots, V_m$ are finite-dimensional. Then $V_1 \times \cdots \times V_m$ is finite-dimensional, and

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

More precisely, if $\mathcal{B}_j$ is a basis of V_j for each j, then the list obtained by concatenating $\iota_1(\mathcal{B}_1), \ldots, \iota_m(\mathcal{B}_m)$ is a basis of the product.

The first proof checks the basis directly. The second never checks a basis of the product: it peels off one factor with a surjection and counts the kernel by rank–nullity, which is how quotients will be measured below.

Proof (by building a basis). ROUTINE

> *Uses.* Definition 1.6.1 (Basis), Proposition 1.6.2 (Bases are independent spanning lists), (4.6.1), Theorem 1.5.5 (Independence is decided at zero)
> *Techniques.* Fix an arbitrary element and chase it (Technique 1.3.13)
> *Hint 1.* Take a basis $\mathbf{b}_{j,1}, \dots, \mathbf{b}_{j,n_j}$ of each V_j. Which vectors of the product would you guess form a basis?
> *Hint 2.* For spanning, use (4.6.1) and expand each $\mathbf{v}_j$ in its basis. For independence, look at the jth component of a linear relation.
> *Sketch.* By (4.6.1) and the linearity of ι_j, each tuple is a combination of the $\iota_j(\mathbf{b}_{j,i})$. The jth component of a relation $\sum_{j,i} c_{j,i}\iota_j(\mathbf{b}_{j,i}) = \mathbf{0}$ is $\sum_i c_{j,i}\mathbf{b}_{j,i} = \mathbf{0}$, so all $c_{j,i} = 0$.

Let $\mathcal{B}_j = \mathbf{b}_{j,1}, \dots, \mathbf{b}_{j,n_j}$ be a basis of V_j, so $n_j = \dim V_j$, and let $\mathcal{L}$ be the concatenation of the lists $\iota_j\big(\mathcal{B}_j\big)$. It has $n_1 + \cdots + n_m$ entries, so it suffices to show that $\mathcal{L}$ is a basis.

Spanning. Let $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ be in the product, and write $\mathbf{v}_j = \sum_i c_{j,i}\mathbf{b}_{j,i}$. By (4.6.1) and the linearity of each ι_j,

$$(\mathbf{v}_1, \dots, \mathbf{v}_m) = \sum_{j=1}^{m} \iota_j(\mathbf{v}_j) = \sum_{j=1}^{m}\sum_{i=1}^{n_j} c_{j,i}\iota_j(\mathbf{b}_{j,i}).$$

Independence. Suppose $\sum_{j,i} c_{j,i}\iota_j(\mathbf{b}_{j,i}) = \mathbf{0}$. The jth component of $\iota_j(\mathbf{b}_{j,i})$ is $\mathbf{b}_{j,i}$, and the jth component of $\iota_k(\mathbf{b}_{k,i})$ is $\mathbf{0}$ for $k \neq j$. So the jth component of the relation reads $\sum_i c_{j,i}\mathbf{b}_{j,i} = \mathbf{0}$, and since $\mathcal{B}_j$ is independent, $c_{j,1} = \cdots = c_{j,n_j} = 0$. This holds for every j. □

Proof (by induction on m, with rank–nullity). STANDARD

> *Uses.* Theorem 4.2.11 (Rank–nullity), Proposition 4.2.6 (Injective maps preserve dimension), (4.6.1), Definition 1.6.23 (Finite-dimensional)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* For $m \geq 2$, the projection Pr_m onto the last factor is surjective. What is its kernel?
> *Hint 2.* The kernel of Pr_m consists of the tuples $(\mathbf{v}_1, \dots, \mathbf{v}_{m-1}, \mathbf{0})$, a copy of the product of the first $m-1$ factors. Apply rank–nullity, after checking that the product is finite-dimensional at all.
> *Sketch.* The vectors $\iota_j(\mathbf{b})$, $\mathbf{b}$ in a basis of V_j, span the product by (4.6.1), so it is finite-dimensional. Pr_m is surjective, and $(\mathbf{v}_1, \dots, \mathbf{v}_{m-1}) \mapsto (\mathbf{v}_1, \dots, \mathbf{v}_{m-1}, \mathbf{0})$ is injective with image $\ker \mathrm{Pr}_m$. Rank–nullity gives $\dim = \dim(V_1 \times \cdots \times V_{m-1}) + \dim V_m$, and induction finishes.

We prove the dimension formula by induction on m; the case $m = 1$ is trivial. Let $m \geq 2$ and $P = V_1 \times \cdots \times V_m$. By (4.6.1), the vectors $\iota_j(\mathbf{b})$, with $\mathbf{b}$ running through a basis of V_j and $j = 1, \dots, m$, span P (expand each $\mathbf{v}_j$ in its basis), so P is finite-dimensional. The projection $\mathrm{Pr}_m : P \to V_m$ is linear and surjective, since $\mathrm{Pr}_m\, \iota_m(\mathbf{v}) = \mathbf{v}$. Its kernel is the set of tuples whose last component is $\mathbf{0}$, which is the image of

$$J : V_1 \times \cdots \times V_{m-1} \to P, \quad (\mathbf{v}_1, \dots, \mathbf{v}_{m-1}) \mapsto (\mathbf{v}_1, \dots, \mathbf{v}_{m-1}, \mathbf{0}).$$

J is linear and injective, so by Proposition 4.2.6(a) and the induction hypothesis, $\dim \ker \mathrm{Pr}_m = \dim(V_1 \times \cdots \times V_{m-1}) = \dim V_1 + \cdots + \dim V_{m-1}$. By rank–nullity (Theorem 4.2.11),

$$\dim P = \dim \ker \mathrm{Pr}_m + \dim \operatorname{im} \mathrm{Pr}_m = (\dim V_1 + \cdots + \dim V_{m-1}) + \dim V_m,$$

as desired. (This proof gives the dimension only; the explicit basis is the first proof's.) □

Exercise 4.6.6. ROUTINE

(a) Show that if V' is a subspace of V and W' is a subspace of W, then $V' \times W'$ is a subspace of $V \times W$.

(b) Prove or disprove: every subspace of $V \times W$ is of the form $V' \times W'$ as in (a).

Solution on page 1343

Maps into and out of a product

A map into a product is the same as a list of maps into the factors, and a map out of a product is the same as a list of maps out of the factors.

Proposition 4.6.7 (Maps into and out of a product).

(a) Let $T_j \in \mathrm{L}\big(U, W_j\big)$ for $j = 1, \dots, m$. Then

$$(T_1, \dots, T_m) : U \to W_1 \times \cdots \times W_m, \quad \mathbf{u} \mapsto (T_1(\mathbf{u}), \dots, T_m(\mathbf{u})),$$

is linear, and every $T \in \mathrm{L}(U, W_1 \times \cdots \times W_m)$ equals $(T_1, \dots, T_m)$ for exactly one list of linear maps, namely $T_j = \mathrm{Pr}_j\, T$.

(b) Let $S_j \in \mathrm{L}\big(V_j, W\big)$ for $j = 1, \dots, m$. Then

$$V_1 \times \cdots \times V_m \to W, \quad (\mathbf{v}_1, \dots, \mathbf{v}_m) \mapsto S_1(\mathbf{v}_1) + \cdots + S_m(\mathbf{v}_m),$$

is linear, and every $S \in \mathrm{L}(V_1 \times \cdots \times V_m, W)$ is of this form for exactly one list of linear maps, namely $S_j = S\iota_j$.

Proof. ROUTINE

Uses. Theorem 2.1.3 (Checking linearity), Proposition 2.1.17 (Compositions of linear maps are linear), (4.6.1)
Techniques. Assemble linear maps from linear pieces (Technique 2.1.26), Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Linearity in both parts is a componentwise computation.
Hint 2. For the converse in (a), compare components of $T(\mathbf{u})$. In (b), apply S to (4.6.1).
Sketch. (a) The jth component of $(T_1, \dots, T_m)(a\mathbf{u} + \mathbf{u}')$ is $aT_j(\mathbf{u}) + T_j(\mathbf{u}')$. A map T into the product is determined by the components $\mathrm{Pr}_j\, T(\mathbf{u})$ of its values. (b) Linearity is a sum of linear computations, and applying S to (4.6.1) gives $S(\mathbf{v}_1, \dots, \mathbf{v}_m) = \sum_j S\iota_j(\mathbf{v}_j)$. Uniqueness: evaluate at $\iota_j(\mathbf{v})$.

(a) For $\mathbf{u}, \mathbf{u}' \in U$ and $a \in F$, the jth component of $(T_1, \dots, T_m)(a\mathbf{u} + \mathbf{u}')$ is $T_j(a\mathbf{u} + \mathbf{u}') = aT_j(\mathbf{u}) + T_j(\mathbf{u}')$, which is the jth component of $a(T_1, \dots, T_m)(\mathbf{u}) + (T_1, \dots, T_m)(\mathbf{u}')$. So the map is linear (Theorem 2.1.3). Now let $T \in \mathrm{L}(U, W_1 \times \cdots \times W_m)$. Each $\mathrm{Pr}_j\, T$ is linear (Proposition 2.1.17), and the jth component of $T(\mathbf{u})$ is $\mathrm{Pr}_j\, T(\mathbf{u})$, so $T = (\mathrm{Pr}_1\, T, \dots, \mathrm{Pr}_m\, T)$. Conversely, if $T = (T_1, \dots, T_m)$, then $T_j(\mathbf{u})$ is the jth component of $T(\mathbf{u})$, that is, $T_j = \mathrm{Pr}_j\, T$. So the list is unique.

(b) Write S' for the map defined by the S_j. For tuples $\mathbf{x} = \big(\mathbf{v}_j\big)_j$, $\mathbf{y} = \big(\mathbf{w}_j\big)_j$ and $a \in F$,

$$S'(a\mathbf{x} + \mathbf{y}) = \sum_{j=1}^{m} S_j\big(a\mathbf{v}_j + \mathbf{w}_j\big) = a\sum_{j=1}^{m} S_j\big(\mathbf{v}_j\big) + \sum_{j=1}^{m} S_j\big(\mathbf{w}_j\big) = aS'(\mathbf{x}) + S'(\mathbf{y}).$$

Also $S'\iota_j(\mathbf{v}) = S_j(\mathbf{v})$, since the other components of $\iota_j(\mathbf{v})$ are $\mathbf{0}$; so $S'\iota_j = S_j$, which shows that the list is determined by the map. Finally, let $S \in \mathrm{L}(V_1 \times \cdots \times V_m, W)$. Applying S to (4.6.1),

$$S(\mathbf{v}_1, \dots, \mathbf{v}_m) = \sum_{j=1}^{m} S\big(\iota_j(\mathbf{v}_j)\big) = \sum_{j=1}^{m} \big(S\iota_j\big)(\mathbf{v}_j),$$

so S is the map defined by the list $S\iota_1, \dots, S\iota_m$ of linear maps. □

The assignments $T \mapsto (\mathrm{Pr}_1\, T, \dots, \mathrm{Pr}_m\, T)$ and $S \mapsto (S\iota_1, \dots, S\iota_m)$ are linear, so the proposition says that

$$\mathrm{L}(U, W_1 \times \cdots \times W_m) \cong \mathrm{L}(U, W_1) \times \cdots \times \mathrm{L}(U, W_m), \quad \mathrm{L}(V_1 \times \cdots \times V_m, W) \cong \mathrm{L}(V_1, W) \times \cdots \times \mathrm{L}(V_m, W).$$

Combining the two for $F^n \times F^p = F^{n+p}$ and $F^m \times F^q = F^{m+q}$: a linear map $F^{n+p} \to F^{m+q}$ is the same as four maps $\mathrm{Pr}_i\, T\iota_j$, and its matrix is the block matrix whose blocks are their matrices (Definition 2.4.33).

4.6.2 Products and direct sums

A sum of subspaces is the image of a single linear map out of a product, and directness becomes injectivity of that map.

Definition 4.6.8 (Addition map). Let $U_1, \dots, U_m$ be subspaces of V. The **addition map** is

$$\Gamma : U_1 \times \cdots \times U_m \to V, \quad \Gamma(\mathbf{u}_1, \dots, \mathbf{u}_m) = \mathbf{u}_1 + \cdots + \mathbf{u}_m.$$

By Proposition 4.6.7(b), applied to the inclusions $U_j \to V$, the map Γ is linear, and by Definition 4.4.1 its image is $U_1 + \cdots + U_m$. More is true: for $\mathbf{v} \in V$, the elements of the fiber $\Gamma^{-1}(\mathbf{v})$ are exactly the decompositions $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_m$ of $\mathbf{v}$ along $U_1, \dots, U_m$, each recorded as the tuple $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ of its components.

Theorem 4.6.9 (Direct sums and the addition map). Let $U_1, \dots, U_m$ be subspaces of V. The sum $U_1 + \cdots + U_m$ is direct if and only if Γ is injective. In that case Γ is an isomorphism from $U_1 \times \cdots \times U_m$ onto $U_1 \oplus \cdots \oplus U_m$. In particular, if $V = U_1 \oplus \cdots \oplus U_m$, then $V \cong U_1 \times \cdots \times U_m$.

Proof. ROUTINE

Uses. Definition 4.4.4 (Direct sum), Definition 2.6.1 (Injective), Definition 4.6.8 (Addition map), Proposition 2.7.7 (Bijective linear maps are invertible)
Hint 1. What are the preimages of $\mathbf{v}$ under Γ?
Hint 2. The preimages of $\mathbf{v}$ are its decompositions. Vectors outside $\operatorname{im}\Gamma$ have no preimages at all.
Sketch. $\Gamma^{-1}(\mathbf{v})$ is the set of decompositions of $\mathbf{v}$. The sum is direct when every $\mathbf{v} \in \operatorname{im}\Gamma$ has exactly one, that is, when every $\mathbf{v} \in V$ has at most one preimage. That is injectivity. An injective linear map is an isomorphism onto its image.

By the remark before the theorem, the decompositions of $\mathbf{v}$ along $U_1, \dots, U_m$ correspond to the elements of $\Gamma^{-1}(\mathbf{v})$. By Definition 4.4.4, the sum is direct exactly when every $\mathbf{v} \in U_1 + \cdots + U_m = \operatorname{im}\Gamma$ has exactly one preimage. Vectors outside $\operatorname{im}\Gamma$ have none, and vectors in $\operatorname{im}\Gamma$ have at least one. So the sum is direct if and only if every $\mathbf{v} \in V$ has at most one preimage under Γ, which is injectivity (Definition 2.6.1). In that case Γ, regarded as a map onto $U_1 \oplus \cdots \oplus U_m$, is linear and bijective, hence an isomorphism (Proposition 2.7.7). □

By Theorem 2.6.11, injectivity of Γ means $\ker\Gamma = \{\mathbf{0}\}$: the only decomposition of $\mathbf{0}$ is the trivial one. That is criterion (II) of Theorem 4.4.7, now read as a statement about a kernel. The point of the reformulation is that kernels can be counted.

Corollary 4.6.10 (Dimension of a sum). Let $U_1, \dots, U_m$ be finite-dimensional subspaces of V. Then

$$\dim(U_1 + \cdots + U_m) \le \dim U_1 + \cdots + \dim U_m,$$

with equality if and only if the sum is direct.

Proof. ROUTINE

Uses. Theorem 4.6.5 (Dimension of a product), Theorem 4.2.11 (Rank–nullity), Theorem 4.6.9 (Direct sums and the addition map), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Apply rank–nullity to Γ.
Hint 2. The domain has dimension $\sum_j \dim U_j$ by Theorem 4.6.5, and the image is the sum. What does $\dim\ker\Gamma = 0$ mean?
Sketch. Rank–nullity for Γ gives $\dim(U_1 + \cdots + U_m) = \sum_j \dim U_j - \dim\ker\Gamma$. Equality holds iff $\ker\Gamma = \{\mathbf{0}\}$ iff Γ is injective iff the sum is direct.

The domain of Γ is finite-dimensional of dimension $\sum_j \dim U_j$ by Theorem 4.6.5, and $\operatorname{im}\Gamma = U_1 + \cdots + U_m$. By rank–nullity (Theorem 4.2.11),

$$\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m - \dim\ker\Gamma \le \dim U_1 + \cdots + \dim U_m.$$

Equality holds if and only if $\dim\ker\Gamma = 0$, that is, $\ker\Gamma = \{\mathbf{0}\}$. By Theorem 2.6.11 and Theorem 4.6.9, this happens exactly when the sum is direct. □

So $\dim\ker\Gamma$ measures how far a sum is from being direct. For two subspaces the kernel can be named, and the count becomes the dimension formula of Theorem 4.4.19.

Proposition 4.6.11 (The kernel of the addition map). Let U and W be subspaces of V, and let $\Gamma : U \times W \to V$, $(\mathbf{u}, \mathbf{w}) \mapsto \mathbf{u} + \mathbf{w}$.

(a) $\ker \Gamma = \{(\mathbf{u}, -\mathbf{u}) \mid \mathbf{u} \in U \cap W\}$, and $\mathbf{u} \mapsto (\mathbf{u}, -\mathbf{u})$ is an isomorphism from $U \cap W$ onto $\ker \Gamma$.

(b) If U and W are finite-dimensional, then $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

Proof. ROUTINE

Uses. Proposition 4.6.7 (Maps into and out of a product), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.6.5 (Dimension of a product), Theorem 4.2.11 (Rank–nullity)
Techniques. Double containment (Technique 1.3.14), Dimension counting (Technique 4.2.24)
Hint 1. If $\mathbf{u} + \mathbf{w} = \mathbf{0}$, then $\mathbf{w} = -\mathbf{u}$. Which subspaces does $\mathbf{u}$ lie in?
Hint 2. For (b), apply rank–nullity to Γ, using (a) for the kernel and Theorem 4.6.5 for the domain.
Sketch. (a) $\mathbf{u} + \mathbf{w} = \mathbf{0}$ forces $\mathbf{u} = -\mathbf{w} \in U \cap W$ and $(\mathbf{u}, \mathbf{w}) = (\mathbf{u}, -\mathbf{u})$; conversely such pairs are killed. The map $\mathbf{u} \mapsto (\mathbf{u}, -\mathbf{u})$ is linear and injective with image $\ker \Gamma$. (b) $\dim U + \dim W = \dim \ker \Gamma + \dim(U + W)$ and $\dim \ker \Gamma = \dim(U \cap W)$.

(a) If $(\mathbf{u}, \mathbf{w}) \in \ker \Gamma$, then $\mathbf{w} = -\mathbf{u}$. So $\mathbf{u} = -\mathbf{w} \in W$, hence $\mathbf{u} \in U \cap W$ and $(\mathbf{u}, \mathbf{w}) = (\mathbf{u}, -\mathbf{u})$. Conversely, if $\mathbf{u} \in U \cap W$, then $(\mathbf{u}, -\mathbf{u}) \in U \times W$ and $\Gamma(\mathbf{u}, -\mathbf{u}) = \mathbf{0}$. This proves the description of $\ker \Gamma$. The map $\varphi : U \cap W \to U \times W$, $\varphi(\mathbf{u}) = (\mathbf{u}, -\mathbf{u})$, is linear by Proposition 4.6.7(a), it is injective because the first component of $\varphi(\mathbf{u})$ is $\mathbf{u}$, and its image is $\ker \Gamma$. So it is an isomorphism onto $\ker \Gamma$.

(b) $U \cap W$ is a subspace of the finite-dimensional space U, hence finite-dimensional (Proposition 4.1.21), and by (a) and Proposition 4.2.6(a), $\dim \ker \Gamma = \dim(U \cap W)$. The domain $U \times W$ has dimension $\dim U + \dim W$ (Theorem 4.6.5) and $\operatorname{im} \Gamma = U + W$. Rank–nullity (Theorem 4.2.11) gives

$$\dim U + \dim W = \dim(U \cap W) + \dim(U + W),$$

as desired. □

This proof builds no basis of $U + W$. It counts instead: the pairs in $U \times W$ that add up to $\mathbf{0}$ form a copy of $U \cap W$, and that copy is exactly what the naive count $\dim U + \dim W$ counts twice.

4.6.3 Cosets

Intuition (Collapsing a subspace to a point). Take $V = \mathbb{R}^2$ and let U be the line through $\mathbf{0}$ spanned by $(2, 1)$. The translates $\mathbf{v} + U$ are the lines parallel to U. They stack up across the plane like the rules on a sheet of paper, and every point of the plane lies on exactly one of them.

Now forget everything about a vector except which of these lines it lies on. Two vectors become indistinguishable exactly when they differ by an element of U, so the whole line U collapses to one point, and so does every line parallel to it. One point per line remains. The lines are labelled by a single number, the value $c = x - 2y$ that is constant along each of them, so the collapsed plane looks like a copy of $\mathbb{R}$ (Figure 4.6.1).

The quotient V/U is this collapsed space. Its elements are the lines themselves, and it becomes a vector space once we say how to add and scale lines.

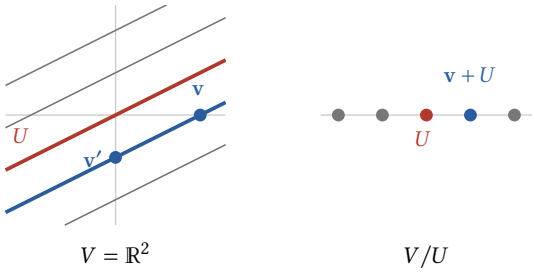


Figure 4.6.1. Left: the line $U = \operatorname{span}((2,1))$ (red) and some of its translates, the lines $x - 2y = c$. The vectors $\mathbf{v} = (2,0)$ and $\mathbf{v}' = (0,-1)$ lie on the same translate (blue), because $\mathbf{v} - \mathbf{v}' = (2,1) \in U$. Right: in V/U each line becomes a single point, drawn at position c; the line U becomes the origin.

Definition 4.6.12 (Coset). Let U be a subspace of V and $\mathbf{v} \in V$. The translate

$$\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\}$$

(Notation 3.1.17) is called a **coset** of U, and $\mathbf{v}$ is a **representative** of it.

A coset is a subset of V, and it has many representatives: by Lemma 4.6.14 below, every element of $\mathbf{v} + U$ is one. Cosets reappear in Section 4.7 as affine subspaces, where their geometry is studied in its own right.

Example 4.6.13 (Cosets).

(a) In $\mathbb{R}^2$ let $U = \operatorname{span}((2,1))$. The coset $(v_1, v_2) + U$ is the line $x - 2y = v_1 - 2v_2$. Indeed, its points $(v_1 + 2t, v_2 + t)$ satisfy $x - 2y = v_1 - 2v_2$; conversely, if $x - 2y = v_1 - 2v_2$, then $(x, y) = (v_1, v_2) + (y - v_2)(2,1)$.

(b) In $\mathbb{F}_2^3$ let $U = \operatorname{span}((1,1,0)) = \{(0,0,0), (1,1,0)\}$. The cosets of U are

$$U, \quad (1,0,0) + U = \{(1,0,0), (0,1,0)\}, \quad (0,0,1) + U = \{(0,0,1), (1,1,1)\}, \quad (1,0,1) + U = \{(1,0,1), (0,1,1)\}.$$

They have two elements each and split the eight vectors of $\mathbb{F}_2^3$ into four pairs.

(c) For $T \in \mathrm{L}(V, W)$, the nonempty fibers $T^{-1}(\mathbf{w})$ are exactly the cosets of $\ker T$ (Corollary 2.6.31). In particular the solution set of a consistent system $A\mathbf{x} = \mathbf{y}$ is a coset of $\ker T_A$ (Proposition 3.1.19).

(d) In $F[x]$ let $U = \{p \mid p(0) = 0\}$, the polynomials with constant term 0. Then $p + U = \{q \in F[x] \mid q(0) = p(0)\}$: if $q = p + u$ with $u \in U$, then $q(0) = p(0) + u(0) = p(0)$ (Proposition 1.2.18), and if $q(0) = p(0)$, then $q - p \in U$. So there is one coset for each scalar.

Lemma 4.6.14 (Equal cosets). Let U be a subspace of V and let $\mathbf{v}, \mathbf{w} \in V$. The following are equivalent.

(I) $\mathbf{v} + U = \mathbf{w} + U$.

(II) $\mathbf{v} - \mathbf{w} \in U$.

(III) $\mathbf{v} \in \mathbf{w} + U$.

(IV) $(\mathbf{v} + U) \cap (\mathbf{w} + U) \neq \emptyset$.

Proof. ROUTINE

Uses. Definition 4.6.12 (Coset), Definition 1.4.1 (Subspace)
Techniques. Proving TFAE (Technique 2.7.1), Double containment (Technique 1.3.14)
Hint 1. Prove the cycle (I) ⇒ (III) ⇒ (IV) ⇒ (II) ⇒ (I). Only the last step takes work.
Hint 2. For (II) ⇒ (I), write $\mathbf{v} + \mathbf{u} = \mathbf{w} + ((\mathbf{v} - \mathbf{w}) + \mathbf{u})$ to get $\mathbf{v} + U \subseteq \mathbf{w} + U$. For the other containment, note that $\mathbf{w} - \mathbf{v} \in U$ as well.
Sketch. (I) ⇒ (III): $\mathbf{v} = \mathbf{v} + \mathbf{0}$. (III) ⇒ (IV): $\mathbf{v}$ is in both cosets. (IV) ⇒ (II): a common element $\mathbf{v} + \mathbf{u} = \mathbf{w} + \mathbf{u}'$ gives $\mathbf{v} - \mathbf{w} = \mathbf{u}' - \mathbf{u} \in U$. (II) ⇒ (I): double containment, using that U is closed under addition and negation.

We prove (I) ⇒ (III) ⇒ (IV) ⇒ (II) ⇒ (I).

(I) ⇒ (III). $\mathbf{v} = \mathbf{v} + \mathbf{0} \in \mathbf{v} + U = \mathbf{w} + U$.

(III) ⇒ (IV). $\mathbf{v} \in \mathbf{v} + U$, and $\mathbf{v} \in \mathbf{w} + U$ by (III).

(IV) ⇒ (II). Let $\mathbf{x}$ be in both cosets, say $\mathbf{x} = \mathbf{v} + \mathbf{u} = \mathbf{w} + \mathbf{u}'$ with $\mathbf{u}, \mathbf{u}' \in U$. Then $\mathbf{v} - \mathbf{w} = \mathbf{u}' - \mathbf{u} \in U$.

(II) ⇒ (I). Let $\mathbf{z} = \mathbf{v} - \mathbf{w} \in U$. For $\mathbf{u} \in U$,

$$\mathbf{v} + \mathbf{u} = \mathbf{w} + (\mathbf{z} + \mathbf{u}) \in \mathbf{w} + U,$$

since $\mathbf{z} + \mathbf{u} \in U$. So $\mathbf{v} + U \subseteq \mathbf{w} + U$. Since $\mathbf{w} - \mathbf{v} = -\mathbf{z} \in U$, the same argument with $\mathbf{v}$ and $\mathbf{w}$ exchanged gives $\mathbf{w} + U \subseteq \mathbf{v} + U$. □

Corollary 4.6.15 (Cosets partition V). Let U be a subspace of V.

(a) Every $\mathbf{v} \in V$ lies in exactly one coset of U, namely $\mathbf{v} + U$.

(b) Two cosets of U are either equal or disjoint.

(c) $\mathbf{v} + U = U$ if and only if $\mathbf{v} \in U$.

Proof. ROUTINE

Uses. Lemma 4.6.14 (Equal cosets)
Hint. Each part is one implication of Lemma 4.6.14. For (c), write $U = \mathbf{0} + U$.
Sketch. (a) $\mathbf{v} \in \mathbf{v} + U$, and $\mathbf{v} \in \mathbf{w} + U$ forces $\mathbf{w} + U = \mathbf{v} + U$ by (III) ⇒ (I). (b) is (IV) ⇒ (I). (c) is (I) ⇔ (II) with $\mathbf{w} = \mathbf{0}$.

(a) $\mathbf{v} = \mathbf{v} + \mathbf{0} \in \mathbf{v} + U$. If also $\mathbf{v} \in \mathbf{w} + U$, then $\mathbf{v} + U = \mathbf{w} + U$ by (III) ⇒ (I) of Lemma 4.6.14.

(b) If two cosets are not disjoint, they are equal by (IV) ⇒ (I).

(c) $U = \mathbf{0} + U$, so by (I) ⇔ (II), $\mathbf{v} + U = U$ if and only if $\mathbf{v} - \mathbf{0} = \mathbf{v} \in U$. □

Remark 4.6.16 (Congruence modulo U). Write $\mathbf{v} \equiv \mathbf{w} (\mathrm{mod}\, U)$ when $\mathbf{v} - \mathbf{w} \in U$. This is an equivalence relation: it is reflexive because $\mathbf{0} \in U$, symmetric because U is closed under negation, and transitive because U is closed under addition. By Lemma 4.6.14 its equivalence classes are exactly the cosets of U. (Section A.3 treats equivalence relations in general, but nothing from there is needed here beyond Lemma 4.6.14.) Notice that each of the three properties used a different part of "U is a subspace".

Exercise 4.6.17. ROUTINE Let $U = \mathrm{span}((1,0,1),(0,1,1)) \subseteq \mathbb{R}^3$.

(a) Show that $U = \{(x,y,z) \mid z = x + y\}$, and that $\mathbf{v} + U$ is the plane $z - x - y = v_3 - v_1 - v_2$.

(b) Which of the cosets $(1,2,3) + U$, $(0,0,1) + U$, $(2,-1,2) + U$, and $(1,1,1) + U$ are equal?

Solution on page 1343

4.6.4 The quotient space

We want to add and scale cosets, and the only handle on a coset is a representative. So, as in the picture, add two lines by picking a point on each, adding the points, and taking the line through the sum:

$$(\mathbf{v} + U) + (\mathbf{w} + U) = (\mathbf{v} + \mathbf{w}) + U, \quad a(\mathbf{v} + U) = a\mathbf{v} + U.$$

The right sides are computed from the chosen representatives, and a coset has many. If two choices could produce different cosets, these formulas would not define operations at all. Figure 4.6.2 shows why they do: replacing $\mathbf{v}$ by $\mathbf{v}' = \mathbf{v} + \mathbf{u}$ and $\mathbf{w}$ by $\mathbf{w}' = \mathbf{w} + \mathbf{u}'$ changes the sum by $\mathbf{u} + \mathbf{u}' \in U$, which moves it along its line but not off it.

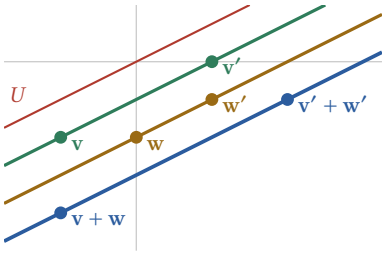


Figure 4.6.2. Adding cosets of $U = \operatorname{span}((2,1))$ (red) in $\mathbb{R}^2$. The green line is $x - 2y = 2$ and the ochre line is $x - 2y = 4$. The representatives $\mathbf{v}$ or $\mathbf{v}'$ of the green line and $\mathbf{w}$ or $\mathbf{w}'$ of the ochre line give different sums, $\mathbf{v} + \mathbf{w} = (-2, -4)$ and $\mathbf{v}' + \mathbf{w}' = (4, -1)$, but both lie on the blue line $x - 2y = 6$. So the sum of the green and ochre lines is the blue line, whichever points we pick.

Technique 4.6.18 (Check that a definition does not depend on the representative). Objects such as cosets are handled through representatives: a coset is named by any one of its elements, and a function on cosets is defined by a formula in that element, $f(\mathbf{v} + U) = g(\mathbf{v})$. Such a formula defines a function only if every name of the same coset gives the same value. So the first thing to prove about it is

$$\mathbf{v} + U = \mathbf{v}' + U \quad \Rightarrow \quad g(\mathbf{v}) = g(\mathbf{v}'),$$

which by Lemma 4.6.14 means: if $\mathbf{v} - \mathbf{v}' \in U$, then $g(\mathbf{v}) = g(\mathbf{v}')$. One says that f is **well defined**.

Mind the direction. Equal inputs must give equal outputs. The converse, that equal outputs force equal inputs, is injectivity, which is a different property that f may or may not have (Exercise 4.6.43).

Reach for it whenever an object is defined through a choice: of a representative, a basis, a preimage, an extension. The canonical example is Theorem 4.6.22. The same check appeared in Exercise 2.6.43, where a map was defined on preimages, and in Proposition 4.2.15, which shows that the rank of a matrix of T does not depend on the bases. It returns when the determinant of an operator is defined through any one of its matrices (Section 5.4).

Used in. Theorem 4.6.22, Exercise 4.6.24, Theorem 4.6.30, Theorem 4.6.34, Exercise 4.6.43, Theorem A.3.10, Exercise A.3.19, Exercise A.4.10.

Non-example 4.6.19 (A formula that defines no function). Let $U = \operatorname{span}((2,1)) \subseteq \mathbb{R}^2$. The formula $f((v_1, v_2) + U) = v_1$ does not define a function on the set of cosets: $(0,0) + U$ and $(2,1) + U$ are the same coset, since $(2,1) - (0,0) \in U$, but the formula gives 0 for the first name and 2 for the second. On the other hand, $g((v_1, v_2) + U) = v_1 - 2v_2$ is well defined: if $(v_1, v_2) - (v_1', v_2') = t(2,1)$, then $(v_1 - 2v_2) - (v_1' - 2v_2') = 2t - 2t = 0$. In the picture, g reads off the label c of the line $x - 2y = c$, which is a property of the line, while f reads off a property of a chosen point on it.

Definition 4.6.20 (Quotient space). Let U be a subspace of V. The **quotient** of V by U is the set of cosets

$$V/U = \{\mathbf{v} + U \mid \mathbf{v} \in V\},$$

with addition and scalar multiplication

$$(\mathbf{v} + U) + (\mathbf{w} + U) = (\mathbf{v} + \mathbf{w}) + U, \quad a(\mathbf{v} + U) = a\mathbf{v} + U.$$

By Theorem 4.6.22 these operations are well defined and make V/U a vector space over F. We read V/U as "V mod U".

Notation 4.6.21. When U is clear from context, we sometimes write $[\mathbf{v}]$ for $\mathbf{v} + U$. Then $[\mathbf{v}] = [\mathbf{w}]$ if and only if $\mathbf{v} - \mathbf{w} \in U$, and the operations read $[\mathbf{v}] + [\mathbf{w}] = [\mathbf{v} + \mathbf{w}]$ and $a[\mathbf{v}] = [a\mathbf{v}]$.

Theorem 4.6.22 (The quotient is a vector space). Let U be a subspace of V.

(a) The operations of Definition 4.6.20 are well defined: if $\mathbf{v}+U=\mathbf{v}'+U$ and $\mathbf{w}+U=\mathbf{w}'+U$, then

$$(\mathbf{v}+\mathbf{w})+U=(\mathbf{v}'+\mathbf{w}')+U \quad \text{and} \quad a\mathbf{v}+U=a\mathbf{v}'+U \quad \text{for all } a\in F.$$

(b) With these operations, V/U is a vector space over F. Its zero vector is $\mathbf{0}+U=U$, and $-(\mathbf{v}+U)=(-\mathbf{v})+U$.

Proof. STANDARD

Uses. Lemma 4.6.14 (Equal cosets), Definition 1.4.1 (Subspace), Definition 1.2.1 (Vector space)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18), Proving from axioms (Technique 1.1.7)
Hint 1. Translate the hypotheses with Lemma 4.6.14: $\mathbf{v}-\mathbf{v}'\in U$ and $\mathbf{w}-\mathbf{w}'\in U$. What must you show about the outputs?
Hint 2. Show that $(\mathbf{v}+\mathbf{w})-(\mathbf{v}'+\mathbf{w}')$ and $a\mathbf{v}-a\mathbf{v}'$ lie in U. Which properties of U does this use?
Hint 3. For (b), write every coset with a representative, use the definitions to turn each axiom into a computation in V, and use the axiom there.
Sketch. (a) $(\mathbf{v}+\mathbf{w})-(\mathbf{v}'+\mathbf{w}')=(\mathbf{v}-\mathbf{v}')+(\mathbf{w}-\mathbf{w}')\in U$ and $a\mathbf{v}-a\mathbf{v}'=a(\mathbf{v}-\mathbf{v}')\in U$, because U is closed under addition and scalar multiplication; conclude with Lemma 4.6.14. (b) Each axiom for cosets is the same axiom for representatives, e.g. $a((\mathbf{v}+U)+(\mathbf{w}+U))=(a\mathbf{v}+a\mathbf{w})+U=a(\mathbf{v}+U)+a(\mathbf{w}+U)$.

(a) By Lemma 4.6.14, the hypotheses say that $\mathbf{v}-\mathbf{v}'\in U$ and $\mathbf{w}-\mathbf{w}'\in U$. Since U is a subspace,

$$(\mathbf{v}+\mathbf{w})-(\mathbf{v}'+\mathbf{w}')=(\mathbf{v}-\mathbf{v}')+(\mathbf{w}-\mathbf{w}')\in U \quad \text{and} \quad a\mathbf{v}-a\mathbf{v}'=a(\mathbf{v}-\mathbf{v}')\in U.$$

By Lemma 4.6.14 again, $(\mathbf{v}+\mathbf{w})+U=(\mathbf{v}'+\mathbf{w}')+U$ and $a\mathbf{v}+U=a\mathbf{v}'+U$.

(b) By (a), the sum of two cosets and a scalar multiple of a coset are well-defined cosets. Each axiom is an identity between cosets. Writing each coset with a representative and applying the definitions turns both sides into cosets of vectors of V, which agree by the corresponding axiom in V. For instance, for $\mathbf{u},\mathbf{v},\mathbf{w}\in V$ and $a,b\in F$,

$$\begin{aligned}((\mathbf{u}+U)+(\mathbf{v}+U))+(\mathbf{w}+U)&=((\mathbf{u}+\mathbf{v})+\mathbf{w})+U=(\mathbf{u}+(\mathbf{v}+\mathbf{w}))+U\\&=(\mathbf{u}+U)+((\mathbf{v}+U)+(\mathbf{w}+U)),\\ a((\mathbf{v}+U)+(\mathbf{w}+U))&=a(\mathbf{v}+\mathbf{w})+U=(a\mathbf{v}+a\mathbf{w})+U=a(\mathbf{v}+U)+a(\mathbf{w}+U),\\ (a+b)(\mathbf{v}+U)&=(a+b)\mathbf{v}+U=(a\mathbf{v}+b\mathbf{v})+U=a(\mathbf{v}+U)+b(\mathbf{v}+U),\end{aligned}$$

and commutativity, $(ab)(\mathbf{v}+U)=a(b(\mathbf{v}+U))$, and $1(\mathbf{v}+U)=\mathbf{v}+U$ are checked the same way. Finally, $(\mathbf{v}+U)+(\mathbf{0}+U)=\mathbf{v}+U$ and $(\mathbf{v}+U)+((-\mathbf{v})+U)=\mathbf{0}+U$, so $\mathbf{0}+U=U$ is a zero vector and $(-\mathbf{v})+U$ is an additive inverse of $\mathbf{v}+U$. □

All the content is in (a), and (a) used exactly that U is closed under addition and under scalar multiplication. Part (b) is bookkeeping. Notice also that U, the subspace we collapsed, is now the origin of V/U.

Example 4.6.23 (Quotient spaces).

(a) Let $U = \operatorname{span}((2,1)) \subseteq \mathbb{R}^2$. By Non-example 4.6.19, $g : \mathbb{R}^2/U \to \mathbb{R}$, $g((v_1, v_2) + U) = v_1 - 2v_2$, is well defined. It is linear, since $g(a(\mathbf{v} + U) + (\mathbf{w} + U)) = g((a\mathbf{v} + \mathbf{w}) + U) = a(v_1 - 2v_2) + (w_1 - 2w_2)$. It is injective: if $v_1 - 2v_2 = 0$, then $\mathbf{v} = v_2(2,1) \in U$, so $\mathbf{v} + U = U$ is the zero coset. It is surjective, since $g((c,0) + U) = c$. So $\mathbb{R}^2/U \cong \mathbb{R}$, as Intuition (Collapsing a subspace to a point) predicted, and g is the coordinate across the stack of lines.

(b) Let $U = \operatorname{span}((1,1,0)) \subseteq \mathbb{F}_2^3$ as in Example 4.6.13(b). Then $\mathbb{F}_2^3/U$ has four elements. For instance $((1,0,0) + U) + ((0,0,1) + U) = (1,0,1) + U$, and with the other representatives $(0,1,0)$ and $(1,1,1)$ we get $(1,0,1) + U$ once more. With $(1,0,0)$ and $(1,1,1)$ we get $(0,1,1) + U$, which is the same coset, as it must be. Every element is its own negative, and Theorem 4.6.27 below gives $\dim\left(\mathbb{F}_2^3/U\right) = 2$.

(c) The extreme cases: V/V has the single element $\mathbf{0} + V = V$, so it is the zero space. In $V/\{\mathbf{0}\}$ each coset $\mathbf{v} + \{\mathbf{0}\} = \{\mathbf{v}\}$ has one element, and $\mathbf{v} \mapsto \{\mathbf{v}\}$ is an isomorphism $V \to V/\{\mathbf{0}\}$.

Exercise 4.6.24. ROUTINE Let $U = \operatorname{span}((1,1)) \subseteq \mathbb{R}^2$. Which of the following formulas define functions on $\mathbb{R}^2/U$? For those that do, decide whether the function is linear.

(a) $f(\mathbf{v} + U) = v_1 + v_2$.

(b) $f(\mathbf{v} + U) = v_1 - v_2$.

(c) $f(\mathbf{v} + U) = (v_1 - v_2)^2$.

(d) $S(\mathbf{v} + U) = (v_2, v_1) + U$.

(e) $S(\mathbf{v} + U) = (v_1, 0) + U$.

Solution on page 1343

The quotient map

Definition 4.6.25 (Quotient map). Let U be a subspace of V. The **quotient map** is $\pi : V \to V/U$, $\pi(\mathbf{v}) = \mathbf{v} + U$.

Proposition 4.6.26 (The quotient map). The quotient map $\pi : V \to V/U$ is linear and surjective, and $\ker \pi = U$.

Proof. ROUTINE

Uses. Definition 4.6.20 (Quotient space), Corollary 4.6.15 (Cosets partition V), Theorem 2.1.3 (Checking linearity)
Hint 1. Linearity is the definition of the operations on V/U, read backwards.
Hint 2. For the kernel, recall that the zero of V/U is U, and use Corollary 4.6.15(c).
Sketch. $\pi(a\mathbf{v} + \mathbf{w}) = (a\mathbf{v} + \mathbf{w}) + U = a(\mathbf{v} + U) + (\mathbf{w} + U)$. Every coset is π of its representatives. $\pi(\mathbf{v}) = U$ iff $\mathbf{v} \in U$.

For $\mathbf{v}, \mathbf{w} \in V$ and $a \in F$, the definition of the operations gives

$$\pi(a\mathbf{v} + \mathbf{w}) = (a\mathbf{v} + \mathbf{w}) + U = a(\mathbf{v} + U) + (\mathbf{w} + U) = a\pi(\mathbf{v}) + \pi(\mathbf{w}),$$

so π is linear (Theorem 2.1.3). Every element of V/U has the form $\mathbf{v} + U = \pi(\mathbf{v})$, so π is surjective. The zero of V/U is U (Theorem 4.6.22), and $\pi(\mathbf{v}) = U$ if and only if $\mathbf{v} \in U$ (Corollary 4.6.15(c)). So $\ker \pi = U$. □

In fact the operations on V/U were chosen to make π linear. One consequence: every subspace of every vector space is the kernel of some linear map. In finite dimensions Section 4.4 gets this from a complement; the quotient map gives it with no choices and no dimension hypothesis.

4.6.5 The dimension of a quotient

In the plane, collapsing a line leaves a line. In general, collapsing U should remove exactly $\dim U$ dimensions. The first proof below finds a basis of V/U inside a basis of V, and the second reads the count off rank–nullity for π.

Theorem 4.6.27 (Dimension of a quotient). Let V be finite-dimensional and let U be a subspace of V. Then V/U is finite-dimensional, and

$$\dim(V/U) = \dim V - \dim U.$$

More precisely, if $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r$ is a basis of V such that $\mathbf{u}_1, \dots, \mathbf{u}_k$ is a basis of U, then $\mathbf{v}_1 + U, \dots, \mathbf{v}_r + U$ is a basis of V/U.

Proof (by extending a basis of U). STANDARD

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis), Lemma 4.6.14 (Equal cosets), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Extend a basis of a subspace (Technique 4.1.18)
Hint 1. Bases of the required kind exist by Theorem 4.1.16. Why should the cosets $\mathbf{u}_i + U$ play no role?
Hint 2. Each $\mathbf{u}_i + U$ is the zero coset. For spanning, expand an arbitrary $\mathbf{v}$ in the basis and apply π.
Hint 3. For independence: $\sum_j b_j(\mathbf{v}_j + U) = U$ means $\sum_j b_j \mathbf{v}_j \in U = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Write that as a linear relation among the basis vectors of V.
Sketch. Extend a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of U to a basis of V. Since $\mathbf{v} - \sum_j b_j \mathbf{v}_j \in U$, every coset is $\sum_j b_j(\mathbf{v}_j + U)$. If $\sum_j b_j(\mathbf{v}_j + U) = U$, then $\sum_j b_j \mathbf{v}_j = \sum_i a_i \mathbf{u}_i$, a relation among basis vectors, so all $b_j = 0$. Hence $\dim(V/U) = r = \dim V - \dim U$.

By Proposition 4.1.21(a), U is finite-dimensional, so it has a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$, and by Theorem 4.1.16 this extends to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r$ of V with $k + r = \dim V$. So it suffices to prove the second statement, for any such basis. Then $\dim(V/U) = r = \dim V - \dim U$.

Spanning. Let $\mathbf{v} \in V$ and write $\mathbf{v} = \sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j$. Then $\mathbf{v} - \sum_j b_j \mathbf{v}_j = \sum_i a_i \mathbf{u}_i \in U$, so by Lemma 4.6.14 and the definition of the operations,

$$\mathbf{v} + U = \left(\sum_{j=1}^{r} b_j \mathbf{v}_j \right) + U = \sum_{j=1}^{r} b_j (\mathbf{v}_j + U).$$

Independence. Suppose $\sum_j b_j(\mathbf{v}_j + U) = \mathbf{0} + U$, that is, $\left(\sum_j b_j \mathbf{v}_j\right) + U = U$. By Corollary 4.6.15(c), $\sum_j b_j \mathbf{v}_j \in U$, so $\sum_j b_j \mathbf{v}_j = \sum_i a_i \mathbf{u}_i$ for some scalars a_i. Then

$$\sum_{i=1}^{k} (-a_i) \mathbf{u}_i + \sum_{j=1}^{r} b_j \mathbf{v}_j = \mathbf{0}$$

is a linear relation among the vectors of a basis of V, so all its coefficients vanish (Theorem 1.5.5). In particular $b_1 = \cdots = b_r = 0$. □

Proof (by rank–nullity for the quotient map). ROUTINE

Uses. Proposition 4.6.26 (The quotient map), Theorem 4.2.11 (Rank–nullity), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Corollary 4.1.11 (Spanning lists are long)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $\pi : V \to V/U$ is linear and surjective. What is its kernel?
Hint 2. Rank–nullity applies because V is finite-dimensional. For the basis statement, a spanning list of length $\dim(V/U)$ is a basis.
Sketch. π is surjective with kernel U, so rank–nullity gives $\dim V = \dim U + \dim(V/U)$, with $V/U = \operatorname{im}\pi$ finite-dimensional. The images $\pi(\mathbf{u}_i) = \mathbf{0}$ and $\pi(\mathbf{v}_j)$ span V/U, so the $r = \dim(V/U)$ cosets $\mathbf{v}_j + U$ span it, hence form a basis.

By Proposition 4.6.26, π is linear and surjective with $\ker \pi = U$. Since V is finite-dimensional, rank–nullity (Theorem 4.2.11(b)) says that $\operatorname{im}\pi = V/U$ is finite-dimensional and $\dim V = \dim U + \dim(V/U)$.

For the second statement, the list $\pi(\mathbf{u}_1), \dots, \pi(\mathbf{u}_k), \pi(\mathbf{v}_1), \dots, \pi(\mathbf{v}_r)$ spans $\operatorname{im}\pi = V/U$ (Proposition 2.1.16(a)). Its first k entries are $\mathbf{0}$, so $\mathbf{v}_1 + U, \dots, \mathbf{v}_r + U$ spans V/U as well. It has length $r = \dim V - \dim U = \dim(V/U)$, so it is a basis (Corollary 4.1.11(b)). □

The second proof is shorter, but it leans on rank–nullity. The first proof uses only the extension of bases, and that matters in Section 4.6.7, where the theorem is used to reprove rank–nullity.

The converse direction needs no finiteness assumption on V: a basis of U together with lifts of a basis of V/U is a basis of V. This is what makes quotients useful for induction on dimension, where a statement about V is split into statements about U and V/U.

Proposition 4.6.28 (Lifting a basis of a quotient). Let U be a subspace of V, let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be a basis of U, and let $\mathbf{v}_1, \dots, \mathbf{v}_r \in V$ be such that $\mathbf{v}_1 + U, \dots, \mathbf{v}_r + U$ is a basis of V/U. Then $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r$ is a basis of V. In particular, if U and V/U are finite-dimensional, then so is V, and $\dim V = \dim U + \dim(V/U)$.

Proof. STANDARD

Uses. Proposition 4.6.26 (The quotient map), Lemma 4.6.14 (Equal cosets), Corollary 1.6.24 (Bases exist in finite dimensions)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. For spanning, expand $\mathbf{v} + U$ in the basis of V/U and pull back with Lemma 4.6.14.
Hint 2. For independence, apply π to a linear relation. Which terms does it kill?
Sketch. If $\mathbf{v} + U = \sum_j b_j(\mathbf{v}_j + U)$, then $\mathbf{v} - \sum_j b_j \mathbf{v}_j \in U = \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Applying π to a relation $\sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j = \mathbf{0}$ kills the $\mathbf{u}_i$ and leaves a relation among the $\mathbf{v}_j + U$, so all $b_j = 0$, and then all $a_i = 0$.

Spanning. Let $\mathbf{v} \in V$. Since $\mathbf{v}_1 + U, \dots, \mathbf{v}_r + U$ spans V/U, there are scalars b_j with

$$\mathbf{v} + U = \sum_{j=1}^{r} b_j(\mathbf{v}_j + U) = \left(\sum_{j=1}^{r} b_j \mathbf{v}_j\right) + U.$$

By Lemma 4.6.14, $\mathbf{v} - \sum_j b_j \mathbf{v}_j \in U$, so it equals $\sum_i a_i \mathbf{u}_i$ for some scalars a_i, and $\mathbf{v} = \sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j$.

Independence. Suppose $\sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j = \mathbf{0}$. Apply the linear map π. Since $\pi(\mathbf{u}_i) = \mathbf{0}$ (as $\ker \pi = U$), we get $\sum_j b_j(\mathbf{v}_j + U) = \mathbf{0} + U$, so all $b_j = 0$ because the cosets $\mathbf{v}_j + U$ are independent. Then $\sum_i a_i \mathbf{u}_i = \mathbf{0}$, so all $a_i = 0$.

For the last statement, U and V/U have bases (Corollary 1.6.24), and every element of V/U is a coset $\mathbf{v} + U$, so the basis of V/U has the form $\mathbf{v}_1 + U, \dots, \mathbf{v}_r + U$. By the first part, V has a basis of length $k + r$. □

Quotients and complements

A complement of U (Proposition 4.4.16) and the quotient V/U have the same dimension. They are related by the quotient map.

Proposition 4.6.29 (Complements and the quotient). Let U and W be subspaces of V, and let $\pi : V \to V/U$ be the quotient map. Then $\pi|_W : W \to V/U$ is injective if and only if $U \cap W = \{\mathbf{0}\}$, and surjective if and only if $U + W = V$. In particular, $\pi|_W$ is an isomorphism if and only if $V = U \oplus W$, so $V/U \cong W$ for every complement W of U.

Proof. STANDARD

Uses. Proposition 4.2.3 (Properties of restriction), Proposition 4.6.26 (The quotient map), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Lemma 4.6.14 (Equal cosets), Theorem 4.4.7 (Criteria for direct sums)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. The kernel of a restriction is the intersection of the subspace with the kernel (Proposition 4.2.3(c)).
Hint 2. $\pi|_W$ is surjective when every coset $\mathbf{v} + U$ contains a vector of W. Translate this with Lemma 4.6.14.
Sketch. $\ker \pi|_W = W \cap \ker \pi = W \cap U$. Surjectivity says that every $\mathbf{v}$ has some $\mathbf{w} \in W$ with $\mathbf{v} - \mathbf{w} \in U$, i.e. $\mathbf{v} \in U + W$. Both together say $V = U \oplus W$ (Theorem 4.4.7).

By Proposition 4.2.3(c) and $\ker \pi = U$, we have $\ker \pi|_W = W \cap U$, so $\pi|_W$ is injective if and only if $U \cap W = \{\mathbf{0}\}$ (Theorem 2.6.11).

The image of $\pi|_W$ consists of the cosets $\mathbf{w} + U$ with $\mathbf{w} \in W$. If $U + W = V$, then every $\mathbf{v}$ is $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$, $\mathbf{w} \in W$, and $\mathbf{v} + U = \mathbf{w} + U$ since $\mathbf{v} - \mathbf{w} \in U$; so $\pi|_W$ is surjective. Conversely, if it is surjective, then for every $\mathbf{v} \in V$ there is $\mathbf{w} \in W$ with $\mathbf{v} + U = \mathbf{w} + U$, so $\mathbf{v} - \mathbf{w} \in U$ (Lemma 4.6.14) and $\mathbf{v} = (\mathbf{v} - \mathbf{w}) + \mathbf{w} \in U + W$.

Hence $\pi|_W$ is bijective if and only if $U \cap W = \{\mathbf{0}\}$ and $U + W = V$, which by Theorem 4.4.7 says that $V = U \oplus W$. A bijective linear map is an isomorphism. □

Intuition (A complement is a choice; the quotient is not). The proposition says that a complement W meets every coset of U in exactly one vector. So a complement amounts to choosing one representative from each coset, in such a way that the chosen vectors form a subspace. In the plane, every line through $\mathbf{0}$ other than U meets each line parallel to U exactly once, and none of these lines is preferred. The quotient is the same space with the choice removed. This is why quotients, not complements, are the tool when a construction must not depend on choices, as in the first isomorphism theorem below, or when complements are hard to produce: in infinite dimensions their existence in general needs Zorn's lemma (Section A.4), while V/U always exists.

With Proposition 4.4.16(b) the proposition also gives a third proof of Theorem 4.6.27: $\dim(V/U) = \dim W = \dim V - \dim U$.

4.6.6 Linear maps out of a quotient

Intuition (Maps that kill U). A linear map out of V/U composed with π is a linear map out of V that sends all of U to $\mathbf{0}$, because π does. Conversely, let $T : V \to W$ be linear with $T(U) = \{\mathbf{0}\}$. If $\mathbf{v}$ and $\mathbf{v}'$ lie in the same coset, then $T(\mathbf{v}) - T(\mathbf{v}') = T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$, so T is constant on each coset and makes sense as a function of the coset. So linear maps out of V/U are the same thing as linear maps out of V that vanish on U.

Theorem 4.6.30 (Universal property of the quotient). Let U be a subspace of V, let $\pi : V \to V/U$ be the quotient map, and let $T \in \mathrm{L}(V, W)$.

(a) If $U \subseteq \ker T$, there is exactly one linear map $\tilde{T} : V/U \to W$ with $\tilde{T} \circ \pi = T$, namely

$$\tilde{T}(\mathbf{v} + U) = T(\mathbf{v}).$$

Its image is $\operatorname{im} \tilde{T} = \operatorname{im} T$, and its kernel is $\ker \tilde{T} = \{\mathbf{v} + U \mid \mathbf{v} \in \ker T\}$.

(b) Conversely, for every $S \in \mathrm{L}(V/U, W)$ the map $S \circ \pi \in \mathrm{L}(V, W)$ vanishes on U.

The map $\tilde{T}$ is the **map induced by** T. The equation $\tilde{T} \circ \pi = T$ is often drawn as a triangle:

$$\begin{array}{ccc} V & \xrightarrow{\pi} & V/U \\ & \searrow^{T} & \downarrow \tilde{T} \\ & & W \end{array}$$

Proof. STANDARD

Uses. Proposition 4.6.26 (The quotient map), Lemma 4.6.14 (Equal cosets), Theorem 2.1.3 (Checking linearity), Definition 4.6.20 (Quotient space)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18), Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Uniqueness first: two maps S, S' with $S \circ \pi = S' \circ \pi$ agree at every $\pi(\mathbf{v})$. Why is that everything?
Hint 2. For existence, the formula $\tilde{T}(\mathbf{v} + U) = T(\mathbf{v})$ is forced. Check that it is well defined: if $\mathbf{v} - \mathbf{v}' \in U$, why is $T(\mathbf{v}) = T(\mathbf{v}')$?
Hint 3. Linearity: compute $\tilde{T}(a(\mathbf{v} + U) + (\mathbf{w} + U))$ by first combining the cosets. The image and kernel come from $\tilde{T}(\mathbf{v} + U) = T(\mathbf{v})$.
Sketch. Since π is surjective, $S \circ \pi = T$ determines S at every coset, which gives uniqueness and the formula. If $\mathbf{v} - \mathbf{v}' \in U \subseteq \ker T$, then $T(\mathbf{v}) = T(\mathbf{v}')$, so the formula is well defined; it is linear because $a(\mathbf{v} + U) + (\mathbf{w} + U) = (a\mathbf{v} + \mathbf{w}) + U$ and T is linear. The values of $\tilde{T}$ are the values of T, and $\tilde{T}(\mathbf{v} + U) = \mathbf{0}$ iff $\mathbf{v} \in \ker T$. (b) $S(\pi(\mathbf{u})) = S(\mathbf{0}) = \mathbf{0}$ for $\mathbf{u} \in U$.

(a) *Uniqueness.* If $S \circ \pi = T$, then $S(\mathbf{v} + U) = S(\pi(\mathbf{v})) = T(\mathbf{v})$ for every $\mathbf{v} \in V$. Every element of V/U is $\mathbf{v} + U$ for some $\mathbf{v}$ (Proposition 4.6.26), so S is determined by T: it must be given by the formula in the statement.

Well defined. Suppose $\mathbf{v}+U=\mathbf{v}'+U$. Then $\mathbf{v}-\mathbf{v}'\in U\subseteq\ker T$ (Lemma 4.6.14), so $T(\mathbf{v})-T(\mathbf{v}')=T(\mathbf{v}-\mathbf{v}')=\mathbf{0}$. Hence the formula defines a function $\tilde{T}:V/U\to W$.

Linear. For $\mathbf{v},\mathbf{w}\in V$ and $a\in F$,

$$\tilde{T}(a(\mathbf{v}+U)+(\mathbf{w}+U))=\tilde{T}((a\mathbf{v}+\mathbf{w})+U)=T(a\mathbf{v}+\mathbf{w})=a\tilde{T}(\mathbf{v}+U)+\tilde{T}(\mathbf{w}+U).$$

By construction $\tilde{T}(\pi(\mathbf{v}))=T(\mathbf{v})$, so $\tilde{T}\circ\pi=T$.

Image and kernel. The values of $\tilde{T}$ are the vectors $T(\mathbf{v})$, $\mathbf{v}\in V$, so $\operatorname{im}\tilde{T}=\operatorname{im}T$. And $\tilde{T}(\mathbf{v}+U)=\mathbf{0}$ if and only if $T(\mathbf{v})=\mathbf{0}$, so $\ker\tilde{T}=\{\mathbf{v}+U\mid\mathbf{v}\in\ker T\}$. (This is well stated: whether $\mathbf{v}\in\ker T$ does not depend on the representative, since $\ker T\supseteq U$.)

(b) $S\circ\pi$ is linear (Proposition 2.1.17), and for $\mathbf{u}\in U$, $(S\circ\pi)(\mathbf{u})=S(\mathbf{0}+U)=\mathbf{0}$, since $\mathbf{u}+U=\mathbf{0}+U$. □

So $T\mapsto\tilde{T}$ and $S\mapsto S\circ\pi$ are inverse bijections between the linear maps $V\to W$ that vanish on U and the linear maps $V/U\to W$. Note how the hypothesis $U\subseteq\ker T$ entered: it is exactly what the well-definedness check needed. When U is not inside $\ker T$, there is no induced map at all. This is the quotient version of Exercise 2.6.43, with π as the surjection.

Technique 4.6.31 (Pass to a quotient). When a construction only depends on vectors up to elements of a subspace U, replace V by V/U. By Theorem 4.6.30, a linear map that kills U is the same as a linear map on V/U, and by Theorem 4.6.27 and Proposition 4.6.28, dimensions split as $\dim V=\dim U+\dim(V/U)$. Two typical uses:

- *Make a map injective.* Dividing out exactly the kernel turns T into an injective map with the same image. This is the first isomorphism theorem (Theorem 4.6.34), and it is the standard way to show that some space is isomorphic to an image.
- *Split a problem in two.* If an operator T maps U into itself, it acts on U and on V/U (Proposition 4.6.32), and a statement about T often follows from the same statement for these two smaller operators, by induction on $\dim V$. Section 7.5 triangularizes operators this way.

The price is the well-definedness check (Technique 4.6.18) for every map defined on cosets.

Used in. Proposition 4.6.32, Theorem 4.6.34, Corollary 4.6.35, Exercise 4.6.37, Exercise 4.6.44, Exercise 4.6.45, Exercise 4.6.46, Exercise 4.6.47, Exercise 4.6.49, Exercise 4.6.50, Exercise 4.6.52, Exercise 4.6.53, Exercise 4.6.55, Exercise 4.8.14, Problem 4.18, Theorem 7.6.32, Exercise 4.6.42.

Operators on a quotient

Proposition 4.6.32 (Operators on a quotient). Let $T\in\mathrm{L}(V)$ and let U be a subspace of V with $T(\mathbf{u})\in U$ for every $\mathbf{u}\in U$ (that is, U is **invariant** under T; see Section 7.2).

(a) There is exactly one $T/U\in\mathrm{L}(V/U)$ with $(T/U)(\mathbf{v}+U)=T(\mathbf{v})+U$ for all $\mathbf{v}\in V$.

(b) Suppose V is finite-dimensional, let $\mathcal{C}=\mathbf{u}_1,\dots,\mathbf{u}_k$ be a basis of U, extend it to a basis $\mathcal{B}=\mathbf{u}_1,\dots,\mathbf{u}_k,\mathbf{v}_1,\dots,\mathbf{v}_r$ of V, and let $\mathcal{D}=\mathbf{v}_1+U,\dots,\mathbf{v}_r+U$ (a basis of V/U by Theorem 4.6.27). Regard $T|_U$ as an operator on U. Then

$$[T]_{\mathcal{B}\mathcal{B}}=\begin{pmatrix}[T|_U]_{\mathcal{C}\mathcal{C}} & B\\ 0 & [T/U]_{\mathcal{D}\mathcal{D}}\end{pmatrix}$$

for some $B\in F^{k\times r}$.

Proof. STANDARD

Uses. Theorem 4.6.30 (Universal property of the quotient), Proposition 4.6.26 (The quotient map), Proposition 2.1.17 (Compositions of linear maps are linear), Definition 2.3.1 (Matrix of a linear map), Definition 2.4.33 (Block matrix)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. For (a), apply Theorem 4.6.30 to the composite $\pi\circ T:V\to V/U$. What is its kernel?

Hint 2. For (b), read the matrix column by column. Where does $T(\mathbf{u}_j)$ live? Expand $T(\mathbf{v}_j)$ in $\mathcal{B}$ and apply π.
Sketch. (a) $\pi \circ T$ is linear and kills U, because $T(U) \subseteq U = \ker \pi$, so it induces T/U, which is unique by Theorem 4.6.30. (b) $T(\mathbf{u}_j) \in U$, so the first k columns have zeros below row k, and their top parts are the columns of $[T|_U]_{\mathcal{C}\mathcal{C}}$. If $T(\mathbf{v}_j) = \sum_i b_{ij}\mathbf{u}_i + \sum_i c_{ij}\mathbf{v}_i$, then $(T/U)(\mathbf{v}_j + U) = \sum_i c_{ij}(\mathbf{v}_i + U)$, so the lower right block is $[T/U]_{\mathcal{D}\mathcal{D}}$.

(a) The composite $\pi \circ T : V \to V/U$ is linear (Proposition 2.1.17). For $\mathbf{u} \in U$, $T(\mathbf{u}) \in U = \ker \pi$, so $(\pi \circ T)(\mathbf{u}) = \mathbf{0}$. By Theorem 4.6.30(a), there is exactly one linear $T/U : V/U \to V/U$ with $(T/U) \circ \pi = \pi \circ T$, that is, $(T/U)(\mathbf{v} + U) = T(\mathbf{v}) + U$.

(b) Column j of $[T]_{\mathcal{B}\mathcal{B}}$ is the coordinate vector of the image of the jth basis vector (Definition 2.3.1). For $j \le k$, the vector $T(\mathbf{u}_j)$ lies in U, so it is a combination of $\mathbf{u}_1, \dots, \mathbf{u}_k$ alone: its last r coordinates are 0, and its first k coordinates form column j of $[T|_U]_{\mathcal{C}\mathcal{C}}$. For $j \le r$, write

$$T(\mathbf{v}_j) = \sum_{i=1}^{k} b_{ij}\mathbf{u}_i + \sum_{i=1}^{r} c_{ij}\mathbf{v}_i.$$

Applying π, which kills the $\mathbf{u}_i$,

$$(T/U)(\mathbf{v}_j + U) = T(\mathbf{v}_j) + U = \sum_{i=1}^{r} c_{ij}(\mathbf{v}_i + U),$$

so (c_{ij}) is $[T/U]_{\mathcal{D}\mathcal{D}}$. Together these give the block form, with $B = (b_{ij})$. □

Example 4.6.33 (A shift and its quotient). Let $N \in \mathrm{L}(F^3)$ be the shift $N(\mathbf{e}_1) = \mathbf{0}$, $N(\mathbf{e}_2) = \mathbf{e}_1$, $N(\mathbf{e}_3) = \mathbf{e}_2$, and let $U = \operatorname{span}(\mathbf{e}_1)$, which N maps into itself (to $\{\mathbf{0}\}$). On F^3/U, with basis $\mathbf{e}_2 + U, \mathbf{e}_3 + U$,

$$(N/U)(\mathbf{e}_2 + U) = \mathbf{e}_1 + U = \mathbf{0} + U, \quad (N/U)(\mathbf{e}_3 + U) = \mathbf{e}_2 + U.$$

So N/U is the shift of a two-dimensional space, and in the standard basis

$$[N] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad [N|_U] = (0), \quad B = (1 \;\; 0), \quad [N/U] = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Collapsing the last vector of the chain $\mathbf{e}_3 \mapsto \mathbf{e}_2 \mapsto \mathbf{e}_1 \mapsto \mathbf{0}$ leaves a chain one step shorter.

4.6.7 The first isomorphism theorem

Intuition (Each slice becomes one vector). Return to the slices of Corollary 2.6.31. A linear map T sends each coset of $\ker T$ to a single vector, and different cosets to different vectors. So T is a bijection from the set of cosets of $\ker T$ onto $\operatorname{im} T$. The quotient $V/\ker T$ makes that set of cosets into a vector space, and the bijection is linear.

Theorem 4.6.34 (First isomorphism theorem). Let $T \in \mathrm{L}(V, W)$. The map

$$\tilde{T} : V/\ker T \to \operatorname{im} T, \quad \tilde{T}(\mathbf{v} + \ker T) = T(\mathbf{v}),$$

is well defined and is an isomorphism. In particular, $V/\ker T \cong \operatorname{im} T$.

The first proof is an application of the universal property. The second proves well-definedness and injectivity at once, with a single chain of equivalences read in both directions.

Proof (by the universal property). ROUTINE
Uses. Theorem 4.6.30 (Universal property of the quotient), Proposition 2.7.7 (Bijective linear maps are invertible), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. Apply Theorem 4.6.30 with $U = \ker T$. What are the image and kernel of the induced map?
Hint 2. The kernel of $\tilde{T}$ is the set of cosets $\mathbf{v} + \ker T$ with $\mathbf{v} \in \ker T$. How many different cosets is that?

Sketch. With $U = \ker T$, Theorem 4.6.30 gives a linear $\tilde{T}$ with $\operatorname{im}\tilde{T} = \operatorname{im} T$ and $\ker\tilde{T} = \{\mathbf{v} + \ker T \mid \mathbf{v} \in \ker T\} = \{\ker T\}$, the zero coset. So $\tilde{T}$ is injective, and as a map onto $\operatorname{im} T$ it is bijective.

Apply Theorem 4.6.30(a) with $U = \ker T$; the hypothesis $U \subseteq \ker T$ holds trivially. It gives a well-defined linear map $\tilde{T} : V/\ker T \to W$ with $\tilde{T}(\mathbf{v} + \ker T) = T(\mathbf{v})$, image $\operatorname{im} T$, and kernel $\{\mathbf{v} + \ker T \mid \mathbf{v} \in \ker T\}$. For $\mathbf{v} \in \ker T$ we have $\mathbf{v} + \ker T = \ker T$, the zero vector of $V/\ker T$ (Corollary 4.6.15(c)). So $\ker\tilde{T}$ is the zero subspace and $\tilde{T}$ is injective (Theorem 2.6.11). Regarded as a map into $\operatorname{im} T$, it is also surjective, so it is a bijective linear map, hence an isomorphism (Proposition 2.7.7). □

Proof (by one chain of equivalences). ROUTINE

Uses. Lemma 4.6.14 (Equal cosets), Definition 2.6.5 (Kernel), Definition 4.6.20 (Quotient space), Proposition 2.7.7 (Bijective linear maps are invertible)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18)
Hint 1. For $\mathbf{v}, \mathbf{v}' \in V$, connect "$\mathbf{v} + \ker T = \mathbf{v}' + \ker T$" with "$T(\mathbf{v}) = T(\mathbf{v}')$" by a chain of equivalences.
Hint 2. Read the chain left to right: that is well-definedness. Read it right to left: that is injectivity.
Sketch. $\mathbf{v} + \ker T = \mathbf{v}' + \ker T$ iff $\mathbf{v} - \mathbf{v}' \in \ker T$ iff $T(\mathbf{v}) = T(\mathbf{v}')$. The forward implications make $\tilde{T}$ well defined, the backward ones make it injective. Linearity and surjectivity onto $\operatorname{im} T$ are immediate.

For $\mathbf{v}, \mathbf{v}' \in V$, by Lemma 4.6.14 and linearity,

$$\mathbf{v} + \ker T = \mathbf{v}' + \ker T \quad\Leftrightarrow\quad \mathbf{v} - \mathbf{v}' \in \ker T \quad\Leftrightarrow\quad T(\mathbf{v} - \mathbf{v}') = \mathbf{0} \quad\Leftrightarrow\quad T(\mathbf{v}) = T(\mathbf{v}').$$

Read from left to right, this says that the formula for $\tilde{T}$ gives the same value for every representative, so $\tilde{T}$ is well defined (Technique 4.6.18). Read from right to left, it says that $\tilde{T}$ is injective. Linearity is the computation

$$\tilde{T}(a(\mathbf{v} + \ker T) + (\mathbf{w} + \ker T)) = \tilde{T}((a\mathbf{v} + \mathbf{w}) + \ker T) = T(a\mathbf{v} + \mathbf{w}) = aT(\mathbf{v}) + T(\mathbf{w}),$$

and $\tilde{T}$ is onto $\operatorname{im} T$ because $T(\mathbf{v}) = \tilde{T}(\mathbf{v} + \ker T)$. So $\tilde{T}$ is a bijective linear map onto $\operatorname{im} T$, hence an isomorphism (Proposition 2.7.7). □

The second proof shows what is special about $U = \ker T$. For a subspace U strictly inside $\ker T$, the chain acquires a one-way step, "$\mathbf{v} - \mathbf{v}' \in U \Rightarrow \mathbf{v} - \mathbf{v}' \in \ker T$". The induced map is still well defined, but it is no longer injective.

Corollary 4.6.35 (Rank–nullity, a third proof). Let V be finite-dimensional and $T \in \mathrm{L}(V, W)$. Then $\operatorname{im} T$ is finite-dimensional, and

$$\dim V = \dim\ker T + \dim\operatorname{im} T.$$

Proof. ROUTINE

Uses. Theorem 4.6.34 (First isomorphism theorem), Theorem 4.6.27 (Dimension of a quotient), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. $\operatorname{im} T$ is isomorphic to a quotient of V. What is the dimension of that quotient?
Hint 2. Use the first proof of Theorem 4.6.27. Using the second would be circular.
Sketch. $V/\ker T$ is finite-dimensional of dimension $\dim V - \dim\ker T$ (Theorem 4.6.27, by extending a basis). The isomorphism $\tilde{T}$ carries it onto $\operatorname{im} T$, so $\operatorname{im} T$ is finite-dimensional of the same dimension.

By Proposition 4.1.21, $\ker T$ is a finite-dimensional subspace of V, and by Theorem 4.6.27 (first proof), $V/\ker T$ is finite-dimensional with $\dim(V/\ker T) = \dim V - \dim\ker T$. By Theorem 4.6.34, $\tilde{T}$ is an injective linear map on $V/\ker T$ with image $\operatorname{im} T$. By Proposition 4.2.6(a), $\operatorname{im} T$ is finite-dimensional and $\dim\operatorname{im} T = \dim(V/\ker T) = \dim V - \dim\ker T$. □

The second proof of Theorem 4.6.27 used rank–nullity, so citing it here would be circular; the first used only Theorem 4.1.16. Unwound, this proof is close to the proof of Theorem 4.2.11 by extending a basis of the kernel: the basis $\mathbf{v}_1 + \ker T, \ldots, \mathbf{v}_r + \ker T$ of $V/\ker T$ is carried by $\tilde{T}$ to $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)$, the basis of $\operatorname{im} T$

found there. What the quotient adds is that the space being counted, $V/\ker T$, no longer depends on the choice of the $\mathbf{v}_j$ (see Remark 4.2.12).

Example 4.6.36 (The first isomorphism theorem at work).

(a) *Evaluation.* For $c \in F$, the evaluation map $\mathrm{ev}_c : F[x] \to F$, $p \mapsto p(c)$ (Example 2.1.9), is linear and surjective (it sends the constant polynomial a to a). Its kernel is $U_c = \{p \mid p(c) = 0\}$. So $F[x]/U_c \cong F$, via $p + U_c \mapsto p(c)$: two polynomials lie in the same coset exactly when they agree at c. An infinite-dimensional space can have a one-dimensional quotient.

(b) *Trace.* The trace $\mathrm{tr} : F^{n\times n} \to F$ is linear and surjective, since $\mathrm{tr}(U_{11}) = 1$ for the matrix unit U_{11} (Lemma 2.4.24). So $F^{n\times n}/\ker \mathrm{tr} \cong F$, and by Theorem 4.6.27 the trace-zero matrices form a subspace of dimension $n^2 - 1$. The coset of A is determined by $\mathrm{tr}\, A$.

(c) *Fibers.* For $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ as in Figure 2.6.1, the cosets of $\ker T_A$ are the lines $x_1 + 2x_2 = t$, and $\widetilde{T_A}$ sends the line $x_1 + 2x_2 = t$ to $t(1, 1)$. So $\mathbb{R}^2/\ker T_A \cong \mathrm{span}((1, 1))$.

(d) *A quotient isomorphic to the whole space.* The backward shift $B : F[x] \to F[x]$,

$$B(a_0 + a_1 x + \cdots + a_n x^n) = a_1 + a_2 x + \cdots + a_n x^{n-1},$$

is linear (it acts on coefficient sequences by deleting the first entry) and surjective, since $B(xp) = p$. Its kernel is the set of constant polynomials, $\mathrm{span}(1)$. So $F[x]/\mathrm{span}(1) \cong F[x]$: dividing out a nonzero subspace gave back a space isomorphic to the original. In finite dimensions this is impossible, because $\dim(V/U) = \dim V - \dim U < \dim V$ when $U \neq \{\mathbf{0}\}$.

Exercise 4.6.37. STANDARD Let $U = \{p \in F[x] \mid p(1) = p(2) = 0\}$.

(a) Show that $F[x]/U \cong F^2$, and that $p + U = q + U$ if and only if $p(1) = q(1)$ and $p(2) = q(2)$.

(b) Show that $1 + U, x + U$ is a basis of $F[x]/U$, and write $x^2 + U$ in this basis. What does your answer say over $\mathbb{F}_2$?

Solution on page 1344

Exercises

Exercise 4.6.38. ROUTINE Let $V = F[x]_{\leq 2}$.

(a) Write down a basis of $V \times V$ and find $\dim(V \times V)$.

(b) Let $D = \{(p, q) \in V \times V \mid p(0) = q(1)\}$. Show that D is a subspace of $V \times V$, find $\dim D$, and find a basis of D.

Solution on page 1344

Exercise 4.6.39. STANDARD Let $T \in \mathrm{L}(V, W)$. The **graph** of T is $G_T = \{(\mathbf{v}, T(\mathbf{v})) \mid \mathbf{v} \in V\} \subseteq V \times W$.

(a) Show that G_T is a subspace of $V \times W$ and that $\mathbf{v} \mapsto (\mathbf{v}, T(\mathbf{v}))$ is an isomorphism $V \to G_T$.

(b) Show that $V \times W = G_T \oplus (\{\mathbf{0}\} \times W)$, and deduce that $(V \times W)/G_T \cong W$.

(c) Let G be any subspace of $V \times W$. Show that $G = G_S$ for some $S \in \mathrm{L}(V, W)$ if and only if $G \cap (\{\mathbf{0}\} \times W) = \{\mathbf{0}\}$ and every $\mathbf{v} \in V$ is the first component of some element of G.

Solution on page 1345

Exercise 4.6.40. ROUTINE For subsets $A, B \subseteq V$ and $a \in F$, let $A + B = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in A, \mathbf{y} \in B\}$ and $aA = \{a\mathbf{x} \mid \mathbf{x} \in A\}$. Let U be a subspace of V.

(a) Show that $(\mathbf{v}+U)+(\mathbf{w}+U)=(\mathbf{v}+\mathbf{w})+U$ as sets.

(b) Show that $a(\mathbf{v}+U)=a\mathbf{v}+U$ as sets when $a\neq 0$.

(c) Compute $0(\mathbf{v}+U)$, and explain why, when $U\neq\{\mathbf{0}\}$, these set operations do not make the set of cosets into a vector space, although they agree with the operations of Definition 4.6.20 whenever the result is a coset.

Solution on page 1345

Exercise 4.6.41. STANDARD Let U be a subspace of V, let $\mathbf{v}_1,\ldots,\mathbf{v}_k\in V$, and let $X=\operatorname{span}(\mathbf{v}_1,\ldots,\mathbf{v}_k)$.

(a) Show that $\mathbf{v}_1+U,\ldots,\mathbf{v}_k+U$ is linearly independent in V/U if and only if $\mathbf{v}_1,\ldots,\mathbf{v}_k$ is linearly independent and $X\cap U=\{\mathbf{0}\}$.

(b) Show that $\mathbf{v}_1+U,\ldots,\mathbf{v}_k+U$ spans V/U if and only if $X+U=V$.

(c) Conclude that $\mathbf{v}_1+U,\ldots,\mathbf{v}_k+U$ is a basis of V/U if and only if $\mathbf{v}_1,\ldots,\mathbf{v}_k$ is a basis of a complement of U.

Solution on page 1346

Exercise 4.6.42. STANDARD Let $V=F^{2\times 2}$ and $S=\left\{A\in V\;\middle|\;A^{\mathsf{T}}=A\right\}$.

(a) Show that $A+S\mapsto a_{12}-a_{21}$ is a well-defined isomorphism $V/S\to F$, where $A=\left(a_{ij}\right)$.

(b) Suppose $1+1\neq 0$ in F. Show that $K=\left\{A\;\middle|\;A^{\mathsf{T}}=-A\right\}$ is a complement of S.

(c) Show that over $\mathbb{F}_2$ the subspace K is not a complement of S, and find a complement.

Solution on page 1346

Exercise 4.6.43. STANDARD Let $T\in \mathrm{L}(V,W)$, and let U be a subspace of V with $\ker T\subseteq U$. A student defines $S:V/U\to W$ by $S(\mathbf{v}+U)=T(\mathbf{v})$, and argues:

"S is well defined: if $S(\mathbf{v}+U)=S(\mathbf{v}'+U)$, then $T(\mathbf{v})=T(\mathbf{v}')$, so $\mathbf{v}-\mathbf{v}'\in\ker T\subseteq U$, hence $\mathbf{v}+U=\mathbf{v}'+U$."

Find the error, and give an example with $V=W=\mathbb{R}^2$ in which the formula does not define a function on V/U. What hypothesis would make the definition correct?

Solution on page 1346

Exercise 4.6.44. STANDARD Let U be a subspace of V, where V may be infinite-dimensional.

(a) Show that $\dim(V/U)=1$ if and only if $U=\ker f$ for some nonzero linear map $f:V\to F$.

(b) Let $f,g:V\to F$ be nonzero linear maps. Show that $\ker f=\ker g$ if and only if $g=cf$ for some nonzero $c\in F$.

Solution on page 1347

Exercise 4.6.45. STANDARD Let $A\in F^{m\times n}$ and $B\in F^{p\times n}$. Show that $B=CA$ for some $C\in F^{p\times m}$ if and only if $\ker T_A\subseteq\ker T_B$.

Solution on page 1347

Exercise 4.6.46. HARD Let $T_1\in\mathrm{L}(V,W_1)$ and $T_2\in\mathrm{L}(V,W_2)$ be surjective, and let $T=(T_1,T_2):V\to W_1\times W_2$.

(a) Show that $\ker T = \ker T_1 \cap \ker T_2$.

(b) Show that T is surjective if and only if $\ker T_1 + \ker T_2 = V$.

(c) Let U, W be subspaces of V. Show that $V \to V/U \times V/W$, $\mathbf{v} \mapsto (\mathbf{v}+U, \mathbf{v}+W)$, has kernel $U \cap W$, and is surjective if and only if $U + W = V$. Deduce that if $U + W = V$, then $V/(U \cap W) \cong V/U \times V/W$.

Solution on page 1348

Exercise 4.6.47. STANDARD Let $U_1, \ldots, U_m$ be subspaces of V such that each V/U_i is finite-dimensional. Here V may be infinite-dimensional.

(a) Show that $V/(U_1 \cap \cdots \cap U_m)$ is finite-dimensional and

$$\dim(V/(U_1 \cap \cdots \cap U_m)) \leq \dim(V/U_1) + \cdots + \dim(V/U_m).$$

(b) Deduce that if V is finite-dimensional, then $\dim(U_1 \cap \cdots \cap U_m) \geq \dim U_1 + \cdots + \dim U_m - (m-1)\dim V$.

(c) Show that any two 4-dimensional subspaces and any 3-dimensional subspace of F^5 have a nonzero vector in common.

Solution on page 1348

Exercise 4.6.48. STANDARD Let $n \geq 1$, and let $Z \subseteq F^{n \times n}$ be the set of matrices whose row sums and column sums are all 0. Let $\mathbf{1} = (1, \ldots, 1) \in F^n$. Use the map

$$\Phi : F^{n \times n} \to F^n \times F^n, \quad \Phi(A) = \left(A\mathbf{1}, A^{\mathsf{T}}\mathbf{1}\right),$$

to show that Z is a subspace of dimension $(n-1)^2$.

Solution on page 1349

Exercise 4.6.49. STANDARD Let U and W be subspaces of V.

(a) Show that $\mathbf{w} \mapsto \mathbf{w} + U$ is a surjective linear map $W \to (U+W)/U$ with kernel $U \cap W$, and deduce the **second isomorphism theorem**: $W/(U \cap W) \cong (U+W)/U$.

(b) Use (a) and Theorem 4.6.27 to give another proof of Theorem 4.4.19.

Solution on page 1349

Exercise 4.6.50. HARD Let U be a subspace of V, and let $\pi : V \to V/U$ be the quotient map. For a subspace X of V with $U \subseteq X$, write $X/U = \{\mathbf{x} + U \mid \mathbf{x} \in X\} \subseteq V/U$.

(a) Show that X/U is a subspace of V/U and that $\pi^{-1}(X/U) = X$.

(b) Show that every subspace Y of V/U equals X/U for exactly one subspace X of V containing U.

(c) (*Third isomorphism theorem.*) Show that $(V/U)/(X/U) \cong V/X$ whenever $U \subseteq X \subseteq V$.

(d) If V is finite-dimensional, show that $\dim(X/U) = \dim X - \dim U$.

Solution on page 1349

Exercise 4.6.51. HARD Let U and W be subspaces of V. For each statement, decide whether it holds when V is finite-dimensional, and whether it holds for all V.

(a) If $U \neq \{\mathbf{0}\}$, then $V/U \ncong V$.

(b) If $V/U \cong V/W$, then $U \cong W$.

(c) If $U \cong W$, then $V/U \cong V/W$.

Solution on page 1350

Exercise 4.6.52. HARD Let V be finite-dimensional, $T \in \mathrm{L}(V)$, and U a subspace with $T(U) \subseteq U$. Write $T_U \in \mathrm{L}(U)$ for $T|_U$ regarded as an operator on U, and T/U for the induced operator on V/U (Proposition 4.6.32).

(a) Show that $\operatorname{rank} T \geq \operatorname{rank} T_U + \operatorname{rank}(T/U)$.

(b) Show that equality holds if and only if $\operatorname{im} T \cap U = T(U)$.

(c) Give an example where the inequality is strict. What does (a) say about block upper triangular matrices?

Solution on page 1351

Exercise 4.6.53. HARD Reprove Frobenius's rank inequality (Exercise 4.3.31(a)) with quotients: for $A \in F^{m\times n}$, $B \in F^{n\times p}$, and $C \in F^{p\times q}$,

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leq \operatorname{rank} B + \operatorname{rank}(ABC).$$

Use the map $\operatorname{im} T_B/\operatorname{im} T_{BC} \to \operatorname{im} T_{AB}/\operatorname{im} T_{ABC}$ induced by T_A.

Solution on page 1351

Exercise 4.6.54. HARD Let F be a field with q elements, let V be an n-dimensional space over F, and let U be a subspace of dimension k. Fix one complement W_0 of U.

(a) For $S \in \mathrm{L}(W_0, U)$, show that $W_S = \{\mathbf{w} + S(\mathbf{w}) \mid \mathbf{w} \in W_0\}$ is a complement of U.

(b) Show that every complement of U is W_S for exactly one $S \in \mathrm{L}(W_0, U)$.

(c) Deduce that U has exactly $q^{k(n-k)}$ complements. How many lines in $\mathbb{F}_3^2$ are complements of a given line?

Solution on page 1352

Exercise 4.6.55. CHALLENGE Let U be a subspace of $F[x]$ with $U \neq \{\mathbf{0}\}$ and $xp \in U$ for every $p \in U$. Let d be the smallest degree of a nonzero element of U.

(a) Show that U contains a polynomial $p = x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$, and that $F[x]_{\leq d-1}$ is a complement of U (read $F[x]_{\leq -1}$ as $\{\mathbf{0}\}$). Conclude that $\dim(F[x]/U) = d$.

(b) Suppose $d \geq 1$. Show that multiplication by x induces an operator X on $F[x]/U$ with $X(q+U) = xq + U$, and find the matrix of X in the basis $1+U, x+U, \ldots, x^{d-1}+U$.

Solution on page 1352

4.7 Affine geometry BEYOND 21-242

The solution set of a linear system $A\mathbf{x} = \mathbf{y}$ with $\mathbf{y} \neq \mathbf{0}$ is never a subspace, but by Proposition 3.1.19 it is the next best thing: the subspace $\ker T_A$, slid over so that it passes through one solution. Lines and planes that miss the origin are sets of this kind, and so is the screen of a computer, which has no preferred origin either. On such a set the sum of two points means nothing (the sum of two solutions of $A\mathbf{x} = \mathbf{y}$ solves $A\mathbf{x} =$

$2\mathbf{y}$), yet some combinations of points, such as midpoints, make perfect sense. This section finds out which combinations these are, and which maps respect them. Its last part, homogeneous coordinates, turns this affine geometry back into linear algebra one dimension up, which is how graphics software moves and projects three-dimensional scenes.

4.7.1 Affine subspaces

Recall the translate $\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\}$ of a subset $U \subseteq V$ (Notation 3.1.17).

Definition 4.7.1 (Affine subspace). Let V be a vector space over F. A subset $E \subseteq V$ is an **affine subspace** of V if $E = \mathbf{v} + U$ for some $\mathbf{v} \in V$ and some subspace U of V.

An affine subspace contains $\mathbf{v} = \mathbf{v} + \mathbf{0}$, so it is never empty. Some authors count $\emptyset$ as an affine subspace as well, which makes statements about intersections shorter; we will say "empty or an affine subspace" where it matters. We use the letters E, E', E_i for affine subspaces and keep A and B for matrices.

Example 4.7.2 (Affine subspaces).

(a) Every subspace U is an affine subspace, since $U = \mathbf{0} + U$. Every single point is one, since $\{\mathbf{v}\} = \mathbf{v} + \{\mathbf{0}\}$.

(b) If the system $A\mathbf{x} = \mathbf{y}$ has a solution $\mathbf{x}_0$, its solution set $\mathbf{x}_0 + \ker T_A$ is an affine subspace of F^n (Proposition 3.1.19). More generally, every nonempty fiber $T^{-1}(\mathbf{w})$ of a linear map $T : V \to W$ is one (Corollary 2.6.31).

(c) In $\mathbb{R}^2$, the line $x_1 + x_2 = 2$ is $(1, 1) + \operatorname{span}((1, -1))$.

(d) In $\mathbb{R}[x]$, the polynomials p with $p(1) = 1$ form the fiber over 1 of the linear map $p \mapsto p(1)$, so they form the affine subspace $1 + \{p \in \mathbb{R}[x] \mid p(1) = 0\}$. Here 1 is the constant polynomial.

The vector $\mathbf{v}$ in $E = \mathbf{v} + U$ is not determined by E, but the subspace U is, and every point of E can play the role of $\mathbf{v}$.

Proposition 4.7.3 (Direction and base points). Let $E = \mathbf{v} + U$ be an affine subspace of V.

(a) $U = \{\mathbf{a} - \mathbf{b} \mid \mathbf{a}, \mathbf{b} \in E\}$. In particular, if also $E = \mathbf{v}' + U'$ for some $\mathbf{v}' \in V$ and some subspace U', then $U' = U$.

(b) $E = \mathbf{a} + U$ for every $\mathbf{a} \in E$.

(c) For $\mathbf{w} \in V$, the conditions $\mathbf{w} + U = E$, $\mathbf{w} \in E$, and $\mathbf{w} - \mathbf{v} \in U$ are equivalent.

Proof. ROUTINE

Uses. Definition 4.7.1 (Affine subspace), Notation 3.1.17 (Translates), Definition 1.4.1 (Subspace)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For (a), write two points of E as $\mathbf{v} + \mathbf{u}$ and $\mathbf{v} + \mathbf{u}'$. For the reverse containment, write $\mathbf{u} \in U$ as a difference of two points of E.
Hint 2. For (b), write $\mathbf{a} = \mathbf{v} + \mathbf{u}_0$ and prove both containments, moving $\mathbf{u}_0$ from one side to the other.
Hint 3. For (c), go around the cycle of the three conditions; (a) and (b) do the work.
Sketch. (a) $(\mathbf{v} + \mathbf{u}) - (\mathbf{v} + \mathbf{u}') = \mathbf{u} - \mathbf{u}' \in U$, and $\mathbf{u} = (\mathbf{v} + \mathbf{u}) - (\mathbf{v} + \mathbf{0})$. The set of differences depends only on E, so it is U however E is written. (b) If $\mathbf{a} = \mathbf{v} + \mathbf{u}_0$, then $\mathbf{a} + \mathbf{u} = \mathbf{v} + (\mathbf{u}_0 + \mathbf{u})$ and $\mathbf{v} + \mathbf{u} = \mathbf{a} + (\mathbf{u} - \mathbf{u}_0)$. (c) $\mathbf{w} + U = E$ gives $\mathbf{w} = \mathbf{w} + \mathbf{0} \in E$; $\mathbf{w} \in E$ gives $\mathbf{w} - \mathbf{v} \in U$ by (a); and $\mathbf{w} - \mathbf{v} \in U$ gives $\mathbf{w} \in E$, so $\mathbf{w} + U = E$ by (b).

(a) The points of E are the vectors $\mathbf{v} + \mathbf{u}$ with $\mathbf{u} \in U$. For two of them, $(\mathbf{v} + \mathbf{u}) - (\mathbf{v} + \mathbf{u}') = \mathbf{u} - \mathbf{u}'$, which lies in U because U is a subspace. Conversely, every $\mathbf{u} \in U$ is the difference $(\mathbf{v} + \mathbf{u}) - (\mathbf{v} + \mathbf{0})$ of two points of E. So U equals the set D of differences of points of E. The set D is defined in terms of E alone, so if also $E = \mathbf{v}' + U'$, the same argument gives $U' = D = U$.

(b) Let $\mathbf{a} \in E$, say $\mathbf{a} = \mathbf{v} + \mathbf{u}_0$ with $\mathbf{u}_0 \in U$. For $\mathbf{u} \in U$,

$$\mathbf{a} + \mathbf{u} = \mathbf{v} + (\mathbf{u}_0 + \mathbf{u}) \in \mathbf{v} + U \quad \text{and} \quad \mathbf{v} + \mathbf{u} = \mathbf{a} + (\mathbf{u} - \mathbf{u}_0) \in \mathbf{a} + U,$$

because $\mathbf{u}_0 + \mathbf{u}$ and $\mathbf{u} - \mathbf{u}_0$ lie in U. So $\mathbf{a} + U = \mathbf{v} + U = E$.

(c) If $\mathbf{w} + U = E$, then $\mathbf{w} = \mathbf{w} + \mathbf{0} \in E$. If $\mathbf{w} \in E$, then $\mathbf{w} - \mathbf{v}$ is a difference of two points of E (as $\mathbf{v} = \mathbf{v} + \mathbf{0} \in E$), so it lies in U by (a). If $\mathbf{w} - \mathbf{v} \in U$, then $\mathbf{w} = \mathbf{v} + (\mathbf{w} - \mathbf{v}) \in E$, and $\mathbf{w} + U = E$ by (b). □

Technique 4.7.4 (Choose a base point). An affine subspace has no origin of its own, but by Proposition 4.7.3(b) any of its points can serve as one. Once a point $\mathbf{a} \in E$ is fixed, $\mathbf{x} \mapsto \mathbf{x} - \mathbf{a}$ carries E onto the subspace $U = E - \mathbf{a}$, and statements about E become statements about U, where linear algebra applies. Every point of E works equally well, so choose the most convenient one, such as a point that two affine subspaces have in common, or a fixed point of a map.

Reach for it whenever an affine question has a linear counterpart that you already know how to answer. The canonical example is Proposition 4.7.10. The same move drives Theorem 4.7.14, Proposition 4.7.17, and Theorem 4.7.37.

Used in. Corollary 4.7.6, Proposition 4.7.10, Theorem 4.7.14, Proposition 4.7.17, Proposition 4.7.20, Theorem 4.7.37, Proposition 4.7.38, Theorem 4.7.39, Theorem 8.4.31, Exercise 8.4.46.

Definition 4.7.5 (Direction, dimension, parallel). Let E be an affine subspace of V. The subspace U with $E = \mathbf{v} + U$, which is unique by Proposition 4.7.3(a), is the **direction** of E, written $\operatorname{dir} E$. If $\operatorname{dir} E$ is finite-dimensional, the **dimension** of E is $\dim E = \dim \operatorname{dir} E$. Affine subspaces of dimension 0, 1, and 2 are called **points**, **lines**, and **planes**; if $\dim V = n$, those of dimension $n - 1$ are called **hyperplanes**. Two affine subspaces E and E' of V are **parallel** if $\operatorname{dir} E = \operatorname{dir} E'$.

A subspace U has direction U, so its dimension is the same whether we regard it as a subspace or as an affine subspace. A line is a set $\mathbf{v} + \operatorname{span}(\mathbf{u}) = \{\mathbf{v} + t\mathbf{u} \mid t \in F\}$ with $\mathbf{u} \neq \mathbf{0}$.

Corollary 4.7.6 (Parallel affine subspaces are equal or disjoint). If E and E' are parallel affine subspaces of V, then $E = E'$ or $E \cap E' = \emptyset$. Given a subspace U of V, every $\mathbf{v} \in V$ lies in exactly one translate of U, namely $\mathbf{v} + U$.

Proof. ROUTINE

Uses. Proposition 4.7.3 (Direction and base points)
Techniques. Choose a base point (Technique 4.7.4)
Hint. Suppose some $\mathbf{c}$ lies in both. Use $\mathbf{c}$ as the base point of each.
Sketch. If $\mathbf{c} \in E \cap E'$, then $E = \mathbf{c} + U = E'$ by Proposition 4.7.3(b), where U is the common direction. For the second claim, $\mathbf{v} \in \mathbf{v} + U$, and a translate $\mathbf{w} + U$ containing $\mathbf{v}$ equals $\mathbf{v} + U$ by the same result.

Let $U = \operatorname{dir} E = \operatorname{dir} E'$. If $\mathbf{c} \in E \cap E'$, then $E = \mathbf{c} + U$ and $E' = \mathbf{c} + U$ by Proposition 4.7.3(b), so $E = E'$. For the second statement, $\mathbf{v} = \mathbf{v} + \mathbf{0} \in \mathbf{v} + U$, and if $\mathbf{v} \in \mathbf{w} + U$, then Proposition 4.7.3(b), applied to the affine subspace $\mathbf{w} + U$ and its point $\mathbf{v}$, gives $\mathbf{w} + U = \mathbf{v} + U$. □

So the translates of a subspace U cut V into disjoint parallel pieces, as the solution sets in Figure 3.1.4 do. They are the elements of the quotient space V/U (Definition 4.6.20), which treats each translate as a single vector.

Non-example 4.7.7 (Sets that are not affine subspaces).

(a) The unit circle $C = \left\{\mathbf{x} \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1\right\}$. If C were an affine subspace, then $(1,0) - (-1,0) = (2,0)$ would lie in $\operatorname{dir} C$ by Proposition 4.7.3(a), and $C = (1,0) + \operatorname{dir} C$ by (b). Then $(3,0) = (1,0) + (2,0) \in C$, which is false.

(b) The union E of the lines $x_2 = 0$ and $x_2 = 1$ in $\mathbb{R}^2$. If E were an affine subspace, then $(0,1) - (0,0) \in \operatorname{dir} E$, so $(0,0) + 2(0,1) = (0,2)$ would lie in E.

(c) The three-point set $\{(0,0), (1,0), (0,1)\} \subseteq \mathbb{F}_2^2$ (Exercise 3.1.48).

Solution sets, revisited

Proposition 3.1.19 says that consistent linear systems have affine solution sets. With dimension theory we can now say how large they are, and show that every affine subspace of F^n arises this way.

Theorem 4.7.8 (Affine subspaces of F^n are solution sets).

(a) Let $A \in F^{m\times n}$ and $\mathbf{y} \in F^m$. If the system $A\mathbf{x} = \mathbf{y}$ is consistent, its solution set is an affine subspace of F^n with direction $\ker T_A$ and dimension $n - \operatorname{rank} A$.

(b) Conversely, let E be an affine subspace of F^n of dimension k. Then E is the solution set of a system $A\mathbf{x} = \mathbf{y}$ with $A \in F^{(n-k)\times n}$, that is, of $n-k$ equations. No system of fewer than $n-k$ equations has solution set E.

Proof. STANDARD

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 4.2.23 (Dimension of solution sets), Proposition 4.4.16 (Complements exist), Theorem 4.4.19 (Dimension of a sum), Proposition 4.4.18 (Every subspace is a kernel and an image), Theorem 4.5.6 (Classification by dimension), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Theorem 4.2.7 (The image of T_A), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Particular plus homogeneous (Technique 3.1.18), Dimension counting (Technique 4.2.24)
Hint 1. (a) is Proposition 3.1.19 plus a dimension count.
Hint 2. For (b), write $E = \mathbf{v} + U$. It suffices to find $A \in F^{(n-k)\times n}$ with $\ker T_A = U$, and then to take $\mathbf{y} = A\mathbf{v}$.
Hint 3. Every subspace is a kernel (Proposition 4.4.18), but of an operator $F^n \to F^n$, whose image is a complement W of U. Follow it by an isomorphism $W \to F^{n-k}$.
Sketch. (a) The solution set is $\mathbf{x}_0 + \ker T_A$, and $\dim \ker T_A = n - \operatorname{rank} A$. (b) Let $E = \mathbf{v} + U$ and let W be a complement of U, so $\dim W = n - k$. Take $T \in \mathrm{L}(F^n)$ with $\ker T = U$ and $\operatorname{im} T = W$, and an isomorphism $\varphi : W \to F^{n-k}$. Then $\varphi \circ T$ has kernel U and equals T_A with $A \in F^{(n-k)\times n}$, and $A\mathbf{x} = A\mathbf{v}$ has solution set $\mathbf{v} + U = E$. With m equations, (a) gives $k = n - \operatorname{rank} A \geq n - m$.

(a) By Proposition 3.1.19, the solution set is $\mathbf{x}_0 + \ker T_A$ for any solution $\mathbf{x}_0$, an affine subspace with direction $\ker T_A$. By Corollary 4.2.23(a), $\dim \ker T_A = n - \operatorname{rank} A$.

(b) Write $E = \mathbf{v} + U$ with $U = \operatorname{dir} E$, so $\dim U = k$. By Proposition 4.4.16, U has a complement W in F^n. Since $U \cap W = \{\mathbf{0}\}$ and $U + W = F^n$, Theorem 4.4.19 gives $\dim W = n - k$. By Proposition 4.4.18 there is $T \in \mathrm{L}(F^n)$ with $\ker T = U$ and $\operatorname{im} T = W$, and by Theorem 4.5.6 there is an isomorphism $\varphi : W \to F^{n-k}$. The map $S : F^n \to F^{n-k}$, $S(\mathbf{x}) = \varphi(T(\mathbf{x}))$, is linear, and since φ is injective, $S(\mathbf{x}) = \mathbf{0}$ if and only if $T(\mathbf{x}) = \mathbf{0}$. So $\ker S = U$. By Proposition 2.3.17, $S = T_A$ for some $A \in F^{(n-k)\times n}$. Put $\mathbf{y} = A\mathbf{v}$. Then $\mathbf{v}$ solves $A\mathbf{x} = \mathbf{y}$, and by Proposition 3.1.19 the solution set is $\mathbf{v} + \ker T_A = \mathbf{v} + U = E$.

Finally, suppose E is the solution set of a system $A'\mathbf{x} = \mathbf{y}'$ with $A' \in F^{m\times n}$. By (a), $k = n - \operatorname{rank} A'$. Since $\operatorname{im} T_{A'}$ is a subspace of F^m, Theorem 4.2.7 and Proposition 4.1.21 give $\operatorname{rank} A' = \dim \operatorname{im} T_{A'} \leq m$. Hence $k \geq n - m$, that is, $m \geq n - k$. □

For $k = n - 1$ this says that the hyperplanes of F^n are exactly the solution sets of single equations $a_1x_1 + \cdots + a_nx_n = c$ with $(a_1, \ldots, a_n) \neq \mathbf{0}$. (Such an equation is consistent, since $x_j = c/a_j$ and $x_i = 0$ for $i \neq j$ solves it when $a_j \neq 0$, and its coefficient matrix has rank 1.)

Example 4.7.9 (A line in $\mathbb{R}^3$ as a solution set). Let E be the line through $(1, 2, 0)$ with direction $\operatorname{span}((1, 1, 1))$. By Theorem 4.7.8(b), two equations cut it out. We need $A \in \mathbb{R}^{2\times 3}$ of rank 2 with $(1, 1, 1) \in \ker T_A$. The rows (a_1, a_2, a_3) must satisfy $a_1 + a_2 + a_3 = 0$ and be independent, so take $(1, -1, 0)$ and $(0, 1, -1)$. Then $A(1, 2, 0) = (-1, 2)$, and E is the solution set of

$$x_1 - x_2 = -1, \quad x_2 - x_3 = 2.$$

Indeed, the solution set contains $(1, 2, 0)$, and its direction $\ker T_A$ contains $(1, 1, 1)$ and has dimension $3 - 2 = 1$, so it equals $\operatorname{span}((1, 1, 1))$ by Proposition 4.1.21(b).

Intersections

Proposition 4.7.10 (Intersections of affine subspaces). Let $(E_i)_{i\in I}$ be a family of affine subspaces of V, indexed by a nonempty set I. Then $\bigcap_{i\in I} E_i$ is either empty or an affine subspace with direction $\bigcap_{i\in I} \operatorname{dir} E_i$.

Proof. ROUTINE

Uses. Proposition 4.7.3 (Direction and base points), Proposition 1.4.16 (Intersections of subspaces)
Techniques. Choose a base point (Technique 4.7.4), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. If the intersection is nonempty, pick a point $\mathbf{c}$ in it. How does each E_i look when $\mathbf{c}$ is used as its base point?
Hint 2. $E_i = \mathbf{c} + \operatorname{dir} E_i$, so $\mathbf{x} \in E_i$ if and only if $\mathbf{x} - \mathbf{c} \in \operatorname{dir} E_i$.
Sketch. With $\mathbf{c}$ in the intersection, $E_i = \mathbf{c} + U_i$ for every i. So $\mathbf{x}$ lies in every E_i if and only if $\mathbf{x} - \mathbf{c}$ lies in every U_i, and the intersection is $\mathbf{c} + \bigcap U_i$, a translate of a subspace.

Suppose $\mathbf{c} \in \bigcap_{i\in I} E_i$, and let $U_i = \operatorname{dir} E_i$. By Proposition 4.7.3(b), $E_i = \mathbf{c} + U_i$ for every i. For $\mathbf{x} \in V$ we have $\mathbf{x} \in \mathbf{c} + U_i$ if and only if $\mathbf{x} - \mathbf{c} \in U_i$ (if $\mathbf{x} = \mathbf{c} + \mathbf{u}$, then $\mathbf{x} - \mathbf{c} = \mathbf{u}$; conversely $\mathbf{x} = \mathbf{c} + (\mathbf{x} - \mathbf{c})$). Hence $\mathbf{x}$ lies in every E_i if and only if $\mathbf{x} - \mathbf{c} \in \bigcap_i U_i$, that is,

$$\bigcap_{i\in I} E_i = \mathbf{c} + \bigcap_{i\in I} U_i.$$

By Proposition 1.4.16, $\bigcap_i U_i$ is a subspace, so the intersection is an affine subspace, and its direction is $\bigcap_i U_i$ by Proposition 4.7.3(a). □

The choice of base point was the whole proof: a point common to all the E_i lets us write all of them as translates from one origin. Without such a point the intersection can be empty even when the directions have a large intersection, as for two distinct parallel lines. In $\mathbb{R}^3$ two lines can also miss each other without being parallel, for example the x_1-axis and $(0,0,1) + \operatorname{span}(\mathbf{e}_2)$; such lines are called **skew**. Exercise 4.7.53 characterizes skew lines, and Exercise 4.7.52 explains why a plane has none.

4.7.2 Affine combinations

Intuition (Which combinations make sense without an origin). Suppose we only know positions relative to some reference point $\mathbf{p}$, as on a screen or in a scene: the point $\mathbf{v}$ has position $\mathbf{v} - \mathbf{p}$. Combining positions with coefficients c_i and converting the result back into a point gives

$$\mathbf{p} + \sum_i c_i(\mathbf{v}_i - \mathbf{p}) = \sum_i c_i \mathbf{v}_i + \left(1 - \sum_i c_i\right)\mathbf{p}.$$

For two reference points $\mathbf{p}$ and $\mathbf{p}'$ the results differ by $\left(1 - \sum_i c_i\right)(\mathbf{p} - \mathbf{p}')$, so the result does not depend on the reference point exactly when the coefficients add up to 1. These are the weighted averages of points. Combinations whose coefficients add up to 0 are meaningful in a different way: $\sum_i c_i(\mathbf{v}_i - \mathbf{p}) = \sum_i c_i \mathbf{v}_i$ does not depend on $\mathbf{p}$, and it is a displacement rather than a point, like the difference $\mathbf{v} - \mathbf{u}$.

Definition 4.7.11 (Affine combination). An **affine combination** of $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ is a linear combination $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ whose coefficients satisfy $c_1 + \cdots + c_k = 1$.

Example 4.7.12 (Affine combinations).

(a) For $\mathbf{u}, \mathbf{v} \in V$ and $t \in F$, $(1-t)\mathbf{u} + t\mathbf{v} = \mathbf{u} + t(\mathbf{v} - \mathbf{u})$. If $\mathbf{u} \neq \mathbf{v}$, these points form the line through $\mathbf{u}$ and $\mathbf{v}$. When $1 + 1 \neq 0$ in F, the case $t = 1/2$ is the **midpoint** $\frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$.

(b) If $k \cdot 1 \neq 0$ in F, the **centroid** $\frac{1}{k}(\mathbf{v}_1 + \cdots + \mathbf{v}_k)$ is an affine combination. Over $\mathbb{F}_3$, three points have no centroid.

(c) $\mathbf{u} - \mathbf{v} + \mathbf{w}$, with coefficients $1, -1, 1$, is the fourth vertex of the parallelogram with consecutive vertices $\mathbf{u}, \mathbf{v}, \mathbf{w}$: its differences with $\mathbf{w}$ and with $\mathbf{u}$ are $\mathbf{u} - \mathbf{v}$ and $\mathbf{w} - \mathbf{v}$.

(d) If $A\mathbf{x}_i = \mathbf{y}$ for $i = 1, \ldots, k$ and $\sum_i c_i = 1$, then $A\big(\sum_i c_i \mathbf{x}_i\big) = \big(\sum_i c_i\big)\mathbf{y} = \mathbf{y}$. So affine combinations of solutions are solutions, while the sum of two solutions is not.

Lemma 4.7.13 (Affine combinations do not depend on the origin). Let $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{p} \in V$ and $c_1, \ldots, c_k \in F$.

(a) If $c_1 + \cdots + c_k = 1$, then $\sum_i c_i \mathbf{v}_i = \mathbf{p} + \sum_i c_i(\mathbf{v}_i - \mathbf{p})$.

(b) If $c_1 + \cdots + c_k = 0$, then $\sum_i c_i \mathbf{v}_i = \sum_i c_i(\mathbf{v}_i - \mathbf{p})$.

Proof. ROUTINE

Uses. Lemma 1.3.9 (Arithmetic of linear combinations)
Hint. Expand $\sum_i c_i(\mathbf{v}_i - \mathbf{p})$.
Sketch. $\sum_i c_i(\mathbf{v}_i - \mathbf{p}) = \sum_i c_i \mathbf{v}_i - \big(\sum_i c_i\big)\mathbf{p}$; substitute $\sum_i c_i = 1$ or 0.

By distributivity and Lemma 1.3.9, $\sum_i c_i(\mathbf{v}_i - \mathbf{p}) = \sum_i c_i \mathbf{v}_i - \big(\sum_i c_i\big)\mathbf{p}$. If $\sum_i c_i = 1$, this is $\sum_i c_i \mathbf{v}_i - \mathbf{p}$, and adding $\mathbf{p}$ gives (a). If $\sum_i c_i = 0$, it is $\sum_i c_i \mathbf{v}_i$, which is (b). □

Subspaces are the nonempty sets closed under linear combinations. Affine subspaces have the analogous description.

Theorem 4.7.14 (Affine subspaces are the sets closed under affine combinations). A nonempty subset $E \subseteq V$ is an affine subspace if and only if every affine combination of elements of E lies in E.

Intuition. The forward direction is a base point computation with Lemma 4.7.13(a). For the converse, the base point tells us what the direction must be: if $E = \mathbf{p} + U$ with $\mathbf{p} \in E$, then $U = \{\mathbf{e} - \mathbf{p} \mid \mathbf{e} \in E\}$. So define U this way and check the subspace criterion. Each closure property of U should come from one affine combination in E.

Proof. STANDARD

Uses. Lemma 4.7.13 (Affine combinations do not depend on the origin), Proposition 4.7.3 (Direction and base points), Theorem 1.4.3 (Subspace criterion)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. For ($\Rightarrow$), write $E = \mathbf{p} + U$ and each $\mathbf{e}_i = \mathbf{p} + \mathbf{u}_i$, and apply Lemma 4.7.13(a).
Hint 2. For ($\Leftarrow$), fix $\mathbf{p} \in E$ and put $U = \{\mathbf{e} - \mathbf{p} \mid \mathbf{e} \in E\}$. Then $E = \mathbf{p} + U$, and it remains to show that U is a subspace.
Hint 3. $c(\mathbf{e} - \mathbf{p}) = (c\mathbf{e} + (1-c)\mathbf{p}) - \mathbf{p}$ and $(\mathbf{e} - \mathbf{p}) + (\mathbf{e}' - \mathbf{p}) = (\mathbf{e} + \mathbf{e}' - \mathbf{p}) - \mathbf{p}$. The vectors in parentheses are affine combinations of points of E.
Sketch. ($\Rightarrow$) $\sum_i c_i(\mathbf{p} + \mathbf{u}_i) = \mathbf{p} + \sum_i c_i \mathbf{u}_i \in \mathbf{p} + U$. ($\Leftarrow$) With $U = E - \mathbf{p}$: $\mathbf{0} = \mathbf{p} - \mathbf{p} \in U$, $c(\mathbf{e} - \mathbf{p}) = (c\mathbf{e} + (1-c)\mathbf{p}) - \mathbf{p} \in U$, and $(\mathbf{e} - \mathbf{p}) + (\mathbf{e}' - \mathbf{p}) = (\mathbf{e} + \mathbf{e}' - \mathbf{p}) - \mathbf{p} \in U$, because $c\mathbf{e} + (1-c)\mathbf{p}$ and $\mathbf{e} + \mathbf{e}' - \mathbf{p}$ are affine combinations of elements of E.

($\Rightarrow$) Let $E = \mathbf{p} + U$ be an affine subspace, let $\mathbf{e}_1, \ldots, \mathbf{e}_k \in E$, and let $\sum_i c_i = 1$. Since $\mathbf{p} = \mathbf{p} + \mathbf{0} \in E$, each $\mathbf{u}_i = \mathbf{e}_i - \mathbf{p}$ lies in U by Proposition 4.7.3(a). By Lemma 4.7.13(a),

$$\sum_i c_i \mathbf{e}_i = \mathbf{p} + \sum_i c_i \mathbf{u}_i \in \mathbf{p} + U = E.$$

($\Leftarrow$) Suppose E is nonempty and closed under affine combinations. Fix $\mathbf{p} \in E$ and put $U = \{\mathbf{e} - \mathbf{p} \mid \mathbf{e} \in E\}$. Then $E = \mathbf{p} + U$, since each $\mathbf{e} \in E$ equals $\mathbf{p} + (\mathbf{e} - \mathbf{p})$. We check the conditions of Theorem 1.4.3 for U.

$\mathbf{0} \in U$, since $\mathbf{0} = \mathbf{p} - \mathbf{p}$.

Scalar multiples. Let $\mathbf{e} \in E$ and $c \in F$. The vector $c\mathbf{e} + (1-c)\mathbf{p}$ is an affine combination of $\mathbf{e}, \mathbf{p} \in E$, so it lies in E, and

$$c(\mathbf{e} - \mathbf{p}) = (c\mathbf{e} + (1-c)\mathbf{p}) - \mathbf{p} \in U.$$

Sums. Let $\mathbf{e}, \mathbf{e}' \in E$. The vector $\mathbf{e} + \mathbf{e}' - \mathbf{p}$ is an affine combination of $\mathbf{e}, \mathbf{e}', \mathbf{p}$, with coefficients $1 + 1 - 1 = 1$, so it lies in E, and

$$(\mathbf{e} - \mathbf{p}) + (\mathbf{e}' - \mathbf{p}) = (\mathbf{e} + \mathbf{e}' - \mathbf{p}) - \mathbf{p} \in U.$$

So U is a subspace, and $E = \mathbf{p} + U$ is an affine subspace. □

Remark 4.7.15 (Two points or three). The proof used affine combinations of two points, for scalar multiples, and of three points, for sums. The three-point combination cannot be dropped in general. When $1 + 1 \neq 0$, Exercise 3.1.48 showed that closure under combinations of two points (the lines through pairs of points) already forces a nonempty set to be affine, by way of midpoints. Over $\mathbb{F}_2$, the only affine combinations of two points $\mathbf{u}, \mathbf{v}$ are $\mathbf{u}$ and $\mathbf{v}$ themselves. The set $\{(0,0), (1,0), (0,1)\}$ is closed under them, but $(1,0) + (0,1) - (0,0) = (1,1)$ is missing. Exercise 4.7.61 shows that over $\mathbb{F}_2$ three points always suffice.

Definition 4.7.16 (Affine hull). The **affine hull** $\operatorname{aff} S$ of a subset $S \subseteq V$ is the set of all affine combinations of finitely many elements of S. For a list, $\operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k) = \operatorname{aff}\{\mathbf{v}_0, \ldots, \mathbf{v}_k\}$.

The empty set has no affine combinations, because an empty list of coefficients adds up to 0, so $\operatorname{aff} \emptyset = \emptyset$.

Proposition 4.7.17 (The affine hull). Let $S \subseteq V$ be nonempty, and let $\mathbf{v}_0, \ldots, \mathbf{v}_k \in V$.

(a) $\operatorname{aff} S$ is an affine subspace containing S, and it is contained in every affine subspace that contains S.

(b) $\operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k) = \mathbf{v}_0 + \operatorname{span}(\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0)$. In particular, $\dim \operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k) \leq k$.

Proof. STANDARD

Uses. Theorem 4.7.14 (Affine subspaces are the sets closed under affine combinations), Lemma 4.7.13 (Affine combinations do not depend on the origin), Theorem 1.6.10 (Every spanning list contains a basis)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. For (a), show that an affine combination of affine combinations of elements of S is again one. List all the elements of S involved once, padding with zero coefficients.
Hint 2. For (b), use $\mathbf{v}_0$ as base point: Lemma 4.7.13(a) with $\mathbf{p} = \mathbf{v}_0$ rewrites $\sum_i c_i \mathbf{v}_i$.
Sketch. (a) If $\mathbf{w}_j = \sum_i a_{ij}\mathbf{s}_i$ with $\sum_i a_{ij} = 1$ and $\sum_j d_j = 1$, then $\sum_j d_j \mathbf{w}_j = \sum_i \big(\sum_j d_j a_{ij}\big)\mathbf{s}_i$, whose coefficients add up to $\sum_j d_j = 1$. So $\operatorname{aff} S$ is closed under affine combinations, hence affine by Theorem 4.7.14, and every affine subspace containing S is closed under them, hence contains $\operatorname{aff} S$. (b) By Lemma 4.7.13(a), $\sum_i c_i \mathbf{v}_i = \mathbf{v}_0 + \sum_{i \geq 1} c_i(\mathbf{v}_i - \mathbf{v}_0)$, and conversely $\mathbf{v}_0 + \sum_{i \geq 1} a_i(\mathbf{v}_i - \mathbf{v}_0) = \big(1 - \sum_i a_i\big)\mathbf{v}_0 + \sum_{i \geq 1} a_i \mathbf{v}_i$.

(a) Each $\mathbf{s} \in S$ is the affine combination $1\mathbf{s}$, so $S \subseteq \operatorname{aff} S$, and $\operatorname{aff} S$ is nonempty. Let $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \operatorname{aff} S$ and $d_1 + \cdots + d_m = 1$. Each $\mathbf{w}_j$ is an affine combination of finitely many elements of S. Let $\mathbf{s}_1, \ldots, \mathbf{s}_N$ list all elements of S that occur in any of these combinations, and pad with zero coefficients, so that $\mathbf{w}_j = \sum_{i=1}^N a_{ij}\mathbf{s}_i$ with $\sum_i a_{ij} = 1$ for every j. Then

$$\sum_j d_j \mathbf{w}_j = \sum_{i=1}^N \left(\sum_j d_j a_{ij} \right) \mathbf{s}_i, \quad \text{and} \quad \sum_i \sum_j d_j a_{ij} = \sum_j d_j \sum_i a_{ij} = \sum_j d_j = 1.$$

So $\operatorname{aff} S$ is closed under affine combinations, and it is an affine subspace by Theorem 4.7.14. If E is an affine subspace with $S \subseteq E$, then E is closed under affine combinations by the same theorem, so it contains every affine combination of elements of S, that is, $\operatorname{aff} S \subseteq E$.

(b) If $c_0 + \cdots + c_k = 1$, then Lemma 4.7.13(a) with $\mathbf{p} = \mathbf{v}_0$ gives

$$\sum_{i=0}^{k} c_i \mathbf{v}_i = \mathbf{v}_0 + \sum_{i=1}^{k} c_i(\mathbf{v}_i - \mathbf{v}_0),$$

because the term with $i = 0$ is $c_0(\mathbf{v}_0 - \mathbf{v}_0) = \mathbf{0}$. So $\operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k) \subseteq \mathbf{v}_0 + \operatorname{span}(\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0)$. Conversely, for $a_1, \ldots, a_k \in F$,

$$\mathbf{v}_0 + \sum_{i=1}^{k} a_i(\mathbf{v}_i - \mathbf{v}_0) = (1 - a_1 - \cdots - a_k)\mathbf{v}_0 + a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k,$$

an affine combination of $\mathbf{v}_0, \ldots, \mathbf{v}_k$. Hence the two sets are equal. The direction $\operatorname{span}(\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0)$ is spanned by a list of length k, which contains a basis (Theorem 1.6.10), so its dimension is at most k. □

For two points $\mathbf{u} \neq \mathbf{v}$, (b) gives $\operatorname{aff}(\mathbf{u}, \mathbf{v}) = \mathbf{u} + \operatorname{span}(\mathbf{v} - \mathbf{u}) = \{(1 - t)\mathbf{u} + t\mathbf{v} \mid t \in F\}$, the line through them. Over $\mathbb{F}_2$ this line consists of $\mathbf{u}$ and $\mathbf{v}$ alone. Three points of $\mathbb{R}^3$ that do not lie on one line have a plane as their affine hull.

4.7.3 Affine independence and barycentric coordinates

Intuition (From linear to affine independence). By Proposition 4.7.17(b), $k + 1$ points have an affine hull of dimension at most k: two points span at most a line, three at most a plane. The points are in general position when they reach this bound, which by Proposition 4.7.17(b) means that the differences $\mathbf{v}_i - \mathbf{v}_0$ are linearly independent. That description singles out $\mathbf{v}_0$. For a definition that treats all points alike, imitate linear independence, which asks that the only linear relation be the trivial one (Theorem 1.5.5). For points, only relations whose coefficients add up to 0 are meaningful (the intuition at the start of Section 4.7.2), so we test only those.

Definition 4.7.18 (Affine independence). A list $\mathbf{v}_0, \ldots, \mathbf{v}_k$ in V is **affinely independent** if the only scalars $c_0, \ldots, c_k \in F$ with

$$c_0\mathbf{v}_0 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \quad \text{and} \quad c_0 + \cdots + c_k = 0$$

are $c_0 = \cdots = c_k = 0$. Otherwise the list is **affinely dependent**, and a relation of this form with some $c_i \neq 0$ is an **affine dependence**.

Example 4.7.19 (Affine independence).

(a) A single point is affinely independent, since $c_0 = 0$ is forced by the coefficient sum. Two points $\mathbf{u}, \mathbf{v}$ are affinely independent if and only if $\mathbf{u} \neq \mathbf{v}$: a relation with coefficients $c, -c$ says $c(\mathbf{u} - \mathbf{v}) = \mathbf{0}$.

(b) The points $(0, 0), (1, 0), (0, 1)$ of $\mathbb{R}^2$ are affinely independent: $c_0(0, 0) + c_1(1, 0) + c_2(0, 1) = (c_1, c_2)$, so a relation has $c_1 = c_2 = 0$, and then $c_0 = 0$. The list is linearly dependent, since it contains $\mathbf{0}$.

(c) The points $(0, 0), (1, 1), (3, 3)$ are affinely dependent: $-2(0, 0) + 3(1, 1) - (3, 3) = \mathbf{0}$, and $-2 + 3 - 1 = 0$. They lie on one line.

An affine dependence is in particular a linear relation, so a linearly independent list is affinely independent; (b) shows that the converse fails. Unlike linear independence, affine independence does not change when all points are moved by the same vector $\mathbf{w}$: by Lemma 4.7.13(b) with $\mathbf{p} = -\mathbf{w}$, $\sum_i c_i(\mathbf{v}_i + \mathbf{w}) = \sum_i c_i\mathbf{v}_i$ whenever $\sum_i c_i = 0$.

Proposition 4.7.20 (Affine independence through differences). A list $\mathbf{v}_0, \ldots, \mathbf{v}_k$ in V is affinely independent if and only if $\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0$ is linearly independent.

Proof (translate the relations). ROUTINE

Uses. Definition 4.7.18 (Affine independence), Theorem 1.5.5 (Independence is decided at zero), Lemma 4.7.13 (Affine combinations do not depend on the origin)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. A linear relation among the differences, $\sum_{i\geq 1} a_i(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{0}$, expands to a relation among $\mathbf{v}_0, \dots, \mathbf{v}_k$. What are its coefficients, and what is their sum?
Hint 2. Conversely, an affine dependence has $c_0 = -(c_1 + \cdots + c_k)$. Use Lemma 4.7.13(b) with $\mathbf{p} = \mathbf{v}_0$.
Sketch. $\sum_{i\geq 1} a_i(\mathbf{v}_i - \mathbf{v}_0) = \big(-\sum_i a_i\big)\mathbf{v}_0 + \sum_{i\geq 1} a_i\mathbf{v}_i$ has coefficient sum 0. So the linear relations among the differences are the affine relations among the points, read with $c_0 = -\sum_{i\geq 1} c_i$, and trivial ones correspond to trivial ones.

(⇒) Suppose $\mathbf{v}_0, \dots, \mathbf{v}_k$ is affinely independent, and let $\sum_{i=1}^{k} a_i(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{0}$. Expanding,

$$(-a_1 - \cdots - a_k)\mathbf{v}_0 + a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0},$$

and the coefficients add up to 0. By affine independence they are all 0; in particular $a_1 = \cdots = a_k = 0$. So the differences are linearly independent by Theorem 1.5.5.

(⇐) Suppose the differences are linearly independent, and let $\sum_{i=0}^{k} c_i\mathbf{v}_i = \mathbf{0}$ with $\sum_{i=0}^{k} c_i = 0$. By Lemma 4.7.13(b) with $\mathbf{p} = \mathbf{v}_0$, $\sum_{i=0}^{k} c_i(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{0}$, and the term with $i = 0$ is $\mathbf{0}$, so $\sum_{i=1}^{k} c_i(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{0}$. Hence $c_1 = \cdots = c_k = 0$, and $c_0 = -(c_1 + \cdots + c_k) = 0$. □

The first approach singles out $\mathbf{v}_0$. The second treats all points alike, at the price of a larger space: it turns each point into a vector one dimension up, so that affine relations become linear relations. For a vector space V, the set $V \times F$ of pairs $(\mathbf{v}, a)$ with the componentwise operations $(\mathbf{v}, a) + (\mathbf{w}, b) = (\mathbf{v} + \mathbf{w}, a + b)$ and $c(\mathbf{v}, a) = (c\mathbf{v}, ca)$ is a vector space, because each axiom holds in each component separately. (Product spaces are studied in Section 4.6.) For $V = F^n$ it is F^{n+1}, with $(\mathbf{v}, a)$ the vector whose last entry is a.

Technique 4.7.21 (Homogenize: add a coordinate 1). To turn an affine problem into a linear one without choosing a base point, replace each point $\mathbf{v}$ by its lift $\hat{\mathbf{v}} = (\mathbf{v}, 1)$ in $V \times F$, or in F^{n+1} if $V = F^n$. Then

$$\sum_i c_i\hat{\mathbf{v}}_i = \left(\sum_i c_i\mathbf{v}_i, \sum_i c_i\right),$$

so the last coordinate records the coefficient sum. Affine combinations of points are the linear combinations of their lifts that land in the hyperplane "last coordinate 1", and affine dependences are the linear relations among the lifts. Row reduction, rank, and matrices then apply directly.

Reach for it when you want to treat all points symmetrically, or when you want affine maps to become matrix multiplication. The canonical example is the second approach to Proposition 4.7.20. Homogeneous coordinates (Section 4.7.6) apply the same idea to maps.
Used in. Proposition 4.7.20, Theorem 4.7.42, Exercise 4.7.58.

Proof (lift to one dimension higher). STANDARD

Uses. Definition 4.7.18 (Affine independence), Proposition 1.5.18 (Operations that preserve independence), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Homogenize: add a coordinate 1 (Technique 4.7.21)
Hint 1. Compute $\sum_i c_i(\mathbf{v}_i, 1)$ in $V \times F$. What does it say about affine dependences?
Hint 2. The lifts are linearly independent exactly when the points are affinely independent. Now subtract the first lift from the others; Proposition 1.5.18 says that independence does not change.
Sketch. $\sum_i c_i\hat{\mathbf{v}}_i = \big(\sum_i c_i\mathbf{v}_i, \sum_i c_i\big)$, so the points are affinely independent if and only if their lifts are linearly independent, if and only if (by Proposition 1.5.18(c)) $(\mathbf{v}_0, 1), (\mathbf{v}_1 - \mathbf{v}_0, 0), \dots, (\mathbf{v}_k - \mathbf{v}_0, 0)$ is. In a relation among these, the last coordinate forces the first coefficient to be 0, so this list is independent if and only if the differences are.

Let $\hat{\mathbf{v}}_i = (\mathbf{v}_i, 1) \in V \times F$. For $c_0, \dots, c_k \in F$,

$$\sum_{i=0}^{k} c_i\hat{\mathbf{v}}_i = \left(\sum_i c_i\mathbf{v}_i, \sum_i c_i\right),$$

so $\sum_i c_i \hat{\mathbf{v}}_i = \mathbf{0}$ if and only if $\sum_i c_i \mathbf{v}_i = \mathbf{0}$ and $\sum_i c_i = 0$. Thus the affine dependences of $\mathbf{v}_0, \dots, \mathbf{v}_k$ are exactly the nontrivial linear relations of $\hat{\mathbf{v}}_0, \dots, \hat{\mathbf{v}}_k$, and the points are affinely independent if and only if their lifts are linearly independent.

Replacing $\hat{\mathbf{v}}_i$ by $\hat{\mathbf{v}}_i + (-1)\hat{\mathbf{v}}_0$ for $i = 1, \dots, k$, one at a time, does not change linear independence (Proposition 1.5.18(c)). The new list is

$$(\mathbf{v}_0, 1), (\mathbf{u}_1, 0), \dots, (\mathbf{u}_k, 0), \quad \text{where } \mathbf{u}_i = \mathbf{v}_i - \mathbf{v}_0.$$

If $a_0(\mathbf{v}_0, 1) + \sum_{i=1}^{k} a_i(\mathbf{u}_i, 0) = \mathbf{0}$, the last coordinate gives $a_0 = 0$, and then $\sum_i a_i \mathbf{u}_i = \mathbf{0}$. So if $\mathbf{u}_1, \dots, \mathbf{u}_k$ is independent, every such relation is trivial. Conversely, a relation $\sum_i a_i \mathbf{u}_i = \mathbf{0}$ gives the relation $0(\mathbf{v}_0, 1) + \sum_i a_i(\mathbf{u}_i, 0) = \mathbf{0}$, so if the new list is independent, all $a_i = 0$. Hence the new list is independent if and only if the differences are, and chaining the equivalences proves the proposition. □

The second approach also gives a mechanical test. For points of F^n, $\mathbf{v}_0, \dots, \mathbf{v}_k$ is affinely independent if and only if the $(n+1) \times (k+1)$ matrix with columns $\hat{\mathbf{v}}_0, \dots, \hat{\mathbf{v}}_k$ has linearly independent columns, which row reduction decides (Theorem 3.5.12).

Theorem 4.7.22 (Characterizations of affine independence). Let $\mathbf{v}_0, \dots, \mathbf{v}_k \in V$. The following are equivalent.

(I) $\mathbf{v}_0, \dots, \mathbf{v}_k$ is affinely independent.

(II) $\mathbf{v}_1 - \mathbf{v}_0, \dots, \mathbf{v}_k - \mathbf{v}_0$ is linearly independent.

(III) $\dim \operatorname{aff}(\mathbf{v}_0, \dots, \mathbf{v}_k) = k$.

(IV) Every point of $\operatorname{aff}(\mathbf{v}_0, \dots, \mathbf{v}_k)$ is an affine combination of $\mathbf{v}_0, \dots, \mathbf{v}_k$ in exactly one way.

(V) No $\mathbf{v}_j$ lies in the affine hull of the other entries of the list.

Proof. STANDARD

Uses. Proposition 4.7.20 (Affine independence through differences), Proposition 4.7.17 (The affine hull), Proposition 1.6.2 (Bases are independent spanning lists), Corollary 4.1.11 (Spanning lists are long)
Techniques. Proving TFAE (Technique 2.7.1), Uniqueness via two candidates (Technique 1.1.8), Contrapositive (Technique 1.5.4)
Hint 1. (I) ⇔ (II) is Proposition 4.7.20. For (II) ⇔ (III), compute the direction of the affine hull with Proposition 4.7.17(b).
Hint 2. For (I) ⇔ (IV), subtract two representations of the same point. What do the coefficients of the difference add up to?
Hint 3. For (I) ⇔ (V), solve an affine dependence for a $\mathbf{v}_j$ with $c_j \neq 0$, and check that the new coefficients add up to 1.
Sketch. The direction of the hull is spanned by the k differences, so it has dimension k exactly when they are a basis of it, that is, independent. Two representations of one point differ by an affine dependence, and conversely an affine dependence added to $\mathbf{v}_0 = 1\mathbf{v}_0$ gives a second representation of $\mathbf{v}_0$. If $\mathbf{v}_j = \sum_{i \neq j} c_i \mathbf{v}_i$ with $\sum c_i = 1$, moving $\mathbf{v}_j$ across gives an affine dependence; if $\sum c_i \mathbf{v}_i = \mathbf{0}$, $\sum c_i = 0$, and $c_j \neq 0$, then $\mathbf{v}_j = \sum_{i \neq j}(-c_i/c_j)\mathbf{v}_i$, and these coefficients add up to 1.

We prove (I) ⇔ (II), (II) ⇔ (III), (I) ⇔ (IV), and (I) ⇔ (V).

(I) ⇔ (II) is Proposition 4.7.20.

(II) ⇔ (III). By Proposition 4.7.17(b), the direction of $\operatorname{aff}(\mathbf{v}_0, \dots, \mathbf{v}_k)$ is $D = \operatorname{span}(\mathbf{v}_1 - \mathbf{v}_0, \dots, \mathbf{v}_k - \mathbf{v}_0)$, which is spanned by a list of length k. If this list is independent, it is a basis of D (Proposition 1.6.2), so $\dim D = k$. If $\dim D = k$, the spanning list of length k is a basis of D by Corollary 4.1.11(b), hence independent.

(I) ⇒ (IV). Suppose $\mathbf{p} = \sum_i c_i \mathbf{v}_i = \sum_i c_i' \mathbf{v}_i$ with $\sum_i c_i = \sum_i c_i' = 1$. Subtracting, $\sum_i (c_i - c_i')\mathbf{v}_i = \mathbf{0}$ and $\sum_i (c_i - c_i') = 0$. By (I), $c_i = c_i'$ for all i.

(IV) ⇒ (I). Suppose $\sum_i c_i \mathbf{v}_i = \mathbf{0}$ with $\sum_i c_i = 0$. Then

$$\mathbf{v}_0 = 1\mathbf{v}_0 + 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_k = (1 + c_0)\mathbf{v}_0 + c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

are two affine combinations (their coefficients add up to 1 and to $1+0=1$) that represent the point $\mathbf{v}_0$ of the affine hull. By (IV) they have the same coefficients, so $c_0 = \cdots = c_k = 0$.

(I) ⇔ (V). We show that (I) fails if and only if (V) fails. If $\mathbf{v}_j = \sum_{i\neq j} c_i\mathbf{v}_i$ with $\sum_{i\neq j} c_i = 1$, then $\sum_{i\neq j} c_i\mathbf{v}_i + (-1)\mathbf{v}_j = \mathbf{0}$ is an affine dependence: its coefficients add up to $1-1=0$, and the coefficient of $\mathbf{v}_j$ is $-1\neq 0$. Conversely, let $\sum_i c_i\mathbf{v}_i = \mathbf{0}$ with $\sum_i c_i = 0$ and $c_j \neq 0$. Then

$$\mathbf{v}_j = \sum_{i\neq j}\left(-\frac{c_i}{c_j}\right)\mathbf{v}_i, \quad \text{and} \quad \sum_{i\neq j}\left(-\frac{c_i}{c_j}\right) = -\frac{1}{c_j}\sum_{i\neq j} c_i = -\frac{1}{c_j}(-c_j) = 1,$$

so $\mathbf{v}_j$ lies in the affine hull of the other entries. □

Condition (V) is the affine form of the linear dependence lemma (Lemma 1.5.10), and (IV) is the affine form of unique coordinates. For $k=0$, (V) holds because the other entries form an empty list, whose affine hull is empty.

Corollary 4.7.23 (Affinely independent lists are short). If $\dim V = n$, then every affinely independent list in V has at most $n+1$ entries. In particular, any $n+2$ points of V are affinely dependent.

Proof. ROUTINE

Uses. Proposition 4.7.20 (Affine independence through differences), Theorem 4.1.10 (Independent lists are short)
Hint. Pass to the differences.
Sketch. If $\mathbf{v}_0,\dots,\mathbf{v}_k$ is affinely independent, the k differences $\mathbf{v}_i - \mathbf{v}_0$ are linearly independent in an n-dimensional space, so $k \le n$.

If $\mathbf{v}_0,\dots,\mathbf{v}_k$ is affinely independent, the k vectors $\mathbf{v}_i - \mathbf{v}_0$ are linearly independent by Proposition 4.7.20, so $k \le n$ by Theorem 4.1.10(a), and the list has $k+1 \le n+1$ entries. A list of $n+2$ points has $k = n+1 > n$, so it is affinely dependent. □

Exercise 4.7.24. ROUTINE Decide whether each list is affinely independent. If it is not, find an affine dependence.

(a) $(1,2,3),(2,3,4),(4,5,6)$ in $\mathbb{R}^3$.

(b) $(1,0,0),(0,1,0),(0,0,1),(1,1,1)$ in $\mathbb{R}^3$.

(c) The same four points, read in $\mathbb{F}_2^3$.

Solution on page 1353

Definition 4.7.25 (Affine frame, barycentric coordinates). Let E be an affine subspace of V. An **affine frame** of E is an affinely independent list $\mathbf{v}_0,\dots,\mathbf{v}_k$ with $\operatorname{aff}(\mathbf{v}_0,\dots,\mathbf{v}_k) = E$. Given an affine frame, each $\mathbf{p}\in E$ can be written as

$$\mathbf{p} = t_0\mathbf{v}_0 + \cdots + t_k\mathbf{v}_k, \quad t_0 + \cdots + t_k = 1,$$

in exactly one way (Theorem 4.7.22(IV)). The scalars $(t_0,\dots,t_k)$ are the **barycentric coordinates** of $\mathbf{p}$ with respect to $\mathbf{v}_0,\dots,\mathbf{v}_k$.

The name comes from physics: if masses $t_0,\dots,t_k \ge 0$ with total 1 sit at $\mathbf{v}_0,\dots,\mathbf{v}_k$, their center of mass, or barycenter, is $\sum_i t_i\mathbf{v}_i$.

Remark 4.7.26 (Frames exist, and coordinates are coordinates). Let $E = \mathbf{p} + U$ with $\dim U = k$, and let $\mathbf{u}_1,\dots,\mathbf{u}_k$ be a basis of U. Then $\mathbf{p}, \mathbf{p}+\mathbf{u}_1,\dots,\mathbf{p}+\mathbf{u}_k$ is an affine frame of E: its differences are the $\mathbf{u}_i$, which are independent, and its affine hull is $\mathbf{p} + \operatorname{span}(\mathbf{u}_1,\dots,\mathbf{u}_k) = E$ by Proposition 4.7.17(b). Conversely, every affine frame of E has $\dim E + 1$ entries, by Theorem 4.7.22(III).

Barycentric coordinates are ordinary coordinates in disguise. By Lemma 4.7.13(a), $\mathbf{p} = \sum_{i=0}^{k} t_i \mathbf{v}_i$ with $\sum_i t_i = 1$ if and only if

$$\mathbf{p} - \mathbf{v}_0 = t_1(\mathbf{v}_1 - \mathbf{v}_0) + \cdots + t_k(\mathbf{v}_k - \mathbf{v}_0) \quad \text{and} \quad t_0 = 1 - t_1 - \cdots - t_k.$$

So $(t_1, \ldots, t_k) = [\mathbf{p} - \mathbf{v}_0]_{\mathcal{B}}$ for the basis $\mathcal{B} = \mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0$ of $\operatorname{dir} E$, and t_0 is determined by the others.

Example 4.7.27 (Barycentric coordinates in a triangle). Let $\mathbf{v}_0 = (1,1)$, $\mathbf{v}_1 = (5,2)$, $\mathbf{v}_2 = (2,5)$ in $\mathbb{R}^2$. The differences $(4,1)$ and $(1,4)$ are independent, so the three points form an affine frame of $\mathbb{R}^2$, a triangle. For $\mathbf{p} = (3,3)$, Remark 4.7.26 asks for $\mathbf{p} - \mathbf{v}_0 = (2,2) = t_1(4,1) + t_2(1,4)$, that is, $4t_1 + t_2 = 2$ and $t_1 + 4t_2 = 2$. So $t_1 = t_2 = \frac{2}{5}$ and $t_0 = \frac{1}{5}$:

$$(3,3) = \frac{1}{5}(1,1) + \frac{2}{5}(5,2) + \frac{2}{5}(2,5).$$

For $\mathbf{q} = (6,0)$, the equations $4t_1 + t_2 = 5$ and $t_1 + 4t_2 = -1$ give $t_1 = \frac{7}{5}$, $t_2 = -\frac{3}{5}$, and $t_0 = \frac{1}{5}$. Check: $\frac{1}{5}(1,1) + \frac{7}{5}(5,2) - \frac{3}{5}(2,5) = \frac{1}{5}(1 + 35 - 6, 1 + 14 - 15) = (6,0)$.

Homogenizing (Technique 4.7.21) gives the same numbers in one step: (t_0, t_1, t_2) is the unique solution of

$$\begin{pmatrix} 1 & 5 & 2 \\ 1 & 2 & 5 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} t_0 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix},$$

whose columns are the lifts $\hat{\mathbf{v}}_i$.

Intuition (Reading barycentric coordinates off a picture). In a triangle in $\mathbb{R}^2$, the coordinate t_0 vanishes exactly on the line through $\mathbf{v}_1$ and $\mathbf{v}_2$ (the points whose representation does not use $\mathbf{v}_0$), equals 1 at $\mathbf{v}_0$, and changes at a constant rate as $\mathbf{p}$ moves, because by Remark 4.7.26 it is an affine function of $\mathbf{p}$. So t_0 measures how far $\mathbf{p}$ has traveled from the line $\mathbf{v}_1\mathbf{v}_2$ toward $\mathbf{v}_0$, as a fraction of the height of $\mathbf{v}_0$ above that line. It is positive on the side of $\mathbf{v}_0$ and negative on the other side, and the three edge lines cut the plane into seven regions with seven sign patterns (Figure 4.7.1, right). For $\mathbf{p}$ inside the triangle, the triangles $\mathbf{p}\mathbf{v}_1\mathbf{v}_2$ and $\mathbf{v}_0\mathbf{v}_1\mathbf{v}_2$ share the base $\mathbf{v}_1\mathbf{v}_2$, so the height description says that t_0 is the ratio of their areas (left). In Example 4.7.27, $\mathbf{q}$ has $t_2 < 0$, and indeed it lies on the far side of the line $\mathbf{v}_0\mathbf{v}_1$ from $\mathbf{v}_2$.

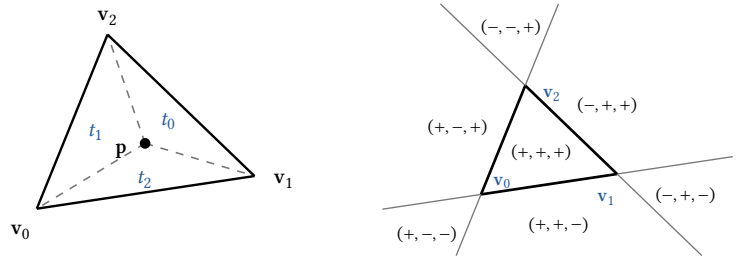
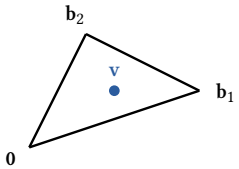


Figure 4.7.1. Barycentric coordinates in a triangle. Left: the point $\mathbf{p}$ with barycentric coordinates $(0.3, 0.4, 0.3)$; for a point inside the triangle, t_i is the fraction of the area taken by the small triangle opposite $\mathbf{v}_i$. Right: the signs of (t_0, t_1, t_2) in the seven regions cut out by the lines through the edges.

When $\mathbf{v}_0 = \mathbf{0}$, Remark 4.7.26 says that (t_1, t_2) are the coordinates of $\mathbf{p}$ in the basis $\mathbf{v}_1, \mathbf{v}_2$, so the change-of-basis component of Section 4.5 also shows barycentric coordinates. Read its two basis vectors as two vertices of a triangle whose third vertex is $\mathbf{0}$.



The triangle with vertices $\mathbf{0}$, $\mathbf{b}_1 = (3,1)$, $\mathbf{b}_2 = (1,2)$ is half of a cell of the $\mathcal{B}$-grid. The point $\mathbf{v} = (1.5,1) = 0.4\mathbf{b}_1 + 0.3\mathbf{b}_2$ has $\mathcal{B}$-coordinates $(t_1,t_2) = (0.4,0.3)$, so its barycentric coordinates with respect to $\mathbf{0},\mathbf{b}_1,\mathbf{b}_2$ are $(0.3,0.4,0.3)$, with $t_0 = 1 - t_1 - t_2$. In the interactive version, drag $\mathbf{v}$: it lies in the triangle exactly when $t_1 \geq 0$, $t_2 \geq 0$, and $t_1 + t_2 \leq 1$. When $\mathbf{b}_1$ and $\mathbf{b}_2$ become parallel, the three points are affinely dependent and the coordinates are undefined. (Interactive in the web version.)

If $(k+1)\cdot 1 \neq 0$ in F, the point whose barycentric coordinates are all $1/(k+1)$ is the centroid $\frac{1}{k+1}(\mathbf{v}_0 + \cdots + \mathbf{v}_k)$. Over $\mathbb{F}_3$ a triangle has no centroid, and Exercise 4.7.57 shows what its medians do instead.

4.7.4 Convexity and simplices

Over $\mathbb{R}$ we can ask for barycentric coordinates to be nonnegative, which describes the solid triangle rather than the plane it spans. Throughout this subsection V is a real vector space; everything works verbatim over $\mathbb{Q}$.

Definition 4.7.28 (Convex combination, convex set, convex hull). Let V be a real vector space. A **convex combination** of $\mathbf{v}_1,\ldots,\mathbf{v}_k$ is an affine combination $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ with $c_i \geq 0$ for all i. A subset $C \subseteq V$ is **convex** if $(1-t)\mathbf{u} + t\mathbf{v} \in C$ for all $\mathbf{u},\mathbf{v} \in C$ and $t \in [0,1]$, that is, if C contains the segment $[\mathbf{u},\mathbf{v}] = \{(1-t)\mathbf{u} + t\mathbf{v} \mid t \in [0,1]\}$ between any two of its points. The **convex hull** $\operatorname{conv} S$ of $S \subseteq V$ is the set of all convex combinations of finitely many elements of S.

Example 4.7.29 (Convex sets). Affine subspaces are convex, by Theorem 4.7.14. So is every half-space $H = \{\mathbf{x} \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n \geq c\}$: writing $\alpha(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n$, if $\alpha(\mathbf{u}) \geq c$ and $\alpha(\mathbf{v}) \geq c$, then $\alpha((1-t)\mathbf{u} + t\mathbf{v}) = (1-t)\alpha(\mathbf{u}) + t\alpha(\mathbf{v}) \geq (1-t)c + tc = c$ for $t \in [0,1]$. An intersection of convex sets is convex, so the solution set of any system of linear inequalities is convex; these are the feasible regions of linear programming. The union of the two coordinate axes in $\mathbb{R}^2$ is not convex, since the midpoint of $(1,0)$ and $(0,1)$ is missing.

Proposition 4.7.30 (Convex sets and convex hulls). Let V be a real vector space, and let $S \subseteq V$.

(a) A subset $C \subseteq V$ is convex if and only if every convex combination of elements of C lies in C.

(b) $\operatorname{conv} S$ is convex, contains S, and is contained in every convex subset of V that contains S.

Proof. STANDARD

Uses. Definition 4.7.28 (Convex combination, convex set, convex hull), Proposition 4.7.17 (The affine hull)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. ($\Leftarrow$) of (a) is the case of two points. For ($\Rightarrow$), induct on the number k of points, splitting off the last one.
Hint 2. If $c_k < 1$, write $\sum_i c_i\mathbf{v}_i = (1-c_k)\mathbf{q} + c_k\mathbf{v}_k$, where $\mathbf{q}$ is a convex combination of $\mathbf{v}_1,\ldots,\mathbf{v}_{k-1}$.
Hint 3. For (b), copy the proof of Proposition 4.7.17(a): over a common list, a convex combination of two convex combinations is a convex combination.
Sketch. Induct on k. For $k=1$ the combination is $\mathbf{v}_1$. If $c_k = 1$, the other coefficients are 0 and the point is $\mathbf{v}_k$. Otherwise $\mathbf{q} = \sum_{i<k}(c_i/(1-c_k))\mathbf{v}_i$ is a convex combination of $k-1$ points, so $\mathbf{q} \in C$, and the point is $(1-c_k)\mathbf{q} + c_k\mathbf{v}_k \in C$. For (b), $(1-t)\sum_i a_i\mathbf{s}_i + t\sum_i b_i\mathbf{s}_i = \sum_i((1-t)a_i + tb_i)\mathbf{s}_i$ is a convex combination.

(a) If C contains every convex combination of its elements, then in particular it contains $(1-t)\mathbf{u} + t\mathbf{v}$ for $\mathbf{u},\mathbf{v} \in C$ and $t \in [0,1]$, so C is convex. Conversely, let C be convex. We show by induction on $k \geq 1$ that every convex combination $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ of elements of C lies in C. For $k = 1$, $c_1 = 1$ and the combination

is $\mathbf{v}_1 \in C$. Let $k \geq 2$, and assume the claim for $k-1$. If $c_k = 1$, then $c_1 + \cdots + c_{k-1} = 0$ with all $c_i \geq 0$, so $c_1 = \cdots = c_{k-1} = 0$ and the combination is $\mathbf{v}_k \in C$. Otherwise $s = 1 - c_k = c_1 + \cdots + c_{k-1} > 0$, and

$$\mathbf{q} = \frac{c_1}{s}\mathbf{v}_1 + \cdots + \frac{c_{k-1}}{s}\mathbf{v}_{k-1}$$

is a convex combination of $k-1$ elements of C (its coefficients are nonnegative and add up to $s/s = 1$), so $\mathbf{q} \in C$ by the induction hypothesis. Then

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = s\mathbf{q} + c_k\mathbf{v}_k = (1 - c_k)\mathbf{q} + c_k\mathbf{v}_k,$$

which lies in C because C is convex and $c_k \in [0, 1]$.

(b) Each $\mathbf{s} \in S$ is the convex combination $1\mathbf{s}$, so $S \subseteq \operatorname{conv} S$. Let $\mathbf{u}, \mathbf{v} \in \operatorname{conv} S$ and $t \in [0, 1]$. As in the proof of Proposition 4.7.17(a), write $\mathbf{u} = \sum_i a_i\mathbf{s}_i$ and $\mathbf{v} = \sum_i b_i\mathbf{s}_i$ over one list $\mathbf{s}_1, \ldots, \mathbf{s}_N$ of elements of S, padding with zero coefficients. Then

$$(1 - t)\mathbf{u} + t\mathbf{v} = \sum_i ((1 - t)a_i + tb_i)\mathbf{s}_i,$$

and the coefficients are nonnegative and add up to $(1 - t) + t = 1$. So $\operatorname{conv} S$ is convex. If C is convex and $S \subseteq C$, then C contains every convex combination of elements of S by (a), that is, $\operatorname{conv} S \subseteq C$. □

Definition 4.7.31 (Simplex). Let $\mathbf{v}_0, \ldots, \mathbf{v}_k$ be affinely independent in a real vector space V. The **k-simplex** with **vertices** $\mathbf{v}_0, \ldots, \mathbf{v}_k$ is

$$\operatorname{conv}(\mathbf{v}_0, \ldots, \mathbf{v}_k) = \{t_0\mathbf{v}_0 + \cdots + t_k\mathbf{v}_k \mid t_i \geq 0 \text{ for all } i,\ t_0 + \cdots + t_k = 1\}.$$

Indexing the coordinates of $\mathbb{R}^{k+1}$ from 0, the **standard k-simplex** is

$$\Delta^k = \left\{(t_0, \ldots, t_k) \in \mathbb{R}^{k+1} \;\middle|\; t_i \geq 0 \text{ for all } i,\ t_0 + \cdots + t_k = 1\right\},$$

the simplex whose vertices are the standard basis vectors of $\mathbb{R}^{k+1}$.

A 0-simplex is a point, a 1-simplex is a segment, a 2-simplex is a solid triangle, and a 3-simplex is a solid tetrahedron. The points of Δ^k are the probability vectors with $k+1$ entries. For instance, a classifier that assigns probabilities to $k+1$ classes outputs a point of Δ^k, and stochastic matrices (Section 11.1) have columns in Δ^{n-1}.

Proposition 4.7.32 (Barycentric coordinates on a simplex). Let σ be the k-simplex with vertices $\mathbf{v}_0, \ldots, \mathbf{v}_k$. The map $\beta : \Delta^k \to \sigma$, $\beta(t_0, \ldots, t_k) = t_0\mathbf{v}_0 + \cdots + t_k\mathbf{v}_k$, is a bijection. Equivalently, a point of $\operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k)$ lies in σ if and only if all of its barycentric coordinates are nonnegative, and β^{-1} sends a point of σ to its barycentric coordinates.

Proof. ROUTINE

Uses. Definition 4.7.31 (Simplex), Theorem 4.7.22 (Characterizations of affine independence), Definition 4.7.25 (Affine frame, barycentric coordinates)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Surjectivity is the definition of σ.
Hint 2. Injectivity is the uniqueness of barycentric coordinates, Theorem 4.7.22(IV).
Sketch. β is onto $\sigma = \operatorname{conv}(\mathbf{v}_0, \ldots, \mathbf{v}_k)$ by definition, and two tuples with the same image are two affine combinations representing one point, so they are equal by Theorem 4.7.22(IV). A point lies in σ exactly when its unique coordinate tuple lies in Δ^k.

By the definition of σ, β is surjective. If $\beta(\mathbf{t}) = \beta(\mathbf{t}')$ for $\mathbf{t}, \mathbf{t}' \in \Delta^k$, then $\sum_i t_i\mathbf{v}_i$ and $\sum_i t_i'\mathbf{v}_i$ are affine combinations representing the same point, so $\mathbf{t} = \mathbf{t}'$ by Theorem 4.7.22(IV). Hence β is a bijection. A point $\mathbf{p} \in \operatorname{aff}(\mathbf{v}_0, \ldots, \mathbf{v}_k)$ has unique barycentric coordinates $\mathbf{t}$, and $\mathbf{p} \in \sigma$ if and only if $\mathbf{p} = \beta(\mathbf{t}')$ for some $\mathbf{t}' \in \Delta^k$. By uniqueness this happens if and only if $\mathbf{t} \in \Delta^k$, that is, if and only if all $t_i \geq 0$. □

In Example 4.7.27, $\mathbf{p} = (3,3)$ lies in the triangle, since its coordinates $\left(\frac{1}{5}, \frac{2}{5}, \frac{2}{5}\right)$ are positive, and $\mathbf{q} = (6,0)$ does not. Renderers test whether a pixel lies in a triangle exactly this way. They also use the coordinates to shade it: if the vertices carry colors $\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^3$ (red, green, and blue intensities), the pixel at $\mathbf{p}$ gets $t_0\mathbf{c}_0 + t_1\mathbf{c}_1 + t_2\mathbf{c}_2$. This is the unique affine map $\mathbb{R}^2 \to \mathbb{R}^3$ with the prescribed values at the vertices (Theorem 4.7.39 below), and the same formula interpolates depths and texture coordinates.

Carathéodory's theorem

A point of $\operatorname{conv} S$ may be given as a convex combination of many points of S. In $\mathbb{R}^2$, though, every point of a convex polygon lies in a triangle formed by three of its vertices. In general a long convex combination can be shortened: when the points it uses are affinely dependent, the dependence gives a way to change the coefficients without changing the point or the coefficient sum, and we change them until one coefficient reaches 0.

Technique 4.7.33 (Slide along a dependence). Suppose $\mathbf{p} = \sum_i c_i\mathbf{v}_i$, where the coefficients are subject to constraints, such as $c_i \geq 0$ and $\sum_i c_i = 1$, and the $\mathbf{v}_i$ satisfy a relation $\sum_i d_i\mathbf{v}_i = \mathbf{0}$ that respects the equality constraints (here $\sum_i d_i = 0$). Then $\mathbf{p} = \sum_i (c_i - \lambda d_i)\mathbf{v}_i$ for every scalar λ, and the equality constraints still hold. Increase λ from 0 until the first inequality constraint becomes tight. At that extremal λ some coefficient is 0, and its vector can be dropped. Repeat as long as a relation exists.

Reach for it to show that a representation can be chosen short, for instance using independent vectors only. It refines the argument of Theorem 1.6.10, where a redundant vector is simply discarded, to situations where the coefficients must keep their signs. The canonical example is Carathéodory's theorem (Theorem 4.7.34).

Used in. Theorem 4.7.34, Exercise 4.7.59.

Theorem 4.7.34 (Carathéodory). Let V be a real vector space with $\dim V = n$, and let $S \subseteq V$. Every point of $\operatorname{conv} S$ is a convex combination of at most $n+1$ elements of S.

Proof. HARD

Uses. Definition 4.7.28 (Convex combination, convex set, convex hull), Corollary 4.7.23 (Affinely independent lists are short), Definition 4.7.18 (Affine independence)
Techniques. Slide along a dependence (Technique 4.7.33), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Take a representation $\mathbf{p} = \sum_{i=1}^{m} c_i\mathbf{v}_i$ with m as small as possible. Then all $c_i > 0$. If $m \geq n+2$, what does Corollary 4.7.23 give?
Hint 2. An affine dependence $\sum_i d_i\mathbf{v}_i = \mathbf{0}$, $\sum_i d_i = 0$, has some $d_i > 0$. Subtract λ times it from the representation.
Hint 3. Take $\lambda = \min\{c_i/d_i \mid d_i > 0\}$.
Sketch. Choose a representation $\mathbf{p} = \sum_{i=1}^{m} c_i\mathbf{v}_i$ with m minimal, so every $c_i > 0$. If $m \geq n+2$, the $\mathbf{v}_i$ are affinely dependent: $\sum_i d_i\mathbf{v}_i = \mathbf{0}$, $\sum_i d_i = 0$, and some $d_i > 0$. With $\lambda = \min_{d_i>0} c_i/d_i$, the coefficients $c_i - \lambda d_i$ are nonnegative, add up to 1, represent $\mathbf{p}$, and one of them is 0, contradicting minimality.

Let $\mathbf{p} \in \operatorname{conv} S$. Among all ways of writing $\mathbf{p}$ as a convex combination of finitely many elements of S, choose one, $\mathbf{p} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$, with m as small as possible. Every c_i is positive, since a term with $c_j = 0$ could be deleted. Suppose, for contradiction, that $m \geq n+2$. By Corollary 4.7.23, the list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is affinely dependent: there are $d_1, \ldots, d_m$, not all 0, with

$$d_1\mathbf{v}_1 + \cdots + d_m\mathbf{v}_m = \mathbf{0}, \quad d_1 + \cdots + d_m = 0.$$

Since the d_i add up to 0 and are not all 0, at least one of them is positive. Let

$$\lambda = \min\left\{\frac{c_i}{d_i} \;\middle|\; d_i > 0\right\},$$

attained at $i = j$. Then $\lambda > 0$. Put $c_i' = c_i - \lambda d_i$. Then

$$\sum_i c_i' \mathbf{v}_i = \sum_i c_i \mathbf{v}_i - \lambda \sum_i d_i \mathbf{v}_i = \mathbf{p}, \quad \sum_i c_i' = \sum_i c_i - \lambda \sum_i d_i = 1.$$

If $d_i \le 0$, then $c_i' \ge c_i > 0$; if $d_i > 0$, then $c_i/d_i \ge \lambda$, so $c_i' = d_i(c_i/d_i - \lambda) \ge 0$. So $\mathbf{p} = \sum_i c_i' \mathbf{v}_i$ is a convex combination, and $c_j' = c_j - (c_j/d_j)d_j = 0$. Deleting the jth term represents $\mathbf{p}$ as a convex combination of $m-1$ elements of S, contradicting the minimality of m. Hence $m \le n+1$. □

The bound $n+1$ cannot be lowered. Let S be the vertex set of an n-simplex. Its centroid has all $n+1$ barycentric coordinates positive, while a convex combination of fewer vertices has a zero coordinate, so by the uniqueness in Proposition 4.7.32 the centroid is not a convex combination of fewer than $n+1$ points of S.

4.7.5 Affine maps

In Section 2.1, the maps $\mathbf{v} \mapsto L(\mathbf{v}) + \mathbf{b}$ were the basic non-examples of linear maps (Non-example 2.1.12), and Proposition 2.1.13 showed that, when $\mathbf{b} \neq \mathbf{0}$, they respect exactly the affine combinations. They are the maps of affine geometry.

Definition 4.7.35 (Affine map). Let V and W be vector spaces over F. A map $f : V \to W$ is **affine** if there are $L \in \mathrm{L}(V, W)$ and $\mathbf{b} \in W$ with $f(\mathbf{v}) = L(\mathbf{v}) + \mathbf{b}$ for all $\mathbf{v} \in V$. We then write $f = L + \mathbf{b}$.

The pair is determined by f: $f(\mathbf{0}) = L(\mathbf{0}) + \mathbf{b} = \mathbf{b}$, and then $L(\mathbf{v}) = f(\mathbf{v}) - f(\mathbf{0})$. We call L the **linear part** of f.

Example 4.7.36 (Affine maps).

(a) Linear maps ($\mathbf{b} = \mathbf{0}$), constant maps ($L = 0$), and translations $\tau_{\mathbf{b}}(\mathbf{v}) = \mathbf{v} + \mathbf{b}$ (with $L = \mathrm{id}_V$) are affine.

(b) The rotation of $\mathbb{R}^2$ by the angle θ about a point $\mathbf{a}$ moves $\mathbf{a}$ to the origin, rotates, and moves back:

$$f(\mathbf{v}) = \mathbf{a} + R_\theta(\mathbf{v} - \mathbf{a}) = R_\theta \mathbf{v} + (\mathbf{a} - R_\theta \mathbf{a}),$$

where R_θ is the rotation about $\mathbf{0}$ (Section 2.1). Its linear part is R_θ wherever the center is. Likewise the scaling by c about $\mathbf{a}$ is $\mathbf{v} \mapsto \mathbf{a} + c(\mathbf{v} - \mathbf{a}) = c\mathbf{v} + (1-c)\mathbf{a}$.

(c) The map $x \mapsto x^2$ on $\mathbb{R}$ is not affine, since it sends $0, 1, 2$ to $0, 1, 4$, and 1 is the midpoint of 0 and 2 while 1 is not the midpoint of 0 and 4. (Affine maps preserve midpoints, by the next theorem.)

Theorem 4.7.37 (Affine maps preserve affine combinations). Let $f : V \to W$ be a map. The following are equivalent.

(I) f is affine.

(II) $f(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = c_1 f(\mathbf{v}_1) + \cdots + c_k f(\mathbf{v}_k)$ for all $k \ge 1$, all $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, and all $c_1, \ldots, c_k \in F$ with $c_1 + \cdots + c_k = 1$.

(III) For every $\mathbf{p} \in V$, the map $g_{\mathbf{p}} : V \to W$, $g_{\mathbf{p}}(\mathbf{v}) = f(\mathbf{p} + \mathbf{v}) - f(\mathbf{p})$, is linear.

(IV) For some $\mathbf{p} \in V$, the map $g_{\mathbf{p}}$ is linear.

In that case $g_{\mathbf{p}}$ is the linear part of f for every $\mathbf{p}$.

Intuition. Condition (II) says that f commutes with weighted averages. Condition (III) says that f becomes linear once positions are measured from $\mathbf{p}$ in V and from $f(\mathbf{p})$ in W. The definition measures from $\mathbf{0}$; the theorem says that $\mathbf{0}$ is not special. To get from (II) to (III), express $\mathbf{p} + \mathbf{u} + \mathbf{v}$ and $\mathbf{p} + c\mathbf{v}$ as affine combinations of $\mathbf{p}$, $\mathbf{p} + \mathbf{u}$, and $\mathbf{p} + \mathbf{v}$.

Proof. STANDARD

Uses. Proposition 2.1.13 (Which combinations an affine map respects), Theorem 2.1.3 (Checking linearity)

Techniques. Proving TFAE (Technique 2.7.1), Choose a base point (Technique 4.7.4)
Hint 1. (I) ⇒ (II) is Proposition 2.1.13.
Hint 2. For (II) ⇒ (III): $\mathbf{p}+\mathbf{u}+\mathbf{v} = (\mathbf{p}+\mathbf{u})+(\mathbf{p}+\mathbf{v})-\mathbf{p}$ and $\mathbf{p}+c\mathbf{v} = c(\mathbf{p}+\mathbf{v})+(1-c)\mathbf{p}$.
Hint 3. For (IV) ⇒ (I), write $f(\mathbf{v}) = f(\mathbf{p}+(\mathbf{v}-\mathbf{p}))$.
Sketch. Applying f to $\mathbf{p}+\mathbf{u}+\mathbf{v} = (\mathbf{p}+\mathbf{u})+(\mathbf{p}+\mathbf{v})-\mathbf{p}$ and to $\mathbf{p}+c\mathbf{v} = c(\mathbf{p}+\mathbf{v})+(1-c)\mathbf{p}$ and subtracting $f(\mathbf{p})$ gives additivity and homogeneity of $g_{\mathbf{p}}$. Conversely, $f(\mathbf{v}) = g_{\mathbf{p}}(\mathbf{v}-\mathbf{p})+f(\mathbf{p}) = g_{\mathbf{p}}(\mathbf{v}) + \big(f(\mathbf{p})-g_{\mathbf{p}}(\mathbf{p})\big)$. If $f = L+\mathbf{b}$, then $g_{\mathbf{p}}(\mathbf{v}) = L(\mathbf{p}+\mathbf{v})-L(\mathbf{p}) = L(\mathbf{v})$.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (IV) ⇒ (I).

(I) ⇒ (II). If $f = L+\mathbf{b}$, then by Proposition 2.1.13, $f\big(\sum_i c_i\mathbf{v}_i\big) - \sum_i c_i f(\mathbf{v}_i) = \big(1-\sum_i c_i\big)\mathbf{b} = \mathbf{0}$ when $\sum_i c_i = 1$.

(II) ⇒ (III). Fix $\mathbf{p}$ and write $g = g_{\mathbf{p}}$. For $\mathbf{u},\mathbf{v}\in V$, the vector $\mathbf{p}+\mathbf{u}+\mathbf{v} = 1(\mathbf{p}+\mathbf{u})+1(\mathbf{p}+\mathbf{v})+(-1)\mathbf{p}$ is an affine combination, as $1+1-1=1$. By (II), $f(\mathbf{p}+\mathbf{u}+\mathbf{v}) = f(\mathbf{p}+\mathbf{u})+f(\mathbf{p}+\mathbf{v})-f(\mathbf{p})$, and subtracting $f(\mathbf{p})$ from both sides gives

$$g(\mathbf{u}+\mathbf{v}) = (f(\mathbf{p}+\mathbf{u})-f(\mathbf{p}))+(f(\mathbf{p}+\mathbf{v})-f(\mathbf{p})) = g(\mathbf{u})+g(\mathbf{v}).$$

For $c\in F$, $\mathbf{p}+c\mathbf{v} = c(\mathbf{p}+\mathbf{v})+(1-c)\mathbf{p}$, so by (II), $f(\mathbf{p}+c\mathbf{v}) = cf(\mathbf{p}+\mathbf{v})+(1-c)f(\mathbf{p})$, and

$$g(c\mathbf{v}) = cf(\mathbf{p}+\mathbf{v})-cf(\mathbf{p}) = cg(\mathbf{v}).$$

So g is linear by Theorem 2.1.3.

(III) ⇒ (IV) holds because V is nonempty.

(IV) ⇒ (I). Let $g = g_{\mathbf{p}}$ be linear. For $\mathbf{v}\in V$,

$$f(\mathbf{v}) = f(\mathbf{p}+(\mathbf{v}-\mathbf{p})) = g(\mathbf{v}-\mathbf{p})+f(\mathbf{p}) = g(\mathbf{v})+(f(\mathbf{p})-g(\mathbf{p})).$$

So $f = g+\mathbf{b}$ with $\mathbf{b} = f(\mathbf{p})-g(\mathbf{p})$, and f is affine.

Finally, if $f = L+\mathbf{b}$, then for every $\mathbf{p}$ and $\mathbf{v}$, $g_{\mathbf{p}}(\mathbf{v}) = (L(\mathbf{p}+\mathbf{v})+\mathbf{b})-(L(\mathbf{p})+\mathbf{b}) = L(\mathbf{v})$, so $g_{\mathbf{p}} = L$. □

Compare Exercise 2.1.39, which assumed only that f respects affine combinations of two points and needed $1+1\neq 0$ to recover additivity. The three-point combination with coefficients $1,1,-1$ makes the division by 2 unnecessary, exactly as in the proof of Theorem 4.7.14.

Proposition 4.7.38 (Affine maps and affine subspaces). Let $f = L+\mathbf{b}: V\to W$ be affine.

(a) If $g = M+\mathbf{c}: W\to X$ is affine, then $g\circ f$ is affine, and $(g\circ f)(\mathbf{v}) = ML(\mathbf{v})+(M(\mathbf{b})+\mathbf{c})$.

(b) f is bijective if and only if L is. In that case f^{-1} is affine, and $f^{-1}(\mathbf{w}) = L^{-1}(\mathbf{w})-L^{-1}(\mathbf{b})$.

(c) If E is an affine subspace of V, then $f(E)$ is an affine subspace of W with direction $\operatorname{im}(L\restriction_{\operatorname{dir}E})$. If E is finite-dimensional, then $\dim f(E)\le\dim E$.

(d) If E' is an affine subspace of W, then $f^{-1}(E')$ is empty or an affine subspace of V with direction $\{\mathbf{u}\in V \mid L(\mathbf{u})\in\operatorname{dir}E'\}$.

Proof. STANDARD

Uses. Definition 4.7.35 (Affine map), Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 2.7.7 (Bijective linear maps are invertible), Proposition 4.7.3 (Direction and base points), Proposition 2.6.8 (Kernels and images are subspaces), Theorem 4.2.11 (Rank–nullity), Theorem 1.4.3 (Subspace criterion)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. (a) is a computation. For (b), write $f = \tau_{\mathbf{b}}\circ L$, where $\tau_{\mathbf{b}}$ is the translation by $\mathbf{b}$, a bijection.
Hint 2. For (c) and (d), choose base points: $\mathbf{p}\in E$ for (c), and a point $\mathbf{q}\in f^{-1}(E')$, if there is one, for (d). The key identity is $f(\mathbf{p}+\mathbf{u}) = f(\mathbf{p})+L(\mathbf{u})$.
Sketch. (a) $M(L(\mathbf{v})+\mathbf{b})+\mathbf{c}$. (b) $\tau_{\mathbf{b}}$ is a bijection with inverse $\tau_{-\mathbf{b}}$, so $f = \tau_{\mathbf{b}}\circ L$ is bijective if and only if L is, and then $f^{-1} = L^{-1}\circ\tau_{-\mathbf{b}}$. (c) $f(\mathbf{p}+U) = f(\mathbf{p})+L(U)$, and $\dim L(U)\le\dim U$ by rank–nullity. (d) With $\mathbf{q}\in f^{-1}(E')$, $f(\mathbf{v})\in E'$ if and only if $L(\mathbf{v}-\mathbf{q}) = f(\mathbf{v})-f(\mathbf{q})\in\operatorname{dir}E'$.

(a) $g(f(\mathbf{v})) = M(L(\mathbf{v})+\mathbf{b})+\mathbf{c} = ML(\mathbf{v})+(M(\mathbf{b})+\mathbf{c})$, and ML is linear by Proposition 2.1.17.

(b) The translation $\tau_{\mathbf{b}} : W \to W, \mathbf{w} \mapsto \mathbf{w} + \mathbf{b}$, is a bijection with inverse $\tau_{-\mathbf{b}}$, and $f = \tau_{\mathbf{b}} \circ L$. If L is bijective, so is f; if f is bijective, so is $L = \tau_{-\mathbf{b}} \circ f$. In that case $f^{-1} = L^{-1} \circ \tau_{-\mathbf{b}}$, that is, $f^{-1}(\mathbf{w}) = L^{-1}(\mathbf{w} - \mathbf{b}) = L^{-1}(\mathbf{w}) - L^{-1}(\mathbf{b})$, and L^{-1} is linear by Proposition 2.7.7. So f^{-1} is affine.

(c) Let $E = \mathbf{p} + U$. For $\mathbf{u} \in U$, $f(\mathbf{p} + \mathbf{u}) = L(\mathbf{p}) + L(\mathbf{u}) + \mathbf{b} = f(\mathbf{p}) + L(\mathbf{u})$. Hence $f(E) = f(\mathbf{p}) + L(U)$, where $L(U) = \operatorname{im}(L|_U)$ is a subspace of W (Proposition 2.6.8, applied to $L|_U$). So $f(E)$ is an affine subspace with direction $\operatorname{im}(L|_U)$. If U is finite-dimensional, rank–nullity for $L|_U$ (Theorem 4.2.11) gives $\dim \operatorname{im}(L|_U) \le \dim U$.

(d) Let $U' = \operatorname{dir} E'$ and $U'' = \{\mathbf{u} \in V \mid L(\mathbf{u}) \in U'\}$. Then U'' is a subspace: $L(\mathbf{0}) = \mathbf{0} \in U'$, and if $L(\mathbf{u}), L(\mathbf{u}') \in U'$ and $c \in F$, then $L(\mathbf{u} + \mathbf{u}') = L(\mathbf{u}) + L(\mathbf{u}') \in U'$ and $L(c\mathbf{u}) = cL(\mathbf{u}) \in U'$. Suppose $f^{-1}(E')$ is nonempty, and let $\mathbf{q} \in f^{-1}(E')$. By Proposition 4.7.3(b), $E' = f(\mathbf{q}) + U'$, so for $\mathbf{v} \in V$ we have $f(\mathbf{v}) \in E'$ if and only if $f(\mathbf{v}) - f(\mathbf{q}) \in U'$. Since $f(\mathbf{v}) - f(\mathbf{q}) = L(\mathbf{v} - \mathbf{q})$, this holds if and only if $\mathbf{v} - \mathbf{q} \in U''$. Hence $f^{-1}(E') = \mathbf{q} + U''$. □

Over $\mathbb{R}$, condition (II) of Theorem 4.7.37 applied to convex combinations shows that an affine map sends segments onto segments, convex sets to convex sets, and $\operatorname{conv} S$ onto $\operatorname{conv} f(S)$. It also preserves ratios along a line: if $\mathbf{p} = (1-t)\mathbf{a} + t\mathbf{b}$, then $f(\mathbf{p}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b})$, so midpoints go to midpoints. Lengths and angles are not preserved (think of a shear), which is why affine geometry has parallel lines and ratios along a line but no distances.

Theorem 4.7.39 (Affine maps are determined by a frame). Let $\mathbf{v}_0, \dots, \mathbf{v}_n$ be an affine frame of V, and let $\mathbf{w}_0, \dots, \mathbf{w}_n \in W$ be arbitrary. There is exactly one affine map $f : V \to W$ with $f(\mathbf{v}_i) = \mathbf{w}_i$ for $i = 0, \dots, n$, namely

$$f(t_0\mathbf{v}_0 + \cdots + t_n\mathbf{v}_n) = t_0\mathbf{w}_0 + \cdots + t_n\mathbf{w}_n \quad \text{whenever } t_0 + \cdots + t_n = 1.$$

Proof. STANDARD

Uses. Theorem 4.7.37 (Affine maps preserve affine combinations), Definition 4.7.25 (Affine frame, barycentric coordinates), Proposition 4.7.17 (The affine hull), Theorem 4.7.22 (Characterizations of affine independence), Theorem 2.2.1 (Linear maps are determined by a basis)
Techniques. Choose a base point (Technique 4.7.4), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Uniqueness: every point is an affine combination of the $\mathbf{v}_i$, and affine maps preserve affine combinations.
Hint 2. Existence: use $\mathbf{v}_0$ as base point. The differences $\mathbf{v}_i - \mathbf{v}_0$ form a basis of V. Which linear map should be the linear part?
Sketch. Every point has barycentric coordinates, and Theorem 4.7.37(II) forces the formula, which gives uniqueness. For existence, the differences $\mathbf{v}_i - \mathbf{v}_0$ are independent and span V, so there is a linear L with $L(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{w}_i - \mathbf{w}_0$; then $f(\mathbf{v}) = \mathbf{w}_0 + L(\mathbf{v} - \mathbf{v}_0)$ is affine and sends $\mathbf{v}_i$ to $\mathbf{w}_i$.

Uniqueness and the formula. Let f be affine with $f(\mathbf{v}_i) = \mathbf{w}_i$ for all i. Every $\mathbf{v} \in V = \operatorname{aff}(\mathbf{v}_0, \dots, \mathbf{v}_n)$ has barycentric coordinates, $\mathbf{v} = \sum_i t_i\mathbf{v}_i$ with $\sum_i t_i = 1$, and by Theorem 4.7.37(II), $f(\mathbf{v}) = \sum_i t_i f(\mathbf{v}_i) = \sum_i t_i\mathbf{w}_i$. So f is determined by the $\mathbf{w}_i$ and is given by the formula.

Existence. By Proposition 4.7.17(b), $V = \mathbf{v}_0 + D$ with $D = \operatorname{span}(\mathbf{v}_1 - \mathbf{v}_0, \dots, \mathbf{v}_n - \mathbf{v}_0)$. For every $\mathbf{x} \in V$, $\mathbf{v}_0 + \mathbf{x} \in V$, so $\mathbf{x} \in D$; hence $D = V$. By Theorem 4.7.22(II), the differences are linearly independent, so they form a basis of V. By Theorem 2.2.1 there is a linear map $L : V \to W$ with $L(\mathbf{v}_i - \mathbf{v}_0) = \mathbf{w}_i - \mathbf{w}_0$ for $i = 1, \dots, n$. Define

$$f(\mathbf{v}) = \mathbf{w}_0 + L(\mathbf{v} - \mathbf{v}_0) = L(\mathbf{v}) + (\mathbf{w}_0 - L(\mathbf{v}_0)),$$

an affine map. Then $f(\mathbf{v}_0) = \mathbf{w}_0$, and $f(\mathbf{v}_i) = \mathbf{w}_0 + (\mathbf{w}_i - \mathbf{w}_0) = \mathbf{w}_i$ for $i \ge 1$. □

Example 4.7.40 (Mapping one triangle onto another). The affine map $f : \mathbb{R}^2 \to \mathbb{R}^2$ with $f(0,0) = (1,1)$, $f(1,0) = (5,2)$, $f(0,1) = (2,5)$ has translation part $f(\mathbf{0}) = (1,1)$ and linear part L with $L(\mathbf{e}_1) = (4,1)$ and $L(\mathbf{e}_2) = (1,4)$:

$$f(\mathbf{x}) = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

By Theorem 4.7.39, f sends the point with barycentric coordinates (t_0, t_1, t_2) with respect to $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2$, which is (t_1, t_2), to the point with the same barycentric coordinates with respect to the triangle of Example 4.7.27. For instance $f\left(\frac{2}{5}, \frac{2}{5}\right) = \left(\frac{8}{5} + \frac{2}{5} + 1, \frac{2}{5} + \frac{8}{5} + 1\right) = (3, 3)$, matching the coordinates $\left(\frac{1}{5}, \frac{2}{5}, \frac{2}{5}\right)$ of $(3, 3)$ found there. Texture mapping works this way: the renderer finds the affine map from a triangle in the texture image to a triangle on the screen.

4.7.6 Homogeneous coordinates

A graphics program moves objects by rotations, scalings, and translations, composed in long chains, and applies the result to every vertex of the scene. Composites of linear maps are products of matrices, which can be computed once and reused. A translation is not linear, so a chain such as "rotate about a point, then translate" is not a single matrix in $F^{n\times n}$. Homogenizing (Technique 4.7.21) repairs this.

Definition 4.7.41 (Homogeneous coordinates). For $\mathbf{x} \in F^n$, let $\hat{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \in F^{n+1}$. A **vector of homogeneous coordinates** of $\mathbf{x}$ is any vector $c\hat{\mathbf{x}} = (c\mathbf{x}, c)$ with $c \in F$, $c \neq 0$. Conversely, a vector $(\mathbf{z}, c) \in F^{n+1}$ with $\mathbf{z} \in F^n$ and $c \neq 0$ is a vector of homogeneous coordinates of exactly one point of F^n, namely $c^{-1}\mathbf{z}$.

The last claim holds because $(\mathbf{z}, c) = c'(\mathbf{x}, 1)$ forces $c' = c$ and $\mathbf{x} = c^{-1}\mathbf{z}$. For example, $(6, -3, 9, 3)$ and $(2, -1, 3, 1)$ are both vectors of homogeneous coordinates of $(2, -1, 3) \in \mathbb{R}^3$. The vectors of homogeneous coordinates of $\mathbf{x}$ are the nonzero vectors on the line $\operatorname{span}(\hat{\mathbf{x}})$ through $\mathbf{0}$ in F^{n+1}, so $\mathbf{x} \mapsto \operatorname{span}(\hat{\mathbf{x}})$ matches the points of F^n with the lines through $\mathbf{0}$ in F^{n+1} that are not contained in the hyperplane $x_{n+1} = 0$. Vectors with last coordinate 0 represent no point of F^n; the aside at the end of the section says what they represent instead.

Theorem 4.7.42 (Affine maps as matrices). Let $f : F^n \to F^m$ be affine, $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ with $A \in F^{m\times n}$ and $\mathbf{b} \in F^m$, and let

$$M_f = \begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix} \in F^{(m+1)\times(n+1)},$$

where 0 is the $1 \times n$ zero row.

(a) $M_f\hat{\mathbf{x}} = \widehat{f(\mathbf{x})}$ for every $\mathbf{x} \in F^n$, and M_f is the only matrix in $F^{(m+1)\times(n+1)}$ with this property.

(b) $f \mapsto M_f$ is a bijection from the set of affine maps $F^n \to F^m$ onto the set of matrices in $F^{(m+1)\times(n+1)}$ whose last row is $(0, \ldots, 0, 1)$.

(c) If $g : F^m \to F^l$ is affine, then $M_{g\circ f} = M_g M_f$.

(d) If $m = n$, then f is bijective if and only if M_f is invertible, and in that case $M_{f^{-1}} = M_f^{-1}$.

Proof. STANDARD

Uses. Theorem 2.4.34 (Block multiplication), Proposition 4.7.38 (Affine maps and affine subspaces), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Corollary 2.2.6 (Agreeing on a spanning list), Theorem 2.3.20 (Linear maps correspond to matrices), Corollary 3.5.25 (The invertible matrix theorem)
Techniques. Homogenize: add a coordinate 1 (Technique 4.7.21), Block decomposition (Technique 2.4.35)
Hint 1. Multiply in blocks (Theorem 2.4.34), splitting $n + 1$ as n and 1, and $m + 1$ as m and 1.
Hint 2. For uniqueness in (a), the vectors $\hat{\mathbf{x}}$ span F^{n+1}: look at $\hat{\mathbf{0}}$ and the $\hat{\mathbf{e}}_j$.
Hint 3. (c) is block multiplication together with Proposition 4.7.38(a). For (d), apply (c) to f and f^{-1}; for the converse, apply M_f to $(\mathbf{a}, 0)$ with $A\mathbf{a} = \mathbf{0}$.

Sketch. Block multiplication gives $M_f\hat{\mathbf{x}} = (A\mathbf{x}+\mathbf{b}, 1)$. Since $\hat{\mathbf{0}} = \mathbf{e}_{n+1}$ and $\hat{\mathbf{e}}_j - \hat{\mathbf{0}} = \mathbf{e}_j$, the $\hat{\mathbf{x}}$ span F^{n+1}, so a matrix is determined by its products with them. $M_gM_f = \left(\begin{smallmatrix} BA & B\mathbf{b}+\mathbf{c} \\ 0 & 1 \end{smallmatrix}\right) = M_{g\circ f}$. If f is bijective, $M_{f^{-1}}M_f = M_{\mathrm{id}} = I_{n+1}$ and likewise in the other order. If M_f is invertible and $A\mathbf{a} = \mathbf{0}$, then $M_f(\mathbf{a}, 0) = \mathbf{0}$, so $\mathbf{a} = \mathbf{0}$ and A is invertible.

(a) By Theorem 2.4.34, with blocks of sizes $m, 1$ and $n, 1$,

$$M_f\hat{\mathbf{x}} = \begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} = \begin{pmatrix} A\mathbf{x}+\mathbf{b} \\ 0\mathbf{x}+1 \end{pmatrix} = \begin{pmatrix} f(\mathbf{x}) \\ 1 \end{pmatrix} = \widehat{f(\mathbf{x})}.$$

For uniqueness, note that $\hat{\mathbf{0}} = \mathbf{e}_{n+1}$ and $\hat{\mathbf{e}}_j - \hat{\mathbf{0}} = \mathbf{e}_j$ for $j \le n$, so the vectors $\hat{\mathbf{x}}$ span F^{n+1}. If $N\hat{\mathbf{x}} = M_f\hat{\mathbf{x}}$ for all $\mathbf{x}$, then the linear maps T_N and T_{M_f} agree on a spanning list, so they are equal (Corollary 2.2.6), and $N = M_f$ (Theorem 2.3.20).

(b) By Proposition 2.3.17, the linear part of an affine map $F^n \to F^m$ is T_A for exactly one $A \in F^{m\times n}$, so each affine map is $\mathbf{x} \mapsto A\mathbf{x}+\mathbf{b}$ for exactly one pair $(A, \mathbf{b})$, since the linear part and $\mathbf{b} = f(\mathbf{0})$ are determined by f. The matrices with last row $(0, \ldots, 0, 1)$ are exactly the $\left(\begin{smallmatrix} A & \mathbf{b} \\ 0 & 1 \end{smallmatrix}\right)$ with $A \in F^{m\times n}$ and $\mathbf{b} \in F^m$, and each of them determines $(A, \mathbf{b})$. So $f \mapsto M_f$ is a bijection.

(c) Let $g(\mathbf{y}) = B\mathbf{y}+\mathbf{c}$. By Theorem 2.4.34,

$$M_gM_f = \begin{pmatrix} B & \mathbf{c} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} BA & B\mathbf{b}+\mathbf{c} \\ 0 & 1 \end{pmatrix},$$

and by Proposition 4.7.38(a), $g \circ f$ is $\mathbf{x} \mapsto BA\mathbf{x} + (B\mathbf{b}+\mathbf{c})$. So $M_gM_f = M_{g\circ f}$.

(d) Note that $M_{\mathrm{id}} = \left(\begin{smallmatrix} I_n & \mathbf{0} \\ 0 & 1 \end{smallmatrix}\right) = I_{n+1}$. If f is bijective, then f^{-1} is affine (Proposition 4.7.38(b)), and by (c), $M_{f^{-1}}M_f = M_{f^{-1}\circ f} = I_{n+1}$ and $M_fM_{f^{-1}} = I_{n+1}$. So M_f is invertible with inverse $M_{f^{-1}}$. Conversely, suppose M_f is invertible, and let $A\mathbf{a} = \mathbf{0}$ with $\mathbf{a} \in F^n$. Then $M_f\left(\begin{smallmatrix} \mathbf{a} \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} A\mathbf{a} \\ 0 \end{smallmatrix}\right) = \mathbf{0}$, so $(\mathbf{a}, 0) = \mathbf{0}$ and $\mathbf{a} = \mathbf{0}$. Hence $\ker T_A = \{\mathbf{0}\}$, A is invertible by Corollary 3.5.25, and f is bijective by Proposition 4.7.38(b). □

The translation by $\mathbf{a}$ has matrix $\left(\begin{smallmatrix} I_n & \mathbf{a} \\ 0 & 1 \end{smallmatrix}\right)$. On F^{n+1} it is a linear map, a shear that slides the hyperplane $x_{n+1} = 1$ along itself by $\mathbf{a}$ and fixes the hyperplane $x_{n+1} = 0$. For $n = 1$: the map $(x, y) \mapsto (x + ay, y)$ of F^2 is linear and moves the line $y = 1$, a copy of F, by a. So translations, which are not linear on F^n, become linear one dimension up.

Exercise 4.7.43. ROUTINE Let $A \in F^{n\times n}$ be invertible and $\mathbf{b} \in F^n$. Verify by block multiplication that

$$\begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}\mathbf{b} \\ 0 & 1 \end{pmatrix}.$$

Use it to invert the matrix of the translation by $\mathbf{a}$, and the matrix $\left(\begin{smallmatrix} 0 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{smallmatrix}\right)$.

Solution on page 1353

Example 4.7.44 (Rotation about a point). Let f be the rotation of $\mathbb{R}^2$ by $90°$ counterclockwise about $\mathbf{a} = (1, 2)$, and $R = R_{\pi/2} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. By Example 4.7.36, $f = \tau_{\mathbf{a}} \circ R \circ \tau_{-\mathbf{a}}$, so by Theorem 4.7.42(c),

$$M_f = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The last column is $\mathbf{a} - R\mathbf{a} = (1, 2) - (-2, 1) = (3, 1)$. As a check, $M_f\widehat{(1, 2)} = (0 - 2 + 3, 1 + 0 + 1, 1)$, so the center is fixed, and $M_f\widehat{(2, 2)} = (1, 3, 1)$: the point one unit to the right of $\mathbf{a}$ goes to the point one unit above it.

The order of the factors matters. The product

$$\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is the matrix of "translate by $\mathbf{a}$, then rotate about $\mathbf{0}$", which is $\mathbf{x} \mapsto R\mathbf{x} + R\mathbf{a}$, a rotation about a different center.

In three dimensions the same matrices are 4×4. A renderer stores each object once, in coordinates convenient for that object, and places it in the scene by a product $M = M_k \cdots M_1$ of rotations, scalings, and translations, computed once per object. Each vertex then costs one 4×4 matrix–vector product, which is why graphics processors are built around 4×4 matrix arithmetic. In the product, M_1 acts first.

Perspective projection

The projection $(x, y, z) \mapsto (x, y, 0)$ onto the xy-plane is linear, but it draws distant objects as large as near ones. A camera, or an eye, projects through a point instead. Fix $d \neq 0$, the **center** $\mathbf{c} = (0, 0, d)$, and the image plane $z = 0$. The **perspective projection** π sends $\mathbf{v} = (x, y, z)$ to the point where the line through $\mathbf{c}$ and $\mathbf{v}$ meets the image plane (Figure 4.7.2).

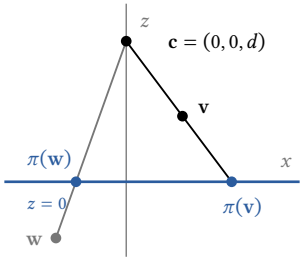


Figure 4.7.2. The perspective projection seen from the side, with the image plane $z = 0$ drawn as the horizontal line. The point $\mathbf{v}$ between the plane and the center is projected outward; the point $\mathbf{w}$ behind the plane is projected inward, so distant points come closer together.

The points of the line through $\mathbf{c}$ and $\mathbf{v}$ are $\mathbf{c} + s(\mathbf{v} - \mathbf{c})$ for $s \in \mathbb{R}$. The third coordinate of such a point is $d + s(z - d)$, which is 0 if and only if $s(d - z) = d$. If $z = d$, there is no such s: the line is parallel to the image plane, or $\mathbf{v} = \mathbf{c}$, and $\mathbf{v}$ has no image. Otherwise $s = d/(d - z)$, and the first two coordinates are sx and sy, so

$$\pi(x, y, z) = \left(\frac{x}{1 - z/d}, \frac{y}{1 - z/d}, 0 \right) \quad (z \neq d).$$

The same formula covers points on the far side of the image plane from the center, and points on the far side of the center from the image plane, where $s < 0$.

The map π is not affine. With $d = 1$, $\pi(0, 0, 0) = \mathbf{0}$ and $\pi\left(1, 0, \frac{1}{2}\right) = (2, 0, 0)$, but the midpoint $\left(\frac{1}{2}, 0, \frac{1}{4}\right)$ goes to $\left(\frac{2}{3}, 0, 0\right)$, not to the midpoint $(1, 0, 0)$ of the images. So by Theorem 4.7.37 no matrix M_f represents π. Even so:

Proposition 4.7.45 (Perspective projection in homogeneous coordinates). Let $d \neq 0$ and

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{pmatrix}.$$

For every $\mathbf{v} = (x, y, z) \in \mathbb{R}^3$ with $z \neq d$, $P\hat{\mathbf{v}} = (x, y, 0, 1 - z/d)$, and this is a vector of homogeneous coordinates of $\pi(\mathbf{v})$.

Proof. ROUTINE

Uses. Definition 2.3.8 (Matrix–vector product), Definition 4.7.41 (Homogeneous coordinates)
Hint 1. Compute $P\hat{\mathbf{v}}$ row by row.
Hint 2. Which point does a vector $(\mathbf{z}, c)$ with $c \neq 0$ represent?
Sketch. The rows of P give x, y, 0, and $-z/d + 1$. The last entry is nonzero because $z \neq d$, and dividing by it gives $\pi(\mathbf{v})$.

Multiplying the rows of P by $\hat{\mathbf{v}} = (x, y, z, 1)$ gives x, y, 0, and $-z/d + 1$. Since $z \neq d$, the last entry $c = 1 - z/d$ is nonzero, and by Definition 4.7.41 the vector represents the point $c^{-1}(x, y, 0) = \pi(\mathbf{v})$. □

The nonlinearity has moved into the freedom of Definition 4.7.41: $P\hat{\mathbf{v}}$ does not have last coordinate 1, and the division by the last coordinate, performed once at the very end, supplies it. The matrix P has rank 3, so the depth z is lost. Renderers therefore use the third row $(0, 0, 1, 0)$ instead of $(0, 0, 0, 0)$, which keeps the number $z/(1 - z/d) = dz/(d - z)$ as a third coordinate. Its derivative in z is $d^2/(d - z)^2 > 0$, so on each side of the plane $z = d$ it increases with z, and it still tells which of two points that project to the same pixel is nearer to the camera.

For a center off the z-axis, or another image plane, conjugate by translations and rotations as in Example 4.7.44; Exercise 4.7.49 carries this out. Under π, parallel lines in space usually stop being parallel, as the rails of a track do in a photograph, and Exercise 4.7.50 finds where their images meet.

ASIDE: *Projective space*

By Definition 4.7.41, a point $\mathbf{x}$ of F^n is the same thing as a line through $\mathbf{0}$ in F^{n+1}, namely $\operatorname{span}(\hat{\mathbf{x}})$, and every line through $\mathbf{0}$ that is not contained in the hyperplane $H = \{x_{n+1} = 0\}$ arises this way exactly once. The **projective space** $\mathbb{P}^n(F)$ is the set of all lines through $\mathbf{0}$ in F^{n+1}, that is, of all 1-dimensional subspaces. Its elements are the points of F^n together with the lines inside H, called **points at infinity**. A point at infinity $\operatorname{span}((\mathbf{w}, 0))$, with $\mathbf{w} \neq \mathbf{0}$, corresponds to a direction in F^n. Over $\mathbb{R}$, the points $\mathbf{x} + s\mathbf{w}$ of a line have the vectors of homogeneous coordinates $(\mathbf{x} + s\mathbf{w}, 1)$ and $\left(s^{-1}\mathbf{x} + \mathbf{w}, s^{-1}\right)$ for $s \neq 0$, and the latter tend to $(\mathbf{w}, 0)$ as $s \to \infty$. So all lines with direction $\operatorname{span}(\mathbf{w})$ share one point at infinity, and in Exercise 4.7.50 the vanishing point of a family of parallel lines is the image of their common point at infinity under P.

A **projective line** is a 2-dimensional subspace of F^{n+1}, viewed as the set of lines through $\mathbf{0}$ that it contains. In the projective plane $\mathbb{P}^2(F)$, two distinct projective lines $U \neq W$ meet in exactly one point: $W \not\subseteq U$ by Proposition 4.1.21(b), so $U + W$ is larger than U and equals F^3, and Theorem 4.4.19 gives $\dim(U \cap W) = 2 + 2 - 3 = 1$. In particular, two parallel lines of F^2, completed by their common point at infinity, meet there. An invertible matrix in $F^{(n+1)\times(n+1)}$ permutes the 1-dimensional subspaces of F^{n+1}, so it acts on $\mathbb{P}^n(F)$, and nonzero multiples of a matrix act in the same way. These are the **projective transformations**. They send projective lines to projective lines, but unlike affine maps they can move points at infinity into F^n, which is how a photograph can show a horizon. Over a field with q elements, $\mathbb{P}^2$ has $\left(q^3 - 1\right)/(q - 1) = q^2 + q + 1$ points; for $q = 2$ it is the seven-point Fano plane.

Exercises

Exercise 4.7.46. ROUTINE Decide which of the following sets are affine subspaces. For each one that is, find its direction and, if it is finite, its dimension.

(a) The solution set in $\mathbb{R}^3$ of $x_1 + x_2 + x_3 = 1$ and $x_1 - x_3 = 2$.

(b) The set of $p \in \mathbb{R}[x]$ with $p(0) = 1$ and $p(1) = 3$.

(c) The set of $\mathbf{x} \in \mathbb{R}^2$ with $x_1 x_2 = 0$.

(d) The set of $(x_1, x_2) \in \mathbb{F}_3^2$ with $x_2 = x_1^2$.

Solution on page 1354

Exercise 4.7.47. ROUTINE Let $\mathbf{v}_0 = (0,0)$, $\mathbf{v}_1 = (4,0)$, $\mathbf{v}_2 = (1,3)$ in $\mathbb{R}^2$. Find the barycentric coordinates of $\mathbf{p} = (2,1)$, $\mathbf{q} = (3,2)$, and $\mathbf{r} = (-1,1)$ with respect to $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2$, and decide which of the three points lie in the triangle $\operatorname{conv}(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$. For each point outside, name the edge line it lies beyond.

Solution on page 1354

Exercise 4.7.48. ROUTINE Work in $\mathbb{R}^2$ with homogeneous coordinates.

(a) Write the 3×3 matrices of the translation T by $(2,-1)$, the rotation R by $90°$ counterclockwise about $\mathbf{0}$, and the scaling S by the factor 3 about $(1,1)$.

(b) Compute $M_T M_R$ and $M_R M_T$, and describe each product as a map of $\mathbb{R}^2$.

(c) Which point of $\mathbb{R}^2$ has $(6,-3,3)$ as a vector of homogeneous coordinates? Give a vector of homogeneous coordinates of the same point with last entry 2.

Solution on page 1355

Exercise 4.7.49. STANDARD Let $d \neq 0$ and $c_1, c_2 \in \mathbb{R}$, and let $\pi_{\mathbf{c}}$ be the perspective projection onto the plane $z = 0$ with center $\mathbf{c} = (c_1, c_2, d)$.

(a) Show that for $\mathbf{v} = (x, y, z)$ with $z \neq d$,

$$\pi_{\mathbf{c}}(\mathbf{v}) = \left(\frac{dx - c_1 z}{d - z}, \frac{dy - c_2 z}{d - z}, 0 \right).$$

(b) Let $\mathbf{a} = (c_1, c_2, 0)$ and let π be the projection with center $(0,0,d)$. Show that $\pi_{\mathbf{c}} = \tau_{\mathbf{a}} \circ \pi \circ \tau_{-\mathbf{a}}$.

(c) Compute the 4×4 matrix $M_{\tau_{\mathbf{a}}} P M_{\tau_{-\mathbf{a}}}$, with P as in Proposition 4.7.45, and check that it sends $\hat{\mathbf{v}}$ to a vector of homogeneous coordinates of $\pi_{\mathbf{c}}(\mathbf{v})$.

Solution on page 1355

Exercise 4.7.50. STANDARD Let $d \neq 0$, and let π and P be as in Proposition 4.7.45. Let $\mathbf{v} = (v_1, v_2, v_3)$ with $v_3 \neq d$, and let $\mathbf{w} = (w_1, w_2, w_3) \neq \mathbf{0}$.

(a) Suppose $w_3 \neq 0$, and let $\mathbf{q}_{\mathbf{w}} = (-dw_1/w_3, -dw_2/w_3, 0)$. Show that $P\begin{pmatrix} \mathbf{w} \\ 0 \end{pmatrix}$ is a vector of homogeneous coordinates of $\mathbf{q}_{\mathbf{w}}$, and that for every $s \in \mathbb{R}$ with $v_3 + sw_3 \neq d$, the point $\pi(\mathbf{v} + s\mathbf{w})$ lies on $\operatorname{aff}(\pi(\mathbf{v}), \mathbf{q}_{\mathbf{w}})$. Conclude that the images of all lines with direction $\operatorname{span}(\mathbf{w})$ lie on lines (or points) through the single point $\mathbf{q}_{\mathbf{w}}$, the **vanishing point** of the direction $\mathbf{w}$.

(b) Suppose $w_3 = 0$. Show that π maps the line $\mathbf{v} + \operatorname{span}(\mathbf{w})$ onto the line $\pi(\mathbf{v}) + \operatorname{span}((w_1, w_2, 0))$. So lines parallel to the image plane have parallel images.

Solution on page 1356

Exercise 4.7.51. STANDARD Prove or disprove: if E_1 and E_2 are affine subspaces of V and $E_1 \cup E_2$ is an affine subspace, then $E_1 \subseteq E_2$ or $E_2 \subseteq E_1$. Does the answer depend on F?

Solution on page 1356

Exercise 4.7.52. STANDARD Let $\dim V = n \geq 2$, and let H and H' be hyperplanes in V. Show that either H and H' are parallel, or $H \cap H'$ is an affine subspace of dimension $n - 2$. Deduce that two lines in a plane are parallel or meet in exactly one point.

Solution on page 1357

Exercise 4.7.53. STANDARD Let $\mathbf{u}, \mathbf{w} \in V$ be nonzero, let $\mathbf{p}, \mathbf{q} \in V$, and let $\ell_1 = \mathbf{p} + \operatorname{span}(\mathbf{u})$ and $\ell_2 = \mathbf{q} + \operatorname{span}(\mathbf{w})$. Show that ℓ_1 and ℓ_2 are disjoint and not parallel if and only if $\mathbf{u}, \mathbf{w}, \mathbf{q} - \mathbf{p}$ is linearly independent. Deduce that if $\dim V = 2$, there are no skew lines in V.

Solution on page 1357

Exercise 4.7.54. ROUTINE Find a system of linear equations whose solution set is $\operatorname{aff}((1,0,2,1),(2,1,1,1),(0,1,3,2))$ in $\mathbb{R}^4$. How many equations are needed?

Solution on page 1358

Exercise 4.7.55. HARD Let $f : \mathbb{Q}^n \to \mathbb{Q}^m$ satisfy

$$f\Big(\frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}\Big) = \frac{1}{2}f(\mathbf{u}) + \frac{1}{2}f(\mathbf{v}) \quad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbb{Q}^n.$$

Show that f is affine. (Over $\mathbb{R}$ the same hypothesis does not imply that f is affine; the counterexamples need a basis of $\mathbb{R}$ over $\mathbb{Q}$, see Section 15.1.)

Solution on page 1358

Exercise 4.7.56. ROUTINE Find the error in the following argument, and decide which half of the claim is true.

Claim. A list $\mathbf{v}_0, \dots, \mathbf{v}_k$ in V is affinely independent if and only if it is linearly independent.

Proof. An affine dependence is a linear relation, so if the list is linearly independent, it is affinely independent. Conversely, let $\sum_i c_i \mathbf{v}_i = \mathbf{0}$ be a linear relation, and put $s = \sum_i c_i$. Replacing c_0 by $c_0 - s$ gives coefficients that add up to 0, so an affinely independent list has $c_0 - s = c_1 = \cdots = c_k = 0$. Then $s = c_0$, so $c_0 = c_0 + c_1 + \cdots + c_k$ is consistent only if all $c_i = 0$.

Solution on page 1359

Exercise 4.7.57. HARD Let $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2$ be affinely independent in V. When $1 + 1 \neq 0$, the **median** from $\mathbf{v}_i$ is the line through $\mathbf{v}_i$ and the midpoint $\mathbf{m}_i$ of the other two vertices.

(a) Suppose $2 \neq 0$ and $3 \neq 0$ in F. Show that the three medians are distinct lines that all pass through the centroid $\mathbf{g} = \frac{1}{3}(\mathbf{v}_0 + \mathbf{v}_1 + \mathbf{v}_2)$, and that $\mathbf{g}$ is their only common point.

(b) Now let $F = \mathbb{F}_3$ and $V = \mathbb{F}_3^2$. Show that the three medians are parallel and pairwise disjoint, and that together they contain every point of $\mathbb{F}_3^2$.

Solution on page 1359

Exercise 4.7.58. STANDARD

(a) Find the affine map $f : \mathbb{R}^2 \to \mathbb{R}^2$ with $f(0,0) = (2,1)$, $f(1,0) = (4,2)$, $f(0,1) = (1,3)$, and its matrix M_f.

(b) Find the affine map $g : \mathbb{R}^2 \to \mathbb{R}^2$ with $g(1,1) = (2,1)$, $g(3,1) = (4,2)$, $g(1,2) = (1,3)$.

(c) Let $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2$ be affinely independent in $\mathbb{R}^2$, let $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^2$, and let $\hat{A}, \hat{B} \in \mathbb{R}^{3\times 3}$ have columns $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}}_i$. Show that $\hat{A}$ is invertible and that the affine map with $\mathbf{a}_i \mapsto \mathbf{b}_i$ has matrix $\hat{B}\hat{A}^{-1}$. Check (b) this way.

Solution on page 1360

Exercise 4.7.59. HARD (Radon's theorem.) Let $\mathbf{v}_1, \dots, \mathbf{v}_{n+2} \in \mathbb{R}^n$. Show that the index set $\{1, \dots, n+2\}$ can be split into two disjoint nonempty parts I and J with

$$\operatorname{conv}\{\mathbf{v}_i \mid i \in I\} \cap \operatorname{conv}\{\mathbf{v}_j \mid j \in J\} \neq \emptyset.$$

Find such a split for the four points $(0,0),(2,0),(0,2),(1,1)$ of $\mathbb{R}^2$, and for $(0,0),(4,0),(0,4),(1,1)$.

Solution on page 1361

Exercise 4.7.60. STANDARD A function $h : \mathbb{R}^n \to \mathbb{R}$ of the form $h(\mathbf{x}) = w_1x_1 + \cdots + w_nx_n + b$ is an affine map, and the rule "label $\mathbf{x}$ by the sign of $h(\mathbf{x})$" is a **linear classifier**, a single neuron of a neural network (Section 16.6).

(a) Show that no affine $h : \mathbb{R}^2 \to \mathbb{R}$ has $h(0,0) > 0$, $h(1,1) > 0$, $h(1,0) < 0$, and $h(0,1) < 0$. (The labels of these four points form the "exclusive or" pattern.)

(b) Show that after adding the feature x_1x_2, that is, after replacing each point $\mathbf{x}$ by $\varphi(\mathbf{x}) = (x_1, x_2, x_1x_2) \in \mathbb{R}^3$, there is an affine $h' : \mathbb{R}^3 \to \mathbb{R}$ with $h' \circ \varphi$ positive at $(0,0),(1,1)$ and negative at $(1,0),(0,1)$.

Solution on page 1361

Exercise 4.7.61. STANDARD Show that a nonempty subset $S \subseteq \mathbb{F}_2^n$ is an affine subspace if and only if $\mathbf{u} + \mathbf{v} + \mathbf{w} \in S$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in S$. Use this to decide whether $\{(0,0,0),(1,1,0),(1,0,1),(0,1,1)\}$ and $\{(1,0,0),(0,1,0),(0,0,1),(1,1,1)\}$ are affine subspaces of $\mathbb{F}_2^3$.

Solution on page 1361

Exercise 4.7.62. STANDARD

(a) Using 3×3 matrices, compute the composite of the rotation by $90°$ counterclockwise about $(0,0)$ followed by the rotation by $90°$ counterclockwise about $(1,0)$. Show that it is a rotation by $180°$, and find its center.

(b) More generally, let $f(\mathbf{v}) = R_\theta\mathbf{v} + \mathbf{b}$ be an affine map of $\mathbb{R}^2$ whose linear part is the rotation by an angle θ that is not a multiple of 2π. Show that f has exactly one fixed point $\mathbf{c}$, and that f is the rotation by θ about $\mathbf{c}$.

Solution on page 1362

Exercise 4.7.63. STANDARD Let F be a field with q elements, and let $0 \le k \le n$.

(a) Show that the number of k-dimensional affine subspaces of F^n is q^{n-k} times the number of k-dimensional subspaces of F^n.

(b) Count the lines in $\mathbb{F}_3^2$ and the planes in $\mathbb{F}_2^3$.

Solution on page 1362

Exercise 4.7.64. HARD Let V and W be real vector spaces, and let $f : V \to W$ satisfy

$$f((1-t)\mathbf{u} + t\mathbf{v}) = (1-t)f(\mathbf{u}) + tf(\mathbf{v}) \quad \text{for all } \mathbf{u}, \mathbf{v} \in V \text{ and } t \in [0,1].$$

That is, f maps each segment onto a segment, at constant speed. Show that f is affine.

Solution on page 1363

Exercise 4.7.65. STANDARD A point $\mathbf{p}$ of a convex set C is an **extreme point** of C if $\mathbf{p} = (1-t)\mathbf{a} + t\mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in C$ and $t \in (0,1)$ implies $\mathbf{a} = \mathbf{b} = \mathbf{p}$. Show that the extreme points of a simplex are exactly its vertices. Conclude that a simplex determines its set of vertices.

Solution on page 1363

4.8 Proof patterns

Chapter 4 attaches a number to every finite-dimensional space, and most of its proofs either compute such a number or use one to force a conclusion: a list that is too long has a relation, and an injective operator on a finite-dimensional space is onto. This section sorts the chapter's arguments by the question they answer, solves two problems in three ways each to compare the moves, and lists the errors that the chapter's exercises warn against. The exercises at the end practice one move at a time, and the problem set (Problems for Chapter 4) mixes them with earlier chapters.

4.8.1 Relations and independent lists

There are three ways to get information about a list from its length.

- *Too many vectors* (Technique 4.1.30). More than $\dim V$ vectors in V satisfy a nontrivial relation. Nobody computes the relation; it exists, and often that is all we need: every square matrix satisfies a nonzero polynomial (Proposition 4.1.31), the inverse of an injective operator is a polynomial in it (Exercise 4.1.56), and a statement about subsets of $\{1, \dots, n\}$ follows from a relation among $n+1$ vectors (Exercise 4.1.57). The smaller the space that contains the vectors, the shorter the relation, so it pays to look for a small one. The powers of an upper triangular 2×2 matrix stay in a space of dimension 3, not 4 (Exercise 4.8.4(a)).
- *The right length.* A list of length $\dim V$ is a basis as soon as it is independent or spanning (Theorem 4.1.10(b), Corollary 4.1.11(b)), so half of the definition of a basis comes for free (Example 4.1.12). The same count squeezes the dimension of a subspace between an independent list inside it and a vector outside it (Example 4.1.22).
- *Kill terms* (Technique 4.1.25). To show that a particular list is independent, apply an operator to a relation so that all but one term dies. Counting says nothing about a particular list; this move does.

The canonical use of the third move is the nilpotent chain (Proposition 4.1.26): if $T^k\mathbf{v} \neq \mathbf{0}$ and $T^{k+1}\mathbf{v} = \mathbf{0}$, then $\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}$ is independent. It returns in every later chapter about operators, so here is why it works. The hypothesis $T^{k+1}\mathbf{v} = \mathbf{0}$ is what makes T^k kill every term except the first, and without it the argument collapses (Exercise 4.8.5). Nothing about $\dim V$ is used, so the fact holds on $F[x]$ as well. And it is a statement about one vector at a time, which is what lets it combine with counting: an independent list in V has at most $\dim V$ entries, so when T is nilpotent, every vector dies within $\dim V$ steps, and that is Corollary 4.1.27. When $k+1 = \dim V$, the chain is a basis in which T is the shift (Exercise 4.5.50), and the vectors $\mathbf{w}$ whose chains span the same space are easy to describe (Exercise 4.8.8). In Section 7.2 the same proof runs with T replaced by $N = T - \lambda\,\mathrm{id}$ (Proposition 7.2.37). The last nonzero vector $N^k\mathbf{v}$ is then an eigenvector, and the span of the chain is the smallest subspace that contains $\mathbf{v}$ and is mapped into itself by T (Corollary 7.2.38). Chains glued together are the Jordan bases of Section 10.3, and whether several chains can be glued is decided by their last vectors alone (Exercise 7.2.51).

Example 4.8.1 (A nilpotent operator dies within n steps, three ways). Let $\dim V = n$, and let $T \in \mathrm{L}(V)$ satisfy $T^m = 0$ for some $m \geq 1$. We show $T^n = 0$.

Chains. Let $\mathbf{v} \neq \mathbf{0}$, and let k be the largest index with $T^k\mathbf{v} \neq \mathbf{0}$; it exists and $k < m$, since $T^m\mathbf{v} = \mathbf{0}$. By Proposition 4.1.26, $\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}$ is an independent list of length $k+1$, so $k+1 \leq n$ (Theorem 4.1.10(a)), and $T^n\mathbf{v} = T^{n-k-1}\big(T^{k+1}\mathbf{v}\big) = \mathbf{0}$. This is the proof of Corollary 4.1.27.

Kernels. The kernels $\ker T^0 \subseteq \ker T \subseteq \ker T^2 \subseteq \cdots$ grow, and once two consecutive ones are equal, all later ones are equal (Exercise 4.2.51(b)). Let k be the first index with $\ker T^k = \ker T^{k+1}$; it exists, since the dimensions of the kernels are at most n and cannot increase forever. Before k every step is strict and raises the dimension by at least 1, so $k \leq \dim \ker T^k \leq n$. After k nothing changes, so $\ker T^k = \ker T^j$ for $j = \max(k, m)$, which contains $\ker T^m = V$. Hence $T^k = 0$ and $T^n = T^{n-k}T^k = 0$.

Restriction and induction on $\dim V$. We induct on n, assuming the statement for all spaces of dimension less than n. For $n = 0$ there is nothing to prove. Let $n \geq 1$. If T were injective, every power of T would

be injective, which is false for $T^m = 0$ on $V \neq \{\mathbf{0}\}$. So $\dim\ker T \geq 1$, and $U = \operatorname{im} T$ has $\dim U \leq n-1$ by Theorem 4.2.11. The operator T maps U into itself, and its restriction $T_U \in \mathrm{L}(U)$ satisfies $T_U^m = 0$. By the induction hypothesis, applied to U, $T_U^{\dim U} = 0$, so $T_U^{n-1} = 0$, and for every $\mathbf{v}$,

$$T^n\mathbf{v} = T^{n-1}(T\mathbf{v}) = T_U^{n-1}(T\mathbf{v}) = \mathbf{0}.$$

The chain proof works one vector at a time, and when $T^{n-1} \neq 0$ it produces a basis. The kernel proof never looks at a vector. It only counts, and the numbers it counts, $\dim\ker T^k - \dim\ker T^{k-1}$, are the data from which the Jordan form is built (Exercise 4.2.52, Section 10.3). The restriction proof is an induction on the dimension, which is the shape of the triangularization proofs of Section 7.5, and it proves a little more: $\operatorname{rank} T^k \leq n-k$ for $k \leq n$. A fourth version runs the chain argument in a quotient, relative to a subspace that T maps into itself (Exercise 4.8.14). The bound n cannot be lowered, as the shift on F^n shows.

Practice: Exercise 4.8.4, Exercise 4.8.8, and Exercise 4.8.9.

4.8.2 Rank–nullity and restriction

Rank–nullity says that the kernel and the image of one map have complementary dimensions, so whichever is easier to find gives the other (Technique 4.2.24). Sometimes the image is easy, because the map is visibly onto a space of known dimension; this is how kernels of nonzero functionals were found to be hyperplanes (Example 4.2.25). Sometimes the kernel is easy, because degrees or a leading coefficient force it to be small, and then the image is everything (Example 4.2.28). The cleanest instance is Corollary 4.2.18: between spaces of the same finite dimension, uniqueness of solutions gives existence (Example 4.2.19). Counting also produces vectors that nobody constructs, in an intersection of large subspaces (Proposition 4.2.26) or in the kernel of a map into a smaller space (Corollary 4.2.21). For matrix equations that are linear in the unknown matrix, the same count applies to the map $X \mapsto AX - XA$ or its relatives (Technique 4.2.29).

When only part of the domain matters, restrict first and then count (Technique 4.2.5). The restriction $T{\restriction_U}$ has kernel $U \cap \ker T$, so

$$\dim\operatorname{im}(T{\restriction_U}) = \dim U - \dim(U \cap \ker T)$$

(Exercise 4.2.44), and with $U = \operatorname{im} S$ this is the exact rank of a composite, Proposition 4.3.5. Restriction to a hyperplane is a small model of the whole move (Exercise 4.8.7). Two cautions. A restriction is a map $U \to W$, and it is an operator on U only when T maps U into itself, as T does with $\operatorname{im} T$ in the third proof of Example 4.8.1. And none of this counting means anything when the dimensions are infinite (Exercise 4.8.6).

4.8.3 Rank inequalities

An inequality between ranks is almost always a containment of subspaces followed by a count. For a product, the two easy containments are $\operatorname{im}(AB) \subseteq \operatorname{im} A$ and $\ker B \subseteq \ker(AB)$; the first bounds $\operatorname{rank}(AB)$ by $\operatorname{rank} A$, and the second, through rank–nullity, by $\operatorname{rank} B$ (Theorem 4.3.3, Remark 4.3.9). For an exact count, restrict A to $\operatorname{im} B$:

$$\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A)$$

((4.3.1)), and every lower bound for $\operatorname{rank}(AB)$ is an upper bound for the subtracted term. The transpose exchanges rows and columns without changing the rank (Theorem 4.3.8), so each statement about column spaces has a mirror statement about row spaces. Sums are handled by images, $\operatorname{im}(A+B) \subseteq \operatorname{im} A + \operatorname{im} B$ (Proposition 4.3.20), or by counting rank-one pieces (Corollary 4.3.24, Technique 4.3.17). Finally, several matrices can be packed into one block matrix, whose rank does not change under invertible block row and column operations (Proposition 4.3.1). The block method needs one fact about block triangular matrices.

Lemma 4.8.2 (Ranks of block triangular matrices). Let $X \in F^{p\times q}$, $Y \in F^{p\times s}$, $Y' \in F^{r\times q}$, and $Z \in F^{r\times s}$.

(a) $\operatorname{rank}\begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} \geq \operatorname{rank} X + \operatorname{rank} Z$ and $\operatorname{rank}\begin{pmatrix} X & 0 \\ Y' & Z \end{pmatrix} \geq \operatorname{rank} X + \operatorname{rank} Z$.
(b) $\operatorname{rank}\begin{pmatrix} X & 0 \\ 0 & Z \end{pmatrix} = \operatorname{rank} X + \operatorname{rank} Z$.

Proof. STANDARD

Uses. Theorem 4.2.7 (The image of T_A), Proposition 2.6.18 (The column space is spanned by the columns), Theorem 1.5.5 (Independence is decided at zero), Theorem 4.1.10 (Independent lists are short), Theorem 4.3.8 (Row rank equals column rank), Proposition 4.4.2 (Sums are the smallest containing subspaces), Theorem 4.4.13 (Bases and dimension of a sum)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Take $\operatorname{rank} X$ independent columns of X and $\operatorname{rank} Z$ independent columns of Z, and look at the corresponding columns of the block matrix.
Hint 2. In a relation among these columns, read the bottom block first: the columns coming from X have zero bottom part. For the lower triangular matrix, transpose.
Sketch. (a) Let $\mathbf{x}_j$ ($j \in J$) be $\operatorname{rank} X$ independent columns of X and $\mathbf{z}_k$ ($k \in K$) be $\operatorname{rank} Z$ independent columns of Z. The columns $(\mathbf{x}_j, \mathbf{0})$ and $(\mathbf{y}_k, \mathbf{z}_k)$ of $\left(\begin{smallmatrix} X & Y \\ 0 & Z \end{smallmatrix}\right)$ are independent: the bottom block of a relation kills the coefficients of the $\mathbf{z}_k$, and then the top block kills the rest. Transposing gives the lower triangular case. (b) With $Y = 0$, the column space lies in the sum of the column spaces of $\left(\begin{smallmatrix} X \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ Z \end{smallmatrix}\right)$, of dimension at most $\operatorname{rank} X + \operatorname{rank} Z$.

(a) Write $M = \left(\begin{smallmatrix} X & Y \\ 0 & Z \end{smallmatrix}\right)$, and write vectors of F^{p+r} as pairs $(\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in F^p$ and $\mathbf{b} \in F^r$. By Theorem 4.2.7, X has a set J of $a = \operatorname{rank} X$ columns $\mathbf{x}_j$ that are independent, and Z has a set K of $b = \operatorname{rank} Z$ columns $\mathbf{z}_k$ that are independent. The corresponding columns of M are $(\mathbf{x}_j, \mathbf{0})$ for $j \in J$ and $(\mathbf{y}_k, \mathbf{z}_k)$ for $k \in K$, where $\mathbf{y}_k$ is the column of Y above $\mathbf{z}_k$. Suppose

$$\sum_{j \in J} \alpha_j (\mathbf{x}_j, \mathbf{0}) + \sum_{k \in K} \beta_k (\mathbf{y}_k, \mathbf{z}_k) = (\mathbf{0}, \mathbf{0}).$$

The bottom block reads $\sum_k \beta_k \mathbf{z}_k = \mathbf{0}$, so every $\beta_k = 0$. Then the top block reads $\sum_j \alpha_j \mathbf{x}_j = \mathbf{0}$, so every $\alpha_j = 0$. So M has $a + b$ independent columns in its column space, and $\operatorname{rank} M \geq a + b$ (Proposition 2.6.18, Theorem 4.1.10(a)). For the lower triangular matrix,

$$\begin{pmatrix} X & 0 \\ Y' & Z \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} X^{\mathsf{T}} & Y'^{\mathsf{T}} \\ 0 & Z^{\mathsf{T}} \end{pmatrix},$$

and by Theorem 4.3.8 and the first case its rank is at least $\operatorname{rank} X^{\mathsf{T}} + \operatorname{rank} Z^{\mathsf{T}} = \operatorname{rank} X + \operatorname{rank} Z$.

(b) By (a) the rank is at least $\operatorname{rank} X + \operatorname{rank} Z$. Every column of $\left(\begin{smallmatrix} X & 0 \\ 0 & Z \end{smallmatrix}\right)$ is a column of $\left(\begin{smallmatrix} X \\ 0 \end{smallmatrix}\right)$ or of $\left(\begin{smallmatrix} 0 \\ Z \end{smallmatrix}\right)$, so the column space is contained in the sum of their column spaces (Proposition 4.4.2), which have dimensions $\operatorname{rank} X$ and $\operatorname{rank} Z$. By Theorem 4.4.13(a), its dimension is at most $\operatorname{rank} X + \operatorname{rank} Z$. □

Example 4.8.3 (Sylvester's inequality, three ways). Let $A \in F^{m \times n}$ and $B \in F^{n \times p}$. We show $\operatorname{rank}(AB) \geq \operatorname{rank} A + \operatorname{rank} B - n$.

By restriction. By (4.3.1), $\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A)$, and the subtracted space lies in $\ker A$, which has dimension $n - \operatorname{rank} A$. So $\operatorname{rank}(AB) \geq \operatorname{rank} B - (n - \operatorname{rank} A)$. This is the solution of Exercise 4.3.30(a).

By kernels. A vector $\mathbf{x} \in F^p$ lies in $\ker(AB)$ exactly when $B\mathbf{x} \in \ker A$. So T_B restricts to a linear map $R : \ker(AB) \to \ker A$, whose kernel is $\ker(AB) \cap \ker B = \ker B$. Rank–nullity for R gives

$$\dim \ker(AB) = \dim \ker B + \dim \operatorname{im} R \leq \dim \ker B + \dim \ker A,$$

and rank–nullity for AB, B, and A turns this into $p - \operatorname{rank}(AB) \leq (p - \operatorname{rank} B) + (n - \operatorname{rank} A)$, which is the claim.

By blocks. Let $M = \left(\begin{smallmatrix} B & I_n \\ 0 & A \end{smallmatrix}\right) \in F^{(n+m) \times (p+n)}$. By Lemma 4.8.2(a), $\operatorname{rank} M \geq \operatorname{rank} B + \operatorname{rank} A$. Clear the blocks with block row and column operations (Theorem 2.4.34):

$$\begin{pmatrix} I_n & 0 \\ -A & I_m \end{pmatrix}\begin{pmatrix} B & I_n \\ 0 & A \end{pmatrix}\begin{pmatrix} I_p & 0 \\ -B & I_n \end{pmatrix} = \begin{pmatrix} B & I_n \\ -AB & 0 \end{pmatrix}\begin{pmatrix} I_p & 0 \\ -B & I_n \end{pmatrix} = \begin{pmatrix} 0 & I_n \\ -AB & 0 \end{pmatrix}.$$

The outer factors are invertible, with inverses obtained by changing the signs of A and B, so M has the rank of the last matrix (Proposition 4.3.1). Exchanging its two block columns is multiplication by an invertible permutation matrix and gives $\begin{pmatrix} I_n & 0 \\ 0 & -AB \end{pmatrix}$, of rank $n + \operatorname{rank}(AB)$ by Lemma 4.8.2(b). So $n + \operatorname{rank}(AB) \geq \operatorname{rank} A + \operatorname{rank} B$.

The restriction proof is the most informative. It identifies the rank lost in the product, $\operatorname{rank} B - \operatorname{rank}(AB) = \dim(\operatorname{im} B \cap \ker A)$, and the slack in Sylvester's bound, $\dim \ker A - \dim(\operatorname{im} B \cap \ker A)$, so it also gives the equality case, $\ker A \subseteq \operatorname{im} B$ (Exercise 4.8.10). The kernel proof proves the inequality $\dim \ker(AB) \leq \dim \ker A + \dim \ker B$, which makes sense for linear maps between any spaces as long as the two kernels are finite-dimensional, and it is the ingredient of the index formula in Problem 4.18. The block proof is mechanical: once the right block matrix is written down, everything is multiplication, and the same computation proves Frobenius's inequality (Exercise 4.8.13). Its price is that the dimension argument hides inside Lemma 4.8.2, and the proof does not say where the rank goes.

Practice: Exercise 4.8.10, Exercise 4.8.13, and Exercise 4.8.11.

4.8.4 Sums and direct sums

To show that a sum $U_1 + \cdots + U_k$ is direct, show that $\mathbf{0}$ has only the trivial decomposition, by sorting the terms of an equation into the summands they belong to (Technique 4.4.8). For two subspaces this is the single condition $U \cap W = \{\mathbf{0}\}$. For more, the induction of Technique 4.4.10 needs each U_k to meet the sum of the previous ones trivially, and pairwise intersections are not enough (Non-example 4.4.6). In finite dimension there are two shortcuts. Counting: the sum is direct exactly when bases of the summands concatenate to an independent list, that is, when the dimensions add up (Theorem 4.4.13, Technique 4.4.14), and for $V = U \oplus W$ any two of $U \cap W = \{\mathbf{0}\}$, $U + W = V$, and $\dim U + \dim W = \dim V$ suffice (Proposition 4.4.21). Idempotents: an operator P with $P^2 = P$ splits $V = \operatorname{im} P \oplus \ker P$ with no computation (Technique 4.4.28). In Problem 4.12 the decomposition $F^n = \operatorname{im} A \oplus \ker A$ comes from a rank count instead (Exercise 4.2.50), and AB then turns out to be the idempotent that splits it.

The formula $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$ received three proofs in Section 4.4.5, and the discussion after them compares them. Extending a basis of $U \cap W$ produces a basis of $U + W$; choosing a complement of $U \cap W$ inside U reduces the formula to the direct case; rank–nullity for a map onto $U + W$ is the one that computes. Read from the point of view of this section, they are Technique 4.1.18, Technique 4.4.14, and Technique 4.2.24, and a fourth proof uses quotients (Exercise 4.6.49). A fifth method, a single row reduction that produces bases of both $U + W$ and $U \cap W$, is Problem 4.6.

4.8.5 Choosing a basis, and transporting

Extend a basis (Technique 4.1.18) when a subspace U must be compared with the whole space or a map must be prescribed on U and left free elsewhere. Complements (Proposition 4.4.16), extensions of maps (Proposition 4.1.19), and functionals separating a vector from a subspace all come from one extended basis.

Choose a convenient basis (Technique 4.5.33) after checking that the statement does not depend on the basis. For statements that survive $A \mapsto GAH$ with G and H invertible, the convenient matrix is the rank normal form $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ (Corollary 4.5.22), as in Exercise 4.8.12. For statements about operators that survive similarity, the basis has to be the same on both sides, and the useful bases come from the operator: a basis adapted to a subspace that T maps into itself (Technique 4.5.36), or a chain $\mathbf{v}, T\mathbf{v}, \ldots$. For instance, an idempotent P of rank r has, in a basis of $\operatorname{im} P$ followed by a basis of $\ker P$, the matrix $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, so $\operatorname{tr} P = r * 1$ in one line, and its commutant can be read off the same matrix (Problem 4.4). The proof of $\operatorname{tr} P = r *$

1 through a column–row factorization (Exercise 4.3.43) chooses no basis of V, but it computes where this proof only looks.

Transport along an isomorphism (Technique 4.5.7) when the space is made of polynomials, matrices, or maps and the question concerns only the linear structure. Coordinates make every n-dimensional space into F^n, and $\mathrm{L}(V,W)$ into $F^{m\times n}$ (Theorem 4.5.8). Only linear notions survive the trip. The degree of a polynomial, the product of two matrices, and the rank of a matrix seen as a vector of $F^{m\times n}$ mean nothing in F^n, and in infinite dimension an isomorphism can even land in a proper subspace (Exercise 4.5.54). To show that two matrices are not similar, transport nothing: compare invariants (Technique 4.5.29).

Practice: Exercise 4.8.12 and Exercise 4.8.4(e).

4.8.6 Quotients and affine arguments BEYOND 21-242

Quotients and affine geometry (Section 4.6, Section 4.7) are not part of 21-242.

A complement of U is a choice, and the quotient V/U is the canonical object that every complement is isomorphic to (Proposition 4.6.29, Intuition (A complement is a choice; the quotient is not)). So proofs that choose a complement can usually be rewritten with a quotient, and the result no longer depends on the choice: rank–nullity becomes the first isomorphism theorem (Corollary 4.6.35), the dimension formula becomes the second (Exercise 4.6.49), and Frobenius's inequality has a quotient proof (Exercise 4.6.53) next to the block proof of Exercise 4.8.13. The price is one check before anything else, that every definition on V/U is independent of the representative (Technique 4.6.18). The second use of quotients is to split an operator that maps U into itself into a piece on U and a piece on V/U (Proposition 4.6.32) and to induct on the dimension (Technique 4.6.31). Applied to the nilpotent chain, this gives a version of Proposition 4.1.26 relative to a subspace (Exercise 4.8.14). Quotients also make sense of rank–nullity in infinite dimension, where $\dim V$ is not a number but the index $\dim\ker T - \dim(W/\operatorname{im} T)$ can be (Problem 4.18).

In affine geometry, a point of an affine subspace can serve as the origin (Technique 4.7.4), or every point can be given an extra coordinate 1 (Technique 4.7.21). The first is shorter when one point is special, and the second treats all points alike. The affine version of too many vectors is that $n+2$ points of an n-dimensional space are affinely dependent (Corollary 4.7.23), and splitting an affine dependence by the signs of its coefficients gives Radon's theorem (Exercise 4.7.59), Carathéodory's theorem (Technique 4.7.33), and Helly's theorem (Problem 4.15).

4.8.7 Reading the chapter's proofs

Where finite dimension enters. The exchange lemma compares an independent list with a finite spanning list, and everything about dimension rests on it (Lemma 4.1.2). Extending an independent list to a basis terminates because independent lists are short (Theorem 4.1.16). Rank–nullity is an equation between numbers, and its corollary that injective operators are surjective fails on $F[x]$ (Example 2.6.4). On the other hand, Proposition 4.1.26 and the restriction formula for $\operatorname{im}(TS)$ use no dimension at all, and the inequality $\dim\ker(AB) \le \dim\ker A + \dim\ker B$ needs only finite-dimensional kernels. For each result, ask which of its steps needed finite dimension.

Inequalities come with their equality cases. Each inequality of the chapter came from a containment of subspaces and a count, and Proposition 4.1.21(b) says that equality of dimensions turns a containment into an equality. So the equality case can be read off the proof. $\operatorname{rank}(TS) = \operatorname{rank} S$ exactly when $\operatorname{im} S \cap \ker T = \{\mathbf{0}\}$ (Proposition 4.3.5), Sylvester's inequality is an equality exactly when $\ker A \subseteq \operatorname{im} B$ (Exercise 4.8.10), and $\dim(U+W) = \dim U + \dim W$ exactly when $U \cap W = \{\mathbf{0}\}$. A proof that drops the subspace it is bounding, instead of writing it down, loses its equality case.

Existence without construction. Counting proves that a relation, a nonzero vector, or a solution exists, and says nothing about what it is: Proposition 4.1.31, Proposition 4.2.26, Corollary 4.2.18. Often the count also

says where to look and how many unknowns to use, after which a small computation finds the object (Exercise 4.1.33, Example 4.2.28).

The field is part of the data. $\mathbb{C}^n$ has dimension $2n$ over $\mathbb{R}$ (Remark 4.1.6, Exercise 4.2.37), and $1 + 1 = 0$ in $\mathbb{F}_2$ makes some sums direct over $\mathbb{Q}$ and not over $\mathbb{F}_2$ (Problem 4.6). The rank of a matrix with rational entries is the same over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, since Gauss–Jordan elimination never leaves $\mathbb{Q}$ and the reduced form is unique (Theorem 3.3.17). For an integer matrix read modulo a prime p, the rank can only drop, so a rank computed modulo p is a lower bound over $\mathbb{Q}$ (Problem 4.17), and it can drop (Exercise 4.2.36).

4.8.8 Common errors

Most of these appear in the chapter as find-the-error exercises.

- *Pairwise for all at once.* Subspaces that meet pairwise in $\{\mathbf{0}\}$ need not form a direct sum (Non-example 4.4.6, Exercise 4.4.37), and there is no inclusion–exclusion formula for three subspaces (Remark 4.4.22, Exercise 4.4.39).
- *Bases that are not adapted.* A basis of V need not contain a basis of a subspace (Exercise 4.1.48), and bases of U and W need not contain a basis of $U \cap W$ (Exercise 4.4.38). Adapted bases have to be built, by extending a basis of the smaller space.
- *A subspace of the wrong space.* The columns of $\operatorname{rref}(A)$ that span its column space are not columns of A, and $\operatorname{im}(\operatorname{rref}(A)) \neq \operatorname{im} A$ in general (Non-example 4.2.8, Exercise 4.3.35). Containments such as $\ker(AB) \subseteq \ker A + \ker B$ compare subspaces of spaces that are different, or are false even when they are the same space (Exercise 4.8.11).
- *Order of factors.* $\operatorname{rank}(AB) \neq \operatorname{rank}(BA)$ in general, even for square matrices (Example 4.3.4, Exercise 4.3.36), although the difference is limited (Problem 4.11).
- *One invariant for all.* Equal ranks do not make matrices similar (Exercise 4.5.45), and equal traces do not either.
- *Infinite dimension.* Rank–nullity, injective if and only if surjective, and "a subspace of the same dimension is everything" all need finite dimension (Exercise 4.8.6, Exercise 4.5.54).
- *A missing hypothesis in a chain.* $T^k\mathbf{v} \neq \mathbf{0}$ alone does not make $\mathbf{v}, \ldots, T^k\mathbf{v}$ independent (Exercise 4.8.5).
- *Extending the wrong list.* A basis whose last vectors map to a basis of the image does not have its first vectors in the kernel (Exercise 4.2.45).
- *Uniqueness of choices that are not unique.* Complements are not unique (Example 4.4.17), so "the complement" of a subspace means nothing. In quotients, well-definedness is the implication from equal cosets to equal values, not its converse (Exercise 4.6.43), and affine independence is not linear independence (Exercise 4.7.56).

Exercises

Exercise 4.8.4. ROUTINE Prove or disprove each statement with the shortest argument you can, and name the move you used.

(a) For every upper triangular $A \in F^{2\times 2}$ there are $c_0, c_1, c_2, c_3 \in F$, not all 0, with $c_0 I_2 + c_1 A + c_2 A^2 + c_3 A^3 = 0$.

(b) There are $A \in F^{2\times 3}$ and $B \in F^{3\times 2}$ with $BA = I_3$.

(c) Any two 4-dimensional subspaces of F^7 have a nonzero vector in common.

(d) If $T \in \mathrm{L}(F^5)$ satisfies $T^3 = 0$, then $\operatorname{rank} T \leq 3$.

(e) For every $A \in F^{2\times 2}$, the matrices A and $A + I_2$ are not similar. (The answer depends on F: decide for which fields it holds.)

Solution on page 1364

Exercise 4.8.5. ROUTINE Find the error in the following argument, and show that the claim is false.

Claim. Let $T \in \mathrm{L}(V)$, $\mathbf{v} \in V$, and $k \geq 1$ with $T^k\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}$ is linearly independent.

Proof. Suppose $a_0\mathbf{v} + a_1T\mathbf{v} + \cdots + a_kT^k\mathbf{v} = \mathbf{0}$. Applying T^k, every term except the first becomes a vector $T^{k+j}\mathbf{v}$ with $j \geq 1$, which is $\mathbf{0}$. So $a_0T^k\mathbf{v} = \mathbf{0}$, and $a_0 = 0$. Applying T^{k-1} to the remaining relation gives $a_1 = 0$ in the same way, and so on.

Solution on page 1364

Exercise 4.8.6. ROUTINE Find the error in the following argument, and show that the claim is false.

Claim. Every injective operator $T \in \mathrm{L}(V)$ is surjective.

Proof. Since T is injective, $\ker T = \{\mathbf{0}\}$. By rank–nullity, $\dim \operatorname{im} T = \dim V - \dim \ker T = \dim V$. So $\operatorname{im} T$ is a subspace of V of the same dimension as V, and therefore $\operatorname{im} T = V$.

Solution on page 1365

Exercise 4.8.7. ROUTINE Let V be finite-dimensional with $\dim V = n \geq 1$, let $T \in \mathrm{L}(V, W)$, and let H be a subspace of V with $\dim H = n - 1$.

(a) Show that $\operatorname{rank}(T\restriction_H)$ is $\operatorname{rank} T$ or $\operatorname{rank} T - 1$, and that it is $\operatorname{rank} T - 1$ if and only if $\ker T \subseteq H$.

(b) Let $A \in F^{m\times n}$ with $n \geq 2$, and let A' be A with column j deleted. Show that $\operatorname{rank} A' = \operatorname{rank} A - 1$ if column j is not a linear combination of the other columns, and $\operatorname{rank} A' = \operatorname{rank} A$ otherwise. Prove this twice: from (a), and directly from the column spaces.

(c) Deduce that deleting k rows and l columns of a matrix, so that at least one row and one column remain, lowers its rank by at most $k + l$.

Solution on page 1365

Exercise 4.8.8. STANDARD Let $T \in \mathrm{L}(V)$, $\mathbf{v} \in V$, and $k \geq 0$ with $T^k\mathbf{v} \neq \mathbf{0}$ and $T^{k+1}\mathbf{v} = \mathbf{0}$, and let $W = \operatorname{span}\big(\mathbf{v}, T\mathbf{v}, \ldots, T^k\mathbf{v}\big)$.

(a) Let $\mathbf{w} = a_0\mathbf{v} + a_1T\mathbf{v} + \cdots + a_kT^k\mathbf{v}$. Show that $\mathbf{w}, T\mathbf{w}, \ldots, T^k\mathbf{w}$ is a basis of W if and only if $a_0 \neq 0$.

(b) Let F have q elements. How many $\mathbf{w} \in W$ satisfy $W = \operatorname{span}\big(\mathbf{w}, T\mathbf{w}, \ldots, T^k\mathbf{w}\big)$?

Solution on page 1366

Exercise 4.8.9. STANDARD Let $n \geq 1$, and let $S \subseteq F^{n\times n}$ be a subspace with $I_n \in S$ and $XY \in S$ for all $X, Y \in S$. Let $A \in S$ be invertible.

(a) Show that $A^{-1} \in S$ by finding a relation among $I_n, A, A^2, \ldots, A^d$, where $d = \dim S$.

(b) Show again that $A^{-1} \in S$, by applying rank–nullity to the map $S \to S$, $X \mapsto AX$.

(c) Deduce that the inverse of an invertible upper triangular matrix is upper triangular.

Solution on page 1366

Exercise 4.8.10. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$.

(a) Show that $\operatorname{rank}(AB) = \operatorname{rank} A + \operatorname{rank} B - n$ if and only if $\ker A \subseteq \operatorname{im} B$.

(b) Suppose $AB = 0$. Show that $\operatorname{rank} A + \operatorname{rank} B \leq n$, with equality if and only if $\operatorname{im} B = \ker A$.

(c) Let $N \in F^{n\times n}$. Show that $\operatorname{rank}\big(N^2\big) = 2\operatorname{rank} N - n$ if and only if $\ker N \subseteq \operatorname{im} N$, and check this on the shift N with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \geq 2$.

Solution on page 1367

Exercise 4.8.11. STANDARD Find the error in the following argument. Is the claim true? Give matrices for which the false step fails, and repair the argument.

Claim. For $A, B \in F^{n\times n}$, $\dim\ker(AB) \le \dim\ker A + \dim\ker B$, and hence $\operatorname{rank}(AB) \ge \operatorname{rank} A + \operatorname{rank} B - n$.

Proof. If $AB\mathbf{x} = \mathbf{0}$, then A kills $B\mathbf{x}$, so either $B\mathbf{x} = \mathbf{0}$ or A kills $\mathbf{x}$. Hence $\ker(AB) \subseteq \ker A \cup \ker B \subseteq \ker A + \ker B$, and by Theorem 4.4.19, $\dim\ker(AB) \le \dim(\ker A + \ker B) \le \dim\ker A + \dim\ker B$. Rank–nullity for AB, A, and B gives the second inequality.

Solution on page 1367

Exercise 4.8.12. HARD Let $A \in F^{m\times n}$ have rank r, and let

$$\Sigma_A = \{X \in F^{n\times m} \mid AXA = A\}.$$

(a) Show that Σ_A is nonempty, and that $\Sigma_A = X_0 + K$ for any $X_0 \in \Sigma_A$, where $K = \{X \mid AXA = 0\}$ is a subspace of $F^{n\times m}$ with $\dim K = mn - r^2$.

(b) Let F have q elements. How many X satisfy $AXA = A$?

(c) For $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$, describe Σ_A.

Solution on page 1368

Exercise 4.8.13. HARD (*Frobenius's inequality by blocks.*) Let $A \in F^{m\times n}$, $B \in F^{n\times p}$, and $C \in F^{p\times q}$, and let

$$M = \begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix} \in F^{(m+n)\times(p+q)}.$$

(a) Compute $\begin{pmatrix} I_m & -A \\ 0 & I_n \end{pmatrix} M \begin{pmatrix} I_p & -C \\ 0 & I_q \end{pmatrix}$.

(b) Deduce $\operatorname{rank}(AB) + \operatorname{rank}(BC) \le \operatorname{rank} B + \operatorname{rank}(ABC)$.

(c) Recover Sylvester's inequality as a special case.

Solution on page 1369

Exercise 4.8.14. STANDARD BEYOND 21-242 Let $T \in \mathrm{L}(V)$, and let U be a subspace of V with $T(\mathbf{u}) \in U$ for all $\mathbf{u} \in U$. Let $\mathbf{v} \in V$ and $k \ge 0$ with $T^k\mathbf{v} \notin U$ and $T^{k+1}\mathbf{v} \in U$.

(a) Show that the cosets $\mathbf{v} + U, T\mathbf{v} + U, \dots, T^k\mathbf{v} + U$ are linearly independent in V/U. Deduce that $\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}$ is linearly independent and that its span meets U only in $\mathbf{0}$.

(b) Let V be finite-dimensional, and suppose $T^m(V) \subseteq U$ for some $m \ge 1$. Show that $T^d(V) \subseteq U$, where $d = \dim V - \dim U$. Which earlier result is the case $U = \{\mathbf{0}\}$?

Solution on page 1369

Problems for Chapter 4

These problems combine the sections of Chapter 4 with each other and with Chapters 1–3. The first three are review prompts: state a main result, prove it without looking, and then answer a question that the text did not ask. The rest are ordered roughly by difficulty. Problems marked "Beyond 21-242" use the quotients, affine geometry, or subspace counting of the sections marked the same way, and problems of

kind "competition" are harder than anything the course expects. Section 4.8 compares the techniques these problems call for.

Review: state and prove

Problem 4.1. STANDARD

(a) State Theorem 4.2.11(b).

(b) Prove it by extending a basis of $\ker T$ to a basis of V.

(c) Give a second proof that lifts a basis of $\operatorname{im} T$ instead: choose $\mathbf{v}_1, \dots, \mathbf{v}_r \in V$ such that $T(\mathbf{v}_1), \dots, T(\mathbf{v}_r)$ is a basis of $\operatorname{im} T$, and show that a basis of $\ker T$ followed by $\mathbf{v}_1, \dots, \mathbf{v}_r$ is a basis of V. Deduce that if V is any vector space and $T \in \mathrm{L}(V, W)$ has finite-dimensional kernel and image, then V is finite-dimensional. In particular, an operator on an infinite-dimensional space has an infinite-dimensional kernel or an infinite-dimensional image.

Solution on page 1370

Problem 4.2. STANDARD For $A \in F^{m\times n}$, sets $I \subseteq \{1, \dots, m\}$ and $J \subseteq \{1, \dots, n\}$ with $|I| = |J| = k$, let $A_{IJ} \in F^{k\times k}$ be the submatrix formed by the entries a_{ij} with $i \in I$ and $j \in J$.

(a) State Theorem 4.3.8.

(b) Prove it.

(c) Let $r = \operatorname{rank} A \geq 1$, let the rows of A with indices in I be linearly independent, and let the columns with indices in J be linearly independent, where $|I| = |J| = r$. Show that A_{IJ} is invertible. Deduce that for $A \neq 0$, $\operatorname{rank} A$ is the largest k for which A has an invertible $k \times k$ submatrix.

Solution on page 1370

Problem 4.3. STANDARD

(a) State Proposition 4.1.26 and Corollary 4.1.27.

(b) Prove both.

(c) Under the hypotheses of Proposition 4.1.26, let $W = \operatorname{span}\left(\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}\right)$. Show that for $1 \leq j \leq k+1$,

$$\ker T^j \cap W = \operatorname{span}\left(T^{k-j+1}\mathbf{v}, \dots, T^k\mathbf{v}\right),$$

so that $\dim\left(\ker T^j \cap W\right) = j$.

Solution on page 1371

Synthesis problems

Problem 4.4. STANDARD Let $P \in F^{n\times n}$ satisfy $P^2 = P$, let $r = \operatorname{rank} P$, and let $\delta : F^{n\times n} \to F^{n\times n}$ be $\delta(X) = PX - XP$.

(a) Show that $\dim \ker \delta = r^2 + (n-r)^2$.

(b) Show that $\operatorname{rank} \delta = 2r(n-r)$, and that $Z \in \operatorname{im} \delta$ if and only if $PZP = 0$ and $(I_n - P)Z(I_n - P) = 0$.

(c) Show that $F^{n\times n} = \ker \delta \oplus \operatorname{im} \delta$, and that the component of X in $\ker \delta$ is $PXP + (I_n - P)X(I_n - P)$.

Solution on page 1372

Problem 4.5. STANDARD Let $S \in \mathrm{L}(F[x])$ be the shift $S(p) = p(x+1)$ (Exercise 2.1.32), and $\Delta = S - \mathrm{id}$. You may use the binomial theorem $(x+c)^j = \sum_{i=0}^{j} \binom{j}{i} c^{j-i} x^i$ in $F[x]$, and that $S^j(p) = p(x+j)$.

(a) Let $n \geq 0$, and suppose that $n! = 1 \cdot 2 \cdots n$ is nonzero in F. Show that for every $p \in F[x]$ of degree exactly n, the translates $p(x), p(x+1), \ldots, p(x+n)$ form a basis of $F[x]_{\leq n}$.

(b) Write 1 as a combination of $x^2, (x+1)^2, (x+2)^2$ over $\mathbb{Q}$.

(c) Let p be a prime. Show that over $\mathbb{F}_p$ the translates $(x+j)^p$, $j \in \mathbb{F}_p$, span a 2-dimensional space. So (a) fails over $\mathbb{F}_p$ for $n = p$.

Solution on page 1373

Problem 4.6. STANDARD (*Zassenhaus's algorithm.*) Let $U = \mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ and $W = \mathrm{span}(\mathbf{w}_1, \ldots, \mathbf{w}_l)$ be subspaces of F^n. Form the matrix $M \in F^{(k+l)\times 2n}$ whose rows are $(\mathbf{u}_i, \mathbf{u}_i)$ for $i = 1, \ldots, k$ and $(\mathbf{w}_j, \mathbf{0})$ for $j = 1, \ldots, l$, written as rows of length $2n$, and let $R = \mathrm{rref}(M)$.

(a) Show that the row space of M is $Z = \{(\mathbf{u}+\mathbf{w}, \mathbf{u}) \mid \mathbf{u} \in U, \mathbf{w} \in W\}$, and that $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a}$ maps Z onto $U + W$ with kernel $\{(\mathbf{0}, \mathbf{u}) \mid \mathbf{u} \in U \cap W\}$.

(b) Show that the left halves of the nonzero rows of R whose left half is nonzero form a basis of $U + W$, and that the right halves of the remaining nonzero rows form a basis of $U \cap W$.

(c) Run the algorithm in F^4 with $\mathbf{u}_1 = (1,1,0,0)$, $\mathbf{u}_2 = (0,1,1,0)$, $\mathbf{w}_1 = (0,0,1,1)$, $\mathbf{w}_2 = (1,0,0,-1)$, for $F = \mathbb{F}_2$, and decide what happens for $F = \mathbb{Q}$.

Solution on page 1374

Problem 4.7. STANDARD Let $A, B \in F^{n\times n}$ satisfy $A^2 = A$ and $B^2 = B$.

(a) Suppose $I_n - A - B$ is invertible. Show that A and B are similar, so in particular $\mathrm{rank}\, A = \mathrm{rank}\, B$.

(b) Show by an example that (a) fails without the invertibility hypothesis.

(c) Suppose $AB = BA = 0$. Show that $A + B$ is idempotent and that $\mathrm{rank}(A+B) = \mathrm{rank}\, A + \mathrm{rank}\, B$.

Solution on page 1375

Problem 4.8. STANDARD Let $\dim V = n \geq 1$, and let $N \in \mathrm{L}(V)$ satisfy $N^k = 0$ for some $k \geq 1$. By Exercise 4.2.52, the jumps $d_j = \dim\ker N^j - \dim\ker N^{j-1}$ satisfy $d_1 \geq d_2 \geq \cdots$, and $\dim\ker N \geq n/k$.

(a) Prove $\dim\ker N \geq n/k$ again, by iterating Sylvester's inequality.

(b) Suppose k divides n and $\dim\ker N = n/k$. Show that $\dim\ker N^j = jn/k$ for $0 \leq j \leq k$; in particular $N^{k-1} \neq 0$. Find such an N.

(c) Suppose $\dim\ker N = 1$. Show that $N^{n-1} \neq 0$, and deduce that V has a basis $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$.

Solution on page 1375

Problem 4.9. STANDARD Let $k \geq 2$, let $V_1, \ldots, V_k$ be finite-dimensional, and let $T_i \in \mathrm{L}(V_i, V_{i+1})$ for $i = 1, \ldots, k-1$. The sequence $V_1 \xrightarrow{T_1} V_2 \xrightarrow{T_2} \cdots \xrightarrow{T_{k-1}} V_k$ is **exact** if T_1 is injective, T_{k-1} is surjective, and $\mathrm{im}\, T_{i-1} = \ker T_i$ for $2 \leq i \leq k-1$.

(a) Show that for an exact sequence, $\sum_{i=1}^{k} (-1)^i \dim V_i = 0$.

(b) What does (a) say for $k = 2$ and $k = 3$?

(c) Let $n_1, \ldots, n_k \geq 0$, and let $r_0 = 0$ and $r_i = n_i - r_{i-1}$ for $i = 1, \ldots, k$. Show that there is an exact sequence with $\dim V_i = n_i$ for all i if and only if $r_i \geq 0$ for all i and $r_k = 0$.

Solution on page 1376

Problem 4.10. STANDARD BEYOND 21-242 Let F have q elements, let $\dim V = n$, and for $0 \le r \le n$ let $N_q(n,r)$ be the number of independent lists of length r in V (Corollary 4.1.35).

(a) Let U and W be subspaces of V with $\dim U + \dim W = n$, and let $r = \dim W$. Show that exactly $N_q(r,r)$ operators $T \in \mathrm{L}(V)$ satisfy $\ker T = U$ and $\operatorname{im} T = W$.

(b) Show that the number of $T \in \mathrm{L}(V)$ with $T^2 = T$ and $\operatorname{rank} T = r$ is

$$N_q \frac{n,r}{N_q}(r,r) \cdot q^{r(n-r)},$$

and check the total over all r against Problem 2.8(b) for $n = 2$ when q is a prime.

(c) Show that the number of $T \in \mathrm{L}(V)$ of rank r is $N_q(n,r)^2 / N_q(r,r)$, and check that the counts for $n = q = 2$ add up to $|\mathrm{L}(V)|$.

Solution on page 1377

Problem 4.11. HARD Let $A, B \in F^{n\times n}$.

(a) Show that $\operatorname{rank}(BA) \ge 2\operatorname{rank}(AB) - n$.

(b) Deduce that $|\operatorname{rank}(AB) - \operatorname{rank}(BA)| \le \lfloor n/2 \rfloor$.

(c) Show that the bound in (b) is attained for every $n \ge 1$.

Solution on page 1378

Problem 4.12. HARD Let $A, B \in F^{n\times n}$ satisfy $A^2B = A$, and let $r = \operatorname{rank} A$.

(a) Show that $\operatorname{rank} A^2 = \operatorname{rank} A$, and deduce that $F^n = \operatorname{im} A \oplus \ker A$.

(b) Show that $(AB - BA)^2 = 0$.

(c) Show that $\operatorname{rank}(AB - BA) \le \min(r, n-r)$, and that for every k with $2k \le n$ there are such A and B with $\operatorname{rank}(AB - BA) = k$.

Solution on page 1378

Problem 4.13. HARD Let $n \ge 1$, and call $N \in F^{n\times n}$ nilpotent if $N^m = 0$ for some $m \ge 1$.

(a) Show that every nilpotent N is similar to a strictly upper triangular matrix, and deduce that $\operatorname{tr} N = 0$.

(b) Show that the nilpotent matrices span the subspace $\{X \in F^{n\times n} \mid \operatorname{tr} X = 0\}$.

(c) Show that the matrices P with $P^2 = P$ span all of $F^{n\times n}$.

Solution on page 1379

Problem 4.14. HARD BEYOND 21-242 Let F have q elements, let $\dim V = n$, let U be a subspace of dimension k, and let $0 \le j \le n-k$. With N_q as in Corollary 4.1.35, show that the number of j-dimensional subspaces W of V with $W \cap U = \{\mathbf{0}\}$ is

$$q^{jk} \cdot N_q \frac{n-k,j}{N_q}(j,j).$$

Check the cases $j = n-k$ and $j = 1$, and the example of lines and planes in $\mathbb{F}_2^3$ meeting a fixed line only in $\mathbf{0}$.

Solution on page 1380

Problem 4.15. HARD BEYOND 21-242 (*Helly's theorem.*) Let V be a real vector space with $\dim V = n$, and let $C_1, \ldots, C_m$ be convex subsets of V with $m \geq n+1$.

(a) Suppose that every $n+1$ of the sets have a common point. Show that all m of them have a common point.

(b) Let $n \geq 1$. Show that "$n+1$" cannot be replaced by "n" in (a).

Solution on page 1381

Problem 4.16. HARD Let $A \in \mathbb{R}^{3\times 2}$ and $B \in \mathbb{R}^{2\times 3}$ satisfy

$$AB = M = \begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}.$$

(a) Show that $BA = 9I_2$.

(b) Find such A and B.

Solution on page 1381

Problem 4.17. HARD For $n \geq 1$, let $M_n \in \mathbb{Q}^{(2n+1)\times(2n+1)}$ be the matrix with $M_n^{\mathsf{T}} = -M_n$ in which every entry in the first n diagonals below the main diagonal is 1 and every other entry below the main diagonal is -1. Find $\operatorname{rank} M_n$.

Solution on page 1382

Problem 4.18. CHALLENGE BEYOND 21-242 For vector spaces V and W, possibly infinite-dimensional, call $T \in \mathrm{L}(V,W)$ **Fredholm** if $\ker T$ and $W/\operatorname{im} T$ are finite-dimensional, and define its **index** $\operatorname{ind} T = \dim\ker T - \dim(W/\operatorname{im} T)$.

(a) Show that if V and W are finite-dimensional, every $T \in \mathrm{L}(V,W)$ is Fredholm with $\operatorname{ind} T = \dim V - \dim W$.

(b) Let M and D be multiplication by x and differentiation on $\mathbb{Q}[x]$. Show that M and D are Fredholm, and compute their indices.

(c) Let $T \in \mathrm{L}(U,V)$ and $S \in \mathrm{L}(V,W)$ be Fredholm. Show that ST is Fredholm and $\operatorname{ind}(ST) = \operatorname{ind} S + \operatorname{ind} T$. Check this for DM and MD on $\mathbb{Q}[x]$.

Solution on page 1383

Chapter 5

Determinants

The determinant from its axioms: permutations, existence and uniqueness, invertibility, multiplicativity, and cofactor expansion.

5.1 What a determinant should be

A square matrix $A \in F^{n\times n}$ gives a linear map $T_A : F^n \to F^n$. Over $\mathbb{R}$ we can watch T_A act on a region of the plane or of space: it stretches the region, shears it, and possibly turns it over. Two features of this action are recorded by a single number, the determinant $\det A$. Its absolute value is the factor by which T_A scales areas (or volumes), and its sign says whether T_A reverses orientation. We will not start from a formula for $\det A$. Instead, this section lists the properties that a signed area must have, takes them as axioms, and shows that for 2×2 matrices they leave exactly one possibility, $ad - bc$. The same axioms determine $\det A$ for $n \times n$ matrices as well, but the proof needs permutations (Section 5.2) and is given in Section 5.3.

5.1.1 Signed area

Let $A \in \mathbb{R}^{2\times 2}$ have columns $\mathbf{a}_1 = A\mathbf{e}_1$ and $\mathbf{a}_2 = A\mathbf{e}_2$. By linearity, $T_A(s\mathbf{e}_1 + t\mathbf{e}_2) = s\mathbf{a}_1 + t\mathbf{a}_2$, so T_A sends the unit square

$$Q = \{s\mathbf{e}_1 + t\mathbf{e}_2 \mid 0 \le s, t \le 1\}$$

onto the parallelogram

$$P(\mathbf{a}_1, \mathbf{a}_2) = \{s\mathbf{a}_1 + t\mathbf{a}_2 \mid 0 \le s, t \le 1\}.$$

The same computation shows that T_A sends each small grid square $\mathbf{p} + hQ$ onto the parallelogram $A\mathbf{p} + hP(\mathbf{a}_1, \mathbf{a}_2)$. So T_A multiplies the area of every grid square by the same factor, the area of $P(\mathbf{a}_1, \mathbf{a}_2)$, and approximating a region by grid squares suggests that T_A multiplies every area by this factor. (Making the approximation precise is a matter for analysis. We use the picture only as motivation; Section 9.2 returns to it.)

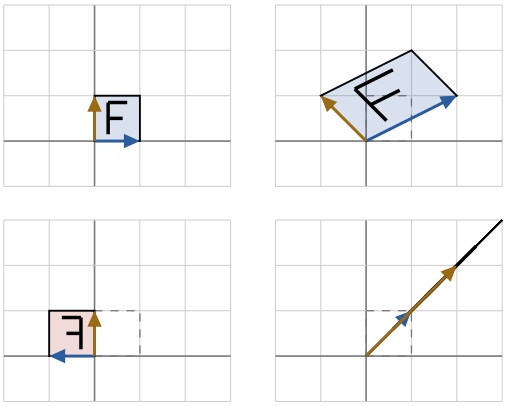


Figure 5.1.1. The unit square (dashed) and a letter F drawn in it, under four linear maps; the arrows are the images of $\mathbf{e}_1$ and $\mathbf{e}_2$. Top row: I_2, and $A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$, which triples areas. Bottom row: $R = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, which preserves areas but reverses orientation, and $B = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, which flattens the square. Blue shading means orientation is preserved, red that it is reversed.

Figure 5.1.1 shows four examples. The map T_A triples areas. The reflection T_R preserves areas but draws the F backwards, and no rotation of the plane turns a backwards F into an F. The map T_B squashes the square onto a segment, so it multiplies areas by 0. This matches the fact that B is not invertible: its second column is twice its first, so its columns are linearly dependent.

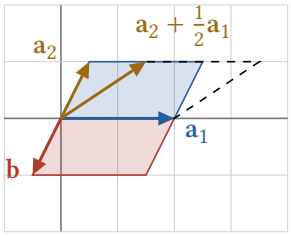
To tell T_R apart from the identity we use orientation. In Q, turning from $\mathbf{e}_1$ to $\mathbf{e}_2$ is a counterclockwise turn through an angle between 0 and π. Under T_A this is still true of the turn from $\mathbf{a}_1 = (2, 1)$ to $\mathbf{a}_2 = (-1, 1)$, but under T_R the turn from $\mathbf{a}_1 = (-1, 0)$ to $\mathbf{a}_2 = (0, 1)$ is clockwise. We say that T_A **preserves orientation** and T_R **reverses** it. Accordingly, the **signed area** of $P(\mathbf{a}_1, \mathbf{a}_2)$ is its area when the turn from $\mathbf{a}_1$ to $\mathbf{a}_2$ is counterclockwise, the negative of its area when the turn is clockwise, and 0 when $\mathbf{a}_1$ and $\mathbf{a}_2$ lie on one line through the origin. The determinant of a real 2×2 matrix should be the signed area of the parallelogram spanned by its columns. Its absolute value is then the area scaling factor of T_A, and its sign tells whether T_A preserves orientation.

Intuition (Signed area is linear in each column). Write $S(\mathbf{a}_1, \mathbf{a}_2)$ for the signed area of $P(\mathbf{a}_1, \mathbf{a}_2)$. We do not yet have a formula for S, but three of its properties can be read off from pictures.

- $S(\mathbf{e}_1, \mathbf{e}_2) = 1$, since $P(\mathbf{e}_1, \mathbf{e}_2)$ is the unit square.
- $S(\mathbf{v}, \mathbf{v}) = 0$ for every $\mathbf{v}$, since $P(\mathbf{v}, \mathbf{v})$ is a segment.
- S is linear in each argument when the other argument is fixed.

The first two are clear. For the third, compute area as base times height. Fix $\mathbf{a}_1 \neq \mathbf{0}$ as the base, and let ℓ be the line through $\mathbf{0}$ and $\mathbf{a}_1$. The area of $P(\mathbf{a}_1, \mathbf{a}_2)$ is the length of $\mathbf{a}_1$ times the distance from $\mathbf{a}_2$ to ℓ. Count that distance as positive when $\mathbf{a}_2$ lies on the side of ℓ that $\mathbf{a}_1$ turns toward when it turns counterclockwise, and negative on the other side. With this signed height $h(\mathbf{a}_2)$, the signed area is the length of $\mathbf{a}_1$ times $h(\mathbf{a}_2)$. If $\mathbf{a}_1$ points along the positive x-axis, then $h(x, y) = y$, a linear function of (x, y). Rotating the whole picture changes neither areas nor orientations, so h is linear whatever the direction of $\mathbf{a}_1$, and so is $S(\mathbf{a}_1, \cdot)$. Finally, reversing the order of $\mathbf{a}_1$ and $\mathbf{a}_2$ reverses the direction of the turn, so $S(\mathbf{a}_2, \mathbf{a}_1) = -S(\mathbf{a}_1, \mathbf{a}_2)$, and linearity in the second argument gives linearity in the first.

Two special cases are worth seeing. Adding a multiple of $\mathbf{a}_1$ to $\mathbf{a}_2$ slides $\mathbf{a}_2$ parallel to ℓ, which changes neither the base nor the height, so $S(\mathbf{a}_1, \mathbf{a}_2 + c\mathbf{a}_1) = S(\mathbf{a}_1, \mathbf{a}_2)$. Moving $\mathbf{a}_2$ across ℓ changes the sign of S. Try both below.



Drag $A\mathbf{e}_2$, or focus it and use the arrow keys. Sliding it parallel to $A\mathbf{e}_1$ leaves the signed area unchanged, and moving it across the line through $A\mathbf{e}_1$ flips the sign; the letter F in the square is mirrored exactly when the sign is negative. The formula for the signed area in terms of the entries comes out of the axioms in Theorem 5.1.13. (Interactive in the web version.)

The signs are what make linearity possible. Unsigned area is not linear in either argument: $P(\mathbf{e}_1, \mathbf{e}_2)$ and $P(-\mathbf{e}_1, \mathbf{e}_2)$ both have area 1, but $P(\mathbf{e}_1 + (-\mathbf{e}_1), \mathbf{e}_2) = P(\mathbf{0}, \mathbf{e}_2)$ has area 0, which is not $1 + 1$. With signs, $S(\mathbf{e}_1, \mathbf{e}_2) + S(-\mathbf{e}_1, \mathbf{e}_2) = 1 + (-1) = 0$, as linearity requires. So any function that agrees with area up to sign and is linear in each argument must take negative values, and the sign it carries is the orientation.

5.1.2 The determinant axioms

The three properties of S make no reference to area or to $\mathbb{R}$, and they make sense for n vectors in F^n over any field. We take them as the definition. For a function D on $F^{n\times n}$ and a matrix A with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, we write $D(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ for $D(A)$. This lets us treat D as a function of n vectors in F^n, the columns of its argument.

Definition 5.1.1 (Determinant function). Let $n \geq 1$, and let $D : F^{n\times n} \to F$.

- D is **normalized** if $D(I_n) = 1$, that is, $D(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$.
- D is **alternating** if $D(A) = 0$ whenever two columns of A are equal, that is, whenever $\mathbf{a}_i = \mathbf{a}_j$ for some $i \neq j$.
- D is **multilinear** if it is linear in each column when the other columns are held fixed. That is, for each $j \in [n]$ and all fixed $\mathbf{a}_i$ with $i \neq j$, the map $F^n \to F$ given by

$$\mathbf{x} \mapsto D\big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{x}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\big)$$

  is linear. Concretely,

$$D(\ldots, c\mathbf{x} + \mathbf{y}, \ldots) = cD(\ldots, \mathbf{x}, \ldots) + D(\ldots, \mathbf{y}, \ldots)$$

  for all $\mathbf{x}, \mathbf{y} \in F^n$ and $c \in F$, where the dots stand for the same fixed columns in all three places.

A **determinant function** on $F^{n\times n}$ is a function $D : F^{n\times n} \to F$ that is normalized, alternating, and multilinear. These three conditions are the **determinant axioms**.

So a determinant function is what the motivation asks for: the unit cube has volume 1, a parallelepiped with two equal edges is flat and has volume 0, and signed volume is linear in each edge. In Section 5.3 we prove that for every n there is exactly one determinant function on $F^{n\times n}$ (Theorem 5.3.4). It is called the **determinant** and written det. Until then, D denotes a function satisfying some of the axioms, and each result below says which ones it uses. Several results need only multilinearity and alternation, so they apply to functions such as $D = 0$ as well.

Example 5.1.2 ($n = 1$). A 1×1 matrix is a scalar, so $D : F \to F$. The alternating condition is empty, since there is only one column. Multilinearity says that D is linear, so $D(a) = D(a \cdot 1) = aD(1)$ for every a, and normalization says $D(1) = 1$. Hence $D(a) = a$ is the only determinant function on $F^{1\times 1}$, and it is one.

Non-example 5.1.3 (One axiom at a time). Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, whose columns are (a, c) and (b, d). Each function below satisfies two of the axioms and fails the third.

(a) $D(A) = ad$ is normalized, and it is multilinear: for a fixed second column, $(a, c) \mapsto da$ is linear, and for a fixed first column, $(b, d) \mapsto ad$ is linear. It is not alternating, since $D\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1$.

(b) $D = 0$ is alternating and multilinear but not normalized.

(c) Over $\mathbb{R}$, the unsigned area $D(A) = |ad - bc|$ is normalized and alternating (if the columns are equal, then $ad - bc = ac - ca = 0$). It is not multilinear: $D(-\mathbf{e}_1, \mathbf{e}_2) = 1$, while linearity in the first column would force $D(-\mathbf{e}_1, \mathbf{e}_2) = -D(\mathbf{e}_1, \mathbf{e}_2) = -1$.

So none of the three axioms follows from the other two.

Remark 5.1.4. Multilinear does not mean linear. A multilinear D is linear in one column at a time, and scaling the whole matrix scales every column, so $D(cA) = c^n D(A)$, not $cD(A)$. Similarly $D(A + B)$ is usually not $D(A) + D(B)$; Exercise 5.1.21 has an example.

> **Exercise 5.1.5.** ROUTINE Let $D : F^{n\times n} \to F$ be multilinear. Show that $D(cA) = c^n D(A)$ for every $c \in F$ and $A \in F^{n\times n}$. In particular $D(-A) = (-1)^n D(A)$, and $D(A) = 0$ whenever A has a zero column.

Solution on page 1384

Remark 5.1.6. We state the axioms for columns because the columns of A are the vectors $A\mathbf{e}_j$, so they describe what T_A does to the standard basis. The same axioms for rows would serve equally well, and Theorem 5.4.14 shows that they lead to the same function.

Multilinearity and alternation also make sense for a function $f : V \times \cdots \times V \to F$ of k vectors in any vector space V. A multilinear such f is a *k-linear form* on V, and an *alternating* k-linear form if it is also alternating. The proofs of Proposition 5.1.10 and Proposition 5.1.17 below use nothing about columns of matrices, so those results hold for alternating k-linear forms with the same proofs. Section 13.4 develops this point of view and uses it to define the determinant of an operator without choosing a basis.

5.1.3 Swapping columns

The axioms say nothing directly about what happens when two columns are interchanged. In the plane, interchanging $\mathbf{a}_1$ and $\mathbf{a}_2$ reverses the direction of the turn from one to the other, so we expect the sign of D to flip. Proving this is our first use of the basic move for multilinear functions.

> **Technique 5.1.7** (Expand by multilinearity). If an argument of a multilinear function is a sum, split it: linearity in slot j turns $D(\ldots, \sum_k c_k \mathbf{x}_k, \ldots)$ into $\sum_k c_k D(\ldots, \mathbf{x}_k, \ldots)$. Doing this in several slots turns one value of D into a sum over all ways of choosing one term in each slot, so m terms in each of r slots give m^r terms. Usually the terms are standard basis vectors, or other vectors on which D is known, and the expansion reduces $D(A)$ to finitely many known values. When D is also alternating, every term that uses the same vector in two slots vanishes, and typically most of the sum disappears.
>
> Reach for it when you need a value of a multilinear function on general inputs and you know its values on special ones, much as a linear map is determined by its values on a basis (Technique 2.2.7). The canonical example is the 2×2 determinant (Theorem 5.1.13). In Section 5.3 the same expansion, carried out in all n columns, produces a formula for D on $F^{n\times n}$.
>
> *Used in.* Lemma 5.1.8, Proposition 5.1.10, Theorem 5.1.13, Proposition 5.1.17, Exercise 5.1.20, Exercise 5.1.33, Exercise 5.1.37, Exercise 5.1.38, Exercise 5.1.40, Exercise 5.2.44, Lemma 5.3.2, Theorem 5.3.3, Exercise 5.3.26, Theorem 5.5.3, Theorem 5.5.14, Theorem 5.5.27, Exercise 5.5.48, Exercise 5.5.52, Exercise 5.6.8, Problem 5.3, Problem 5.6, Proposition 6.1.22, Proposition 6.1.29, Corollary 6.3.6, Exercise 7.3.53, Exercise 6.1.44, Exercise 6.1.57, Exercise 6.3.56.

The simplest expansion has just two terms, one of which vanishes.

Lemma 5.1.8 (Column replacement). Let $D : F^{n\times n} \to F$ be multilinear and alternating, let $i \neq j$, and let $c \in F$. If A' is obtained from A by adding c times column j to column i, then $D(A') = D(A)$.

Proof. ROUTINE

Uses. Definition 5.1.1 (Determinant function)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Column i of A' is a sum, $\mathbf{a}_i + c\mathbf{a}_j$. Expand $D(A')$ in column i.
Hint 2. One of the two terms is $D(A)$. In the other, which vector appears in two columns?
Sketch. Linearity in column i gives $D(A') = D(A) + cD(A'')$, where A'' is A with column i replaced by $\mathbf{a}_j$. Then A'' has equal columns i and j, so $D(A'') = 0$.

Let A'' be the matrix obtained from A by replacing column i with $\mathbf{a}_j$. The matrices A, A', and A'' agree outside column i, where they have $\mathbf{a}_i$, $\mathbf{a}_i + c\mathbf{a}_j$, and $\mathbf{a}_j$. By linearity in column i,

$$D(A') = D(A) + cD(A'').$$

Columns i and j of A'' are both $\mathbf{a}_j$, and $i \neq j$, so $D(A'') = 0$ because D is alternating. Hence $D(A') = D(A)$.□

This is the algebraic form of the shear in Intuition (Signed area is linear in each column): sliding one edge of a parallelogram parallel to another edge does not change its signed area. Section 5.4 combines this lemma with the other column operations to compute determinants by elimination.

Definition 5.1.9 (Antisymmetric). A function $D : F^{n\times n} \to F$ is **antisymmetric** if interchanging two columns negates its value: whenever A' is obtained from A by interchanging columns i and j for some $i \neq j$, we have $D(A') = -D(A)$.

For instance, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$ is antisymmetric, since interchanging the columns gives $bc - ad$. The function $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad$ is not: interchanging the columns of I_2 changes its value from 1 to 0, not to -1.

Every result in this subsection fixes all columns except two, so we set up a shorthand. Fix $i \neq j$ and the columns $\mathbf{a}_k$ for $k \notin \{i, j\}$, and for $\mathbf{x}, \mathbf{y} \in F^n$ let $f(\mathbf{x}, \mathbf{y})$ be the value of D at the matrix with $\mathbf{x}$ in column i, $\mathbf{y}$ in column j, and $\mathbf{a}_k$ in every other column k. If D is multilinear, then f is linear in each of $\mathbf{x}$ and $\mathbf{y}$ when the other is fixed; if D is alternating, then $f(\mathbf{x}, \mathbf{x}) = 0$; and D is antisymmetric exactly when $f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y})$ for every such choice of i, j, and the fixed columns.

Proposition 5.1.10 (Alternating implies antisymmetric). If $D : F^{n\times n} \to F$ is multilinear and alternating, then D is antisymmetric.

Proof (by expanding). ROUTINE

Uses. Definition 5.1.1 (Determinant function), Definition 5.1.9 (Antisymmetric)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Only columns i and j change, so work with the two-variable function f above. You want $f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y})$.
Hint 2. Alternation only says something when two columns are equal. Which matrix with two equal columns involves both $\mathbf{x}$ and $\mathbf{y}$?
Hint 3. Expand $0 = f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$ in both slots.
Sketch. Expanding $0 = f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$ gives four terms. Alternation kills $f(\mathbf{x}, \mathbf{x})$ and $f(\mathbf{y}, \mathbf{y})$, leaving $f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) = 0$.

Fix $i \neq j$ and the other columns, and let f be as above. Let $\mathbf{x}, \mathbf{y} \in F^n$. Since D is alternating, $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = 0$. Expanding by linearity in the first slot and then in the second,

$$\begin{aligned} 0 &= f(\mathbf{x}, \mathbf{x} + \mathbf{y}) + f(\mathbf{y}, \mathbf{x} + \mathbf{y}) \\ &= f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y}) \\ &= f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}), \end{aligned}$$

where the last step uses $f(\mathbf{x},\mathbf{x}) = f(\mathbf{y},\mathbf{y}) = 0$, again because D is alternating. Hence $f(\mathbf{y},\mathbf{x}) = -f(\mathbf{x},\mathbf{y})$. As i, j, the other columns, $\mathbf{x}$, and $\mathbf{y}$ were arbitrary, D is antisymmetric. □

The second proof avoids the four-term expansion. It performs the swap with moves whose effect on D we already know: three column replacements, which do not change D (Lemma 5.1.8), and one sign change. In the plane, the three replacements are shears, and together they rotate the pair $(\mathbf{x},\mathbf{y})$ to $(\mathbf{y},-\mathbf{x})$.

Proof (by column replacements). ROUTINE

> *Uses.* Lemma 5.1.8 (Column replacement), Definition 5.1.1 (Determinant function), Definition 5.1.9 (Antisymmetric)
> *Hint 1.* Lemma 5.1.8 says that $f(\mathbf{x}+c\mathbf{y},\mathbf{y}) = f(\mathbf{x},\mathbf{y})$ and $f(\mathbf{x},\mathbf{y}+c\mathbf{x}) = f(\mathbf{x},\mathbf{y})$. Try to reach $f(\mathbf{y},\pm\mathbf{x})$ by such moves.
> *Hint 2.* Start with $f(\mathbf{x},\mathbf{y}) = f(\mathbf{x}+\mathbf{y},\mathbf{y})$. Now subtract the first slot from the second.
> *Sketch.* Three column replacements take $(\mathbf{x},\mathbf{y})$ to $(\mathbf{x}+\mathbf{y},\mathbf{y})$, then to $(\mathbf{x}+\mathbf{y},-\mathbf{x})$, then to $(\mathbf{y},-\mathbf{x})$. None of them changes f, and pulling out the factor -1 gives $f(\mathbf{x},\mathbf{y}) = -f(\mathbf{y},\mathbf{x})$.

Fix $i \neq j$ and the other columns, let f be as above, and let $\mathbf{x},\mathbf{y} \in F^n$. Each of the first three steps below adds a multiple of one slot to the other, so it does not change f by Lemma 5.1.8:

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= f(\mathbf{x}+\mathbf{y},\mathbf{y}) && \text{(add slot 2 to slot 1)} \\ &= f(\mathbf{x}+\mathbf{y},\mathbf{y}-(\mathbf{x}+\mathbf{y})) = f(\mathbf{x}+\mathbf{y},-\mathbf{x}) && \text{(subtract slot 1 from slot 2)} \\ &= f(\mathbf{x}+\mathbf{y}+(-\mathbf{x}),-\mathbf{x}) = f(\mathbf{y},-\mathbf{x}) && \text{(add slot 2 to slot 1)} \\ &= -f(\mathbf{y},\mathbf{x}) && \text{(linearity in slot 2).} \end{aligned}$$

So $f(\mathbf{y},\mathbf{x}) = -f(\mathbf{x},\mathbf{y})$, and D is antisymmetric. □

The first proof is the more economical one. The second shows that a swap is built from shears and a single reflection, which is how Section 5.4 analyzes column operations.

Antisymmetry looks like a reasonable alternative axiom: it is exactly the orientation reversal we observed in the plane. For multilinear functions the two conditions are equivalent over most fields, by Proposition 5.1.11 below. But the converse fails over fields in which $1+1=0$ (as soon as $n \geq 2$), and that failure is why we did not choose antisymmetry.

Proposition 5.1.11 (Antisymmetric implies alternating when $1+1 \neq 0$). Suppose $1+1 \neq 0$ in F. If $D : F^{n\times n} \to F$ is antisymmetric, then D is alternating. Consequently, a multilinear function on $F^{n\times n}$ is alternating if and only if it is antisymmetric.

Proof. ROUTINE

> *Uses.* Definition 5.1.9 (Antisymmetric), Proposition 1.1.17 (No zero divisors), Proposition 5.1.10 (Alternating implies antisymmetric)
> *Hint 1.* Suppose columns i and j of A are equal. What does interchanging them do to A?
> *Hint 2.* It does not change A, so antisymmetry gives $D(A) = -D(A)$. Rewrite this as $(1+1)D(A) = 0$.
> *Sketch.* If $\mathbf{a}_i = \mathbf{a}_j$ with $i \neq j$, interchanging these columns leaves A unchanged, so $D(A) = -D(A)$ and $(1+1)D(A) = 0$. Since $1+1 \neq 0$, $D(A) = 0$. The "consequently" combines this with Proposition 5.1.10.

Let $A \in F^{n\times n}$ with $\mathbf{a}_i = \mathbf{a}_j$ for some $i \neq j$. Interchanging columns i and j of A gives A again, so antisymmetry gives $D(A) = -D(A)$. Adding $D(A)$ to both sides, $D(A) + D(A) = 0$, that is, $(1+1)D(A) = 0$ by distributivity. Since $1+1 \neq 0$ and a field has no zero divisors (Proposition 1.1.17), $D(A) = 0$. So D is alternating.

For the last sentence: a multilinear alternating function is antisymmetric by Proposition 5.1.10, and an antisymmetric function is alternating by what we just proved. □

Note that this direction does not use multilinearity at all. The condition $1+1 \neq 0$ says that $\operatorname{char} F \neq 2$ (Definition 1.1.28). It holds over $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{F}_p$ for odd p, and fails over $\mathbb{F}_2$, where $-1 = 1$ and antisymmetry says only that interchanging two columns changes nothing.

Example 5.1.12 (Antisymmetric but not alternating, over $\mathbb{F}_2$). On $\mathbb{F}_2^{2\times 2}$, define

$$D\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad + bc + ab.$$

For a fixed second column (b, d), the map $(a, c) \mapsto ad + bc + ab = (d + b)a + bc$ is linear, and for a fixed first column, $(b, d) \mapsto (c + a)b + ad$ is linear, so D is multilinear. Interchanging the columns gives $bc + ad + ba$, the same value, and since $-1 = 1$ in $\mathbb{F}_2$ this means D is antisymmetric. Also $D(I_2) = 1 + 0 + 0 = 1$. But D is not alternating, since $D\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 0 + 0 + 1 = 1$.

So over $\mathbb{F}_2$ the axioms "normalized, antisymmetric, multilinear" are satisfied by at least two different functions: this D, and $ad - bc = ad + bc$, which we will see is a determinant function (Theorem 5.1.13). With alternation in place of antisymmetry, the 2×2 case has exactly one solution over every field. This is the reason the definition uses alternation, and an instance of the warning in Aside (Characteristic) that arguments dividing by $1 + 1$ fail in characteristic 2.

5.1.4 The 2×2 determinant

We now show that the axioms determine D completely on $F^{2\times 2}$.

Intuition (Strategy). The axioms tell us the value of D on very few matrices: $D(\mathbf{e}_1, \mathbf{e}_1) = D(\mathbf{e}_2, \mathbf{e}_2) = 0$ by alternation, $D(\mathbf{e}_2, \mathbf{e}_1) = -D(\mathbf{e}_1, \mathbf{e}_2)$ by Proposition 5.1.10, and $D(\mathbf{e}_1, \mathbf{e}_2) = D(I_2) = 1$ by normalization. Multilinearity lets us reach every other matrix from these four. Write each column as a combination of $\mathbf{e}_1$ and $\mathbf{e}_2$ and expand (Technique 5.1.7): $D(A)$ becomes a combination of the four known values, with coefficients that are products of entries of A. Since the expansion uses only multilinearity and alternation, we keep $D(I_2)$ as an unknown until the end; this gives a stronger statement for free.

Theorem 5.1.13 (The 2×2 determinant).

(a) Let $D : F^{2\times 2} \to F$ be multilinear and alternating. Then for all $a, b, c, d \in F$,

$$D\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ad - bc)D(I_2).$$

(b) The function $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$ is a determinant function on $F^{2\times 2}$.

Consequently, there is exactly one determinant function on $F^{2\times 2}$, namely $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc$.

Proof, part 1 of 2: the axioms force $ad - bc$. STANDARD

Uses. Definition 5.1.1 (Determinant function), Proposition 5.1.10 (Alternating implies antisymmetric), Definition 1.6.8 (Standard bases)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Write each column of A as a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$.
Hint 2. Expand $D(a\mathbf{e}_1 + c\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2)$ by multilinearity, first in column 1, then in column 2. You get four terms.
Hint 3. The four terms involve $D(\mathbf{e}_1, \mathbf{e}_1)$, $D(\mathbf{e}_1, \mathbf{e}_2)$, $D(\mathbf{e}_2, \mathbf{e}_1)$, and $D(\mathbf{e}_2, \mathbf{e}_2)$. Evaluate them with alternation and Proposition 5.1.10.
Sketch. The columns are $a\mathbf{e}_1 + c\mathbf{e}_2$ and $b\mathbf{e}_1 + d\mathbf{e}_2$. Expanding gives $abD(\mathbf{e}_1, \mathbf{e}_1) + adD(\mathbf{e}_1, \mathbf{e}_2) + cbD(\mathbf{e}_2, \mathbf{e}_1) + cdD(\mathbf{e}_2, \mathbf{e}_2)$. Alternation kills the first and last terms, and antisymmetry turns $D(\mathbf{e}_2, \mathbf{e}_1)$ into $-D(\mathbf{e}_1, \mathbf{e}_2) = -D(I_2)$.

(a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Its columns are $a\mathbf{e}_1 + c\mathbf{e}_2$ and $b\mathbf{e}_1 + d\mathbf{e}_2$. By linearity in the first column,

$$D(A) = aD(\mathbf{e}_1, b\mathbf{e}_1 + d\mathbf{e}_2) + cD(\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2),$$

and by linearity in the second column,

$$D(A) = abD(\mathbf{e}_1, \mathbf{e}_1) + adD(\mathbf{e}_1, \mathbf{e}_2) + cbD(\mathbf{e}_2, \mathbf{e}_1) + cdD(\mathbf{e}_2, \mathbf{e}_2).$$

Since D is alternating, $D(\mathbf{e}_1, \mathbf{e}_1) = D(\mathbf{e}_2, \mathbf{e}_2) = 0$. By Proposition 5.1.10, $D(\mathbf{e}_2, \mathbf{e}_1) = -D(\mathbf{e}_1, \mathbf{e}_2)$, and $D(\mathbf{e}_1, \mathbf{e}_2) = D(I_2)$. Therefore

$$D(A) = adD(I_2) - bcD(I_2) = (ad - bc)D(I_2),$$

as desired.

Part (a) shows that there is at most one determinant function on $F^{2\times 2}$: if $D(I_2) = 1$, then $D(A) = ad - bc$. It remains to check that this candidate satisfies the axioms. For $n = 2$ this is a direct computation. For general n the analogous check is the existence half of Theorem 5.3.4, and it is where permutations enter.

Proof, part 2 of 2: $ad - bc$ satisfies the axioms. ROUTINE

> *Uses.* Definition 5.1.1 (Determinant function), Definition 2.1.1 (Linear map)
> *Hint 1.* Check the three axioms separately. For multilinearity, fix one column and look at $ad - bc$ as a function of the entries of the other column.
> *Hint 2.* For a fixed second column (b, d), the value is $da - bc$, which is a fixed linear combination of the entries a and c of the first column.
> *Sketch.* For a fixed second column, $(a, c) \mapsto da - bc$ is linear; for a fixed first column, $(b, d) \mapsto ad - cb$ is linear. Equal columns give $ac - ca = 0$, and I_2 gives $1 \cdot 1 - 0 \cdot 0 = 1$.

(b) Let $\Delta\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

Normalized. $\Delta(I_2) = 1 \cdot 1 - 0 \cdot 0 = 1$.

Alternating. If the two columns are equal, then $b = a$ and $d = c$, so $\Delta = ac - ac = 0$.

Multilinear. Fix the second column (b, d). For first columns (a, c) and (a', c') and $s \in F$,

$$\begin{aligned}\Delta\begin{pmatrix} sa + a' & b \\ sc + c' & d \end{pmatrix} &= (sa + a')d - b(sc + c') \\ &= s(ad - bc) + (a'd - bc') = s\Delta\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \Delta\begin{pmatrix} a' & b \\ c' & d \end{pmatrix},\end{aligned}$$

so Δ is linear in the first column. Similarly, for a fixed first column (a, c),

$$a(sd + d') - (sb + b')c = s(ad - bc) + (ad' - b'c),$$

so Δ is linear in the second column. Hence Δ is a determinant function.

Finally, if D is any determinant function on $F^{2\times 2}$, then $D(I_2) = 1$ and (a) gives $D = \Delta$. So Δ is the only one.□

Definition 5.1.14 (The 2×2 determinant). For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$, the **determinant** of A is $\det A = ad - bc$. By Theorem 5.1.13 it is the unique determinant function on $F^{2\times 2}$, and every multilinear alternating $D : F^{2\times 2} \to F$ is a scalar multiple of it: $D = D(I_2)\det$.

For the matrices of Figure 5.1.1, $\det A = 2 \cdot 1 - (-1) \cdot 1 = 3$, $\det R = -1$, and $\det B = 1 \cdot 2 - 2 \cdot 1 = 0$, as the pictures suggested. In general, the formula is the signed area of the motivation.

Example 5.1.15 (det is signed area). Let $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$ be nonzero, and write them in polar form, $\mathbf{a}_1 = r(\cos\alpha, \sin\alpha)$ and $\mathbf{a}_2 = s(\cos\beta, \sin\beta)$ with $r, s > 0$. Then

$$\det(\mathbf{a}_1, \mathbf{a}_2) = rs(\cos\alpha \sin\beta - \sin\alpha \cos\beta) = rs\sin(\beta - \alpha).$$

The parallelogram $P(\mathbf{a}_1, \mathbf{a}_2)$ has base r and height $s|\sin(\beta - \alpha)|$, so $|\det(\mathbf{a}_1, \mathbf{a}_2)|$ is its area. Also $\sin(\beta - \alpha) > 0$ exactly when $\beta - \alpha$ lies strictly between 0 and π modulo 2π, that is, when $\mathbf{a}_2$ is reached from $\mathbf{a}_1$ by a counterclockwise turn of less than a half-turn. So $\det(\mathbf{a}_1, \mathbf{a}_2)$ is the signed area $S(\mathbf{a}_1, \mathbf{a}_2)$ of Intuition (Signed area is linear in each column). (If $\mathbf{a}_1$ or $\mathbf{a}_2$ is $\mathbf{0}$, both are 0.) Together with part (b) of Theorem 5.1.13, this proves the three properties of signed area that we read off from pictures. The general statement that $|\det A|$ is the factor by which T_A scales volume needs a theory of volume; see Section 9.2.

Exercise 5.1.16. ROUTINE Let $A = \begin{pmatrix} 3 & 5 \\ 4 & 2 \end{pmatrix}$. Compute $\det A$ when A is regarded as a matrix over $\mathbb{Q}$, over $\mathbb{F}_7$, and over $\mathbb{F}_2$ (reading each integer as its remainder). For which primes p is $\det A = 0$ over $\mathbb{F}_p$?

Solution on page 1384

Intuition (What the proof did, and where it leads). The expansion in Theorem 5.1.13 produced one term for each way of choosing $\mathbf{e}_1$ or $\mathbf{e}_2$ in each column. The two terms that chose the same vector twice died, and the two survivors correspond to the two orders of $\mathbf{e}_1, \mathbf{e}_2$: the identity order, with value $D(I_2)$, and the swapped order, with value $-D(I_2)$.

For an $n \times n$ matrix, write column j as $\sum_i a_{ij}\mathbf{e}_i$ and expand all n columns. This gives n^n terms, each a product of entries times some $D\big(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}\big)$. Alternation kills every term with a repeated index, leaving one term for each ordering $i_1, \ldots, i_n$ of $1, \ldots, n$. Each surviving value can be computed by interchanging columns until the order is $1, \ldots, n$, picking up a factor -1 per interchange. For example, two interchanges sort $\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1$:

$$D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) = -D(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2) = D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3).$$

Two questions remain. First, could one ordering be sorted both by an even and by an odd number of interchanges? If so, the axioms would force $D(I_n) = -D(I_n)$, so $D(I_n) = 0$ when $1 + 1 \neq 0$, and no determinant function could exist. Second, does the formula produced by the expansion actually satisfy the axioms? Section 5.2 answers the first question (no), and Section 5.3 answers the second (yes).

5.1.5 Dependent columns

In Figure 5.1.1, the dependent columns of B span a flat parallelogram, and its signed area is 0. The same holds in every dimension: if the columns of A are linearly dependent, then one column is a combination of the others, and expanding that column (Technique 5.1.7) produces only terms with a repeated column.

Proposition 5.1.17 (Dependent columns). Let $D : F^{n\times n} \to F$ be multilinear and alternating. If the columns of $A \in F^{n\times n}$ are linearly dependent, then $D(A) = 0$.

Proof (by expanding). ROUTINE

Uses. Definition 1.5.1 (Linear independence), Definition 5.1.1 (Determinant function), Exercise 5.1.5
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Use a nontrivial dependence relation to write one column $\mathbf{a}_k$ as a linear combination of the other columns.
Hint 2. Expand $D(A)$ in column k. Each term has some $\mathbf{a}_j$ with $j \neq k$ in column k. Where else does $\mathbf{a}_j$ appear?
Sketch. Some $\mathbf{a}_k$ equals $\sum_{j\neq k} b_j\mathbf{a}_j$. Linearity in column k gives $D(A) = \sum_{j\neq k} b_j D\big(A_j\big)$, where A_j has $\mathbf{a}_j$ in columns j and k. Each $D\big(A_j\big)$ is 0.

Since the columns are linearly dependent, there are scalars $c_1, \ldots, c_n$, not all 0, with $c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n = \mathbf{0}$. Choose k with $c_k \neq 0$. Then

$$\mathbf{a}_k = \sum_{j\neq k} b_j\mathbf{a}_j, \quad \text{where } b_j = -c_k^{-1}c_j.$$

If $n = 1$, the sum is empty, so $\mathbf{a}_1 = \mathbf{0}$ and $D(A) = 0$ by Exercise 5.1.5. Otherwise, for $j \neq k$ let A_j be the matrix obtained from A by replacing column k with $\mathbf{a}_j$. By linearity in column k,

$$D(A) = \sum_{j\neq k} b_j D\big(A_j\big).$$

For each $j \neq k$, columns j and k of A_j are both $\mathbf{a}_j$, so $D\big(A_j\big) = 0$ because D is alternating. Hence $D(A) = 0$. □

The second proof replaces the expansion by an invariance argument: change A, without changing $D(A)$, until D is obviously 0.

Proof (by column replacement). ROUTINE

> *Uses.* Definition 1.5.1 (Linear independence), Lemma 5.1.8 (Column replacement), Exercise 5.1.5
> *Hint 1.* Again write $\mathbf{a}_k = \sum_{j \neq k} b_j \mathbf{a}_j$.
> *Hint 2.* Use Lemma 5.1.8 to subtract $b_j \mathbf{a}_j$ from column k, one j at a time. What is column k at the end?
> *Sketch.* Subtracting $b_j \mathbf{a}_j$ from column k for each $j \neq k$ does not change D and turns column k into $\mathbf{0}$. A matrix with a zero column has $D = 0$.

As in the first proof, write $\mathbf{a}_k = \sum_{j \neq k} b_j \mathbf{a}_j$. For each $j \neq k$ in turn, add $-b_j$ times column j to column k. Columns other than k never change, so each step adds $-b_j \mathbf{a}_j$ to column k, and by Lemma 5.1.8 it does not change the value of D. After all the steps, column k is $\mathbf{a}_k - \sum_{j \neq k} b_j \mathbf{a}_j = \mathbf{0}$. So $D(A)$ equals the value of D at a matrix with a zero column, which is 0 by Exercise 5.1.5. □

The first proof is self-contained. The second generalizes better in practice: computing $D(A)$ by column operations that either preserve D or change it in a known way, until the matrix is simple, is how determinants are evaluated in Section 5.4.

Corollary 5.1.18 (Non-invertible matrices). Let $D : F^{n \times n} \to F$ be multilinear and alternating. If $A \in F^{n \times n}$ is not invertible, then $D(A) = 0$. Equivalently, if $D(A) \neq 0$, then A is invertible.

Proof. ROUTINE

> *Uses.* Proposition 5.1.17 (Dependent columns), Corollary 3.5.25 (The invertible matrix theorem), Theorem 3.5.12 (First Fundamental Theorem, completed)
> *Hint 1.* Relate invertibility of a square matrix to linear independence of its columns.
> *Hint 2.* For $A \in F^{n \times n}$, Theorem 3.5.12 says the columns are independent if and only if $\operatorname{rank} A = n$, and Corollary 3.5.25 says A is invertible if and only if $\operatorname{rank} A = n$.
> *Sketch.* If A is not invertible, then $\operatorname{rank} A \neq n$ (Corollary 3.5.25), so the columns of A are dependent (Theorem 3.5.12), and $D(A) = 0$ by Proposition 5.1.17.

Suppose $A \in F^{n \times n}$ is not invertible. By Corollary 3.5.25, $\operatorname{rank} A \neq n$. By Theorem 3.5.12, the columns of A are then not linearly independent. So $D(A) = 0$ by Proposition 5.1.17. The second statement is the contrapositive of the first. □

The converse is false for $D = 0$, which is multilinear and alternating but vanishes on every matrix. So any proof that $\det A \neq 0$ for invertible A must use normalization. For $n = 2$ it follows from the formula (Exercise 5.1.25), and for general n it is Theorem 5.4.7.

Exercises

Exercise 5.1.19. ROUTINE For each function $D : \mathbb{Q}^{2 \times 2} \to \mathbb{Q}$ below, where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, decide which of the three determinant axioms it satisfies. Then decide which of your answers change if $\mathbb{Q}$ is replaced by $\mathbb{F}_2$.

(a) $D(A) = a + d$.

(b) $D(A) = ab - cd$.

(c) $D(A) = bc - ad$.

(d) $D(A) = (a - b)(d - c)$.

Solution on page 1384

Exercise 5.1.20. ROUTINE Suppose D is a determinant function on $F^{3 \times 3}$. Using only the axioms and the results of this section, find $D(A)$ for

(a) $A = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 3 & 0 \\ 5 & 0 & 0 \end{pmatrix}$, (b) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, (c) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution on page 1385

Exercise 5.1.21. ROUTINE Find the error in the following argument.

Claim. If D is a determinant function on $F^{n\times n}$, then $D(A+B) = D(A) + D(B)$ for all $A, B \in F^{n\times n}$.

Proof. Column j of $A+B$ is $\mathbf{a}_j + \mathbf{b}_j$. Since D is linear in each column,

$$D(A+B) = D(\mathbf{a}_1 + \mathbf{b}_1, ..., \mathbf{a}_n + \mathbf{b}_n) = D(\mathbf{a}_1, ..., \mathbf{a}_n) + D(\mathbf{b}_1, ..., \mathbf{b}_n) = D(A) + D(B).$$

Solution on page 1385

Exercise 5.1.22. ROUTINE Find the error in the following argument.

Claim. If $D : F^{n\times n} \to F$ is multilinear and alternating and $D(A) = 0$, then the columns of A are linearly dependent.

Proof. By Proposition 5.1.17, a matrix with dependent columns has $D(A) = 0$. Here $D(A) = 0$, so the columns of A are dependent.

Solution on page 1385

Exercise 5.1.23. ROUTINE Let $a, b, c, d > 0$ be real numbers with $ad > bc$. Figure 5.1.2 shows the parallelogram P spanned by (a, c) and (b, d) inside the rectangle $[0, a+b] \times [0, c+d]$.

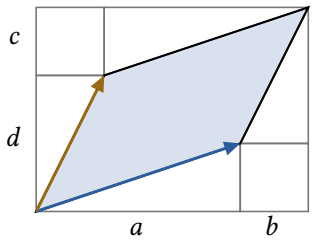


Figure 5.1.2. The parallelogram spanned by (a, c) and (b, d), drawn for $(a, b, c, d) = (3, 1, 1, 2)$. The rest of the rectangle splits into four right triangles and two rectangles.

(a) By computing the areas of the other pieces, show that P has area $ad - bc$.

(b) Explain why the hypothesis $ad > bc$ means that the turn from (a, c) to (b, d) is counterclockwise, so that the answer agrees with the signed area $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Solution on page 1386

Exercise 5.1.24. ROUTINE The **permanent** of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\operatorname{per} A = ad + bc$.

(a) Show that per is normalized and multilinear, and that interchanging the two columns of A does not change per A.

(b) For which fields F is per a determinant function on $F^{2\times 2}$?

Solution on page 1386

Exercise 5.1.25. STANDARD Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in F^{2\times 2}$.

(a) Show that $\det A = 0$ if and only if the columns of A are linearly dependent.

(b) Deduce that A is invertible if and only if $\det A \neq 0$, and that in this case

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Solution on page 1386

Exercise 5.1.26. ROUTINE

(a) Compute $\det\begin{pmatrix} 1+i & 2 \\ i & 1-i \end{pmatrix}$ over $\mathbb{C}$.

(b) Find all $z \in \mathbb{C}$ for which the columns of $\begin{pmatrix} z & i \\ i & z \end{pmatrix}$ are linearly dependent, and for each such z write down a dependence relation.

Solution on page 1387

Exercise 5.1.27. STANDARD Show that $\det(AB) = \det A \det B$ for all $A, B \in F^{2\times 2}$.

Solution on page 1387

Solution on page 1387

Exercise 5.1.28. ROUTINE For $\theta \in \mathbb{R}$, let $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ be the rotation by θ, and for $t \in \mathbb{R}$ let $S_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$. Show that

$$\det(R_\theta A) = \det A = \det(AS_t) \quad \text{for all } A \in \mathbb{R}^{2\times 2},$$

and explain what each equality says about the parallelogram spanned by the columns of A.

Solution on page 1388

Exercise 5.1.29. STANDARD Find all multilinear alternating functions $D : F^{2\times 2} \to F$ with $D\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 6$ when (a) $F = \mathbb{Q}$; (b) $F = \mathbb{F}_2$; (c) $F = \mathbb{F}_3$.

Solution on page 1388

Exercise 5.1.30. STANDARD For which fields F is the function $A \mapsto (\det A)^2$ a determinant function on $F^{2\times 2}$?

Solution on page 1388

Exercise 5.1.31. STANDARD Prove or disprove each statement.

(a) If $D : F^{2\times 2} \to F$ is multilinear and alternating and $D(A) = 0$ for some invertible A, then $D = 0$.

(b) If $D : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is alternating and $D(cA) = c^2 D(A)$ for all $c \in \mathbb{R}$ and all A, then D is a scalar multiple of det.

(c) If $D : F^{3\times 3} \to F$ is multilinear and alternating, then $D(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = D(\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_1)$ for all $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in F^3$.

Solution on page 1389

Exercise 5.1.32. STANDARD For $\mathbf{x}, \mathbf{y} \in F^3$, define the **cross product**

$$\mathbf{x} \times \mathbf{y} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1),$$

and define $D : F^{3\times 3} \to F$ by

$$D(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (\mathbf{x} \times \mathbf{y})_1 z_1 + (\mathbf{x} \times \mathbf{y})_2 z_2 + (\mathbf{x} \times \mathbf{y})_3 z_3.$$

Show that D is a determinant function on $F^{3\times 3}$. (Over $\mathbb{R}$ this is the scalar triple product. So determinant functions exist for $n = 3$; Section 5.3 treats all n.)

Solution on page 1389

Exercise 5.1.33. STANDARD Let $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix}$, and let $D : F^{3\times 3} \to F$ be multilinear and alternating.

(a) Show that $D(A) = 25D(I_3)$ by expanding each column in the standard basis.

(b) Regard A as a matrix over $\mathbb{F}_5$. Find a nontrivial linear dependence among its columns, and explain why this is consistent with (a).

(c) Regard A as a matrix over $\mathbb{Q}$. Show that A is invertible. (A determinant function on $\mathbb{Q}^{3\times 3}$ exists by Exercise 5.1.32.)

Solution on page 1389

Exercise 5.1.34. STANDARD Let Alt_n be the set of multilinear alternating functions $F^{n\times n} \to F$.

(a) Show that Alt_n is a subspace of the vector space of all functions $F^{n\times n} \to F$.

(b) Show that $\mathrm{Alt}_2 = \mathrm{span}(\det)$, and that $\dim \mathrm{Alt}_2 = 1$.

Solution on page 1390

Exercise 5.1.35. STANDARD Let $D : F^{n\times n} \to F$ be multilinear, and suppose that $D(A) = 0$ whenever two *adjacent* columns of A are equal, that is, whenever $\mathbf{a}_i = \mathbf{a}_{i+1}$ for some i. Show that D is alternating.

Solution on page 1390

Exercise 5.1.36. STANDARD Let $k > n$, and let $f : F^n \times \cdots \times F^n \to F$ (with k factors) be multilinear and alternating, in the sense of Remark 5.1.6. Show that $f = 0$. Interpret the result geometrically for $k = 3$ and $n = 2$.

Solution on page 1391

Exercise 5.1.37. STANDARD For $\mathbf{x}, \mathbf{y} \in F^3$ and $1 \le i < j \le 3$, let $f_{ij}(\mathbf{x}, \mathbf{y}) = x_i y_j - x_j y_i$.

(a) Show that f_{12}, f_{13}, and f_{23} are multilinear and alternating functions of two vectors in F^3.

(b) Show that they are linearly independent, as elements of the vector space of functions $F^3 \times F^3 \to F$.

(c) Show that every multilinear alternating $f : F^3 \times F^3 \to F$ satisfies $f = \sum_{i<j} f(\mathbf{e}_i, \mathbf{e}_j) f_{ij}$.

So for two vectors in F^3 there is a three-dimensional space of "signed areas", not a single one up to scaling. This is one reason determinants are only defined for square matrices.

Solution on page 1391

Exercise 5.1.38. STANDARD

(a) Show that $\det(A + tI_2) = t^2 + (\operatorname{tr} A)t + \det A$ for all $A \in F^{2\times 2}$ and $t \in F$.

(b) Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the rotation by $\pi/2$. Find all $t \in \mathbb{C}$ for which $A + tI_2$ is not invertible over $\mathbb{C}$, and all $t \in \mathbb{R}$ for which it is not invertible over $\mathbb{R}$.

Solution on page 1392

Exercise 5.1.39. STANDARD Let p be a prime.

(a) Count the matrices $A \in \mathbb{F}_p^{2\times 2}$ with $\det A \neq 0$.

(b) For $p = 3$, show that exactly half of them have $\det A = 1$.

(c) How many $A \in \mathbb{F}_2^{2\times 2}$ have $\det A = 1$?

Solution on page 1392

Exercise 5.1.40. HARD Let $n \geq 2$, and let $D : F^{n\times n} \to F$ be multilinear and alternating. Write $\mathbf{a}_{n+1} = \mathbf{a}_1$.

(a) Show that when $n = 3$,

$$D(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_2 + \mathbf{a}_3, \mathbf{a}_3 + \mathbf{a}_1) = 2D(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3).$$

(b) Show that in general

$$D(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_2 + \mathbf{a}_3, \ldots, \mathbf{a}_n + \mathbf{a}_{n+1}) = \left(1 + (-1)^{n-1}\right) D(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

(c) For even n, explain the answer to (b) directly by finding a linear dependence among the columns $\mathbf{a}_k + \mathbf{a}_{k+1}$.

Solution on page 1392

Exercise 5.1.41. HARD Let $D : F^{n\times n} \to F$ satisfy the following two conditions.

(i) Scaling a column scales D: $D(\ldots, c\mathbf{x}, \ldots) = cD(\ldots, \mathbf{x}, \ldots)$ for every column position, every $c \in F$, and all fixed other columns.

(ii) Column replacement preserves D: if A' is obtained from A by adding a multiple of one column to a different column, then $D(A') = D(A)$.

Show that D is multilinear and alternating. Since conversely every multilinear alternating function satisfies (i) and (ii) (by Lemma 5.1.8), multilinearity and alternation may be replaced by (i) and (ii) in the definition of a determinant function. These are the two properties of volume that are visible in the shear picture.

Solution on page 1393

Exercise 5.1.42. CHALLENGE Let $D : \mathbb{R}^{2\times 2} \to \mathbb{R}$ satisfy $D(AB) = D(A)D(B)$ for all A, B, and $D\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} = t$ for all $t \in \mathbb{R}$. Show that $D = \det$.

Solution on page 1394

5.2 Permutations

Let $D : F^{n\times n} \to F$ be multilinear and alternating. In Intuition (What the proof did, and where it leads) we wrote column j of A as $\sum_i a_{ij}\mathbf{e}_i$ and expanded $D(A)$ in all n columns. A term of the expansion chooses one index in each column, so it is indexed by a function $f : [n] \to [n]$, and it equals

$$a_{f(1)1}a_{f(2)2}\cdots a_{f(n)n} D\left(\mathbf{e}_{f(1)}, \mathbf{e}_{f(2)}, \ldots, \mathbf{e}_{f(n)}\right).$$

Alternation kills the term whenever f takes some value twice, so only the bijections f survive. This section studies these bijections, the permutations of $[n]$. To evaluate a surviving value $D\left(\mathbf{e}_{f(1)}, \ldots, \mathbf{e}_{f(n)}\right)$ we interchange two columns at a time until they are in order, and each interchange contributes a factor

-1. The main result, Theorem 5.2.18, says that the parity of the number of interchanges does not depend on how we sort. This gives every permutation σ a sign $\operatorname{sgn}(\sigma) = \pm 1$, and Theorem 5.2.31 shows that $D\big(\mathbf{e}_{\sigma(1)}, \dots, \mathbf{e}_{\sigma(n)}\big) = \operatorname{sgn}(\sigma) D(I_n)$.

5.2.1 Permutations

Throughout, $n \geq 1$ and $[n] = \{1, 2, \dots, n\}$.

Definition 5.2.1 (Permutation). A **permutation** of $[n]$ is a bijection $\sigma : [n] \to [n]$. The set of all permutations of $[n]$ is denoted S_n.

Since $[n]$ is finite, a function $f : [n] \to [n]$ is injective if and only if it is surjective, and hence if and only if it is bijective (Section A.1). So f is a permutation exactly when the list $f(1), f(2), \dots, f(n)$ has no repeated entries, and then this list contains each element of $[n]$ exactly once. In this sense a permutation is a rearrangement of $1, 2, \dots, n$. We record a permutation in **two-line notation**, writing each i above its image:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

The bottom row is the rearrangement.

Example 5.2.2 (S_1, S_2, and S_3). S_1 contains only the identity map of $[1]$, and S_2 contains the identity and the permutation $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$, which interchanges 1 and 2. The six elements of S_3 are

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

In the expansion of a 3×3 determinant these are the six orderings of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ whose terms survive alternation. For instance, the fourth gives the term $a_{21}a_{32}a_{13}D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$.

Non-example 5.2.3 (A function that is not a permutation). Define $f : [3] \to [3]$ by $f(1) = 2$, $f(2) = 2$, and $f(3) = 1$. Then f is not injective, so $f \notin S_3$. In the expansion of $D(A)$ for $A \in F^{3\times 3}$, the term indexed by f is $a_{21}a_{22}a_{13}D(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1)$, which is 0 because $\mathbf{e}_2$ appears in two columns. Every function that is not injective is killed in the same way.

Proposition 5.2.4 (Counting permutations). S_n has exactly $n!$ elements.

Proof. ROUTINE

Uses. Definition 5.2.1 (Permutation), injective maps $[n] \to [n]$ are bijective (Section A.1)
Hint 1. A permutation is determined by the list $\sigma(1), \dots, \sigma(n)$. Given the first k entries, how many choices are there for the next?
Hint 2. Show by induction on k that there are $n(n-1)\cdots(n-k+1)$ lists of k distinct elements of $[n]$.
Sketch. The permutations correspond to the lists of n distinct elements of $[n]$. There are n choices for the first entry, then $n-1$ for the second, and so on, which gives $n!$ lists; induction on the length makes this precise.

A permutation σ is determined by the list $\sigma(1), \dots, \sigma(n)$, and by the remark above, the lists that arise are exactly the lists of n distinct elements of $[n]$. We show by induction on k, for $0 \leq k \leq n$, that there are $n(n-1)\cdots(n-k+1)$ lists of k distinct elements of $[n]$, with the empty product 1 for $k = 0$ (the empty list). This is clear for $k = 0$. If $k < n$, a list of $k+1$ distinct elements consists of a list of k distinct elements followed by one of the $n-k$ elements not in it, and different choices give different lists. So there are $n(n-1)\cdots(n-k+1)\cdot(n-k)$ such lists. Taking $k = n$ gives $n!$. □

The factorial grows quickly: S_{10} already has $3,628,800$ elements. So an expansion over S_n is a tool for proving things about determinants, not for computing them; Section 5.4 computes determinants by elimination.

Composition and inverses

Permutations are functions $[n] \to [n]$, so we can compose them. We write $\sigma\tau$ for the composite $\sigma \circ \tau$, so that $(\sigma\tau)(i) = \sigma(\tau(i))$ for $i \in [n]$: first τ, then σ. We write id for the identity permutation, σ^k for the composite of k copies of σ, and $\sigma^0 = \mathrm{id}$.

Proposition 5.2.5 (Composition and inverses). Let $\sigma, \tau, \rho \in S_n$.

(a) $\mathrm{id} \in S_n$, $\sigma\tau \in S_n$, and $\sigma^{-1} \in S_n$.

(b) $(\sigma\tau)\rho = \sigma(\tau\rho)$, $\sigma\,\mathrm{id} = \mathrm{id}\,\sigma = \sigma$, and $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \mathrm{id}$.

(c) $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$.

Proof. ROUTINE

Uses. Definition 5.2.1 (Permutation), composition of functions is associative; inverse functions (Section A.1)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. A function is a bijection exactly when it has a two-sided inverse. Find one for $\sigma\tau$.
Hint 2. Try $\tau^{-1}\sigma^{-1}$, and regroup using associativity of composition.
Sketch. The identity and σ^{-1} are bijections. Regrouping gives $(\sigma\tau)(\tau^{-1}\sigma^{-1}) = \sigma\sigma^{-1} = \mathrm{id}$ and $(\tau^{-1}\sigma^{-1})(\sigma\tau) = \mathrm{id}$, so $\sigma\tau$ is a bijection with inverse $\tau^{-1}\sigma^{-1}$. Part (b) is associativity of composition and the definition of an inverse function.

(a) The identity map is a bijection, and the inverse of a bijection σ is a bijection, with inverse σ. Since composition of functions is associative,

$$(\sigma\tau)(\tau^{-1}\sigma^{-1}) = \sigma(\tau\tau^{-1})\sigma^{-1} = \sigma\sigma^{-1} = \mathrm{id}, \quad (\tau^{-1}\sigma^{-1})(\sigma\tau) = \tau^{-1}(\sigma^{-1}\sigma)\tau = \tau^{-1}\tau = \mathrm{id}.$$

So $\sigma\tau$ has a two-sided inverse, and it is a bijection.

(b) These are the associativity of composition and the defining properties of the identity map and of the inverse function.

(c) By the computation in (a), $\tau^{-1}\sigma^{-1}$ is a two-sided inverse of $\sigma\tau$. A bijection has only one inverse function (if g and g' are both inverses of f, then $g = g(fg') = (gf)g' = g'$), so $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$. □

Remark 5.2.6. Parts (a) and (b) say that S_n with composition is a *group*: a set with an associative operation that has an identity element and inverses. It is called the *symmetric group* of degree n. We will not need any group theory, but Section 14.1 develops this point of view.

Example 5.2.7 (Composing and inverting). Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix}.$$

To compute $(\sigma\tau)(i)$, apply τ first: $(\sigma\tau)(1) = \sigma(\tau(1)) = \sigma(2) = 5$, $(\sigma\tau)(2) = \sigma(1) = 3$, and so on. We get

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 4 & 1 \end{pmatrix}, \quad \tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}.$$

So $\sigma\tau \neq \tau\sigma$: composition of permutations is not commutative. To invert σ, read the two-line notation from the bottom up. Since σ sends $1,2,3,4,5$ to $3,5,4,1,2$, its inverse sends $3,5,4,1,2$ to $1,2,3,4,5$; sorting the columns,

$$\sigma^{-1} = \begin{pmatrix} 3 & 5 & 4 & 1 & 2 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix}.$$

Sums over S_n appear throughout this chapter. The following fact lets us substitute $\rho\sigma$, $\sigma\rho$, or σ^{-1} for the summation variable σ.

Proposition 5.2.8 (Reindexing sums over S_n). Fix $\rho \in S_n$. Each of the maps $S_n \to S_n$ given by $\sigma \mapsto \rho\sigma$, by $\sigma \mapsto \sigma\rho$, and by $\sigma \mapsto \sigma^{-1}$ is a bijection. Consequently, for every function $g : S_n \to F$,

$$\sum_{\sigma \in S_n} g(\sigma) = \sum_{\sigma \in S_n} g(\rho\sigma) = \sum_{\sigma \in S_n} g(\sigma\rho) = \sum_{\sigma \in S_n} g\big(\sigma^{-1}\big).$$

Proof. ROUTINE

Uses. Proposition 5.2.5 (Composition and inverses)
Hint 1. To show that $\sigma \mapsto \rho\sigma$ is a bijection, write down its inverse.
Hint 2. For the sums: if $\varphi : S_n \to S_n$ is a bijection, how do the terms of $\sum_\sigma g(\varphi(\sigma))$ compare with those of $\sum_\sigma g(\sigma)$?
Sketch. The inverses are $\sigma \mapsto \rho^{-1}\sigma$, $\sigma \mapsto \sigma\rho^{-1}$, and $\sigma \mapsto \sigma^{-1}$ itself. If φ is a bijection of S_n, then $\sum_\sigma g(\varphi(\sigma))$ has the same terms as $\sum_\sigma g(\sigma)$ in a different order.

By Proposition 5.2.5, $\rho^{-1}(\rho\sigma) = \sigma$ and $\rho\big(\rho^{-1}\sigma\big) = \sigma$ for every $\sigma \in S_n$, so $\sigma \mapsto \rho^{-1}\sigma$ is a two-sided inverse of $\sigma \mapsto \rho\sigma$. In the same way $\sigma \mapsto \sigma\rho^{-1}$ is a two-sided inverse of $\sigma \mapsto \sigma\rho$, and $\sigma \mapsto \sigma^{-1}$ is its own inverse because $\big(\sigma^{-1}\big)^{-1} = \sigma$. So all three maps are bijections.

Now let $\varphi : S_n \to S_n$ be a bijection. As σ runs through S_n, the element $\varphi(\sigma)$ runs through S_n and takes each value exactly once. So the sum $\sum_\sigma g(\varphi(\sigma))$ has exactly the terms of $\sum_\sigma g(\sigma)$, in another order, and since addition in F is commutative and associative, the two sums are equal. Applying this to the three bijections gives the displayed equalities. □

Exercise 5.2.9. ROUTINE Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$. Compute σ^2, σ^3, and σ^{-1}, and find the least $m \geq 1$ with $\sigma^m = \mathrm{id}$.

Solution on page 1394

5.2.2 Transpositions and cycles

In Intuition (What the proof did, and where it leads) we sorted the columns of $D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$ by interchanging two columns at a time. The permutations that interchange two elements are the building blocks of this section.

Definition 5.2.10 (Transposition). Let $k, l \in [n]$ with $k \neq l$. The **transposition** $\tau = (k\ l)$ is the permutation that interchanges k and l and fixes everything else:

$$\tau(k) = l, \quad \tau(l) = k, \quad \tau(i) = i \text{ for } i \notin \{k, l\}.$$

It is an **adjacent transposition** if $l = k + 1$.

Transpositions are also called **swaps**. Note that $(k\ l) = (l\ k)$, and that interchanging twice changes nothing: $(k\ l)^2 = \mathrm{id}$, so $(k\ l)^{-1} = (k\ l)$.

Remark 5.2.11 (Composing with a transposition). Let $\sigma \in S_n$ and $\tau = (k\ l)$. Then $(\sigma\tau)(k) = \sigma(l)$, $(\sigma\tau)(l) = \sigma(k)$, and $(\sigma\tau)(i) = \sigma(i)$ for all other i. So the bottom row of $\sigma\tau$ is the bottom row of σ with the entries in *positions* k and l interchanged. On the other hand, $\tau\sigma$ applies τ to every entry of the bottom row of σ, so its bottom row is that of σ with the *values* k and l interchanged. For example, with $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$,

$$\sigma(2\ 4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 5 & 2 \end{pmatrix}, \quad (2\ 4)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 1 & 4 \end{pmatrix}.$$

Either way, the bottom row changes by interchanging two of its entries.

Every rearrangement of $1, \ldots, n$ can be sorted by interchanging two entries at a time, so we expect every permutation to be a product of transpositions. To turn this into a proof we need a measure of progress. Say that σ **moves** i if $\sigma(i) \neq i$, and let $m(\sigma)$ be the number of points that σ moves. If $\sigma(k) = j$ with $j \neq k$, then $(k\ j)\sigma$ sends k to k, so composing with one transposition fixes at least one more point. It may fix two, if also $\sigma(j) = k$, so the measure drops by an amount we do not control. This calls for strong induction.

Technique 5.2.12 (Strong induction and well-ordering). To prove a statement $P(m)$ for every integer $m \geq 0$, it suffices to prove $P(m)$ for each m under the assumption that $P(k)$ holds for every $k < m$. When $m = 0$ there is nothing to assume, so the argument must handle that case outright, but no separate base case is needed. Reach for strong induction when a reduction step replaces an object by a smaller one but you cannot say by how much it shrinks, so that induction from $m - 1$ to m does not apply. Often the main decision is the measure of size, such as the number of points a permutation moves or the degree of a polynomial.

An equivalent form is the **well-ordering principle**: every nonempty set of nonnegative integers has a least element. It justifies arguments that begin "let k be the least positive integer such that ...", and proofs by minimal counterexample. The canonical example is Proposition 5.2.13, by induction on the number of moved points. Section A.2 compares the forms of induction, and Section 7.1 uses strong induction on the degree to divide polynomials.

Used in. Proposition 5.2.13, Proposition 5.2.16, Lemma 5.3.14, Theorem 7.1.14, Proposition 7.5.8, Theorem 7.5.13, Lemma 8.2.4, Exercise A.2.9, Exercise 7.1.68.

Proposition 5.2.13 (Products of transpositions). Every permutation $\sigma \in S_n$ is a product of transpositions: there are transpositions $\tau_1, \ldots, \tau_r \in S_n$, with $r \geq 0$, such that $\sigma = \tau_1 \tau_2 \cdots \tau_r$. (For $\sigma = \mathrm{id}$ we take $r = 0$, the empty product.)

Proof. STANDARD

Uses. Definition 5.2.10 (Transposition), Proposition 5.2.5 (Composition and inverses)
Techniques. Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Measure progress by the number $m(\sigma)$ of points that σ moves. Can you compose σ with one transposition so that the result fixes more points?
Hint 2. If $\sigma(k) = j \neq k$, look at $\rho = (k\ j)\sigma$. Which points does ρ fix?
Hint 3. ρ fixes k and every point that σ fixes, so $m(\rho) < m(\sigma)$, but possibly $m(\rho) = m(\sigma) - 2$. Use strong induction, and recover σ from ρ using $(k\ j)^2 = \mathrm{id}$.
Sketch. Strong induction on $m(\sigma)$. If $\sigma = \mathrm{id}$, take the empty product. Otherwise pick k with $\sigma(k) = j \neq k$. Then $\rho = (k\ j)\sigma$ fixes k and every point fixed by σ (a fixed point i of σ cannot be j, by injectivity), so $m(\rho) < m(\sigma)$. Write ρ as a product of transpositions, and $\sigma = (k\ j)\rho$.

We prove by strong induction on $m \geq 0$ that every $\sigma \in S_n$ with $m(\sigma) = m$ is a product of transpositions. Let $\sigma \in S_n$ with $m(\sigma) = m$, and assume the statement for all permutations that move fewer than m points.

If $m = 0$, then $\sigma = \mathrm{id}$, the empty product. Otherwise choose k with $\sigma(k) \neq k$, let $j = \sigma(k)$, and let $\tau = (k\ j)$ and $\rho = \tau\sigma$. Then $\rho(k) = \tau(j) = k$. Now let i be a point fixed by σ. Then $i \neq k$, since σ moves k. Also $i \neq j$: otherwise $\sigma(j) = j = \sigma(k)$ with $j \neq k$, contradicting injectivity of σ. So τ fixes i, and $\rho(i) = \tau(\sigma(i)) = \tau(i) = i$. Thus ρ fixes every point that σ fixes, and it also fixes k, which σ moves. Hence $m(\rho) \leq m - 1$.

By the induction hypothesis, $\rho = \tau_1 \cdots \tau_r$ for some transpositions τ_i. Since $\tau^2 = \mathrm{id}$, we get $\sigma = \tau^2\sigma = \tau\rho = \tau\tau_1 \cdots \tau_r$, a product of transpositions. □

Example 5.2.14 (Sorting by transpositions). Run the proof on $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$, always choosing the least moved point k. By Remark 5.2.11, composing with a transposition on the left interchanges two values in the bottom row. Since $\sigma(1) = 3$, the first step interchanges the values 1 and 3, then the values 2 and 5, then 3 and 4:

$$(1\ 3)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix},$$

$$(2\ 5)(1\ 3)\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 3 & 5 \end{pmatrix},$$

$$(3\ 4)(2\ 5)(1\ 3)\sigma = \mathrm{id}.$$

Multiplying on the left by $(1\ 3)(2\ 5)(3\ 4)$, the inverse of $(3\ 4)(2\ 5)(1\ 3)$, gives $\sigma = (1\ 3)(2\ 5)(3\ 4)$.

Cycles

Following where a permutation sends one element at a time gives a more economical notation. For $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$ we have $1 \mapsto 3 \mapsto 4 \mapsto 1$ and $2 \mapsto 5 \mapsto 2$: the elements of $[5]$ travel around two loops.

Definition 5.2.15 (Cycle). Let $k \geq 2$, and let $a_1, \ldots, a_k \in [n]$ be distinct. The **cycle** $(a_1\ a_2\ \ldots\ a_k) \in S_n$ is the permutation that sends

$$a_1 \mapsto a_2 \mapsto \cdots \mapsto a_k \mapsto a_1$$

and fixes every element of $[n]$ outside $\{a_1, \ldots, a_k\}$. Its **length** is k, and we also call it a k-**cycle**. Two cycles $(a_1\ \ldots\ a_k)$ and $(b_1\ \ldots\ b_l)$ are **disjoint** if $\{a_1, \ldots, a_k\} \cap \{b_1, \ldots, b_l\} = \emptyset$.

A 2-cycle is a transposition. A k-cycle has k names, one for each starting point: $(a_1\ a_2\ \ldots\ a_k) = (a_2\ \ldots\ a_k\ a_1)$. Its inverse runs the loop backwards, $(a_1\ a_2\ \ldots\ a_k)^{-1} = (a_k\ \ldots\ a_2\ a_1)$. Disjoint cycles γ and δ commute. Indeed, if γ moves x, then δ fixes both x and $\gamma(x)$, so $\gamma\delta(x) = \gamma(x) = \delta\gamma(x)$; the case where δ moves x is symmetric, and if neither moves x, both sides are x. Cycles that are not disjoint need not commute: $(1\ 2)(2\ 3) = (1\ 2\ 3)$, while $(2\ 3)(1\ 2) = (1\ 3\ 2)$.

Proposition 5.2.16 (Cycle decomposition). Let $\sigma \in S_n$. For $i \in [n]$, the **orbit** of i under σ is $O(i) = \{\sigma^m(i) \mid m \geq 0\}$.

(a) There is a least positive integer k with $\sigma^k(i) = i$. For this k, the elements $i, \sigma(i), \ldots, \sigma^{k-1}(i)$ are distinct, and they are all the elements of $O(i)$.

(b) Any two orbits are equal or disjoint.

(c) σ is the product of the cycles $\big(i\ \sigma(i)\ \ldots\ \sigma^{k-1}(i)\big)$, one for each orbit with $k \geq 2$ elements, and these cycles are pairwise disjoint.

Proof. STANDARD

Uses. Definition 5.2.15 (Cycle), Proposition 5.2.5 (Composition and inverses)
Techniques. Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Follow $i, \sigma(i), \sigma^2(i), \ldots$. These lie in the finite set $[n]$, so two of them coincide. Why does that force some positive power of σ to fix i?
Hint 2. If $\sigma^p(i) = \sigma^q(i)$ with $p < q$, apply σ^{-1} p times. Then take the least k with $\sigma^k(i) = i$; a repetition among $i, \ldots, \sigma^{k-1}(i)$ would produce a smaller one.
Hint 3. For (b), if $j = \sigma^p(i)$, show that i is also of the form $\sigma^q(j)$. For (c), check the product at each $x \in [n]$: only the cycle of the orbit of x moves x.
Sketch. Finiteness gives $\sigma^p(i) = \sigma^q(i)$ with $p < q$, and cancelling σ^p gives $\sigma^{q-p}(i) = i$. Taking k least makes $i, \ldots, \sigma^{k-1}(i)$ distinct, and division with remainder shows that every $\sigma^m(i)$ is among them. If $j = \sigma^p(i)$, then $i = \sigma^{k-p}(j)$, so each of i, j is in the orbit of the other and the orbits agree. On each orbit σ acts as the corresponding cycle, and cycles of different orbits move disjoint sets of points, so their product agrees with σ everywhere.

(a) The elements $i, \sigma(i), \sigma^2(i), \ldots$ all lie in the finite set $[n]$, so $\sigma^p(i) = \sigma^q(i)$ for some $0 \leq p < q$. Applying σ^{-1} to both sides p times gives $\sigma^{q-p}(i) = i$, with $q - p \geq 1$. So the set of positive integers m with $\sigma^m(i) = i$ is nonempty, and by the well-ordering principle it has a least element k. If $\sigma^p(i) = \sigma^q(i)$ with $0 \leq p < q \leq k - 1$, the same argument gives $\sigma^{q-p}(i) = i$ with $0 < q - p < k$, contradicting the choice of k. So $i, \sigma(i), \ldots, \sigma^{k-1}(i)$

are distinct. Finally, let $m \geq 0$, and write $m = qk + r$ with $q \geq 0$ and $0 \leq r < k$. Since $\sigma^k(i) = i$, applying σ^k q times to i gives i, so $\sigma^m(i) = \sigma^r\left(\sigma^{qk}(i)\right) = \sigma^r(i)$. Hence $O(i) = \left\{i, \sigma(i), \ldots, \sigma^{k-1}(i)\right\}$.

(b) Let $j \in O(i)$. By (a), $j = \sigma^p(i)$ with $0 \leq p < k$, where k is as in (a). For every $m \geq 0$, $\sigma^m(j) = \sigma^{m+p}(i) \in O(i)$, so $O(j) \subseteq O(i)$. Also $\sigma^{k-p}(j) = \sigma^k(i) = i$, so $i \in O(j)$, and the same argument with the roles of i and j exchanged gives $O(i) \subseteq O(j)$. So $O(j) = O(i)$ whenever $j \in O(i)$. If two orbits $O(i)$ and $O(i')$ have a common element j, then $O(i) = O(j) = O(i')$.

(c) Let $O_1, \ldots, O_t$ be the distinct orbits with at least two elements. For each s choose $i_s \in O_s$, let $k_s = |O_s|$, and let $\gamma_s = \left(i_s\ \sigma(i_s)\ \ldots\ \sigma^{k_s-1}(i_s)\right)$. By (a), the entries of γ_s are distinct, so γ_s is a cycle, and the set of its entries is O_s. By (b), $\gamma_1, \ldots, \gamma_t$ are pairwise disjoint. We show that $\sigma = \gamma_1 \cdots \gamma_t$ by evaluating both sides at each $x \in [n]$.

If x lies in none of $O_1, \ldots, O_t$, then its orbit $O(x)$ has one element, so $\sigma(x) = x$; and every γ_s fixes x, so the product does too. Otherwise $x \in O_s$ for exactly one s, say $x = \sigma^p(i_s)$ with $0 \leq p < k_s$. Then $\gamma_s(x) = \sigma^{p+1}(i_s) = \sigma(x)$; for $p = k_s - 1$ this reads $\gamma_s(x) = i_s = \sigma^{k_s}(i_s) = \sigma(x)$. Every other $\gamma_{s'}$ fixes the elements of O_s, in particular x and $\sigma(x)$. So when we apply $\gamma_1 \cdots \gamma_t$ to x, the factors to the right of γ_s fix x, then γ_s sends x to $\sigma(x)$, and the factors to the left of γ_s fix $\sigma(x)$. Hence $(\gamma_1 \cdots \gamma_t)(x) = \sigma(x)$. □

Writing σ as a product of disjoint cycles, with the fixed points omitted, is **cycle notation**. To find it, start at the least element, follow it around its loop until it returns, and then start again at the least element not yet used. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} = (1\ 3\ 4)(2\ 5), \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix} = (1\ 2)(3\ 5\ 4).$$

We write $c(\sigma)$ for the number of orbits of σ, counting each fixed point as an orbit with one element. So $c(\mathrm{id}) = n$, and $c((1\ 3\ 4)(2\ 5)) = 2$ in S_5, while $c((1\ 2)(3\ 4\ 5)) = 3$ in S_6, the orbits being $\{1, 2\}$, $\{3, 4, 5\}$, and $\{6\}$.

Lemma 5.2.17 (Cycles as products of transpositions). Let $k \geq 2$ and let $a_1, \ldots, a_k \in [n]$ be distinct. Then

$$(a_1\ a_2\ \ldots\ a_k) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{k-1}\ a_k).$$

In particular, a k-cycle is a product of $k - 1$ transpositions.

Proof. ROUTINE

Uses. Definition 5.2.15 (Cycle), Definition 5.2.10 (Transposition)
Hint 1. Evaluate the right side at each point. The rightmost factor acts first.
Hint 2. Follow a_i for $i < k$: which factors move it, and in what order? Then follow a_k.
Sketch. For $i < k$, the factors to the right of $(a_i\ a_{i+1})$ do not involve a_i, that factor sends a_i to a_{i+1}, and the factors to its left do not involve a_{i+1}. The point a_k is sent to a_{k-1}, then a_{k-2}, and so on down to a_1. Points outside $\{a_1, \ldots, a_k\}$ are fixed.

Let $\pi = (a_1\ a_2)(a_2\ a_3)\cdots(a_{k-1}\ a_k)$. To compute $\pi(x)$ we apply the factors from right to left, starting with $(a_{k-1}\ a_k)$.

If $x \notin \{a_1, \ldots, a_k\}$, every factor fixes x, so $\pi(x) = x$. Let $1 \leq i < k$. The factors applied before $(a_i\ a_{i+1})$ are the $\left(a_j\ a_{j+1}\right)$ with $j > i$, which do not involve a_i; then $(a_i\ a_{i+1})$ sends a_i to a_{i+1}; and the factors applied after it are the $\left(a_j\ a_{j+1}\right)$ with $j < i$, which do not involve a_{i+1}. So $\pi(a_i) = a_{i+1}$. Finally, $(a_{k-1}\ a_k)$ sends a_k to a_{k-1}, then $(a_{k-2}\ a_{k-1})$ sends it to a_{k-2}, and so on, until $(a_1\ a_2)$ sends a_2 to a_1. So $\pi(a_k) = a_1$. Thus π agrees with $(a_1\ \ldots\ a_k)$ at every point. □

With Proposition 5.2.16, the lemma gives a second proof of Proposition 5.2.13, and this one counts. If the orbits of σ with at least two elements have sizes $k_1, \ldots, k_t$, then σ is a product of $\sum_s (k_s - 1)$ transpositions. The orbits partition $[n]$ and $c(\sigma) - t$ of them are fixed points, so $k_1 + \cdots + k_t + (c(\sigma) - t) = n$, and the number of transpositions is

$$\sum_{s=1}^{t}(k_s - 1) = n - c(\sigma).$$

For $\sigma = (1\ 3\ 4)(2\ 5)$ this gives $\sigma = (1\ 3)(3\ 4)(2\ 5)$, with $5 - 2 = 3$ transpositions.

5.2.3 Parity

A permutation has many factorizations into transpositions. Since $(1\ 2)^2 = \mathrm{id}$, the identity is both the empty product and a product of two transpositions. A direct check shows that

$$(1\ 3) = (1\ 2)(2\ 3)(1\ 2),$$

so $(1\ 3)$ is a product of one transposition and also of three. Substituting this into $\sigma = (1\ 3)(3\ 4)(2\ 5)$ writes that σ as a product of five transpositions as well as of three. In each example the lengths have the same parity, and Theorem 5.2.18 below says that this always happens. Determinants depend on it: if some σ were a product of an even number of transpositions and also of an odd number, then sorting the columns of $D\left(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}\right)$ in two ways would give both $D(I_n)$ and $-D(I_n)$. So $D(I_n) = 0$ whenever $1 + 1 \neq 0$, and no determinant function would exist.

Theorem 5.2.18 (Parity is well defined). Let $\sigma \in S_n$, and suppose that $\sigma = \tau_1 \cdots \tau_r$ and $\sigma = \tau'_1 \cdots \tau'_s$, where all τ_i and τ'_j are transpositions. Then r and s are both even or both odd. In particular, the identity is not a product of an odd number of transpositions.

Intuition (Looking for an invariant). Comparing two factorizations directly is hopeless, since nothing relates the transpositions in one of them to those in the other. Instead, look for a quantity $Q(\sigma)$ that is computed from σ alone, without reference to any factorization, and that changes sign whenever we multiply by one more transposition:

$$Q(\rho\tau) = -Q(\rho) \quad \text{for every } \rho \in S_n \text{ and every transposition } \tau.$$

If $\sigma = \tau_1 \cdots \tau_r$, applying this r times gives $Q(\sigma) = (-1)^r Q(\mathrm{id})$. So if Q takes values in $\mathbb{Q}$ (or in any field in which $1 \neq -1$) and $Q(\mathrm{id}) \neq 0$, then $(-1)^r = Q(\sigma)/Q(\mathrm{id})$ is determined by σ, and all factorizations of σ have lengths of the same parity.

An obvious candidate is $Q(\sigma) = D\left(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}\right)$ for a determinant function D. It does change sign, by Proposition 5.1.10 and Remark 5.2.11. But we do not know yet that a determinant function exists for $n \geq 3$, and the proof of existence in Section 5.3 uses this theorem. So we need a Q that we can write down explicitly. Approach 1 below uses a product of differences, and Approach 2 counts the pairs that σ puts out of order.

Technique 5.2.19 (Parity via an invariant). To show that the number of steps in a process has a determined parity, attach to each state a quantity that every step changes in a controlled way: it is multiplied by -1, or it changes by an odd integer. After r steps the quantity has been multiplied by $(-1)^r$, or has changed by an amount of the same parity as r, so the parity of r is determined by the initial and final states, whichever steps were taken. The same idea shows that a state cannot be reached: find a quantity that every allowed move preserves and that differs between the two states.

Reach for it when a statement is about parity, or about which configurations can be reached by repeating moves of one kind: swaps, slides, rotations. The canonical example is Lemma 5.2.20, where a product of differences $x_j - x_i$ changes sign under every transposition of the variables. Exercise 5.2.47 applies it to a classic puzzle, and Section 8.4 uses the sign of the determinant to separate rotations from reflections.
Used in. Lemma 5.2.20, Lemma 5.2.22, Theorem 5.2.18, Exercise 5.2.47, Exercise 8.4.44, Exercise 5.2.48.

Approach 1: a product of differences

For $\mathbf{x} = (x_1, \ldots, x_n) \in F^n$, let

$$\Delta(\mathbf{x}) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

For instance, $\Delta(x_1, x_2, x_3) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$. By Proposition 1.1.17, $\Delta(\mathbf{x}) \neq 0$ exactly when the entries of $\mathbf{x}$ are distinct. (This product reappears as the Vandermonde determinant in Theorem 5.5.21.) For $\sigma \in S_n$, write

$$\mathbf{x}^\sigma = \left(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}\right)$$

for the vector whose i-th entry is $x_{\sigma(i)}$, that is, the entries of $\mathbf{x}$ rearranged by σ.

Take $n = 3$ and $\tau = (1\ 2)$. Then $\mathbf{x}^\tau = (x_2, x_1, x_3)$, and

$$\Delta(\mathbf{x}^\tau) = (x_1 - x_2)(x_3 - x_2)(x_3 - x_1) = -(x_2 - x_1)(x_3 - x_1)(x_3 - x_2) = -\Delta(\mathbf{x}).$$

The factor $x_2 - x_1$ changed sign, and the two factors involving x_3 traded places. The same happens for every transposition.

Lemma 5.2.20 (A transposition negates Δ). Let $n \ge 2$, let $\tau \in S_n$ be a transposition, and let $\mathbf{x} \in F^n$. Then $\Delta(\mathbf{x}^\tau) = -\Delta(\mathbf{x})$.

Proof. STANDARD

Uses. Definition 5.2.10 (Transposition), Proposition 1.1.16 (Sign rules)
Techniques. Parity via an invariant (Technique 5.2.19)
Hint 1. Write $\tau = (k\ l)$ with $k < l$ and $\mathbf{y} = \mathbf{x}^\tau$, so $y_k = x_l$, $y_l = x_k$, and $y_i = x_i$ otherwise. Sort the factors $y_j - y_i$ of $\Delta(\mathbf{y})$ according to how many of i, j lie in $\{k, l\}$.
Hint 2. Factors that avoid k and l are unchanged, and the factor for the pair (k, l) changes sign. Every other factor involves exactly one of k, l, together with a third index m.
Hint 3. For each $m \notin \{k, l\}$, multiply the two factors that involve m. Show that their product is unchanged, separately in the cases $m < k$, $k < m < l$, and $m > l$.
Sketch. With $\tau = (k\ l)$, $k < l$, and $\mathbf{y} = \mathbf{x}^\tau$, the factors of $\Delta(\mathbf{y})$ not involving k or l are those of $\Delta(\mathbf{x})$, and $y_l - y_k = -(x_l - x_k)$. For each $m \notin \{k, l\}$ the two factors involving m trade places, with two sign changes when $k < m < l$, so their product is unchanged. So exactly one sign changes overall.

Write $\tau = (k\ l)$ with $k < l$, and let $\mathbf{y} = \mathbf{x}^\tau$. Then $y_k = x_{\tau(k)} = x_l$, $y_l = x_k$, and $y_i = x_i$ for $i \notin \{k, l\}$. Each factor $y_j - y_i$ (where $i < j$) of $\Delta(\mathbf{y})$ belongs to exactly one of the following groups, and we compare each group with the corresponding factors of $\Delta(\mathbf{x})$.

Factors with $i, j \notin \{k, l\}$. Here $y_j - y_i = x_j - x_i$.

The factor with $(i, j) = (k, l)$. Here $y_l - y_k = x_k - x_l = -(x_l - x_k)$.

The two factors involving a fixed $m \notin \{k, l\}$. These are the factors for the pairs $\{m, k\}$ and $\{m, l\}$. Let $P_m(\mathbf{x})$ be the product of the corresponding two factors of $\Delta(\mathbf{x})$, each of the form $x_j - x_i$ with $i < j$.

- If $m < k$, then $P_m(\mathbf{y}) = (y_k - y_m)(y_l - y_m) = (x_l - x_m)(x_k - x_m) = P_m(\mathbf{x})$.
- If $k < m < l$, then $P_m(\mathbf{y}) = (y_m - y_k)(y_l - y_m) = (x_m - x_l)(x_k - x_m)$. Pulling a factor -1 out of each difference (Proposition 1.1.16), this equals $(x_l - x_m)(x_m - x_k) = P_m(\mathbf{x})$.
- If $m > l$, then $P_m(\mathbf{y}) = (y_m - y_k)(y_m - y_l) = (x_m - x_l)(x_m - x_k) = P_m(\mathbf{x})$.

Multiplying over all groups and reordering factors by commutativity, $\Delta(\mathbf{y}) = -\Delta(\mathbf{x})$, as desired. □

Approach 2: counting inversions

The second approach measures how far the bottom row of σ is from being sorted.

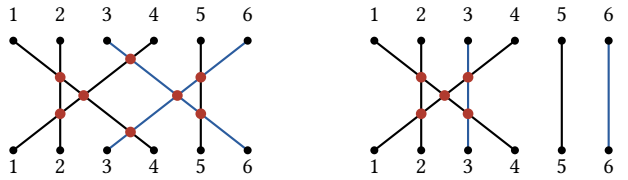
Definition 5.2.21 (Inversion). An **inversion** of $\sigma \in S_n$ is a pair (i, j) with $1 \le i < j \le n$ and $\sigma(i) > \sigma(j)$. We write $\mathrm{inv}(\sigma)$ for the number of inversions of σ.

In the bottom row of the two-line notation, an inversion is a pair of positions whose entries are out of order, the larger entry coming first. For $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$ the inversions are

$$(1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5),$$

coming from $3 > 1, 3 > 2, 5 > 4, 5 > 1, 5 > 2, 4 > 1$, and $4 > 2$. So $\mathrm{inv}(\sigma) = 7$. The identity has no inversions, the adjacent transposition $(k\ k+1)$ has exactly one, namely $(k, k+1)$, and the reversal $i \mapsto n+1-i$ has all $n(n-1)/2$ pairs as inversions.

There is a picture. Join each i on a top row to $\sigma(i)$ on a bottom row by a straight segment. Two segments starting at $i < j$ cross exactly when they end in the opposite order, that is, when $\sigma(i) > \sigma(j)$. So $\mathrm{inv}(\sigma)$ is the number of crossings, counted pair by pair. Figure 5.2.1 shows this for a permutation $\pi \in S_6$ with 8 inversions, together with the effect of multiplying it by a transposition on the right.



π, with 8 crossings $\qquad$ $\pi(3\ 6)$, with 5 crossings

Figure 5.2.1. Left: $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 6 & 1 & 5 & 3 \end{pmatrix}$, drawn by joining i on the top row to $\pi(i)$ on the bottom row. Each crossing (dot) is an inversion. Right: $\pi(3\ 6)$, in which the strands from 3 and 6 (blue) have exchanged their endpoints. The two blue strands no longer cross. Of the strands that start between 3 and 6, the one from 4 still crosses one blue strand, and the one from 5 has lost both of its crossings with blue strands. The strands from 1 and 2 cross as many blue strands as before. The total dropped by $1 + 2 = 3$, an odd number.

Lemma 5.2.22 (A transposition changes inv by an odd number). Let $\sigma \in S_n$ and let τ be a transposition. Then $\mathrm{inv}(\sigma\tau) - \mathrm{inv}(\sigma)$ is odd. If τ is an adjacent transposition, then $\mathrm{inv}(\sigma\tau) = \mathrm{inv}(\sigma) \pm 1$.

Proof. STANDARD

Uses. Definition 5.2.21 (Inversion), Remark 5.2.11 (Composing with a transposition)
Techniques. Parity via an invariant (Technique 5.2.19)
Hint 1. Write $\tau = (k\ l)$ with $k < l$. By Remark 5.2.11, the bottom row of $\sigma\tau$ is that of σ with the entries in positions k and l interchanged. Which pairs of positions can change status?
Hint 2. The pair (k, l) always changes status. For a third position m, look at the two pairs that m forms with k and with l, and consider $m < k$, $k < m < l$, and $m > l$ separately.
Hint 3. When $k < m < l$, let N and N' be the numbers of inverted pairs among the two pairs, before and after. Show that $N + N' = 2$.
Sketch. Let $\tau = (k\ l)$, $k < l$. The bottom rows of σ and $\sigma\tau$ differ by interchanging the entries a_k and a_l at positions k and l. Pairs avoiding k and l keep their status, and (k, l) flips. If $m < k$ or $m > l$, the two pairs formed by m with k and l make the same two comparisons, of a_m with a_k and with a_l, before and after. If $k < m < l$, the number N of inverted pairs among them becomes $2 - N$. So the total changes by ± 1 plus an even number. For adjacent k, l there is no m in between.

Write $\tau = (k\ l)$ with $k < l$, and let $a_i = \sigma(i)$ and $b_i = (\sigma\tau)(i)$. By Remark 5.2.11, $b_k = a_l$, $b_l = a_k$, and $b_i = a_i$ for $i \notin \{k, l\}$. For positions $i < j$, say the pair (i, j) is *inverted in a* if $a_i > a_j$, and *inverted in b* if $b_i > b_j$; so $\mathrm{inv}(\sigma)$ and $\mathrm{inv}(\sigma\tau)$ count the pairs inverted in a and in b. We compare them group by group.

Pairs (i, j) with $i, j \notin \{k, l\}$. These have the same entries in a and b, so the same status.

The pair (k,l). It is inverted in b if and only if $a_l > a_k$, that is, if and only if it is not inverted in a. So this pair contributes $+1$ or -1 to $\operatorname{inv}(\sigma\tau) - \operatorname{inv}(\sigma)$.

The two pairs that a fixed $m \notin \{k,l\}$ *forms with* k *and with* l. If $m < k$, then in a the pair (m,k) is inverted if and only if $a_m > a_k$, and (m,l) if and only if $a_m > a_l$; in b, the pair (m,k) is inverted if and only if $a_m > a_l$, and (m,l) if and only if $a_m > a_k$. These are the same two conditions, so the number of inverted pairs among the two is the same in a and in b. The case $m > l$ is the same with the inequalities reversed. If $k < m < l$, the pairs are (k,m) and (m,l). In a they are inverted if and only if $a_k > a_m$ and $a_m > a_l$, respectively, and in b if and only if $a_l > a_m$ and $a_m > a_k$. Let N and N' be the numbers of these pairs that are inverted in a and in b. Since the entries are distinct, exactly one of $a_k > a_m$ and $a_m > a_k$ holds, and exactly one of $a_m > a_l$ and $a_l > a_m$ holds. Hence $N + N' = 2$, and $N' - N = 2 - 2N$ is even.

Adding up, $\operatorname{inv}(\sigma\tau) - \operatorname{inv}(\sigma)$ is ± 1 plus a sum of even numbers, so it is odd. If $l = k+1$, there is no m with $k < m < l$, so the pair (k,l) is the only group that contributes to $\operatorname{inv}(\sigma\tau) - \operatorname{inv}(\sigma)$, and $\operatorname{inv}(\sigma\tau) = \operatorname{inv}(\sigma) \pm 1$. □

Two proofs of the theorem

Each lemma supplies the quantity Q of Intuition (Looking for an invariant): $Q(\sigma) = \Delta(\mathbf{x}^\sigma)$ for a fixed $\mathbf{x}$ with distinct entries, and $Q(\sigma) = (-1)^{\operatorname{inv}(\sigma)}$.

Proof of Theorem 5.2.18 (via the product of differences). STANDARD

> *Uses.* Lemma 5.2.20 (A transposition negates Δ), Proposition 5.2.13 (Products of transpositions)
> *Techniques.* Parity via an invariant (Technique 5.2.19)
> *Hint 1.* Lemma 5.2.20 says what one transposition does to Δ. What does a product of r transpositions do?
> *Hint 2.* You will need that rearranging by ρ and then by τ is rearranging by $\rho\tau$: check that $(\mathbf{x}^\rho)^\tau = \mathbf{x}^{\rho\tau}$.
> *Hint 3.* Show $\Delta(\mathbf{x}^\sigma) = (-1)^r\Delta(\mathbf{x})$ by induction on r, then evaluate at a vector with $\Delta(\mathbf{x}) \neq 0$ over a field in which $1 \neq -1$.
> *Sketch.* Since $(\mathbf{x}^\rho)^\tau = \mathbf{x}^{\rho\tau}$, Lemma 5.2.20 and induction on r give $\Delta(\mathbf{x}^\sigma) = (-1)^r\Delta(\mathbf{x})$ whenever σ is a product of r transpositions. Take $\mathbf{x} = (1,2,\ldots,n) \in \mathbb{Q}^n$, so $\Delta(\mathbf{x}) \neq 0$. Then two factorizations of lengths r and s give $(-1)^r\Delta(\mathbf{x}) = (-1)^s\Delta(\mathbf{x})$, so $(-1)^r = (-1)^s$.

Let F be any field and $\mathbf{x} \in F^n$. First, for $\rho, \pi \in S_n$, the i-th entry of $(\mathbf{x}^\rho)^\pi$ is the $\pi(i)$-th entry of $\mathbf{x}^\rho$, which is $x_{\rho(\pi(i))}$. So

$$(\mathbf{x}^\rho)^\pi = \mathbf{x}^{\rho\pi}.$$

We show by induction on $r \geq 0$ that if $\sigma = \tau_1 \cdots \tau_r$ is a product of r transpositions, then $\Delta(\mathbf{x}^\sigma) = (-1)^r\Delta(\mathbf{x})$. For $r = 0$, $\sigma = \mathrm{id}$ and $\mathbf{x}^{\mathrm{id}} = \mathbf{x}$. For $r \geq 1$ there is a transposition in S_n, so $n \geq 2$. Let $\rho = \tau_1 \cdots \tau_{r-1}$, so that $\sigma = \rho\tau_r$ and, by the induction hypothesis, $\Delta(\mathbf{x}^\rho) = (-1)^{r-1}\Delta(\mathbf{x})$. Applying Lemma 5.2.20 to the vector $\mathbf{x}^\rho$,

$$\Delta(\mathbf{x}^\sigma) = \Delta\big((\mathbf{x}^\rho)^{\tau_r}\big) = -\Delta(\mathbf{x}^\rho) = (-1)^r\Delta(\mathbf{x}).$$

Now take $F = \mathbb{Q}$ and $\mathbf{x} = (1,2,\ldots,n)$. Then $\Delta(\mathbf{x}) = \prod_{i<j}(j-i)$ is a product of positive integers (the empty product 1 if $n = 1$), so $\Delta(\mathbf{x}) \neq 0$. If $\sigma = \tau_1 \cdots \tau_r = \tau'_1 \cdots \tau'_s$, then

$$(-1)^r\Delta(\mathbf{x}) = \Delta(\mathbf{x}^\sigma) = (-1)^s\Delta(\mathbf{x}),$$

and dividing by $\Delta(\mathbf{x})$ gives $(-1)^r = (-1)^s$. Since $1 \neq -1$ in $\mathbb{Q}$, the integers r and s have the same parity. □

Proof of Theorem 5.2.18 (by counting inversions). ROUTINE

> *Uses.* Lemma 5.2.22 (A transposition changes inv by an odd number), Definition 5.2.21 (Inversion)
> *Techniques.* Parity via an invariant (Technique 5.2.19)
> *Hint 1.* Start from the identity, which has no inversions, and multiply by the transpositions one at a time on the right.
> *Hint 2.* Show by induction on r that if $\sigma = \tau_1 \cdots \tau_r$, then $\operatorname{inv}(\sigma) - r$ is even.
> *Sketch.* $\operatorname{inv}(\mathrm{id}) = 0$, and by Lemma 5.2.22 each additional factor on the right changes inv by an odd number. So $\operatorname{inv}(\tau_1 \cdots \tau_r) \equiv r$ modulo 2, and any two factorizations of σ have lengths with the parity of $\operatorname{inv}(\sigma)$.

We show by induction on $r \geq 0$ that if $\sigma = \tau_1 \cdots \tau_r$ with transpositions τ_i, then $\operatorname{inv}(\sigma) - r$ is even. For $r = 0$, $\sigma = \mathrm{id}$ and $\operatorname{inv}(\mathrm{id}) = 0$. For $r \geq 1$, let $\rho = \tau_1 \cdots \tau_{r-1}$, so that $\sigma = \rho\tau_r$. Then

$$\operatorname{inv}(\sigma) - r = (\operatorname{inv}(\sigma) - \operatorname{inv}(\rho)) + (\operatorname{inv}(\rho) - (r-1)) - 1.$$

The first term is odd by Lemma 5.2.22 and the second is even by the induction hypothesis, so the sum is even.

Hence if $\sigma = \tau_1 \cdots \tau_r = \tau'_1 \cdots \tau'_s$, both r and s have the same parity as $\operatorname{inv}(\sigma)$. □

Intuition (Comparing the two proofs). Both proofs follow Technique 5.2.19, and the invariants are closer than they look. With $\mathbf{x} = (1, \ldots, n)$, the factor $\sigma(j) - \sigma(i)$ of $\Delta(\mathbf{x}^\sigma)$ is negative exactly when (i, j) is an inversion, and the absolute values $|\sigma(j) - \sigma(i)|$ are the numbers $j - i$ in another order (as $\{i, j\}$ runs over all two-element sets, so does $\{\sigma(i), \sigma(j)\}$). So $\Delta(\mathbf{x}^\sigma) = (-1)^{\operatorname{inv}(\sigma)}\Delta(\mathbf{x})$. The arguments are still different. Approach 1 never asks which factors are negative. Its lemma is an algebraic identity, valid for every vector over every field, and the sign falls out of regrouping. That is the pattern of Theorem 5.2.31 below, where an alternating function plays the role of Δ. Approach 2 is combinatorial and gives a formula you can read off from the two-line notation. Its sharper statement for adjacent transpositions also shows that $\operatorname{inv}(\sigma)$ is the number of interchanges of neighbors that sorting σ requires (Exercise 5.2.43).

A third proof counts orbits instead. By Lemma 5.2.17, σ is a product of $n - c(\sigma)$ transpositions, and one can show that multiplying by a transposition changes c by exactly 1. Exercise 5.2.50 works this out.

5.2.4 The sign of a permutation

Definition 5.2.23 (Sign of a permutation). A permutation $\sigma \in S_n$ is **even** if it is a product of an even number of transpositions, and **odd** if it is a product of an odd number of transpositions. The **sign** of σ is $\operatorname{sgn}(\sigma) = 1$ if σ is even and $\operatorname{sgn}(\sigma) = -1$ if σ is odd.

By Proposition 5.2.13 every permutation is even or odd, and by Theorem 5.2.18 it is not both. So sgn is well defined, and

$$\operatorname{sgn}(\sigma) = (-1)^r \quad \text{whenever } \sigma \text{ is a product of } r \text{ transpositions.}$$

The sign is an integer, 1 or -1. When it multiplies an element of a field F, we read it as 1 or -1 in F; if $1 + 1 = 0$ in F, these are the same element.

Corollary 5.2.24 (Formulas for the sign). Let $\sigma \in S_n$.

(a) $\operatorname{sgn}(\sigma) = (-1)^{\operatorname{inv}(\sigma)}$.

(b) For every field F and every $\mathbf{x} \in F^n$, $\Delta(\mathbf{x}^\sigma) = \operatorname{sgn}(\sigma)\Delta(\mathbf{x})$. In particular,

$$\operatorname{sgn}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Proof. ROUTINE

Uses. Theorem 5.2.18 (Parity is well defined), Definition 5.2.23 (Sign of a permutation), Proposition 5.2.13 (Products of transpositions)
Hint 1. Both proofs of Theorem 5.2.18 established more than the theorem. What did each show about a product of r transpositions?
Hint 2. For the product formula, take $F = \mathbb{Q}$ and $\mathbf{x} = (1, 2, \ldots, n)$ in (b).
Sketch. Write $\sigma = \tau_1 \cdots \tau_r$. The second proof showed $\operatorname{inv}(\sigma) \equiv r$, and the first showed $\Delta(\mathbf{x}^\sigma) = (-1)^r\Delta(\mathbf{x})$ for every $\mathbf{x}$ over every field. Both say the claimed formulas, since $(-1)^r = \operatorname{sgn}(\sigma)$. With $\mathbf{x} = (1, \ldots, n)$, (b) becomes the product formula after dividing by $\Delta(\mathbf{x}) \neq 0$.

Write $\sigma = \tau_1 \cdots \tau_r$ with transpositions τ_i (Proposition 5.2.13), so that $\operatorname{sgn}(\sigma) = (-1)^r$.

(a) The second proof of Theorem 5.2.18 showed that $\operatorname{inv}(\sigma) - r$ is even, so $(-1)^{\operatorname{inv}(\sigma)} = (-1)^r = \operatorname{sgn}(\sigma)$.

(b) The first proof showed that $\Delta(\mathbf{x}^\sigma) = (-1)^r \Delta(\mathbf{x})$ for every field F and every $\mathbf{x} \in F^n$, which is the first claim. Taking $F = \mathbb{Q}$ and $\mathbf{x} = (1, 2, \ldots, n)$, we have $\Delta(\mathbf{x}^\sigma) = \prod_{i<j}(\sigma(j) - \sigma(i))$ and $\Delta(\mathbf{x}) = \prod_{i<j}(j - i) \neq 0$, and dividing gives the product formula. □

For $\sigma = (1\ 3\ 4)(2\ 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$ we found $\mathrm{inv}(\sigma) = 7$ and $\sigma = (1\ 3)(3\ 4)(2\ 5)$. Both give $\mathrm{sgn}(\sigma) = -1$, as they must. Table 5.2.1 lists the signs of the six permutations of $[3]$.

σ	Cycles	$\mathrm{inv}(\sigma)$	$\mathrm{sgn}(\sigma)$
$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	id	0	1
$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	$(2\ 3)$	1	-1
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	$(1\ 2)$	1	-1
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$(1\ 2\ 3)$	2	1
$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$(1\ 3\ 2)$	2	1
$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	$(1\ 3)$	3	-1

Table 5.2.1. The six permutations of $[3]$. The three even ones are the identity and the two 3-cycles; the three odd ones are the transpositions.

Proposition 5.2.25 (Multiplicativity of the sign).

(a) $\mathrm{sgn}(\mathrm{id}) = 1$, and $\mathrm{sgn}(\tau) = -1$ for every transposition τ.

(b) $\mathrm{sgn}(\sigma\rho) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\rho)$ for all $\sigma, \rho \in S_n$.

(c) $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ for all $\sigma \in S_n$.

Proof. ROUTINE

Uses. Definition 5.2.23 (Sign of a permutation), Theorem 5.2.18 (Parity is well defined), Proposition 5.2.5 (Composition and inverses)
Hint 1. Use the definition: factor σ and ρ into transpositions. What is a factorization of $\sigma\rho$?
Hint 2. For (c), apply (b) to $\sigma\sigma^{-1} = \mathrm{id}$.
Sketch. Concatenating factorizations of σ (with r factors) and ρ (with s factors) gives one of $\sigma\rho$ with $r + s$ factors, so $\mathrm{sgn}(\sigma\rho) = (-1)^{r+s}$. Then $\mathrm{sgn}(\sigma)\,\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\mathrm{id}) = 1$.

(a) The identity is the product of zero transpositions, and a transposition is a product of one.

(b) Write $\sigma = \tau_1 \cdots \tau_r$ and $\rho = \tau'_1 \cdots \tau'_s$ with transpositions τ_i, τ'_j. Then $\sigma\rho = \tau_1 \cdots \tau_r \tau'_1 \cdots \tau'_s$ is a product of $r + s$ transpositions, so by Definition 5.2.23 (which is legitimate by Theorem 5.2.18),

$$\mathrm{sgn}(\sigma\rho) = (-1)^{r+s} = (-1)^r (-1)^s = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\rho).$$

(c) By (b) and (a), $\mathrm{sgn}(\sigma)\,\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma\sigma^{-1}) = \mathrm{sgn}(\mathrm{id}) = 1$. Multiplying both sides by $\mathrm{sgn}(\sigma)$ and using $\mathrm{sgn}(\sigma)^2 = 1$ gives $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$. □

With the definition through transpositions, (b) takes one line. Proving it directly from the inversion count is harder (Exercise 5.2.49). This is the reason to define sgn through transpositions and to treat $(-1)^{\mathrm{inv}(\sigma)}$ as a way of computing it. In the language of Section 14.1, (b) says that sgn is a homomorphism from S_n to the group $\{1, -1\}$ under multiplication. For example, if σ is odd, then σ^{-1} is odd by (c), while σ^2 is even since $\mathrm{sgn}(\sigma^2) = \mathrm{sgn}(\sigma)^2 = 1$.

Proposition 5.2.26 (Sign of a cycle). A k-cycle has sign $(-1)^{k-1}$. More generally, $\mathrm{sgn}(\sigma) = (-1)^{n-c(\sigma)}$ for every $\sigma \in S_n$, where $c(\sigma)$ is the number of orbits of σ.

Proof. ROUTINE

Uses. Lemma 5.2.17 (Cycles as products of transpositions), Proposition 5.2.16 (Cycle decomposition), Definition 5.2.23 (Sign of a permutation)
Hint 1. Lemma 5.2.17 writes a k-cycle as a product of how many transpositions?
Hint 2. For general σ, combine Proposition 5.2.16 with the count $n - c(\sigma)$ that follows Lemma 5.2.17.
Sketch. A k-cycle is a product of $k-1$ transpositions. Decomposing σ into the cycles of its orbits and each cycle into transpositions writes σ as a product of $n - c(\sigma)$ transpositions.

By Lemma 5.2.17 a k-cycle is a product of $k-1$ transpositions, so its sign is $(-1)^{k-1}$. For general σ, Proposition 5.2.16 and Lemma 5.2.17 write σ as a product of $n - c(\sigma)$ transpositions, as computed after Lemma 5.2.17. So $\operatorname{sgn}(\sigma) = (-1)^{n-c(\sigma)}$. □

So a cycle of odd length is an even permutation and a cycle of even length is odd. For $\sigma = (1\ 3\ 4)(2\ 5) \in S_5$, $c(\sigma) = 2$, and again $\operatorname{sgn}(\sigma) = (-1)^3 = -1$.

Proposition 5.2.27 (Half of all permutations are even). For $n \geq 2$, exactly $n!/2$ elements of S_n are even and $n!/2$ are odd. Equivalently, $\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) = 0$.

Proof. ROUTINE

Uses. Proposition 5.2.8 (Reindexing sums over S_n), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.4 (Counting permutations)
Hint 1. Find a bijection from the even permutations to the odd ones.
Hint 2. Multiply by a fixed transposition, say $(1\ 2)$, on the right.
Sketch. $\varphi(\sigma) = \sigma(1\ 2)$ is a bijection of S_n that reverses signs, and $\varphi \circ \varphi = \mathrm{id}$, so it restricts to a bijection from the even permutations to the odd ones. With $|S_n| = n!$ the two sets have $n!/2$ elements each.

Let $\tau = (1\ 2)$, which exists since $n \geq 2$, and let $\varphi : S_n \to S_n$ be $\varphi(\sigma) = \sigma\tau$, a bijection by Proposition 5.2.8. By Proposition 5.2.25, $\operatorname{sgn}(\varphi(\sigma)) = -\operatorname{sgn}(\sigma)$, so φ maps the set S_n^+ of even permutations into the set S_n^- of odd permutations, and S_n^- into S_n^+. Since $\tau^2 = \mathrm{id}$, we have $\varphi(\varphi(\sigma)) = \sigma$, so the restrictions $S_n^+ \to S_n^-$ and $S_n^- \to S_n^+$ of φ are inverse to each other. Hence $|S_n^+| = |S_n^-|$. As every permutation is even or odd but not both, $|S_n^+| + |S_n^-| = n!$ by Proposition 5.2.4, so each set has $n!/2$ elements, and $\sum_\sigma \operatorname{sgn}(\sigma) = |S_n^+| - |S_n^-| = 0$. □

Pairing each σ with $\sigma(k\ l)$, which has the opposite sign, is the standard way to make the terms of a sum over S_n cancel in pairs.

Exercise 5.2.28. ROUTINE

(a) Show that σ^2 and $\sigma\rho\sigma^{-1}\rho^{-1}$ are even for all $\sigma, \rho \in S_n$.

(b) Deduce that $(1\ 2)$ is not the square of any permutation, and that neither is $(1\ 2)(3\ 4\ 5)$.

Solution on page 1395

5.2.5 Permutation matrices

We return to the values that survive in the expansion of Intuition (What the proof did, and where it leads). They are values of D at matrices whose columns are the standard basis vectors in some order.

Definition 5.2.29 (Permutation matrix). For $\sigma \in S_n$, the **permutation matrix** $E_\sigma \in F^{n \times n}$ is the matrix with columns $\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, \ldots, \mathbf{e}_{\sigma(n)}$. That is, its (i, j) entry is 1 if $i = \sigma(j)$, and 0 otherwise.

Each column of E_σ is a standard basis vector, and since σ is a bijection, each standard basis vector is exactly one of the columns. So E_σ has exactly one entry 1 in each row and in each column, and zeros elsewhere. We have $E_{\mathrm{id}} = I_n$, and for a transposition $(k\ l)$, $E_{(k\ l)}$ is I_n with columns k and l interchanged, which is the elementary matrix that interchanges rows k and l (Definition 3.2.2). For example, in S_3,

$$E_{(1\,2\,3)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_{(1\,3\,2)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_{(1\,3)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The first has columns $\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1$, because (1 2 3) sends $1, 2, 3$ to $2, 3, 1$.

Proposition 5.2.30 (Algebra of permutation matrices). Let $\sigma, \rho \in S_n$, and let $A \in F^{m\times n}$ have columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

(a) AE_σ is the matrix with columns $\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(n)}$.

(b) $E_\sigma E_\rho = E_{\sigma\rho}$.

(c) E_σ is invertible, and $E_\sigma^{-1} = E_{\sigma^{-1}} = E_\sigma^{\mathsf{T}}$.

Proof. ROUTINE

Uses. Definition 5.2.29 (Permutation matrix), Definition 2.4.1 (Matrix product), Definition 2.3.8 (Matrix–vector product), Definition 2.5.12 (Invertible), Definition 2.4.21 (Transpose)
Hint 1. Column j of a product AB is A times column j of B, and $A\mathbf{e}_k$ is column k of A.
Hint 2. For (b), apply (a) with $A = E_\sigma$. For (c), use (b) with $\rho = \sigma^{-1}$, and compare entries for the transpose.
Sketch. Column j of AE_σ is $A\mathbf{e}_{\sigma(j)} = \mathbf{a}_{\sigma(j)}$. With $A = E_\sigma$, column j of $E_\sigma E_\rho$ is column $\rho(j)$ of E_σ, which is $\mathbf{e}_{\sigma(\rho(j))}$. Then $E_\sigma E_{\sigma^{-1}} = E_{\mathrm{id}} = I_n = E_{\sigma^{-1}} E_\sigma$, and the (i,j) entry of E_σ^{T} is 1 exactly when $j = \sigma(i)$, that is, $i = \sigma^{-1}(j)$.

(a) By Definition 2.4.1, column j of AE_σ is A times column j of E_σ, which is $A\mathbf{e}_{\sigma(j)} = \mathbf{a}_{\sigma(j)}$.

(b) Apply (a) with $A = E_\sigma$, whose columns are $\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}$: column j of $E_\sigma E_\rho$ is column $\rho(j)$ of E_σ, which is $\mathbf{e}_{\sigma(\rho(j))} = \mathbf{e}_{(\sigma\rho)(j)}$. So $E_\sigma E_\rho = E_{\sigma\rho}$.

(c) By (b), $E_\sigma E_{\sigma^{-1}} = E_{\mathrm{id}} = I_n$ and $E_{\sigma^{-1}} E_\sigma = E_{\mathrm{id}} = I_n$, so $E_{\sigma^{-1}}$ is a two-sided inverse of E_σ. For the transpose, the (i,j) entry of E_σ^{T} is the (j,i) entry of E_σ, which is 1 if $j = \sigma(i)$ and 0 otherwise. Since $j = \sigma(i)$ if and only if $i = \sigma^{-1}(j)$, this is the (i,j) entry of $E_{\sigma^{-1}}$. □

Part (a) says that multiplying by E_σ on the right rearranges the columns, putting column $\sigma(j)$ in position j. Multiplication of a vector on the left goes the other way: $E_\sigma \mathbf{x} = \sum_j x_j \mathbf{e}_{\sigma(j)}$, so the entry x_j moves to position $\sigma(j)$, and $E_\sigma \mathbf{x} = \mathbf{x}^{\sigma^{-1}}$ in the notation of Approach 1: a product of differences. Keep the two descriptions apart.

Intuition. Write $\sigma = \tau_1 \cdots \tau_r$ with transpositions τ_i. By Proposition 5.2.30, $AE_\sigma = AE_{\tau_1} \cdots E_{\tau_r}$, so the columns of AE_σ are obtained from those of A by r interchanges of two columns, and each interchange negates D. So $D(AE_\sigma) = (-1)^r D(A)$. The right side involves a factorization of σ, and Theorem 5.2.18 is exactly what makes it independent of the factorization.

Theorem 5.2.31 (Alternating functions and permutations). Let $D : F^{n\times n} \to F$ be multilinear and alternating, and let $\sigma \in S_n$. For all $\mathbf{a}_1, \ldots, \mathbf{a}_n \in F^n$,

$$D\big(\mathbf{a}_{\sigma(1)}, \mathbf{a}_{\sigma(2)}, \ldots, \mathbf{a}_{\sigma(n)}\big) = \operatorname{sgn}(\sigma) D(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n).$$

That is, $D(AE_\sigma) = \operatorname{sgn}(\sigma) D(A)$ for every $A \in F^{n\times n}$. In particular $D(E_\sigma) = \operatorname{sgn}(\sigma) D(I_n)$, and if D is a determinant function, then $D(E_\sigma) = \operatorname{sgn}(\sigma)$.

Proof. STANDARD

Uses. Proposition 5.1.10 (Alternating implies antisymmetric), Proposition 5.2.13 (Products of transpositions), Definition 5.2.23 (Sign of a permutation), Proposition 5.2.30 (Algebra of permutation matrices)
Hint 1. Factor σ into transpositions and induct on the number r of factors.
Hint 2. If $\sigma = \rho\tau$ with $\tau = (k\ l)$, put $\mathbf{b}_i = \mathbf{a}_{\rho(i)}$. How is the list $\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(n)}$ related to $\mathbf{b}_1, \ldots, \mathbf{b}_n$?
Hint 3. It is the same list with the entries in positions k and l interchanged. Use Proposition 5.1.10.

Sketch. By induction on r, if σ is a product of r transpositions then $D\big(\mathbf{a}_{\sigma(1)},\dots\big) = (-1)^r D(\mathbf{a}_1,\dots)$: writing $\sigma = \rho\tau$ with $\tau = (k\ l)$ and $\mathbf{b}_i = \mathbf{a}_{\rho(i)}$, the list $\mathbf{a}_{\sigma(i)} = \mathbf{b}_{\tau(i)}$ is $\mathbf{b}$ with positions k and l interchanged, which negates D. Since $(-1)^r = \operatorname{sgn}(\sigma)$, done. The matrix forms follow from Proposition 5.2.30.

We show by induction on $r \geq 0$: if σ is a product of r transpositions, then for all $\mathbf{a}_1,\dots,\mathbf{a}_n \in F^n$,

$$D\big(\mathbf{a}_{\sigma(1)},\dots,\mathbf{a}_{\sigma(n)}\big) = (-1)^r D(\mathbf{a}_1,\dots,\mathbf{a}_n).$$

For $r = 0$, $\sigma = \mathrm{id}$ and there is nothing to prove. For $r \geq 1$, write $\sigma = \rho\tau$, where ρ is a product of $r-1$ transpositions and $\tau = (k\ l)$. Put $\mathbf{b}_i = \mathbf{a}_{\rho(i)}$ for $i \in [n]$. Then $\mathbf{a}_{\sigma(i)} = \mathbf{a}_{\rho(\tau(i))} = \mathbf{b}_{\tau(i)}$. Since τ interchanges k and l and fixes everything else, the list $\mathbf{b}_{\tau(1)},\dots,\mathbf{b}_{\tau(n)}$ is the list $\mathbf{b}_1,\dots,\mathbf{b}_n$ with the entries in positions k and l interchanged. By Proposition 5.1.10 and then the induction hypothesis,

$$D\big(\mathbf{a}_{\sigma(1)},\dots,\mathbf{a}_{\sigma(n)}\big) = -D(\mathbf{b}_1,\dots,\mathbf{b}_n) = -D\big(\mathbf{a}_{\rho(1)},\dots,\mathbf{a}_{\rho(n)}\big) = (-1)^r D(\mathbf{a}_1,\dots,\mathbf{a}_n).$$

By Proposition 5.2.13, every σ is a product of some number r of transpositions, and $(-1)^r = \operatorname{sgn}(\sigma)$ by Definition 5.2.23. This proves the displayed identity. By Proposition 5.2.30, the columns of AE_σ are $\mathbf{a}_{\sigma(1)},\dots,\mathbf{a}_{\sigma(n)}$, so $D(AE_\sigma) = \operatorname{sgn}(\sigma)D(A)$. Taking $A = I_n$ gives $D(E_\sigma) = \operatorname{sgn}(\sigma)D(I_n)$, and a determinant function has $D(I_n) = 1$. □

This answers the first question raised in Intuition (What the proof did, and where it leads): each surviving value is determined, $D\big(\mathbf{e}_{\sigma(1)},\dots,\mathbf{e}_{\sigma(n)}\big) = \operatorname{sgn}(\sigma)D(I_n)$, however we sort. Substituting into the expansion expresses $D(A)$ as $D(I_n)$ times a sum, over $\sigma \in S_n$, of signed products of entries of A. Section 5.3 carries out this substitution (Theorem 5.3.3) and answers the second question by checking that the resulting formula satisfies the axioms.

Remark 5.2.32 (Two comments on the theorem). (1) If $1 + 1 = 0$ in F, then $\operatorname{sgn}(\sigma) = 1$ in F for every σ, and the theorem says that $D(AE_\sigma) = D(A)$: rearranging the columns does not change D at all. This matches Example 5.1.12, where interchanging two columns changed nothing over $\mathbb{F}_2$.

(2) The induction in the proof does not use Theorem 5.2.18. So if we had, over $\mathbb{Q}$, a multilinear alternating $D : \mathbb{Q}^{n\times n} \to \mathbb{Q}$ with $D(I_n) \neq 0$, constructed without permutations, then two factorizations of σ of lengths r and s would give $(-1)^r D(I_n) = D(E_\sigma) = (-1)^s D(I_n)$, a third proof of Theorem 5.2.18. In this sense the parity theorem is exactly what the existence of determinants requires. The construction in Section 5.3 uses sgn, so it cannot be used this way (Exercise 5.2.38).

Exercises

Exercise 5.2.33. ROUTINE Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 1 & 3 & 4 \end{pmatrix}, \quad \tau = (1\ 4)(2\ 6) \in S_6.$$

Compute $\sigma\tau$, $\tau\sigma$, and σ^{-1}, in two-line and in cycle notation, and find the signs of σ, τ, $\sigma\tau$, $\tau\sigma$, and σ^{-1}.

Solution on page 1395

Exercise 5.2.34. ROUTINE

(a) Let $1 \leq k < l \leq n$. Show that $\operatorname{inv}((k\ l)) = 2(l-k) - 1$.

(b) Find inv of the n-cycle $(1\ 2\ \dots\ n)$, and check the result against Proposition 5.2.26.

(c) Let $\omega \in S_n$ be the reversal, $\omega(i) = n + 1 - i$. Compute $\operatorname{sgn}(\omega)$ in two ways, from $\operatorname{inv}(\omega)$ and from the cycle notation of ω, and deduce that $(-1)^{n(n-1)/2} = (-1)^{\lfloor n/2 \rfloor}$. For which n is ω even?

Solution on page 1395

Exercise 5.2.35. ROUTINE Find the error in the following argument.

Claim. If $\sigma \in S_n$ and τ is a transposition, then $\operatorname{inv}(\sigma\tau) = \operatorname{inv}(\sigma) \pm 1$.

Proof. By Remark 5.2.11, the bottom row of $\sigma\tau$ is that of σ with two entries interchanged. Interchanging two entries changes their relative order and nothing else. So exactly one pair of positions changes from inverted to not inverted, or the other way.

Solution on page 1396

Exercise 5.2.36. STANDARD Let $\sigma \in S_n$.

(a) Let $A \in F^{n\times m}$. Show that row $\sigma(i)$ of $E_\sigma A$ is row i of A, for every i.

(b) Let $A \in F^{n\times n}$. Show that the $(\sigma(i), \sigma(j))$ entry of $E_\sigma A E_\sigma^{\mathsf{T}}$ is a_{ij}, for all i, j.

(c) The *adjacency matrix* of a graph with vertices $1, \ldots, n$ is the $n \times n$ matrix with (i, j) entry 1 if i and j are joined by an edge, and 0 otherwise. Explain what (b) says about adjacency matrices when the vertices of a graph are renamed by σ.

Solution on page 1396

Exercise 5.2.37. STANDARD Prove or disprove each statement, for all n and all $\sigma, \tau \in S_n$.

(a) The product of two odd permutations is odd.

(b) If $\sigma^2 = \mathrm{id}$, then σ is the identity or a transposition.

(c) $\operatorname{inv}\left(\sigma^{-1}\right) = \operatorname{inv}(\sigma)$.

(d) $\operatorname{inv}(\sigma\tau) = \operatorname{inv}(\sigma) + \operatorname{inv}(\tau)$.

(e) $\operatorname{sgn}\left(\sigma\tau\sigma^{-1}\right) = \operatorname{sgn}(\tau)$.

Solution on page 1397

Exercise 5.2.38. STANDARD Find the errors in the following argument.

Claim. Theorem 5.2.18.

Proof. Let det be the determinant function on $F^{n\times n}$ (Theorem 5.3.4). If $\sigma = \tau_1 \cdots \tau_r$ is a product of transpositions, then by Proposition 5.2.30 the columns of E_σ are obtained from those of I_n by r interchanges of two columns, and each interchange multiplies det by -1 (Proposition 5.1.10). So $\det E_\sigma = (-1)^r \det I_n = (-1)^r$. Since $\det E_\sigma$ depends only on σ, so does $(-1)^r$, and the parity of r is determined by σ.

Solution on page 1397

Exercise 5.2.39. STANDARD

(a) Let $\sigma \in S_n$ and let $(a_1\ \ldots\ a_k)$ be a cycle. Show that

$$\sigma(a_1\ a_2\ \ldots\ a_k)\sigma^{-1} = (\sigma(a_1)\ \sigma(a_2)\ \ldots\ \sigma(a_k)).$$

(b) Let $n \geq 2$ and $k \neq l$ in $[n]$. Show that $(k\ l) = \rho(1\ 2)\rho^{-1}$ for some $\rho \in S_n$.

(c) Compute $\sigma(1\ 2\ 3)(4\ 5)\sigma^{-1}$ for $\sigma = (1\ 5\ 2)(3\ 4)$ without multiplying permutations.

Solution on page 1398

Exercise 5.2.40. STANDARD Let $n \geq 3$. Show that every even permutation in S_n is a product of 3-cycles. Conclude that a permutation is even if and only if it is a product of 3-cycles.

Solution on page 1398

Exercise 5.2.41. STANDARD The **order** of $\sigma \in S_n$ is the least $m \geq 1$ with $\sigma^m = \mathrm{id}$.

(a) Show that the order of a k-cycle is k, and that if $\sigma = \gamma_1 \cdots \gamma_t$ is a product of disjoint cycles of lengths $k_1, \ldots, k_t$, then the order of σ is the least common multiple of $k_1, \ldots, k_t$.

(b) Let $\sigma = (1\ 2\ 3)(4\ 5\ 6\ 7) \in S_7$. Find the least $m \geq 1$ with $E_\sigma^m = I_7$.

Solution on page 1398

Exercise 5.2.42. STANDARD

(a) Find $\sigma \in S_5$ with $\sigma^2 = (1\ 2\ 3\ 4\ 5)$.

(b) Find $\sigma \in S_4$ with $\sigma^2 = (1\ 2)(3\ 4)$.

(c) Show that there is no $\sigma \in S_n$ (for any $n \geq 5$) with $\sigma^2 = (1\ 2)(3\ 4\ 5)$.

Solution on page 1399

Exercise 5.2.43. STANDARD Show that $\mathrm{inv}(\sigma)$ is the least number of adjacent transpositions whose product is σ. That is, σ is a product of $\mathrm{inv}(\sigma)$ adjacent transpositions, and not of fewer.

Solution on page 1399

Exercise 5.2.44. STANDARD A matrix is **monomial** if it has exactly one nonzero entry in each row and in each column.

(a) Let $A \in F^{n\times n}$ be monomial. Show that there are $\sigma \in S_n$ and nonzero $c_1, \ldots, c_n \in F$ such that column j of A is $c_j \mathbf{e}_{\sigma(j)}$. Deduce that if $D : F^{n\times n} \to F$ is multilinear and alternating, then $D(A) = \mathrm{sgn}(\sigma) c_1 c_2 \cdots c_n D(I_n)$.

(b) Over $\mathbb{C}$, find $D(A)$ in terms of $D(I_4)$ for

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ i & 0 & 0 & 0 \\ 0 & 0 & 1-i & 0 \end{pmatrix}.$$

(c) Over $\mathbb{F}_3$, find $D(B)$ in terms of $D(I_3)$ for $B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 0 & 0 \end{pmatrix}$.

Solution on page 1400

Exercise 5.2.45. STANDARD

(a) Let $2 \leq k \leq n$. Show that S_n contains exactly $n!/((n-k)!\,k)$ cycles of length k.

(b) Sort the 24 elements of S_4 by their cycle notation: the identity, transpositions, 3-cycles, 4-cycles, and products of two disjoint transpositions. Count each kind, and use the counts to check Proposition 5.2.27 for $n = 4$.

Solution on page 1400

Exercise 5.2.46. HARD Let $n \geq 2$, let F be a field, and let $f : S_n \to F$ satisfy $f(\sigma) \neq 0$ and $f(\sigma\rho) = f(\sigma) f(\rho)$ for all $\sigma, \rho \in S_n$. Show that either $f(\sigma) = 1$ for all σ, or $f(\sigma) = \mathrm{sgn}(\sigma)$ for all σ.

Solution on page 1401

Exercise 5.2.47 (The 15 puzzle). HARD In the 15 puzzle, fifteen square tiles numbered $1, \ldots, 15$ and one empty space sit in a 4×4 frame. A move slides a tile that shares an edge with the empty space into it.

In the solved position the tiles read $1, \ldots, 15$ row by row, from the top left, with the empty space in the bottom right corner. Show that no sequence of moves turns the solved position into the position in which tiles 14 and 15 are interchanged and all other tiles, and the empty space, are where they were.

Solution on page 1401

Exercise 5.2.48. HARD The numbers $1, \ldots, n$ (where $n \geq 3$) are written in a row in some order. A move chooses three consecutive entries x, y, z and replaces them by z, x, y. For which starting orders can the row be sorted into $1, 2, \ldots, n$ by a sequence of moves?

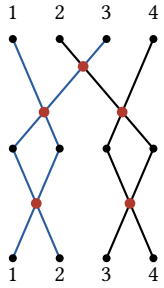
Solution on page 1401

Exercise 5.2.49. HARD This exercise proves $\operatorname{sgn}(\sigma\rho) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\rho)$ directly from inversions, without using Theorem 5.2.18.

For a two-element set $\{i, j\} \subseteq [n]$, say that $\pi \in S_n$ **reverses** $\{i, j\}$ if $\pi(i) - \pi(j)$ and $i - j$ have opposite signs. So $\operatorname{inv}(\pi)$ is the number of two-element subsets that π reverses.

(a) Show that $\sigma\rho$ reverses $\{i, j\}$ if and only if exactly one of the following holds: ρ reverses $\{i, j\}$; σ reverses $\{\rho(i), \rho(j)\}$.

(b) Deduce that $\operatorname{inv}(\sigma\rho) - \operatorname{inv}(\sigma) - \operatorname{inv}(\rho)$ is even. Figure 5.2.2 shows an example.



ρ (top) followed by σ (bottom)

Figure 5.2.2. Stacking the diagrams of $\rho = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$ and $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ gives a diagram of $\sigma\rho = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}$. The blue strands from 1 and 3 cross once in each half, so they end in their original order. There are $3 + 2 = 5$ crossings but only $\operatorname{inv}(\sigma\rho) = 1$ inversion, and $5 - 1$ is even.

(c) Use (b) and Exercise 5.2.34 to give another proof of Theorem 5.2.18.

Solution on page 1402

Exercise 5.2.50. HARD BEYOND 21-242 This exercise gives a third proof of Theorem 5.2.18, by counting orbits. Recall that $c(\sigma)$ is the number of orbits of σ, fixed points included.

(a) Let $\sigma \in S_n$ and let $\tau = (a\ b)$. Show that if a and b lie in the same orbit of σ, then $c(\sigma\tau) = c(\sigma) + 1$, and otherwise $c(\sigma\tau) = c(\sigma) - 1$.

(b) Deduce that if σ is a product of r transpositions, then $c(\sigma) - (n - r)$ is even, and use this to prove Theorem 5.2.18 and $\operatorname{sgn}(\sigma) = (-1)^{n-c(\sigma)}$ without Lemma 5.2.20 or Lemma 5.2.22.

(c) Show that the least number of transpositions whose product is σ is $n - c(\sigma)$.

Solution on page 1403

Exercise 5.2.51. CHALLENGE BEYOND 21-242 Let p be an odd prime and let $a \in \mathbb{F}_p$ with $a \neq 0$. Multiplication by a permutes the nonzero elements $1, 2, \ldots, p-1$ of $\mathbb{F}_p$; let $\mu_a \in S_{p-1}$ be the permutation with $\mu_a(k) = ak$, the product computed in $\mathbb{F}_p$.

(a) Show that $a^{p-1} = 1$ in $\mathbb{F}_p$.

(b) Show that $\operatorname{sgn}(\mu_a) = a^{(p-1)/2}$, where the right side is computed in $\mathbb{F}_p$ and is therefore 1 or -1. (This is **Zolotarev's lemma**.)

(c) Check (b) for $p = 7$ and $a = 2, 3$ by finding the cycle notation of μ_2 and μ_3.

Solution on page 1403

5.3 Existence and uniqueness

Theorem 5.1.13 showed that the determinant axioms leave exactly one function on $F^{2\times 2}$. Intuition (What the proof did, and where it leads) outlined the same argument for $n \times n$ matrices and left two questions open. The first, whether an ordering of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ could be sorted both by an even and by an odd number of interchanges, was settled in Section 5.2: it cannot (Theorem 5.2.18), and $D\big(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}\big) = \operatorname{sgn}(\sigma) D(I_n)$ (Theorem 5.2.31). The second, whether the resulting formula satisfies the axioms, is answered here, and this section finishes the argument. Carrying out the expansion shows that a determinant function must be given by an explicit formula, the Leibniz formula, so there is at most one. Checking that the formula satisfies the axioms shows that there is one. The formula has $n!$ terms and is a poor way to compute, but it displays $\det A$ as an explicit function of the entries of A, and the rest of the section reads several facts off it: the 3×3 formula, the determinants of triangular and block-diagonal matrices, and the fact that $\det A$ is a polynomial in the entries.

Throughout, $A \in F^{n\times n}$ has entries a_{ij} and columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, so that $\mathbf{a}_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i$.

5.3.1 The axioms force a formula

Intuition (Strategy). As in the 2×2 case, we break $D(A)$ into values of D at matrices of 0s and 1s. Write each column in the standard basis and expand by multilinearity in every column (Technique 5.1.7). Each column contributes n terms, so there are n^n terms, one for each way of choosing a row index in every column. Written with n nested sums, $\sum_{i_1} \cdots \sum_{i_n}$, this quickly becomes unreadable. The better bookkeeping indexes a choice of row indices by a single object: a function $f : [n] \to [n]$, where $f(j)$ is the row chosen in column j. Alternation then kills every term whose f takes some value twice, the survivors are indexed by permutations, and Theorem 5.2.31 evaluates them.

Notation 5.3.1. We write $[n]^{[n]}$ for the set of all functions $f : [n] \to [n]$. It has n^n elements, since each of $f(1), \ldots, f(n)$ can be any element of $[n]$. For $f \in [n]^{[n]}$, let $E_f \in F^{n\times n}$ be the matrix with columns $\mathbf{e}_{f(1)}, \mathbf{e}_{f(2)}, \ldots, \mathbf{e}_{f(n)}$. When $f = \sigma$ is a permutation, E_σ is the permutation matrix of Definition 5.2.29.

For example, the function $f : [3] \to [3]$ of Non-example 5.2.3, with $f(1) = f(2) = 2$ and $f(3) = 1$, gives the matrix E_f with columns $\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1$.

Lemma 5.3.2 (Expanding every column). Let $D : F^{n\times n} \to F$ be multilinear. Then for every $A \in F^{n\times n}$,

$$D(A) = \sum_{f \in [n]^{[n]}} a_{f(1)1} a_{f(2)2} \cdots a_{f(n)n} D\big(E_f\big).$$

Proof. STANDARD

Uses. Definition 5.1.1 (Determinant function), Definition 1.6.8 (Standard bases), induction on the number k of expanded columns
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Expand one column at a time. After the first k columns have been expanded, what does a single term look like?
Hint 2. For $0 \le k \le n$ and $g : [k] \to [n]$, let A_g be the matrix with columns $\mathbf{e}_{g(1)}, \dots, \mathbf{e}_{g(k)}, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n$. Prove by induction on k that $D(A) = \sum_g a_{g(1)1} \cdots a_{g(k)k} D\big(A_g\big)$, the sum running over all $g : [k] \to [n]$.
Hint 3. For the inductive step, expand column $k+1$ of A_g as $\sum_i a_{i,k+1}\mathbf{e}_i$. A function $[k+1] \to [n]$ is the same thing as a function $g : [k] \to [n]$ together with a value $i \in [n]$ at $k+1$.
Sketch. Induct on the number k of columns already expanded. The statement for k writes $D(A)$ as the sum, over $g : [k] \to [n]$, of $a_{g(1)1} \cdots a_{g(k)k}$ times D at the matrix whose first k columns are $\mathbf{e}_{g(1)}, \dots, \mathbf{e}_{g(k)}$ and whose other columns are those of A. Expanding column $k+1$ of each such matrix by linearity gives the statement for $k+1$, because extending g by a value at $k+1$ produces each function $[k+1] \to [n]$ exactly once. For $k = n$ these matrices are the E_f.

For $0 \le k \le n$ and a function $g : [k] \to [n]$, let A_g be the matrix with columns $\mathbf{e}_{g(1)}, \dots, \mathbf{e}_{g(k)}, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n$. For $k = 0$ there is exactly one function $g : \varnothing \to [n]$, and $A_g = A$. We show by induction on k that

$$D(A) = \sum_{g:[k]\to[n]} a_{g(1)1} \cdots a_{g(k)k} D\big(A_g\big),$$

the sum running over all functions $g : [k] \to [n]$. For $k = 0$ the sum has the single term $D(A)$, whose coefficient is the empty product 1.

Suppose the identity holds for some $k < n$, and fix $g : [k] \to [n]$. For $i \in [n]$, let $g_i : [k+1] \to [n]$ agree with g on $[k]$ and send $k+1$ to i. Column $k+1$ of A_g is $\mathbf{a}_{k+1} = \sum_{i=1}^{n} a_{i,k+1}\mathbf{e}_i$, and replacing it by $\mathbf{e}_i$ gives A_{g_i}. So by linearity in column $k+1$,

$$D\big(A_g\big) = \sum_{i=1}^{n} a_{i,k+1} D\big(A_{g_i}\big).$$

Substituting this into the identity for k,

$$D(A) = \sum_{g:[k]\to[n]} \sum_{i=1}^{n} a_{g(1)1} \cdots a_{g(k)k} a_{i,k+1} D\big(A_{g_i}\big).$$

Every function $h : [k+1] \to [n]$ equals g_i for exactly one pair (g, i), namely $g = h|_{[k]}$ and $i = h(k+1)$, and the coefficient of that term is $a_{h(1)1} \cdots a_{h(k+1),k+1}$. So the double sum equals $\sum_h a_{h(1)1} \cdots a_{h(k+1),k+1} D(A_h)$, the sum over all $h : [k+1] \to [n]$, which is the identity for $k+1$. For $k = n$ we have $A_f = E_f$, and the identity is the claim. □

The lemma uses only multilinearity: it is Technique 5.1.7 carried out in all n columns at once. Alternation now removes most of the n^n terms.

Theorem 5.3.3 (The Leibniz formula). Let $D : F^{n\times n} \to F$ be multilinear and alternating. Then for every $A \in F^{n\times n}$,

$$D(A) = D(I_n) \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

In particular, every determinant function D on $F^{n\times n}$ satisfies

$$D(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Proof. ROUTINE

Uses. Lemma 5.3.2 (Expanding every column), Definition 5.1.1 (Determinant function), Definition 5.2.1 (Permutation), Theorem 5.2.31 (Alternating functions and permutations), an injective function $[n] \to [n]$ is a bijection (Section A.1)
Techniques. Expand by multilinearity (Technique 5.1.7)

Hint 1. Start from Lemma 5.3.2. For which f does alternation force $D\big(E_f\big) = 0$?
Hint 2. If f is not injective, two columns of E_f are the same standard basis vector. The injective functions $[n] \to [n]$ are exactly the permutations.
Hint 3. For a permutation σ, Theorem 5.2.31 gives $D(E_\sigma)$.
Sketch. By Lemma 5.3.2, $D(A) = \sum_f a_{f(1)1} \cdots a_{f(n)n} D\big(E_f\big)$. If $f(p) = f(q)$ with $p \neq q$, then E_f has two equal columns and $D\big(E_f\big) = 0$. The remaining f are injective, hence permutations, and $D(E_\sigma) = \operatorname{sgn}(\sigma) D(I_n)$.

By Lemma 5.3.2,

$$D(A) = \sum_{f \in [n]^{[n]}} a_{f(1)1} \cdots a_{f(n)n} D\big(E_f\big).$$

Suppose f is not injective, say $f(p) = f(q)$ with $p \neq q$. Then columns p and q of E_f are both $\mathbf{e}_{f(p)}$, so $D\big(E_f\big) = 0$ because D is alternating, and the term of f is 0. An injective function from the finite set $[n]$ to itself is a bijection, so the remaining terms are indexed by the permutations $\sigma \in S_n$. For these, $D(E_\sigma) = \operatorname{sgn}(\sigma) D(I_n)$ by Theorem 5.2.31. Hence

$$D(A) = \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) D(I_n),$$

which is the claim after pulling out the common factor $D(I_n)$. If D is a determinant function, then $D(I_n) = 1$. □

For $n = 2$, the permutations are id and (1 2), with signs 1 and -1, and the formula reads $D(A) = (a_{11}a_{22} - a_{21}a_{12})D(I_2)$. This is part (a) of Theorem 5.1.13. For $n = 1$ it reads $D(a) = aD(1)$, as in Example 5.1.2.

Theorem 5.3.3 proves uniqueness: any two determinant functions on $F^{n \times n}$ are given by the same formula, so they are equal. It does not prove existence. Its hypothesis is that D is multilinear and alternating, and its conclusion describes such a D if there is one. For $n = 2$ we checked separately that $ad - bc$ satisfies the axioms (part (b) of Theorem 5.1.13). For general n that check is the next subsection.

5.3.2 The formula satisfies the axioms

By Theorem 5.3.3, if a determinant function on $F^{n \times n}$ exists, it is the function $L : F^{n \times n} \to F$ given by

$$L(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \tag{5.3.1}$$

(The name L is temporary, until we know that L is a determinant function.) For $\sigma \in S_n$ we write

$$t_\sigma(A) = \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

for the term of σ, so that $L(A) = \sum_\sigma t_\sigma(A)$. Realize that $t_\sigma(A)$ contains exactly one entry from each column of A, namely $a_{\sigma(j)j}$ from column j. This makes normalization and multilinearity routine to check. Alternation is the delicate axiom: when two columns of A are equal, no single term needs to vanish, and the terms have to cancel against each other.

Theorem 5.3.4 (Existence and uniqueness of the determinant). Let $n \geq 1$. The function L of (5.3.1) is a determinant function on $F^{n \times n}$, and it is the only one. Moreover, every multilinear alternating function $D : F^{n \times n} \to F$ satisfies $D(A) = D(I_n)L(A)$ for all $A \in F^{n \times n}$.

Proof, part 1 of 3: uniqueness. ROUTINE

Uses. Theorem 5.3.3 (The Leibniz formula), Definition 5.1.1 (Determinant function)
Hint 1. This half is already done. Which result expresses every multilinear alternating D through the formula (5.3.1)?
Hint 2. Theorem 5.3.3 says that $D(A) = D(I_n)L(A)$. What is $D(I_n)$ for a determinant function?
Sketch. By Theorem 5.3.3, every multilinear alternating D equals $D(I_n)L$, and a determinant function has $D(I_n) = 1$, so it equals L.

Let D be multilinear and alternating. By Theorem 5.3.3, $D(A) = D(I_n)L(A)$ for every A, which is the last sentence of the theorem. If D is a determinant function, then $D(I_n) = 1$, so $D = L$. Hence there is at most one determinant function, and if there is one, it is L. It remains to show that L is a determinant function.

Proof, part 2 of 3: L is normalized and multilinear. ROUTINE

> *Uses.* Definition 5.1.1 (Determinant function), Definition 5.2.23 (Sign of a permutation)
> *Hint 1.* For normalization, the (i, j) entry of I_n is 1 if $i = j$ and 0 otherwise. Which terms of $L(I_n)$ can be nonzero?
> *Hint 2.* For linearity in column k, each term contains exactly one entry of column k, namely $a_{\sigma(k)k}$. Separate it from the other factors.
> *Sketch.* In $L(I_n)$, a term with $\sigma \neq \mathrm{id}$ contains a factor $a_{\sigma(j)j} = 0$ for some j with $\sigma(j) \neq j$, so only $t_{\mathrm{id}}(I_n) = 1$ survives. If column k is $\mathbf{x}$, then $L = \sum_\sigma b_\sigma x_{\sigma(k)}$, where $b_\sigma = \operatorname{sgn}(\sigma) \prod_{j \neq k} a_{\sigma(j)j}$ does not depend on $\mathbf{x}$, and such a function of $\mathbf{x}$ is linear.

Normalized. The (i, j) entry of I_n is 1 if $i = j$ and 0 otherwise. If $\sigma \neq \mathrm{id}$, choose j with $\sigma(j) \neq j$; then the factor $a_{\sigma(j)j}$ of $t_\sigma(I_n)$ is 0, so $t_\sigma(I_n) = 0$. Hence $L(I_n) = t_{\mathrm{id}}(I_n) = \operatorname{sgn}(\mathrm{id}) \cdot 1 \cdot \cdots \cdot 1 = 1$.

Multilinear. Fix $k \in [n]$ and the columns $\mathbf{a}_j$ for $j \neq k$. For $\mathbf{x} \in F^n$, let $A_{\mathbf{x}}$ be the matrix with $\mathbf{x}$ in column k and $\mathbf{a}_j$ in column j for every $j \neq k$. The only factor of $t_\sigma(A_{\mathbf{x}})$ that comes from column k is $x_{\sigma(k)}$, so

$$L(A_{\mathbf{x}}) = \sum_{\sigma \in S_n} b_\sigma x_{\sigma(k)}, \quad \text{where } b_\sigma = \operatorname{sgn}(\sigma) \prod_{j \neq k} a_{\sigma(j)j}$$

does not depend on $\mathbf{x}$. Hence, for $\mathbf{x}, \mathbf{y} \in F^n$ and $c \in F$,

$$\begin{aligned} L\big(A_{c\mathbf{x}+\mathbf{y}}\big) &= \sum_\sigma b_\sigma \big(c x_{\sigma(k)} + y_{\sigma(k)}\big) \\ &= c \sum_\sigma b_\sigma x_{\sigma(k)} + \sum_\sigma b_\sigma y_{\sigma(k)} = cL(A_{\mathbf{x}}) + L\big(A_{\mathbf{y}}\big). \end{aligned}$$

So L is linear in column k, and k was arbitrary.

> *Intuition (Why alternation needs care).* Suppose columns k and l of A are equal, where $k \neq l$, and let $\tau = (k\ l)$. The permutation $\sigma\tau$ uses the row $\sigma(l)$ in column k and the row $\sigma(k)$ in column l, and otherwise agrees with σ. Since columns k and l are equal, the product of entries in $t_{\sigma\tau}(A)$ is the same as in $t_\sigma(A)$, while the sign is opposite. So $t_{\sigma\tau}(A) = -t_\sigma(A)$.
>
> The quickest way to use this is to reindex the whole sum: by Proposition 5.2.8, $L(A) = \sum_\sigma t_{\sigma\tau}(A) = -\sum_\sigma t_\sigma(A) = -L(A)$. But this only shows $(1 + 1)L(A) = 0$, which gives $L(A) = 0$ when $1 + 1 \neq 0$ (compare Proposition 5.1.11). Over $\mathbb{F}_2$ it says nothing, since $-L(A) = L(A)$ there.
>
> The fix is to cancel the terms two at a time instead of summing them all at once. The permutations σ and $\sigma\tau$ are different, one is even and the other odd, and their terms are negatives of each other. Grouping the $n!$ terms into $n!/2$ such pairs writes $L(A)$ as a sum of zeros, and no division is involved.

Technique 5.3.5 (Cancel terms in pairs). To show that a finite sum $\sum_{x \in X} t(x)$ is 0, find a bijection $\varphi : X \to X$ with $\varphi(\varphi(x)) = x$ and $\varphi(x) \neq x$ for every x, such that $t(\varphi(x)) = -t(x)$. Then X splits into two-element sets $\{x, \varphi(x)\}$, and the two terms in each set cancel. To write this down, choose a set $P \subseteq X$ containing exactly one element of each pair, so that φ maps P bijectively onto $X \setminus P$; then $\sum_{x \in X} t(x) = \sum_{x \in P}(t(x) + t(\varphi(x))) = 0$. If φ has fixed points, and $t(\varphi(x)) = -t(x)$ holds for the other x, the same argument shows that the sum equals the sum over the fixed points, which is often easy to evaluate.

This is not the same as the shortcut "reindexing by φ shows $S = -S$, so $S = 0$". The shortcut divides by $1 + 1$, which is impossible in characteristic 2, while pairing never divides. Reach for pairing when the terms of a sum come with signs and a natural operation flips the sign: for sums over S_n, composing with a fixed transposition. The canonical example is the alternation of the Leibniz formula in Theorem 5.3.4.

Counting is a special case: an involution with no fixed points shows that a finite set has an even number of elements, and in general $|X|$ has the same parity as the number of fixed points of φ (Exercise 5.3.40).
Used in. Theorem 5.3.4, Exercise 5.3.27, Exercise 5.3.39, Exercise 5.3.40, Exercise 5.6.13, Problem 5.1, Problem 5.15, Proposition 7.3.2.

Proof of Theorem 5.3.4, part 3 of 3: L is alternating. STANDARD

Uses. Remark 5.2.11 (Composing with a transposition), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.27 (Half of all permutations are even), Proposition 5.2.8 (Reindexing sums over S_n), Proposition 5.2.13 (Products of transpositions), Theorem 5.2.18 (Parity is well defined)
Techniques. Cancel terms in pairs (Technique 5.3.5)
Hint 1. Suppose columns $k \neq l$ of A are equal, and let $\tau = (k\ l)$. Compare the terms $t_\sigma(A)$ and $t_{\sigma\tau}(A)$.
Hint 2. $(\sigma\tau)(k) = \sigma(l)$, $(\sigma\tau)(l) = \sigma(k)$, and $\sigma\tau$ agrees with σ elsewhere. Because columns k and l are equal, the products of entries agree, while the signs are opposite.
Hint 3. Split S_n into the even and the odd permutations. The map $\sigma \mapsto \sigma\tau$ is a bijection from the even ones to the odd ones. Pair each even σ with $\sigma\tau$.
Sketch. Let columns $k \neq l$ be equal and $\tau = (k\ l)$. The permutation $\sigma\tau$ differs from σ only by exchanging the rows used in columns k and l, and these columns are equal, so the products of entries agree; since $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn}(\sigma)$, $t_{\sigma\tau}(A) = -t_\sigma(A)$. The map $\sigma \mapsto \sigma\tau$ is a bijection from the even permutations to the odd ones, so $L(A) = \sum_{\sigma \text{ even}}(t_\sigma(A) + t_{\sigma\tau}(A)) = 0$.

Let $A \in F^{n\times n}$ with $\mathbf{a}_k = \mathbf{a}_l$ for some $k \neq l$, so that $n \geq 2$ and $a_{ik} = a_{il}$ for every $i \in [n]$. Let $\tau = (k\ l)$.

Step 1: $t_{\sigma\tau}(A) = -t_\sigma(A)$ for every $\sigma \in S_n$. By Remark 5.2.11, $(\sigma\tau)(k) = \sigma(l)$, $(\sigma\tau)(l) = \sigma(k)$, and $(\sigma\tau)(j) = \sigma(j)$ for $j \notin \{k, l\}$. So the product of entries in $t_{\sigma\tau}(A)$ is

$$\begin{aligned}\prod_{j=1}^{n} a_{(\sigma\tau)(j)j} &= a_{\sigma(l)k}a_{\sigma(k)l}\prod_{j\notin\{k,l\}} a_{\sigma(j)j} \\ &= a_{\sigma(l)l}a_{\sigma(k)k}\prod_{j\notin\{k,l\}} a_{\sigma(j)j} = \prod_{j=1}^{n} a_{\sigma(j)j},\end{aligned}$$

where the second equality uses $a_{\sigma(l)k} = a_{\sigma(l)l}$ and $a_{\sigma(k)l} = a_{\sigma(k)k}$, which hold because columns k and l are equal. By Proposition 5.2.25, $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = -\operatorname{sgn}(\sigma)$. Hence $t_{\sigma\tau}(A) = -t_\sigma(A)$.

Step 2: pairing. Let S_n^+ and S_n^- be the sets of even and of odd permutations. By Proposition 5.2.13 and Theorem 5.2.18, S_n is the disjoint union of S_n^+ and S_n^-. Let $\varphi(\sigma) = \sigma\tau$. As in the proof of Proposition 5.2.27, φ maps S_n^+ into S_n^- and S_n^- into S_n^+, because $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn}(\sigma)$, and $\varphi(\varphi(\sigma)) = \sigma\tau^2 = \sigma$. So the restrictions $S_n^+ \to S_n^-$ and $S_n^- \to S_n^+$ of φ are inverse to each other, and in particular φ maps S_n^+ bijectively onto S_n^-. Reindexing the sum over S_n^- along this bijection (the argument of Proposition 5.2.8 applies to any bijection between finite sets) and then using Step 1,

$$\begin{aligned}L(A) &= \sum_{\sigma\in S_n^+} t_\sigma(A) + \sum_{\rho\in S_n^-} t_\rho(A) = \sum_{\sigma\in S_n^+} t_\sigma(A) + \sum_{\sigma\in S_n^+} t_{\sigma\tau}(A) \\ &= \sum_{\sigma\in S_n^+} (t_\sigma(A) + t_{\sigma\tau}(A)) = \sum_{\sigma\in S_n^+} 0 = 0.\end{aligned}$$

So L is alternating. By Part 2, L is a determinant function, and by Part 1 it is the only one. □

Step 2 never divides by $1 + 1$, so the proof works over every field. Over $\mathbb{F}_2$, where $\operatorname{sgn}(\sigma) = 1$ for every σ, Step 1 says that the two terms of each pair are *equal*, and they cancel because $x + x = 0$ in $\mathbb{F}_2$.

Definition 5.3.6 (Determinant). The **determinant** is the unique determinant function $\det : F^{n\times n} \to F$ of Theorem 5.3.4. By Theorem 5.3.3,

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$$

for every $A \in F^{n\times n}$, and every multilinear alternating $D : F^{n\times n} \to F$ satisfies $D = D(I_n)\det$. As for D, we also write $\det(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ for the determinant of the matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

For $n = 2$ this is the function $ad - bc$ of Definition 5.1.14. From now on every result about multilinear alternating functions in Section 5.1 and Section 5.2 applies to det. For instance, $\det A = 0$ whenever the columns of A are linearly dependent (Proposition 5.1.17), and $\det(AE_\sigma) = \operatorname{sgn}(\sigma)\det A$ (Theorem 5.2.31).

Exercise 5.3.7. ROUTINE Let Alt_n be the vector space of multilinear alternating functions $F^{n\times n} \to F$ (Exercise 5.1.34). Show that $\operatorname{Alt}_n = \operatorname{span}(\det)$ and $\dim \operatorname{Alt}_n = 1$. Deduce that if $D \in \operatorname{Alt}_n$ and $D(A) \neq 0$ for some A, then $D = c\det$ for some $c \neq 0$, and $\det A \neq 0$.

Solution on page 1404

Exercise 5.3.8. STANDARD Show that for every $A \in F^{n\times n}$,

$$\det A = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Here each term takes one entry from each *row*: row i contributes the entry in column $\sigma(i)$.

Solution on page 1404

5.3.3 The 3×3 determinant

For $n = 3$ the Leibniz formula has $3! = 6$ terms, one for each permutation in Table 5.2.1. Listing the even permutations first,

$$\begin{aligned}\det A = {} & a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ & - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13}.\end{aligned} \tag{5.3.2}$$

Expanding the triple product of Exercise 5.1.32 gives the same six terms, as it must by Theorem 5.3.4. Each term takes one entry from each column, and since σ is a bijection, the rows $\sigma(1), \sigma(2), \sigma(3)$ are also different. Figure 5.3.1 marks the chosen entries. In general a term of the Leibniz formula is a choice of n entries, one in each row and one in each column, like n rooks on an $n \times n$ chessboard of which no two attack each other, together with the sign of the corresponding permutation.

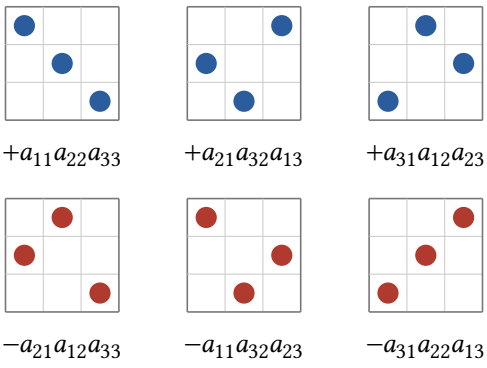


Figure 5.3.1. The six terms of (5.3.2). Each grid stands for a 3×3 matrix, and the dots mark the entries $a_{\sigma(1)1}, a_{\sigma(2)2}, a_{\sigma(3)3}$ multiplied in the term of σ. Top row (blue): the even permutations id, (1 2 3), (1 3 2). Bottom row (red): the odd permutations (1 2), (2 3), (1 3).

In the top row the dots run along the main diagonal and along its two parallels, wrapped around from the right edge to the left. In the bottom row they run along the other diagonal and its two wrapped parallels. Writing the first two columns again to the right of A unwraps these lines, as in Figure 5.3.2.

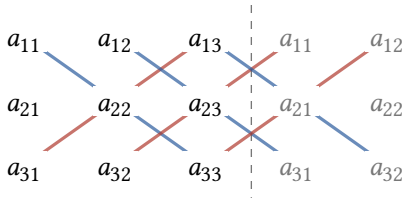


Figure 5.3.2. The rule of Sarrus. The three columns of A are followed by its first two columns again (grey). Products along the blue diagonals are added, and products along the red diagonals are subtracted.

Corollary 5.3.9 (Rule of Sarrus). Let $A \in F^{3\times 3}$, and write its three columns followed by its first two columns again, as in Figure 5.3.2. Then $\det A$ is the sum of the products along the three diagonals that run down and to the right from a_{11}, a_{12}, a_{13}, minus the sum of the products along the three diagonals that run up and to the right from a_{31}, a_{32}, a_{33}.

Proof. ROUTINE

Uses. Definition 5.3.6 (Determinant)
Hint 1. List the six diagonal products and compare them with the six terms of (5.3.2).
Hint 2. The diagonal down from a_{12} passes through a_{23} and then a_{31}, in the repeated first column.
Sketch. The down-right diagonals give $a_{11}a_{22}a_{33}$, $a_{12}a_{23}a_{31}$, $a_{13}a_{21}a_{32}$, the terms of the even permutations. The up-right diagonals give $a_{31}a_{22}a_{13}$, $a_{32}a_{23}a_{11}$, $a_{33}a_{21}a_{12}$, the terms of the transpositions.

The diagonals running down and to the right from a_{11}, a_{12}, a_{13} have products

$$a_{11}a_{22}a_{33}, \quad a_{12}a_{23}a_{31}, \quad a_{13}a_{21}a_{32},$$

and the diagonals running up and to the right from a_{31}, a_{32}, a_{33} have products

$$a_{31}a_{22}a_{13}, \quad a_{32}a_{23}a_{11}, \quad a_{33}a_{21}a_{12}.$$

After reordering the factors, the first three are the terms of id, (1 3 2), and (1 2 3) in (5.3.2), which enter with sign $+$, and the last three are the terms of (1 3), (2 3), and (1 2), which enter with sign $-$. So the rule computes exactly (5.3.2). □

Example 5.3.10 (A 3×3 determinant). Let $A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & 0 \\ -2 & 1 & 5 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$. The diagonals running down to the right have products $2\cdot 4\cdot 5 = 40$, $(-1)\cdot 0\cdot(-2) = 0$, and $3\cdot 1\cdot 1 = 3$. The diagonals running up to the right have products $(-2)\cdot 4\cdot 3 = -24$, $1\cdot 0\cdot 2 = 0$, and $5\cdot 1\cdot(-1) = -5$. So

$$\det A = (40 + 0 + 3) - (-24 + 0 - 5) = 43 + 29 = 72.$$

Since $\det A \neq 0$, A is invertible by Corollary 5.1.18.

Remark 5.3.11 (Sarrus does not generalize). The rule of Sarrus works because for $n = 3$ the six diagonals happen to be the six permutations. For $n = 4$, the same picture (the four columns followed by the first three again) has four diagonals running down and four running up, so it produces 8 products, while the Leibniz formula has $4! = 24$ terms. The signs go wrong as well (Exercise 5.3.25). For $n \geq 4$ there is no shortcut of this kind. Section 5.4 computes determinants by elimination, and Section 5.5 by expanding along a row or column.

5.3.4 Triangular and block-diagonal matrices

A term of the Leibniz formula vanishes as soon as one of its n entries is 0. So a matrix with a pattern of zeros can have very few nonzero terms, and the formula becomes a practical tool.

Technique 5.3.12 (Read structure off the Leibniz expansion). A term of the Leibniz formula is $\operatorname{sgn}(\sigma)$ times n entries of A, one in each column and one in each row, and it vanishes if one of these entries is 0. So to evaluate or analyze $\det A$ for a matrix with a pattern of zeros, ask which permutations σ satisfy $a_{\sigma(j)j} \neq 0$ for every j. Often the pattern confines σ to a small, explicit set, and only those terms need to be computed. The same question asked about entries that depend on a parameter, instead of entries that vanish, tells you how $\det A$ depends on the parameter (Technique 5.3.18).

Reach for it when the matrix is sparse or structured (triangular, block-diagonal, banded, mostly zeros), or when a question asks which products of entries can occur in $\det A$. Expect to prove a small lemma about permutations along the way, such as Lemma 5.3.14. The canonical example is Proposition 5.3.15, where only the identity survives.

Used in. Proposition 5.3.15, Proposition 5.3.17, Proposition 5.3.19, Exercise 5.3.24, Exercise 5.3.29, Exercise 5.3.33, Exercise 5.3.34, Exercise 5.3.38, Theorem 5.4.26, Theorem 5.5.3, Exercise 5.6.6, Exercise 5.6.13, Problem 5.14, Proposition 7.3.18, Theorem 7.3.24, Theorem 7.5.22.

Recall (Definition 3.3.31) that $A \in F^{n\times n}$ is upper triangular if $a_{ij} = 0$ whenever $i > j$, lower triangular if $a_{ij} = 0$ whenever $i < j$, and diagonal if it is both. We call A **triangular** if it is upper or lower triangular. The entries $a_{11}, \ldots, a_{nn}$ are the **diagonal entries** of A, and $\operatorname{diag}(d_1, \ldots, d_n)$ denotes the diagonal matrix with diagonal entries $d_1, \ldots, d_n$.

Example 5.3.13 (Triangular and not). The matrices

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are upper triangular, lower triangular, and neither, respectively. The transpose of an upper triangular matrix is lower triangular. I_n and the zero matrix are diagonal, and every elementary matrix that adds a multiple of one row to another (Definition 3.2.2) is triangular. The third matrix has determinant -1, while the product of its diagonal entries is 0, so the hypothesis of Proposition 5.3.15 below cannot be dropped.

For an upper triangular A, the entry $a_{\sigma(j)j}$ can be nonzero only if $\sigma(j) \leq j$. So the question of Technique 5.3.12 becomes: which permutations satisfy $\sigma(j) \leq j$ for every j?

Lemma 5.3.14 (Permutations with $\sigma(j) \leq j$). Let $\sigma \in S_n$. If $\sigma(j) \leq j$ for every $j \in [n]$, then $\sigma = \mathrm{id}$. Likewise, if $\sigma(j) \geq j$ for every $j \in [n]$, then $\sigma = \mathrm{id}$.

Proof (by induction). ROUTINE

Uses. Definition 5.2.1 (Permutation), Proposition 5.2.5 (Composition and inverses)
Techniques. Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Under the first hypothesis, what can $\sigma(1)$ be? Then $\sigma(2)$?
Hint 2. Show by induction on j that $\sigma(j) = j$: the value $\sigma(j)$ is at most j, and it differs from $\sigma(1), \ldots, \sigma(j-1)$.
Hint 3. For the second hypothesis, show that σ^{-1} satisfies the first.
Sketch. If $\sigma(j) \leq j$ for all j, then $\sigma(1) = 1$, and once $\sigma(i) = i$ for all $i < j$, injectivity rules out $\sigma(j) \in \{1, \ldots, j-1\}$, leaving $\sigma(j) = j$. If $\sigma(j) \geq j$ for all j, then $\sigma^{-1}(i) \leq i$ for all i, so $\sigma^{-1} = \mathrm{id}$.

Suppose $\sigma(j) \leq j$ for every j. We show by strong induction on j that $\sigma(j) = j$. Let $j \in [n]$, and assume $\sigma(i) = i$ for every $i < j$. Since σ is injective, $\sigma(j) \neq \sigma(i) = i$ for each $i < j$, so $\sigma(j) \notin \{1, \ldots, j-1\}$. Together with $\sigma(j) \leq j$ this gives $\sigma(j) = j$. Hence $\sigma = \mathrm{id}$.

Now suppose $\sigma(j) \geq j$ for every j. Let $i \in [n]$ and $j = \sigma^{-1}(i)$. Then $\sigma^{-1}(i) = j \leq \sigma(j) = i$. So σ^{-1} satisfies the first hypothesis, $\sigma^{-1} = \mathrm{id}$, and $\sigma = \mathrm{id}$. □

Proof (by summing). ROUTINE

Uses. Definition 5.2.1 (Permutation)
Hint 1. Compare $\sum_j \sigma(j)$ with $\sum_j j$.

Hint 2. Since σ is a bijection of $[n]$, the numbers $\sigma(1), \dots, \sigma(n)$ are $1, \dots, n$ in some order, so $\sum_j (j - \sigma(j)) = 0$. What do the hypotheses say about the signs of the terms?
Sketch. $\sum_{j=1}^{n}(j - \sigma(j)) = 0$ because $\sigma(1), \dots, \sigma(n)$ is a rearrangement of $1, \dots, n$. Under either hypothesis all terms have the same sign, so all of them are 0.

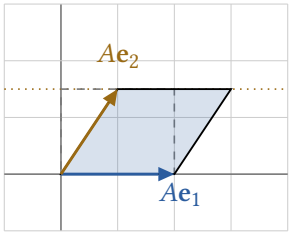
Since σ is a bijection $[n] \to [n]$, the list $\sigma(1), \dots, \sigma(n)$ is a rearrangement of $1, \dots, n$, so $\sum_j \sigma(j) = \sum_j j$, that is,

$$\sum_{j=1}^{n}(j - \sigma(j)) = 0.$$

Under the first hypothesis every term $j - \sigma(j)$ is a nonnegative integer, and a sum of nonnegative integers is 0 only if every term is 0. So $\sigma(j) = j$ for every j. Under the second hypothesis every term is a nonpositive integer, and the same argument applied to the negatives of the terms gives the same conclusion. □

The second proof handles both hypotheses at once and uses only that σ rearranges $1, \dots, n$. The first uses injectivity one value at a time, which is the kind of argument Technique 5.3.12 usually calls for.

Intuition (Triangular matrices in the plane). Let $A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Then $A\mathbf{e}_1 = (a_{11}, 0)$ lies on the horizontal axis, and $A\mathbf{e}_2 = (a_{12}, a_{22})$ lies on the horizontal line at height a_{22}. The parallelogram spanned by them has base $|a_{11}|$ and height $|a_{22}|$, so its area is $|a_{11}a_{22}|$ whatever a_{12} is. In terms of the axioms, when $a_{11} \neq 0$, changing a_{12} slides $A\mathbf{e}_2$ parallel to $A\mathbf{e}_1$, which is a column replacement (Lemma 5.1.8). Sliding all the way to $(0, a_{22})$ turns the parallelogram into a rectangle, the image of the unit square under $\operatorname{diag}(a_{11}, a_{22})$. This is the second proof below, in the smallest case.



The upper triangular matrix $A = \begin{pmatrix} 2 & 1 \\ 0 & 3/2 \end{pmatrix}$, with $\det A = 2 \cdot 3/2 = 3$. Drag $A\mathbf{e}_2$ left or right, or focus it and use the arrow keys: A stays upper triangular, and neither the base nor the height changes, so neither does $\det A$. Moving $A\mathbf{e}_1$ off the horizontal axis makes A not triangular. (Interactive in the web version.)

Proposition 5.3.15 (Triangular matrices). Let $A \in F^{n\times n}$ be upper triangular or lower triangular. Then

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

More generally, $D(A) = a_{11}a_{22}\cdots a_{nn}D(I_n)$ for every multilinear alternating $D : F^{n\times n} \to F$.

Proof (via the Leibniz formula). ROUTINE

Uses. Definition 5.3.6 (Determinant), Lemma 5.3.14 (Permutations with $\sigma(j) \leq j$), Definition 3.3.31 (Triangular matrices), Theorem 5.3.4 (Existence and uniqueness of the determinant)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. For an upper triangular A, when can the entry $a_{\sigma(j)j}$ be nonzero?
Hint 2. Only when $\sigma(j) \leq j$. If $\sigma \neq \mathrm{id}$, Lemma 5.3.14 gives a j with $\sigma(j) > j$.
Hint 3. For the statement about D, use the last sentence of Theorem 5.3.4.
Sketch. Let A be upper triangular. If $\sigma \neq \mathrm{id}$, then $\sigma(j) > j$ for some j by Lemma 5.3.14, so $a_{\sigma(j)j} = 0$ and the term of σ vanishes. Only $\sigma = \mathrm{id}$ survives, giving $a_{11}\cdots a_{nn}$. For lower triangular A, use $\sigma(j) < j$ instead. Then $D = D(I_n)\det$.

Suppose A is upper triangular, and let $\sigma \in S_n$ with $\sigma \neq \mathrm{id}$. By Lemma 5.3.14, $\sigma(j) \leq j$ fails for some j, that is, $\sigma(j) > j$. Then $a_{\sigma(j)j} = 0$ because A is upper triangular, so the term of σ in the Leibniz formula is 0. Only the term of the identity remains:

$$\det A = \operatorname{sgn}(\mathrm{id}) a_{11}a_{22}\cdots a_{nn} = a_{11}a_{22}\cdots a_{nn}.$$

If A is lower triangular and $\sigma \neq \mathrm{id}$, then Lemma 5.3.14 gives some j with $\sigma(j) < j$, so $a_{\sigma(j)j} = 0$, and the same conclusion follows. Finally, every multilinear alternating D equals $D(I_n)\det$ by Theorem 5.3.4. □

Proof (by column replacement). STANDARD

Uses. Lemma 5.1.8 (Column replacement), Definition 5.1.1 (Determinant function), Exercise 5.1.5, Definition 3.3.31 (Triangular matrices)
Hint 1. Column j of an upper triangular matrix is $a_{jj}\mathbf{e}_j$ plus a combination of $\mathbf{e}_1, \ldots, \mathbf{e}_{j-1}$. If $a_{11} \neq 0$, how can you clear the first row outside the diagonal without changing D?
Hint 2. Work from left to right. Once columns $1, \ldots, j-1$ have become $a_{11}\mathbf{e}_1, \ldots, a_{j-1,j-1}\mathbf{e}_{j-1}$, subtract suitable multiples of them from column j. This needs $a_{ii} \neq 0$ for $i < j$.
Hint 3. If some diagonal entry is 0, let k be the least such index and stop after processing column k. What is column k then?
Sketch. If all diagonal entries are nonzero, column replacements turn columns $1, 2, \ldots, n$ in turn into $a_{jj}\mathbf{e}_j$ without changing D, so $D(A) = D(\operatorname{diag}(a_{11}, \ldots, a_{nn})) = a_{11}\cdots a_{nn}D(I_n)$. If $a_{kk} = 0$ with k least, the same replacements in columns $2, \ldots, k$ make column k equal to $\mathbf{0}$, so $D(A) = 0$. For lower triangular matrices, work from right to left.

Let D be multilinear and alternating, and let A be upper triangular. Column j of A is

$$\mathbf{a}_j = a_{jj}\mathbf{e}_j + \sum_{i<j} a_{ij}\mathbf{e}_i.$$

For $j \geq 2$, consider the following step, which may be performed when columns $1, \ldots, j-1$ of the current matrix are $a_{11}\mathbf{e}_1, \ldots, a_{j-1,j-1}\mathbf{e}_{j-1}$, column j is still $\mathbf{a}_j$, and $a_{11}, \ldots, a_{j-1,j-1}$ are nonzero: for each $i < j$, add $-a_{ij}a_{ii}^{-1}$ times column i to column j. Since column i is $a_{ii}\mathbf{e}_i$, this subtracts $a_{ij}\mathbf{e}_i$ from column j. After the step, column j is $a_{jj}\mathbf{e}_j$, the other columns are unchanged, and D has not changed, by Lemma 5.1.8.

Column 1 of A is already $a_{11}\mathbf{e}_1$. Suppose first that all diagonal entries are nonzero. Performing the step for $j = 2, 3, \ldots, n$ in turn, we reach $\operatorname{diag}(a_{11}, \ldots, a_{nn})$, whose columns are $a_{jj}\mathbf{e}_j$. Pulling the scalar a_{jj} out of each column j by multilinearity,

$$D(A) = D(\operatorname{diag}(a_{11}, \ldots, a_{nn})) = a_{11}a_{22}\cdots a_{nn}D(I_n).$$

Suppose instead that some diagonal entry is 0, and let k be the least index with $a_{kk} = 0$. If $k = 1$, column 1 of A is $\mathbf{0}$. Otherwise, perform the step for $j = 2, \ldots, k$; this uses only $a_{11}, \ldots, a_{k-1,k-1}$, which are nonzero. Afterwards column k is $a_{kk}\mathbf{e}_k = \mathbf{0}$. In both cases $D(A)$ equals the value of D at a matrix with a zero column, which is 0 by Exercise 5.1.5. And $a_{11}\cdots a_{nn}D(I_n) = 0$ as well, since the factor a_{kk} is 0.

If A is lower triangular, the same argument works with the columns processed from right to left. Column n is $a_{nn}\mathbf{e}_n$, column j is $a_{jj}\mathbf{e}_j$ plus a combination of $\mathbf{e}_{j+1}, \ldots, \mathbf{e}_n$, and the step for $j = n-1, \ldots, 1$ subtracts multiples of the columns to its right; if some diagonal entry is 0, take k to be the largest such index. □

The first proof is shorter, given the Leibniz formula. The second uses neither the formula nor permutations, only the axioms. It is the first instance of computing a determinant by column operations, which Section 5.4 turns into an algorithm.

In particular, $\det \operatorname{diag}(d_1, \ldots, d_n) = d_1d_2\cdots d_n$. If a triangular matrix has no zero diagonal entries, its determinant is nonzero, so it is invertible by Corollary 5.1.18. If it has a zero diagonal entry, its determinant is 0, and Theorem 5.4.7 will show that it is not invertible.

Exercise 5.3.16. ROUTINE

(a) Compute the determinants of

$$\begin{pmatrix} 2 & 7 & -1 & 4 \\ 0 & -1 & 3 & 5 \\ 0 & 0 & 3 & 8 \\ 0 & 0 & 0 & 1/2 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$$

and

$$\begin{pmatrix} i & 0 & 0 \\ 2 & 1+i & 0 \\ -1 & 5 & 1-i \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

(b) A matrix $N \in F^{n\times n}$ is **strictly upper triangular** if its (i,j) entry is 0 whenever $i \geq j$. Show that $\det N = 0$ and $\det(cI_n + N) = c^n$ for every $c \in F$.

Solution on page 1405

The same reading of the Leibniz formula handles matrices built from square blocks along the diagonal.

Proposition 5.3.17 (Block-diagonal matrices). Let $A \in F^{k\times k}$ and $B \in F^{m\times m}$, and let

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in F^{(k+m)\times(k+m)},$$

where the zero blocks are $k \times m$ and $m \times k$ (Definition 2.4.33). Then $\det M = \det A \det B$.

Proof (via uniqueness). STANDARD

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Definition 5.1.1 (Determinant function), Proposition 2.1.17 (Compositions of linear maps are linear), Definition 2.4.33 (Block matrix)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Fix B and regard $\det\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$ as a function of $X \in F^{k\times k}$. Which axioms does it satisfy?
Hint 2. It is multilinear and alternating in the columns of X, so by Theorem 5.3.4 it equals its value at I_k times $\det X$.
Hint 3. Then show that $Y \mapsto \det\begin{pmatrix} I_k & 0 \\ 0 & Y \end{pmatrix}$ is a determinant function on $F^{m\times m}$.
Sketch. For fixed B, $X \mapsto \det\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$ is multilinear and alternating in the columns of X, whose columns enter as $(\mathbf{x}_j, \mathbf{0})$. So it equals $\det X \cdot \det\begin{pmatrix} I_k & 0 \\ 0 & B \end{pmatrix}$. In turn $Y \mapsto \det\begin{pmatrix} I_k & 0 \\ 0 & Y \end{pmatrix}$ is multilinear, alternating, and 1 at I_m, so it is det by uniqueness.

Fix B, and for $X \in F^{k\times k}$ let $D(X) = \det\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$. For $j \leq k$, column j of $\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$ is $\iota(\mathbf{x}_j)$, where $\iota : F^k \to F^{k+m}$, $\mathbf{x} \mapsto (\mathbf{x}, \mathbf{0})$, is linear; the other columns do not depend on X. So when all columns of X but the j-th are fixed, $\mathbf{x} \mapsto D(\ldots, \mathbf{x}, \ldots)$ is the composite of ι with a map that is linear because det is linear in column j, and it is linear by Proposition 2.1.17. If two columns $\mathbf{x}_i, \mathbf{x}_j$ of X are equal, then columns i and j of $\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$ are equal, so $D(X) = 0$. Thus D is multilinear and alternating, and Theorem 5.3.4 gives

$$\det M = D(A) = D(I_k) \det A = \det\begin{pmatrix} I_k & 0 \\ 0 & B \end{pmatrix} \cdot \det A.$$

Now for $Y \in F^{m\times m}$ let $E(Y) = \det\begin{pmatrix} I_k & 0 \\ 0 & Y \end{pmatrix}$. For $j \leq m$, column $k+j$ of this matrix is $(\mathbf{0}, \mathbf{y}_j)$, the image of $\mathbf{y}_j$ under the linear map $\mathbf{y} \mapsto (\mathbf{0}, \mathbf{y})$, and the first k columns do not depend on Y. The argument above shows that E is multilinear and alternating, and $E(I_m) = \det I_{k+m} = 1$. So E is a determinant function on $F^{m\times m}$, and $E = \det$ by Theorem 5.3.4. Hence $\det M = \det B \det A$, as desired. □

Proof (via the Leibniz formula). STANDARD

Uses. Definition 5.3.6 (Determinant), Definition 5.2.23 (Sign of a permutation), Proposition 5.2.13 (Products of transpositions), Definition 5.2.10 (Transposition)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)

Hint 1. Write c_{ij} for the entries of M and $n = k + m$. If the term of σ is nonzero, where can σ send $1, \dots, k$?
Hint 2. Such a σ maps $[k]$ onto itself and $\{k+1, \dots, n\}$ onto itself, so it is made of a permutation α of $[k]$ and a permutation β of $[m]$. Show that $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\alpha)\operatorname{sgn}(\beta)$ by factoring α and β into transpositions.
Hint 3. Then the sum over these σ is a sum over pairs (α, β), and it factors.
Sketch. A nonzero term needs $\sigma(j) \le k$ for every $j \le k$, so σ maps $[k]$ onto $[k]$ and the rest onto the rest. Such σ correspond to pairs $(\alpha, \beta) \in S_k \times S_m$, and factoring α and β into transpositions gives a factorization of σ, so $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\alpha)\operatorname{sgn}(\beta)$. The term of σ is $\operatorname{sgn}(\alpha)\prod a_{\alpha(j)j} \cdot \operatorname{sgn}(\beta)\prod b_{\beta(i)i}$, and the double sum factors as $\det A \det B$.

Let $n = k + m$, and write c_{ij} for the (i, j) entry of M. So $c_{ij} = a_{ij}$ for $i, j \le k$, $c_{k+i,k+j} = b_{ij}$ for $i, j \le m$, and $c_{ij} = 0$ otherwise.

Let $\sigma \in S_n$ have a nonzero term. For $j \le k$, the entry $c_{\sigma(j)j}$ is nonzero, so $\sigma(j) \le k$. Thus σ maps $[k]$ into $[k]$, hence onto $[k]$, being injective, and since σ is a bijection it maps $K = \{k+1, \dots, n\}$ onto K. Call such permutations *block-preserving*. A block-preserving σ determines $\alpha \in S_k$ and $\beta \in S_m$ by

$$\alpha(j) = \sigma(j) \quad (j \le k), \qquad \beta(i) = \sigma(k+i) - k \quad (i \le m),$$

and conversely every pair $(\alpha, \beta) \in S_k \times S_m$ arises from exactly one block-preserving σ, namely the one given by these formulas.

Signs. Write $\alpha = \tau_1 \cdots \tau_r$ and $\beta = \tau_1' \cdots \tau_s'$ as products of transpositions (Proposition 5.2.13). If $\tau_q = (p\ p')$ in S_k, let $\rho_q = (p\ p')$ in S_n, and if $\tau_q' = (p\ p')$ in S_m, let $\rho_q' = (k+p\ k+p')$ in S_n. Let $\hat\alpha = \rho_1 \cdots \rho_r$ and $\hat\beta = \rho_1' \cdots \rho_s'$. Each ρ_q fixes every point of K and acts on $[k]$ as τ_q does, so $\hat\alpha$ agrees with α on $[k]$ and fixes K. Similarly, $\hat\beta$ fixes $[k]$ and sends $k+i$ to $k+\beta(i)$. Hence $\hat\alpha\hat\beta = \sigma$: for $j \le k$, $\hat\alpha\big(\hat\beta(j)\big) = \alpha(j) = \sigma(j)$, and for $i \le m$, $\hat\alpha\big(\hat\beta(k+i)\big) = \hat\alpha(k+\beta(i)) = k+\beta(i) = \sigma(k+i)$. So σ is a product of $r+s$ transpositions, and $\operatorname{sgn}(\sigma) = (-1)^{r+s} = \operatorname{sgn}(\alpha)\operatorname{sgn}(\beta)$.

The sum. For block-preserving σ, the entries in its term are $c_{\sigma(j)j} = a_{\alpha(j)j}$ for $j \le k$ and $c_{\sigma(k+i),k+i} = c_{k+\beta(i),k+i} = b_{\beta(i)i}$ for $i \le m$. Since all other terms vanish,

$$\begin{aligned}
\det M &= \sum_{\alpha \in S_k} \sum_{\beta \in S_m} \operatorname{sgn}(\alpha)\operatorname{sgn}(\beta) \prod_{j=1}^{k} a_{\alpha(j)j} \prod_{i=1}^{m} b_{\beta(i)i} \\
&= \left(\sum_{\alpha \in S_k} \operatorname{sgn}(\alpha) \prod_{j=1}^{k} a_{\alpha(j)j} \right) \left(\sum_{\beta \in S_m} \operatorname{sgn}(\beta) \prod_{i=1}^{m} b_{\beta(i)i} \right) \\
&= \det A \det B. \qquad \square
\end{aligned}$$

The first proof is an instance of a pattern we will use repeatedly: to identify a function with det, check that it is multilinear and alternating, and evaluate it at the identity. The second proof never used the upper right block of M, a point that Section 5.4 takes up for block-triangular matrices. Applying the proposition repeatedly, a block-diagonal matrix with any number of square diagonal blocks has determinant equal to the product of the determinants of the blocks.

5.3.5 The determinant as a polynomial in the entries

The Leibniz formula computes $\det A$ from the entries of A using only additions, subtractions, and multiplications: it is a sum of $n!$ terms, each ± 1 times a product of n entries. No division occurs. This has two useful consequences.

Technique 5.3.18 (det is a polynomial in the entries). By the Leibniz formula, $\det A$ is a polynomial in the n^2 entries of A with integer coefficients ± 1, each of whose monomials has degree n and uses one entry from each row and each column. So if the entries of a matrix depend polynomially on a parameter t, then so does its determinant, and the degree and the leading coefficients can be read off the formula with Technique 5.3.12. Similarly, any map that respects sums and products can be applied to

det A entry by entry: for example, complex conjugation (Exercise 5.3.28) and reduction modulo a prime (Technique 5.3.21).

Reach for it when a question asks how det varies with a parameter, as for $\det(A + tI_n)$ or $\det(A + tB)$, for how many values of t a matrix can fail to be invertible, or how det interacts with an operation on the entries. The canonical example is Proposition 5.3.19.

Used in. Proposition 5.3.19, Proposition 5.3.20, Exercise 5.3.28, Exercise 5.6.9, Exercise 5.6.12, Problem 5.11, Problem 5.16, Problem 5.17, Problem 5.18, Proposition 7.3.22, Lemma 7.6.21.

Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t). Let $A \in F^{n\times n}$. There are scalars $c_0, c_1, \ldots, c_{n-1} \in F$, depending only on A, with $c_{n-1} = \operatorname{tr} A$ and $c_0 = \det A$, such that

$$\det(A + tI_n) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

for every $t \in F$.

Proof. STANDARD

Uses. Definition 5.3.6 (Determinant), Definition 2.4.26 (Trace), distributivity in F

Techniques. Read structure off the Leibniz expansion (Technique 5.3.12), det is a polynomial in the entries (Technique 5.3.18)

Hint 1. The (i, j) entry of $A + tI_n$ is a_{ij} if $i \neq j$ and $a_{jj} + t$ if $i = j$. In the term of σ, which factors contain t?

Hint 2. Only the factors $a_{jj} + t$ with $\sigma(j) = j$. Show that a permutation $\sigma \neq \mathrm{id}$ fixes at most $n - 2$ points.

Hint 3. So only the identity contributes to t^n and t^{n-1}. Expand $(a_{11} + t)\cdots(a_{nn} + t)$. For c_0, set $t = 0$.

Sketch. The term of σ has a factor $a_{jj} + t$ for each fixed point j of σ and constant factors otherwise, so expanding it gives powers of t up to the number of fixed points. If $\sigma \neq \mathrm{id}$ moves j, it also moves $\sigma(j)$, so it fixes at most $n - 2$ points. Hence t^n and t^{n-1} come only from $(a_{11} + t)\cdots(a_{nn} + t) = t^n + (a_{11} + \cdots + a_{nn})t^{n-1} + \cdots$. Setting $t = 0$ gives $c_0 = \det A$.

The (i, j) entry of $A + tI_n$ is a_{ij} if $i \neq j$ and $a_{jj} + t$ if $i = j$. For $\sigma \in S_n$, let $\operatorname{Fix}(\sigma) = \{j \in [n] \mid \sigma(j) = j\}$ and $f_\sigma = |\operatorname{Fix}(\sigma)|$. The term of σ in $\det(A + tI_n)$ is

$$T_\sigma(t) = \operatorname{sgn}(\sigma) \prod_{j \notin \operatorname{Fix}(\sigma)} a_{\sigma(j)j} \prod_{j \in \operatorname{Fix}(\sigma)} (a_{jj} + t).$$

Expanding the second product by distributivity, choosing either a_{jj} or t from each factor, gives

$$\prod_{j \in \operatorname{Fix}(\sigma)} (a_{jj} + t) = \sum_{S \subseteq \operatorname{Fix}(\sigma)} \left(\prod_{j \in S} a_{jj} \right) t^{f_\sigma - |S|}.$$

So $T_\sigma(t) = \sum_{p=0}^{f_\sigma} e_{\sigma,p} t^p$, where

$$e_{\sigma,p} = \operatorname{sgn}(\sigma) \prod_{j \notin \operatorname{Fix}(\sigma)} a_{\sigma(j)j} \sum_{S \subseteq \operatorname{Fix}(\sigma), |S| = f_\sigma - p} \prod_{j \in S} a_{jj}$$

does not depend on t. Set $e_{\sigma,p} = 0$ for $f_\sigma < p \leq n$.

If $\sigma \neq \mathrm{id}$, choose j with $\sigma(j) \neq j$, and let $i = \sigma(j)$. Then $\sigma(i) \neq i$, since otherwise $\sigma(i) = i = \sigma(j)$ with $i \neq j$, contradicting injectivity. So σ moves the two points i and j, and $f_\sigma \leq n - 2$. Hence $e_{\sigma,n} = e_{\sigma,n-1} = 0$ for $\sigma \neq \mathrm{id}$. For $\sigma = \mathrm{id}$ we have $f_{\mathrm{id}} = n$, and the subsets $S = \emptyset$ and $S = \{j\}$ show that $e_{\mathrm{id},n} = 1$ and $e_{\mathrm{id},n-1} = a_{11} + \cdots + a_{nn} = \operatorname{tr} A$.

Summing over σ, $\det(A + tI_n) = \sum_{p=0}^{n} c_p t^p$ with $c_p = \sum_\sigma e_{\sigma,p}$, and these scalars depend only on A. By the previous paragraph, $c_n = 1$ and $c_{n-1} = \operatorname{tr} A$. Finally, setting $t = 0$ gives $c_0 = \det(A + 0I_n) = \det A$. □

In particular, $A + tI_n$ fails to be invertible for at most n values of $t \in F$. Indeed, if $A + tI_n$ is not invertible, then $\det(A + tI_n) = 0$ by Corollary 5.1.18, so t is a root of the polynomial $x^n + c_{n-1}x^{n-1} + \cdots + c_0 \in F[x]$, which has degree n and therefore at most n roots (Proposition 7.1.22). For $n = 2$ the proposition is Exercise 5.1.38.

In Section 7.3, the values of t for which $\det(A + tI_n) = 0$ become, after a change of sign, the eigenvalues of A.

The proposition is about the coefficients $c_p = \sum_\sigma e_{\sigma,p}$ built in the proof, not about "the" coefficients of the function $t \mapsto \det(A + tI_n)$. Over a small field the function does not determine its coefficients: over $\mathbb{F}_2$, $t^3 = t^3 + t^2 + t$ for both elements t. From Section 7.1 on, $\det(A + xI_n)$ is a polynomial in $F[x]$ with the coefficients c_p.

The second consequence concerns matrices with integer entries. We regard such a matrix as an element of $\mathbb{Q}^{n\times n}$, and for a prime p we write $\overline{A} \in \mathbb{F}_p^{n\times n}$ for the matrix with entries $a_{ij} \bmod p$.

Proposition 5.3.20 (Reduction modulo p). Let $A \in \mathbb{Q}^{n\times n}$ have integer entries, and let p be a prime. Then $\det A$ is an integer, and $(\det A) \bmod p = \det \overline{A}$, the determinant of $\overline{A}$ computed in $\mathbb{F}_p$.

Proof. ROUTINE

Uses. Definition 5.3.6 (Determinant), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Definition 1.1.19 (Integers modulo n)
Techniques. det is a polynomial in the entries (Technique 5.3.18)
Hint 1. The Leibniz formula uses only sums, negatives, and products of the entries.
Hint 2. By Lemma 1.1.21, taking remainders modulo p respects sums and products of two integers. Extend this to sums and products of many integers, and to negatives.
Sketch. Each term of $\det A$ is $\pm$ a product of integers, so $\det A \in \mathbb{Z}$. Reduction modulo p respects sums, products, and negatives (Lemma 1.1.21, extended by induction), so reducing the Leibniz sum for A term by term gives the Leibniz sum for $\overline{A}$ in $\mathbb{F}_p$.

Each term $\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n}$ is $\pm$ a product of integers, so $\det A$ is an integer. Write $r(x) = x \bmod p$ for $x \in \mathbb{Z}$. By part (a) of Lemma 1.1.21, $r(x + y) = r(x) + r(y)$ and $r(xy) = r(x)r(y)$, with the operations on the right taken in $\mathbb{F}_p$. By induction on the number of terms or factors, r respects sums and products of any finite number of integers. It also respects negatives: $r(x) + r(-x) = r(x + (-x)) = r(0) = 0$, so $r(-x) = -r(x)$ in $\mathbb{F}_p$. Applying r to the Leibniz formula for A therefore gives

$$r(\det A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) r\big(a_{\sigma(1)1}\big)\cdots r\big(a_{\sigma(n)n}\big),$$

where $\operatorname{sgn}(\sigma) = \pm 1$ is now read in $\mathbb{F}_p$. The right side is the Leibniz formula for $\overline{A}$, so $r(\det A) = \det \overline{A}$. □

Technique 5.3.21 (Reduce modulo p). To show that a matrix with integer entries has nonzero determinant, and hence is invertible (Corollary 5.1.18), it is enough to find one prime p for which $\det \overline{A} \neq 0$ in $\mathbb{F}_p$: then $(\det A) \bmod p \neq 0$ by Proposition 5.3.20, so $\det A \neq 0$. Choose p so that $\overline{A}$ is simple, for instance triangular with nonzero diagonal entries or a permutation matrix, and there is no need to compute $\det A$ itself. The converse inference is not available: $\det \overline{A} = 0$ only says that p divides $\det A$.

Reach for it when an integer matrix has a recognizable pattern modulo a small prime, most often modulo 2, where the entries are just their parities and $\operatorname{sgn}(\sigma) = 1$ for every σ. The canonical example is Example 5.3.22.

Used in. Exercise 5.3.23, Exercise 5.3.36, Exercise 5.3.40, Exercise 5.6.12, Problem 5.15.

Example 5.3.22 (Odd diagonal, even elsewhere). Let $A \in \mathbb{Q}^{n\times n}$ have integer entries, with a_{ii} odd for every i and a_{ij} even for $i \neq j$. Then $\overline{A} = I_n$ in $\mathbb{F}_2^{n\times n}$, so by Proposition 5.3.20, $(\det A) \bmod 2 = \det I_n = 1$. So $\det A$ is odd, in particular nonzero, and A is invertible by Corollary 5.1.18. For instance,

$$\det\begin{pmatrix} 3 & 2 & -4 \\ 6 & 1 & 2 \\ 0 & 8 & 5 \end{pmatrix} = -285,$$

which is odd, as predicted.

Exercises

Exercise 5.3.23. ROUTINE Let

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 3 & 1 & 1 \\ 1 & 4 & 1 \end{pmatrix}.$$

(a) Compute $\det A$ over $\mathbb{Q}$ with the rule of Sarrus.

(b) Find $\det \overline{A}$ for the reductions $\overline{A}$ of A modulo 2, 3, 5, and 7.

(c) Over $\mathbb{F}_5$, find a nonzero $\mathbf{x} \in \mathbb{F}_5^3$ with $\overline{A}\mathbf{x} = \mathbf{0}$, and explain why this is consistent with (b).

Solution on page 1405

Exercise 5.3.24. ROUTINE Let

$$B = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & 4 \\ 1 & 0 & 5 & 0 \end{pmatrix}.$$

Each column of B has exactly two nonzero entries. Find the permutations $\sigma \in S_4$ whose terms in the Leibniz formula are nonzero, and compute $\det B$.

Solution on page 1405

Exercise 5.3.25. ROUTINE Find the errors in the following argument, and give a matrix for which the claim fails.

Claim. For $A \in F^{4\times 4}$, write the four columns of A followed by its first three columns again. Then $\det A$ is the sum of the products along the four diagonals running down and to the right from $a_{11}, a_{12}, a_{13}, a_{14}$, minus the sum of the products along the four diagonals running up and to the right from $a_{41}, a_{42}, a_{43}, a_{44}$.

Proof. By the Leibniz formula, $\det A$ is a signed sum of products of four entries, one from each row and one from each column. Each of the eight diagonals is such a product. The diagonals running down correspond to cyclic shifts of the columns, which have sign $+1$, and the diagonals running up have sign -1, exactly as in the rule of Sarrus.

Solution on page 1406

Exercise 5.3.26. STANDARD

(a) Find the error in the following argument.

Claim. $\det(A + B) = \det A + \det B$ for all $A, B \in F^{n\times n}$.

Proof. By the Leibniz formula,

$$\begin{aligned} \det(A+B) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} \left(a_{\sigma(j)j} + b_{\sigma(j)j}\right) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \left(\prod_{j=1}^{n} a_{\sigma(j)j} + \prod_{j=1}^{n} b_{\sigma(j)j} \right) = \det A + \det B. \end{aligned}$$

(b) For $S \subseteq [n]$, let C_S be the matrix whose column j is $\mathbf{a}_j$ if $j \in S$ and $\mathbf{b}_j$ if $j \notin S$. Show that $\det(A + B) = \sum_{S \subseteq [n]} \det C_S$.

Solution on page 1406

Exercise 5.3.27. STANDARD The axioms concern columns. This exercise shows, directly from the Leibniz formula, that det behaves the same way with respect to rows.

(a) Show that det is linear in each row of A when the other rows are held fixed.

(b) Show that $\det A = 0$ if two rows of A are equal.

(c) Let $d_1, \ldots, d_n \in F$ and $\Lambda = \operatorname{diag}(d_1, \ldots, d_n)$. Show that $\det(\Lambda A) = \det(A\Lambda) = d_1 \cdots d_n \det A$.

Solution on page 1407

Exercise 5.3.28. ROUTINE For $z = x + yi \in \mathbb{C}$ with $x, y \in \mathbb{R}$, let $\overline{z} = x - yi$, and for $A \in \mathbb{C}^{n\times n}$ let $\overline{A}$ be the matrix with entries $\overline{a_{ij}}$.

(a) Show that $\det \overline{A} = \overline{\det A}$. Deduce that a matrix with real entries has a real determinant.

(b) Suppose every entry of $A \in \mathbb{C}^{n\times n}$ is purely imaginary, that is, $\overline{A} = -A$. Show that $\det A$ is real if n is even and purely imaginary if n is odd.

Solution on page 1407

Exercise 5.3.29. STANDARD Let $A \in F^{n\times n}$, and suppose there are sets $R, C \subseteq [n]$ with $|R| + |C| > n$ such that $a_{ij} = 0$ for all $i \in R$ and $j \in C$. (So A has an $|R| \times |C|$ block of zeros, after rearranging rows and columns.)

(a) Show that $\det A = 0$.

(b) Deduce from (a) that an upper triangular matrix with a zero diagonal entry has determinant 0.

(c) Show that the hypothesis $|R| + |C| > n$ cannot be weakened to $|R| + |C| = n$, for any $n \geq 2$ and any $|R| = r$ with $1 \leq r \leq n - 1$.

Solution on page 1408

Exercise 5.3.30. STANDARD Prove or disprove each statement.

(a) If every entry of $A \in \mathbb{R}^{n\times n}$ is nonzero, then $\det A \neq 0$.

(b) If $A \in F^{n\times n}$ has exactly n nonzero entries and $\det A \neq 0$, then each row and each column of A contains exactly one nonzero entry.

(c) If $A \in F^{n\times n}$ is upper triangular, then $\det A^{\mathsf{T}} = \det A$.

(d) If $N \in F^{n\times n}$ is strictly upper triangular (Exercise 5.3.16), then $\det(A + N) = \det A$ for every $A \in F^{n\times n}$.

Solution on page 1408

Exercise 5.3.31. STANDARD The **permanent** of $A \in F^{n\times n}$ is the Leibniz formula without signs:

$$\operatorname{per} A = \sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

(a) Show that per is normalized and multilinear, and that $\operatorname{per}(AE_\pi) = \operatorname{per} A$ for every $\pi \in S_n$.

(b) Let $n \geq 2$. Show that per is a determinant function on $F^{n\times n}$ if and only if $1 + 1 = 0$ in F, and that in this case per = det.

(c) Let $A \in \mathbb{Q}^{n\times n}$ have all entries in $\{0, 1\}$. Show that per A is the number of permutations σ with $a_{\sigma(j)j} = 1$ for every j, and that $\det A$ and per A are both even or both odd.

Solution on page 1409

Exercise 5.3.32. STANDARD Let $A \in \mathbb{Q}^{n\times n}$ have every entry equal to 1 or -1. Show that $\det A$ is an integer divisible by 2^{n-1}.

Solution on page 1409

Exercise 5.3.33. ROUTINE Let $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(a) Compute $\det(A + tI_3)$ from the Leibniz formula, listing the nonzero terms, and check the coefficients against Proposition 5.3.19.

(b) Find all t with $\det(A + tI_3) = 0$, for $t \in \mathbb{C}$, for $t \in \mathbb{R}$, and for $t \in \mathbb{F}_2$ (regarding A as a matrix over $\mathbb{F}_2$).

Solution on page 1410

Exercise 5.3.34. STANDARD Let $n \geq 2$ and $A \in F^{n\times n}$. Show that the coefficient $c_{n-2} = \sum_\sigma e_{\sigma,n-2}$ constructed in the proof of Proposition 5.3.19 is

$$c_{n-2} = \sum_{1\leq i<j\leq n} \big(a_{ii}a_{jj} - a_{ij}a_{ji}\big),$$

the sum of the determinants of the 2×2 matrices formed by rows i, j and columns i, j of A.

Solution on page 1410

Exercise 5.3.35. STANDARD Let V be a vector space with a basis $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$. Show that there is exactly one multilinear alternating function $f : V \times \cdots \times V \to F$ of n vectors (Remark 5.1.6) with $f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = 1$, and that it is given by

$$f(\mathbf{w}_1, \ldots, \mathbf{w}_n) = \det\big([\mathbf{w}_1]_\mathcal{B}, \ldots, [\mathbf{w}_n]_\mathcal{B}\big).$$

Solution on page 1411

Exercise 5.3.36. STANDARD

(a) Show, without computing its determinant, that

$$A = \begin{pmatrix} 2 & 1 & 4 & 6 \\ 3 & 2 & 0 & 2 \\ 4 & 6 & 2 & 1 \\ 0 & 2 & 5 & 4 \end{pmatrix}$$

is invertible over $\mathbb{Q}$.

(b) Let $B = \begin{pmatrix} 4 & 1 & 3 \\ 0 & 2 & 6 \\ 3 & 3 & 1 \end{pmatrix}$. Explain why reducing B modulo 2 cannot show that B is invertible, and show that B is invertible by reducing modulo 3.

Solution on page 1411

Exercise 5.3.37. STANDARD Let $A \in F^{k\times k}$ and $B \in F^{m\times m}$. Show that

$$\det\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix} = (-1)^{km} \det A \det B,$$

where the upper left zero block is $k \times m$ and the lower right one is $m \times k$.

Solution on page 1412

Exercise 5.3.38. HARD

(a) Find a 3×3 matrix with all entries in $\{0, 1\}$ whose determinant (over $\mathbb{Q}$) is 2.

(b) Show that no 3×3 matrix with all entries in $\{0, 1\}$ has determinant greater than 2.

Solution on page 1412

Exercise 5.3.39. HARD This exercise constructs a determinant function by induction on n, without permutations. For $A \in F^{n \times n}$ with $n \geq 2$ and $i \in [n]$, let $A^{(i)} \in F^{(n-1)\times(n-1)}$ be A with row i and column 1 deleted. Define $D_1(a) = a$ and, for $n \geq 2$,

$$D_n(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} D_{n-1}\left(A^{(i)}\right).$$

(a) Show by induction on n that D_n is normalized and multilinear.

(b) Show by induction on n that $D_n(A) = 0$ whenever two adjacent columns of A are equal. Deduce that D_n is a determinant function, so $D_n = \det$.

(c) Explain why, with Remark 5.2.32, this gives a third proof of Theorem 5.2.18.

Solution on page 1413

Exercise 5.3.40. CHALLENGE Let n be even, and let $A \in \mathbb{Q}^{n \times n}$ have $a_{ii} = 0$ for every i and $a_{ij} \in \{1, -1\}$ for $i \neq j$. Show that A is invertible. Show by an example that the conclusion can fail for odd n.

Solution on page 1414

5.4 Determinants and invertibility

Corollary 5.1.18 says that a matrix that is not invertible has determinant 0. This section proves the converse, so that $\det A \neq 0$ exactly when A is invertible. The converse cannot follow from multilinearity and alternation alone, since the zero function has both properties, so the proof has to use normalization. Normalization only tells us that $\det I_n = 1$, and we get from I_n to an arbitrary invertible matrix by multiplying by elementary matrices (Theorem 3.5.19). So we first work out what one elementary matrix does to the determinant, and then build up. The same analysis proves $\det(AB) = \det A \det B$ and $\det A^{\mathsf{T}} = \det A$. From these we get the row versions of all the column properties, a fast way to compute determinants by row reduction, the determinant of a linear operator, and the determinants of block-triangular matrices.

5.4.1 Elementary matrices

Recall the elementary $n \times n$ matrices (Definition 3.2.2). For $i \neq j$, E_{ij} swaps rows i and j of I_n; for $c \neq 0$, E_{ic} scales row i of I_n by c; and E_{icj} adds c times row i of I_n to row j, so it is I_n with an extra entry c in position (j, i). Multiplying by E on the left performs the row operation (Proposition 3.2.7). Multiplying on the right performs a column operation (Proposition 3.2.28): AE_{ij} is A with columns i and j swapped, AE_{ic} is A with column i scaled by c, and AE_{icj} is A with c times column j added to column i.

Intuition (Why columns first). Our axioms are statements about columns, and Section 5.1 already found the effect of each column operation on a multilinear alternating function. A swap negates it (Proposition 5.1.10), scaling a column by c multiplies it by c (multilinearity), and adding a multiple of one column to another does not change it (Lemma 5.1.8). So the products we can handle now are the products AE, in which E acts on the columns of A. The products EA, which act on rows, have to wait for the product formula (Theorem 5.4.10), and the row versions of the axioms for $\det A^{\mathsf{T}} = \det A$ (Theorem 5.4.14).

The lecture passes between the two kinds of products with a trick worth remembering. By Proposition 2.4.23, $AE = \left(E^{\mathsf{T}}A^{\mathsf{T}}\right)^{\mathsf{T}}$. Here $E^{\mathsf{T}}A^{\mathsf{T}}$ performs a row operation on A^{T}, whose rows are the columns of A, and the outer transpose turns these rows back into columns. This is the first proof of Proposition 3.2.28.

Proposition 5.4.1 (Determinants of elementary matrices). Let $1 \le i, j \le n$ with $i \ne j$, and let $c \in F$. Then $\det E_{ij} = -1$, $\det E_{icj} = 1$, and $\det E_{ic} = c$ if $c \ne 0$. In particular, every elementary matrix has a nonzero determinant.

Proof. ROUTINE

Uses. Definition 3.2.2 (Elementary matrices), Definition 5.1.1 (Determinant function), Proposition 5.1.10 (Alternating implies antisymmetric)
Hint 1. Each elementary matrix is I_n with a small change. Describe the change in terms of the columns of I_n, and use the axiom that fits it.
Hint 2. E_{ij} is I_n with columns i and j interchanged, and E_{ic} is I_n with column i multiplied by c. Column i of E_{icj} is $\mathbf{e}_i + c\mathbf{e}_j$.
Hint 3. For E_{icj}, expand by linearity in column i. One of the two resulting matrices has two equal columns.
Sketch. Swapping rows i and j of I_n gives the same matrix as swapping its columns i and j, so $\det E_{ij} = -\det I_n = -1$ by antisymmetry. E_{ic} is I_n with column i scaled by c, so $\det E_{ic} = c$ by linearity in column i. Column i of E_{icj} is $\mathbf{e}_i + c\mathbf{e}_j$, so linearity in column i gives $\det I_n + c \det I'$, where I' has $\mathbf{e}_j$ in columns i and j; the value is $1 + c \cdot 0 = 1$.

Write $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for the columns of I_n.

Swaps. Swapping rows i and j of I_n moves the 1 in position (i,i) to position (j,i) and the 1 in position (j,j) to position (i,j), and changes nothing else. So column i of E_{ij} is $\mathbf{e}_j$, column j is $\mathbf{e}_i$, and every other column k is $\mathbf{e}_k$. That is, E_{ij} is I_n with columns i and j interchanged, and $\det E_{ij} = -\det I_n = -1$ by Proposition 5.1.10.

Scalings. E_{ic} is I_n with column i replaced by $c\mathbf{e}_i$. By linearity in column i, $\det E_{ic} = c \det I_n = c$.

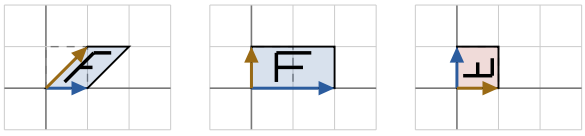
Replacements. E_{icj} is I_n with an extra entry c in position (j,i), so its column i is $\mathbf{e}_i + c\mathbf{e}_j$ and its other columns are those of I_n. Let I' be I_n with column i replaced by $\mathbf{e}_j$. By linearity in column i,

$$\det E_{icj} = \det I_n + c \det I' = 1 + c \cdot 0 = 1,$$

since columns i and j of I' are both $\mathbf{e}_j$, and det is alternating.

Finally, $-1 \ne 0$ because $1 \ne 0$, and $c \ne 0$ for a scaling, so all three values are nonzero. □

Over $\mathbb{R}$ these values are the signed areas observed in Section 3.2.6: a shear keeps the area of the unit square, the stretch E_{1c} multiplies its signed area by c (reflecting it when $c < 0$), and the swap keeps the area but reverses orientation.



$E_{2,1,1}$, $\det = 1$ $E_{1,2}$, $\det = 2$ E_{12}, $\det = -1$

The plane transformer, starting at the shear $E_{2,1,1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Sliding $A\mathbf{e}_2$ along the line $x_2 = 1$ gives the shears E_{2c1}, and the readout stays at $\det A = 1$. With $A\mathbf{e}_2$ back at $\mathbf{e}_2$, moving $A\mathbf{e}_1$ along the x_1-axis gives the scalings E_{1c} for $c \ne 0$, with $\det A = c$. Typing the entries $0, 1, 1, 0$ gives the swap E_{12}: the F is mirrored and $\det A = -1$. (Interactive in the web version.)

Lemma 5.4.2 (Multiplying by an elementary matrix). Let $A, E \in F^{n \times n}$, and suppose E is elementary. Then $\det(AE) = \det A \det E$.

Proof. ROUTINE

Uses. Proposition 3.2.28 (Column operations are right multiplications), Proposition 5.1.10 (Alternating implies antisymmetric), Lemma 5.1.8 (Column replacement), Definition 5.1.1 (Determinant function), Proposition 5.4.1 (Determinants of elementary matrices)
Hint 1. By Proposition 3.2.28, AE is A with one column operation performed. Treat the three types of E separately.
Hint 2. In each case, compare the factor by which the column operation changes $\det A$ with the value of $\det E$ in Proposition 5.4.1.
Sketch. AE_{ij} is A with columns i and j swapped, so its determinant is $-\det A$. AE_{ic} is A with column i scaled by c, so its determinant is $c \det A$. AE_{icj} is A with c times column j added to column i, so its determinant is $\det A$. The factors -1, c, and 1 are the determinants of the three types.

We consider the three types in turn, using Proposition 3.2.28 to describe AE and Proposition 5.4.1 for $\det E$.

Swaps. AE_{ij} is A with columns i and j interchanged, so by Proposition 5.1.10,

$$\det\big(AE_{ij}\big) = -\det A = \det A \det E_{ij}.$$

Scalings. AE_{ic} is A with column i multiplied by c, so by linearity in column i,

$$\det(AE_{ic}) = c \det A = \det A \det E_{ic}.$$

Replacements. AE_{icj} is A with c times column j added to column i, where $i \neq j$. By Lemma 5.1.8,

$$\det\big(AE_{icj}\big) = \det A = \det A \det E_{icj}.$$

□

Remark 5.4.3 (The values are no coincidence). In each case of the proof, $\det(AE) = c_E \det A$, where the factor c_E depends only on the column operation of E and not on A. Taking $A = I_n$ gives $\det E = c_E \det I_n = c_E$, because $E = I_n E$ is the result of performing its own column operation on I_n. So the lemma can be proved without computing $\det E$ first, and the computation of Proposition 5.4.1 then comes for free. (This is how LADW argues in its Lemma 3.6. We followed the lecture, which computes $\det E$ first.)

Corollary 5.4.4 (Products of elementary matrices). Let $A \in F^{n\times n}$, and let $E_1, \ldots, E_k \in F^{n\times n}$ be elementary. Then

$$\det(AE_1 \cdots E_k) = \det A \det E_1 \cdots \det E_k.$$

In particular, $\det(E_1 \cdots E_k) = \det E_1 \cdots \det E_k$, and this is nonzero.

Proof. ROUTINE

Uses. Lemma 5.4.2 (Multiplying by an elementary matrix), Proposition 5.4.1 (Determinants of elementary matrices), Proposition 1.1.17 (No zero divisors), induction on k
Hint 1. Induct on k. Group the product as $(AE_1 \cdots E_{k-1})E_k$.
Hint 2. For the last sentence, take $A = I_n$, and recall that a field has no zero divisors.
Sketch. By Lemma 5.4.2 applied to the matrix $AE_1 \cdots E_{k-1}$ and the induction hypothesis, $\det(AE_1 \cdots E_k) = \det(AE_1 \cdots E_{k-1}) \det E_k = \det A \det E_1 \cdots \det E_k$. With $A = I_n$ this is the second formula; each factor is nonzero by Proposition 5.4.1, and so is the product.

We induct on k. For $k = 0$ the product of elementary matrices is empty, that is, equal to I_n, and both sides equal $\det A$. For $k \geq 1$, apply Lemma 5.4.2 to the matrix $AE_1 \cdots E_{k-1}$ and to E_k, and then the induction hypothesis:

$$\begin{aligned}\det(AE_1 \cdots E_k) &= \det(AE_1 \cdots E_{k-1}) \det E_k \\ &= \det A \det E_1 \cdots \det E_{k-1} \det E_k.\end{aligned}$$

Taking $A = I_n$ gives $\det(E_1 \cdots E_k) = \det E_1 \cdots \det E_k$. Each $\det E_l$ is nonzero by Proposition 5.4.1, and a product of nonzero elements of a field is nonzero (Proposition 1.1.17, and induction on the number of factors). □

Exercise 5.4.5. ROUTINE Let $A \in \mathbb{Q}^{3\times 3}$ have $\det A = 5$, and let $G = E_{13}E_{2,4}E_{3,-7,1}$, a product of 3×3 elementary matrices. Find $\det G$ and $\det(AG)$. Which column operations turn A into AG, and in which order are they performed?

Solution on page 1415

5.4.2 Invertibility

We can now carry out the plan from the beginning of the section. The idea of building every invertible matrix out of elementary ones is used several times in this section, so we name it.

Technique 5.4.6 (Reduce to elementary matrices). Every invertible matrix is a product of elementary matrices (Theorem 3.5.19). So to prove a statement about all invertible matrices, it is often enough to

1. prove it for the three types of elementary matrices, by a direct computation; and
2. show that it survives multiplication: if it holds for A and E is elementary, then it holds for AE (or for EA).

Induction on the number of factors then gives the statement for every invertible matrix. A statement about all square matrices needs a separate argument for the matrices that are not invertible. Often both sides of the claimed identity are 0 there. Otherwise, write $A = G^{-1}\operatorname{rref}(A)$ (Theorem 3.3.10) and use that $\operatorname{rref}(A)$ has a zero row.

Reach for it when a statement involves products, inverses, or transposes, and is easy to check on elementary matrices, which differ from I_n in at most four entries. The fullest example is the first proof of Theorem 5.4.10; the first proof of Theorem 5.4.14 has the same shape. Remember the case split: a matrix that is not invertible is never a product of elementary matrices, since such products are invertible (Corollary 3.2.11). Forgetting this is a common error (Exercise 5.4.35).

Used in. Corollary 3.5.22, Exercise 3.6.10, Problem 3.7, Theorem 5.4.7, Theorem 5.4.10, Theorem 5.4.14, Exercise 5.4.35, Exercise 5.6.11, Exercise 5.4.49.

Theorem 5.4.7 (Invertibility and the determinant). Let $A \in F^{n\times n}$. The following are equivalent.

(I) A is invertible.
(II) $\det A \neq 0$.

Proof. STANDARD

Uses. Corollary 5.1.18 (Non-invertible matrices), Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Corollary 5.4.4 (Products of elementary matrices)
Techniques. Reduce to elementary matrices (Technique 5.4.6), Contrapositive (Technique 1.5.4)
Hint 1. One direction was proved in Section 5.1. For the other, which description of an invertible matrix makes its determinant computable?
Hint 2. Write an invertible A as a product of elementary matrices.
Hint 3. By Corollary 5.4.4, $\det A$ is then a product of determinants of elementary matrices. Can such a product be 0?
Sketch. (II) $\Rightarrow$ (I) is the contrapositive of Corollary 5.1.18. For (I) $\Rightarrow$ (II), write $A = E_1\cdots E_k$ with elementary E_i. Then $\det A = \det E_1 \cdots \det E_k$, a product of the nonzero scalars -1, c, and 1, which is nonzero.

(II) $\Rightarrow$ (I). We prove the contrapositive. If A is not invertible, then $\det A = 0$ by Corollary 5.1.18.

(I) $\Rightarrow$ (II). Suppose A is invertible. By Theorem 3.5.19 there are elementary matrices $E_1, \ldots, E_k$ with $A = E_1\cdots E_k$. By Corollary 5.4.4, $\det A = \det E_1 \cdots \det E_k$, and this is nonzero. □

The two directions use different axioms. (II) $\Rightarrow$ (I) used only multilinearity and alternation. (I) $\Rightarrow$ (II) used normalization, through the computation $\det E_{ic} = c \det I_n = c$ and its relatives in Proposition 5.4.1. This is as it must be: the zero function is multilinear and alternating, and (I) $\Rightarrow$ (II) fails for it. Combining the theorem with the Fundamental Theorems gives several equivalent ways to say that a determinant vanishes.

Corollary 5.4.8 (Vanishing determinants). Let $A \in F^{n\times n}$. The following are equivalent.
(a) $\det A = 0$.
(b) A is not invertible.
(c) The columns of A are linearly dependent.
(d) $A\mathbf{x} = \mathbf{0}$ for some nonzero $\mathbf{x} \in F^n$.

Proof. ROUTINE

Uses. Theorem 5.4.7 (Invertibility and the determinant), Corollary 3.5.25 (The invertible matrix theorem), Theorem 3.5.12 (First Fundamental Theorem, completed), Definition 2.3.8 (Matrix–vector product), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. (a) ⇔ (b) is Theorem 5.4.7. Relate both invertibility and independence of the columns to the rank.
Hint 2. For (c) ⇔ (d), write $A\mathbf{x}$ as a linear combination of the columns of A.
Sketch. (a) ⇔ (b) by Theorem 5.4.7. By Corollary 3.5.25 and Theorem 3.5.12, A is invertible if and only if $\operatorname{rank} A = n$, if and only if its columns are independent, which gives (b) ⇔ (c). Since $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$, a nonzero solution of $A\mathbf{x} = \mathbf{0}$ is the same thing as a nontrivial linear relation among the columns.

(a) ⇔ (b) is Theorem 5.4.7. By Corollary 3.5.25, A is invertible if and only if $\operatorname{rank} A = n$, and by Theorem 3.5.12 this holds if and only if the columns of A are linearly independent. Negating both sides gives (b) ⇔ (c). Finally, $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ by Definition 2.3.8, so a nonzero $\mathbf{x}$ with $A\mathbf{x} = \mathbf{0}$ is exactly a nontrivial linear relation among the columns of A. By Theorem 1.5.5, such a relation exists if and only if the columns are linearly dependent, which is (c) ⇔ (d). □

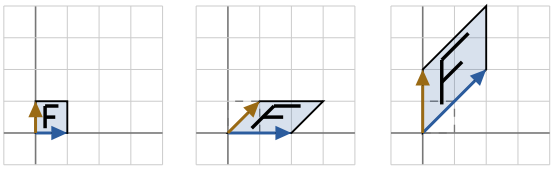
For det, the implication (c) ⇒ (a) is Proposition 5.1.17, which holds for every multilinear alternating function. The implication (a) ⇒ (c) is the converse that Exercise 5.1.22 warned about, and it needed normalization. Together with Proposition 5.3.19, the corollary also says that $A + tI_n$ is invertible for all but at most n scalars t, and that the exceptional t are exactly the roots of $\det(A + tI_n)$. Their negatives reappear as eigenvalues in Section 7.3.

Exercise 5.4.9. ROUTINE Let $A(t) = \begin{pmatrix} 1 & t & 0 \\ t & 1 & t \\ 0 & t & 1 \end{pmatrix}$. Find all t for which $A(t)$ is invertible, when t ranges over $\mathbb{R}$, over $\mathbb{Q}$, over $\mathbb{F}_3$, and over $\mathbb{F}_7$.

Solution on page 1415

5.4.3 The product formula

Intuition (Areas multiply). Over $\mathbb{R}$, the absolute value of $\det A$ is the factor by which T_A scales areas in the plane (or volumes in space), and its sign records whether T_A preserves orientation (Section 5.1). Composition multiplies these factors. If T_B doubles areas and T_A triples them, then $T_A \circ T_B$, which is T_{AB} (Theorem 2.4.2), multiplies them by six. If both reverse orientation, the composite preserves it, matching $(-1)(-1) = 1$. So signed area predicts $\det(AB) = \det A \det B$. Figure 5.4.1 shows an example.



Q: area 1 $T_B(Q)$: area 2 $T_{AB}(Q)$: area 4

Figure 5.4.1. The unit square Q with a letter F (left), its image under T_B for $B = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ (middle), and its image under $T_A \circ T_B = T_{AB}$ for $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ (right), where $AB = \begin{pmatrix} 2 & 0 \\ 2 & 2 \end{pmatrix}$. The arrows are the images of $\mathbf{e}_1$ (blue) and $\mathbf{e}_2$ (ochre). Here $\det B = 2$ and $\det A = 2$: T_B doubles the area of Q, and T_A, which rotates by 45° and stretches by $\sqrt{2}$, doubles it again, so $\det(AB) = 4$.

Theorem 5.4.10 (Multiplicativity). For all $A, B \in F^{n\times n}$,

$$\det(AB) = \det A \det B.$$

Proof (by factoring into elementary matrices). STANDARD

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Corollary 5.4.4 (Products of elementary matrices), Theorem 5.4.7 (Invertibility and the determinant), Theorem 3.5.23 (One-sided inverses of square matrices)
Techniques. Reduce to elementary matrices (Technique 5.4.6)
Hint 1. Suppose first that B is invertible, and factor it into elementary matrices. Note that Corollary 5.4.4 allows an arbitrary matrix on the left.
Hint 2. If B is not invertible, show that AB is not invertible either. What are both sides then?
Hint 3. If AB had an inverse C, then $(CA)B = I_n$. What does a one-sided inverse of a square matrix give?
Sketch. If $B = E_1 \cdots E_k$, then Corollary 5.4.4 gives $\det(AB) = \det A \det E_1 \cdots \det E_k = \det A \det B$, the last step by the same corollary with I_n in place of A. If B is not invertible, neither is AB, since an inverse C of AB would make CA a left inverse of B. So both sides are 0 by Theorem 5.4.7.

Case 1: B is invertible. By Theorem 3.5.19, $B = E_1 \cdots E_k$ for some elementary matrices $E_1, \ldots, E_k$. Applying Corollary 5.4.4 first to A and then to I_n,

$$\begin{aligned} \det(AB) &= \det(AE_1 \cdots E_k) = \det A \det E_1 \cdots \det E_k \\ &= \det A \det(E_1 \cdots E_k) = \det A \det B. \end{aligned}$$

Case 2: B is not invertible. Then $\det B = 0$ by Theorem 5.4.7, so the right side is 0. We show that AB is not invertible, so that the left side is 0 as well. Suppose instead that AB has an inverse C. Then $(CA)B = C(AB) = I_n$, so B has a left inverse. A square matrix with a one-sided inverse is invertible (Theorem 3.5.23), which contradicts our assumption on B. Hence AB is not invertible, and $\det(AB) = 0 = \det A \det B$. □

The lecture splits into cases according to whether A and B are both invertible, and factors both. Since Corollary 5.4.4 allows any matrix on the left, only B needs to be factored. The second proof avoids factoring altogether. It rests on a move that was already used for block-diagonal matrices in Proposition 5.3.17.

Technique 5.4.11 (Recognize the determinant by its axioms). By Theorem 5.3.4, every multilinear alternating function $D : F^{n\times n} \to F$ equals $D(I_n)\det$. So to prove an identity of the form $D(B) = c \det B$, where $D(B)$ is some expression built from B, check two things about D as a function of the columns of B: that it is linear in each column, and that it vanishes when two columns are equal. Then evaluate at I_n to find $c = D(I_n)$. No formula for det and no case analysis is needed.

Reach for it when an expression is visibly linear in each column of a matrix: $\det(AB)$ for a fixed A, $\det\begin{pmatrix} X & 0 \\ 0 & B \end{pmatrix}$ for a fixed B, or a function of n vectors in a space with a basis. The canonical example is the second proof of Theorem 5.4.10. The same move proved Proposition 5.3.17, and Section 13.4 uses it to define the determinant of an operator without choosing a basis.
Used in. Theorem 5.4.10, Exercise 5.6.8, Problem 5.2, Proposition 7.3.2.

Proof of Theorem 5.4.10 (by uniqueness). STANDARD

> *Uses.* Theorem 5.3.4 (Existence and uniqueness of the determinant), Definition 2.4.1 (Matrix product), Proposition 2.3.13 (Linearity of the product), Proposition 2.1.17 (Compositions of linear maps are linear), Definition 5.1.1 (Determinant function)
> *Techniques.* Recognize the determinant by its axioms (Technique 5.4.11)
> *Hint 1.* Fix A, and regard $D(B) = \det(AB)$ as a function of the columns $\mathbf{b}_1, \dots, \mathbf{b}_n$ of B. What are the columns of AB?
> *Hint 2.* Column j of AB is $A\mathbf{b}_j$. Check that D is multilinear and alternating, and apply Theorem 5.3.4.
> *Sketch.* Since column j of AB is $A\mathbf{b}_j$, $D(B) = \det(A\mathbf{b}_1, \dots, A\mathbf{b}_n)$. In each column this is det composed with the linear map $\mathbf{x} \mapsto A\mathbf{x}$, so it is linear, and equal columns of B give equal columns of AB. So D is multilinear and alternating, and $D(B) = D(I_n) \det B = \det A \det B$.

Fix A, and define $D : F^{n\times n} \to F$ by $D(B) = \det(AB)$. By Definition 2.4.1, column j of AB is $A\mathbf{b}_j$, so

$$D(\mathbf{b}_1, \dots, \mathbf{b}_n) = \det(A\mathbf{b}_1, \dots, A\mathbf{b}_n).$$

Multilinear. Fix j and all columns of B except the j-th. Then $\mathbf{x} \mapsto D(\dots, \mathbf{x}, \dots)$, with $\mathbf{x}$ in column j, is the composite of $\mathbf{x} \mapsto A\mathbf{x}$, which is linear (Proposition 2.3.13), and the map that puts a vector $\mathbf{y}$ in column j of the matrix with columns $A\mathbf{b}_k$ ($k \neq j$) and takes det, which is linear because det is linear in column j. So it is linear by Proposition 2.1.17.

Alternating. If $\mathbf{b}_i = \mathbf{b}_j$ for some $i \neq j$, then $A\mathbf{b}_i = A\mathbf{b}_j$, so AB has two equal columns and $D(B) = 0$.

By Theorem 5.3.4, $D(B) = D(I_n) \det B = \det(AI_n) \det B = \det A \det B$ for every B. □

For $n = 2$ the second proof is the first solution of Exercise 5.1.27.

> *Intuition (Two proofs compared).* The two proofs are the two basic strategies of this chapter. The first, Technique 5.4.6, builds B out of simple pieces. It needs the computation for elementary matrices, the factorization theorem that comes out of Gauss–Jordan elimination (Theorem 3.5.19), and a separate argument for matrices that are not invertible. The second, Technique 5.4.11, never looks inside B. It notices that $B \mapsto \det(AB)$ satisfies two of the three axioms, so by uniqueness it is a multiple of det, and it finds the multiple by setting $B = I_n$. It uses neither elimination nor a case split.
>
> The second proof is shorter, and it works verbatim with any multilinear alternating function in place of det, which is why the basis-free theory of Section 13.4 is built on it. It also gives (I) ⇒ (II) in Theorem 5.4.7 without elementary matrices: if A is invertible, then $1 = \det I_n = \det\left(AA^{-1}\right) = \det A \det A^{-1}$, so $\det A \neq 0$. The first proof explains where the numbers come from: the factor that each elementary matrix contributes is the effect of its column operation. Its pattern also adapts to statements that are not about columns, such as $\det A^{\mathsf{T}} = \det A$ below, where the axioms give no direct handle.

Exercise 5.4.12. ROUTINE Let $A, B \in F^{n\times n}$. Show that $\det(AB) = \det(BA)$, and that $\det\left(A^k\right) = (\det A)^k$ for every integer $k \geq 0$. Show also that if A is invertible, then $\det\left(A^{-1}\right) = (\det A)^{-1}$, and deduce that $\det\left(A^k\right) = (\det A)^k$ for negative k as well, where $A^{-k} = \left(A^{-1}\right)^k$.

Solution on page 1415

So $\det(AB) = \det(BA)$ even when $AB \neq BA$. For example, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ have $AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and $BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, which are different matrices, both with determinant 1. In the language of Section 14.1, Theorem 5.4.7 and Theorem 5.4.10 say that det is a homomorphism from the group of invertible $n \times n$ matrices under multiplication to the group $F \setminus \{0\}$ under multiplication.

5.4.4 The transpose

The axioms single out the columns of a matrix. The transpose exchanges rows and columns (Definition 2.4.21), and the next theorem says that the determinant does not notice the exchange. First, the elementary matrices.

Lemma 5.4.13 (Transposes of elementary matrices). If $E \in F^{n\times n}$ is elementary, then E^{T} is an elementary matrix of the same type, and $\det E^{\mathsf{T}} = \det E$. Explicitly,

$$E_{ij}^{\mathsf{T}} = E_{ij}, \quad E_{ic}^{\mathsf{T}} = E_{ic}, \quad E_{icj}^{\mathsf{T}} = E_{jci}.$$

Proof. ROUTINE

Uses. Definition 3.2.2 (Elementary matrices), Definition 2.4.21 (Transpose), Proposition 5.4.1 (Determinants of elementary matrices)
Hint 1. Transposing reflects the entries in the diagonal. Which off-diagonal entries does each elementary matrix have?
Hint 2. E_{icj} has a single off-diagonal entry, c, in position (j,i). Which elementary matrix has c in position (i,j)?
Sketch. E_{ic} is diagonal. The off-diagonal entries of E_{ij} are 1s in positions (i,j) and (j,i), a symmetric pattern. E_{icj} is I_n plus c in position (j,i), so its transpose is I_n plus c in position (i,j), which is E_{jci}. Each pair has the same determinant by Proposition 5.4.1.

Transposing a matrix moves the entry in position (k,l) to position (l,k) (Definition 2.4.21), so it fixes the diagonal and the identity matrix. E_{ic} is diagonal, so $E_{ic}^{\mathsf{T}} = E_{ic}$. By the proof of Proposition 5.4.1, E_{ij} agrees with I_n except that the entries in positions (i,i) and (j,j) are 0 and those in positions (i,j) and (j,i) are 1. This pattern is unchanged by the reflection, so $E_{ij}^{\mathsf{T}} = E_{ij}$. Finally, E_{icj} is I_n with an extra entry c in position (j,i), so E_{icj}^{T} is I_n with an extra entry c in position (i,j). Since $j \neq i$, this is the elementary matrix E_{jci}, which adds c times row j to row i. In each case E and E^{T} have the same type, so they have the same determinant (-1, c, or 1) by Proposition 5.4.1. □

Theorem 5.4.14 (The transpose). For every $A \in F^{n\times n}$,

$$\det A^{\mathsf{T}} = \det A.$$

Proof (via elementary matrices). STANDARD

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Proposition 2.4.23 (Transpose of a product), Lemma 5.4.13 (Transposes of elementary matrices), Corollary 5.4.4 (Products of elementary matrices), Theorem 5.4.7 (Invertibility and the determinant), Definition 2.5.12 (Invertible)
Techniques. Reduce to elementary matrices (Technique 5.4.6), Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. First dispose of the matrices that are not invertible: show that A^{T} is invertible exactly when A is.
Hint 2. For invertible A, factor $A = E_1 \cdots E_k$. What is A^{T}, by Proposition 2.4.23?
Hint 3. $A^{\mathsf{T}} = E_k^{\mathsf{T}} \cdots E_1^{\mathsf{T}}$ is again a product of elementary matrices. Compare the two products of determinants using Lemma 5.4.13.
Sketch. If B is an inverse of A, then B^{T} is an inverse of A^{T} by Proposition 2.4.23, and conversely since $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$. So if A is not invertible, both determinants are 0. If $A = E_1 \cdots E_k$, then $A^{\mathsf{T}} = E_k^{\mathsf{T}} \cdots E_1^{\mathsf{T}}$, and Corollary 5.4.4 with Lemma 5.4.13 gives $\det A^{\mathsf{T}} = \prod \det E_i^{\mathsf{T}} = \prod \det E_i = \det A$.

First, A is invertible if and only if A^{T} is. Indeed, if $AB = BA = I_n$, then by Proposition 2.4.23,

$$B^{\mathsf{T}} A^{\mathsf{T}} = (AB)^{\mathsf{T}} = I_n^{\mathsf{T}} = I_n,$$
$$A^{\mathsf{T}} B^{\mathsf{T}} = (BA)^{\mathsf{T}} = I_n,$$

so A^{T} is invertible. Applying this to A^{T}, whose transpose is A, gives the converse.

If A is not invertible, then neither is A^{T}, and $\det A = 0 = \det A^{\mathsf{T}}$ by Theorem 5.4.7. Suppose that A is invertible, and write $A = E_1 E_2 \cdots E_k$ with elementary E_i (Theorem 3.5.19). By Proposition 2.4.23 and induction on k,

$$A^{\mathsf{T}} = E_k^{\mathsf{T}} \cdots E_2^{\mathsf{T}} E_1^{\mathsf{T}},$$

and each E_i^{T} is elementary by Lemma 5.4.13. So by Corollary 5.4.4 (with I_n in place of A), Lemma 5.4.13, and commutativity of multiplication in F,

$$\det A^{\mathsf{T}} = \det E_k^{\mathsf{T}} \cdots \det E_1^{\mathsf{T}} = \det E_k \cdots \det E_1$$

$$= \det E_1 \cdots \det E_k = \det A. \qquad \square$$

Proof (via the Leibniz formula). STANDARD

> *Uses.* Definition 5.3.6 (Determinant), Definition 2.4.21 (Transpose), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.8 (Reindexing sums over S_n)
> *Hint 1.* Write out the Leibniz formula for A^{T}, whose (i,j) entry is a_{ji}.
> *Hint 2.* You get $\sum_\sigma \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$. Reorder each product by the second index: if $j = \sigma(i)$, then $i = \sigma^{-1}(j)$.
> *Hint 3.* Then use $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$, and reindex the sum by $\sigma \mapsto \sigma^{-1}$.
> *Sketch.* The term of σ in $\det A^{\mathsf{T}}$ is $\operatorname{sgn}(\sigma) \prod_i a_{i\sigma(i)}$. Reordering the factors by $j = \sigma(i)$ turns it into $\operatorname{sgn}(\sigma^{-1}) \prod_j a_{\sigma^{-1}(j)j}$, the term of σ^{-1} in $\det A$. As σ runs over S_n, so does σ^{-1}, so the two sums agree.

Let $B = A^{\mathsf{T}}$, so that $b_{ij} = a_{ji}$ (Definition 2.4.21). By the Leibniz formula (Definition 5.3.6),

$$\begin{aligned} \det A^{\mathsf{T}} &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \end{aligned}$$

Fix σ. As i runs over $[n]$, the index $j = \sigma(i)$ runs over $[n]$, taking each value once, and $i = \sigma^{-1}(j)$. Reordering the factors by j, which is allowed since multiplication in F is commutative,

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}.$$

Also $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ by Proposition 5.2.25. So the term of σ in $\det A^{\mathsf{T}}$ is $g(\sigma^{-1})$, where

$$g(\rho) = \operatorname{sgn}(\rho) a_{\rho(1)1} \cdots a_{\rho(n)n}$$

is the term of ρ in $\det A$. By Proposition 5.2.8,

$$\det A^{\mathsf{T}} = \sum_{\sigma \in S_n} g(\sigma^{-1}) = \sum_{\sigma \in S_n} g(\sigma) = \det A. \qquad \square$$

The second proof is the computation of Exercise 5.3.8.

> *Intuition (What the two proofs see).* The lecture calls the fact $\det E^{\mathsf{T}} = \det E$ dumb luck, and in a sense it is: the axioms never mention rows. But there is a reason behind it. E^{T} performs the same kind of operation as E, on rows where E acts on columns, and the three kinds of operations scale signed volume by -1, c, and 1 whichever way they act. The second proof gives a more symmetric reason. A term of the Leibniz formula picks one entry from each row and each column, like non-attacking rooks on a chessboard (Figure 5.3.1). Transposing reflects the board in its diagonal, which turns the term of σ into the term of σ^{-1}, a permutation with the same sign.
>
> As for strategy, the first proof is Technique 5.4.6 once more, with the same case split as the first proof of Theorem 5.4.10. Technique 5.4.11 does not apply directly. To use it on $D(A) = \det A^{\mathsf{T}}$, we would need to know that D is multilinear and alternating in the columns of A, that is, that det is multilinear and alternating in rows, which is what we are trying to establish. Exercise 5.3.27 proves exactly this from the Leibniz formula, and gives a third proof.

Every statement about the columns of a determinant now has a version for rows (Technique 3.2.27).

> **Corollary 5.4.15** (Rows). Let $A \in F^{n \times n}$.
> (a) $\det A$ is multilinear and alternating as a function of the rows of A.
> (b) Swapping two rows of A negates $\det A$, multiplying a row by c multiplies $\det A$ by c, and adding a multiple of one row to a different row does not change $\det A$.
> (c) $\det A = 0$ if and only if the rows of A are linearly dependent.

Proof. ROUTINE

Uses. Theorem 5.4.14 (The transpose), Definition 2.4.21 (Transpose), Definition 5.1.1 (Determinant function), Proposition 5.1.10 (Alternating implies antisymmetric), Lemma 5.1.8 (Column replacement), Corollary 5.4.8 (Vanishing determinants)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. Row i of A, written as a column, is column i of A^{T}. Apply the column statements to A^{T}.
Hint 2. For (b) you can also use Theorem 5.4.10 with the elementary matrices of Proposition 5.4.1.
Sketch. The rows of A, written as columns, are the columns of A^{T}, and $\det A = \det A^{\mathsf{T}}$. So each statement about the rows of A is the corresponding statement about the columns of A^{T}: the axioms for (a), Proposition 5.1.10, multilinearity, and Lemma 5.1.8 for (b), and Corollary 5.4.8 for (c).

For $\mathbf{r}_1, \dots, \mathbf{r}_n \in F^n$, let R be the matrix whose i-th row is $\mathbf{r}_i^{\mathsf{T}}$. Its transpose has columns $\mathbf{r}_1, \dots, \mathbf{r}_n$, so by Theorem 5.4.14,

$$\det R = \det(\mathbf{r}_1, \dots, \mathbf{r}_n).$$

(a) The right side is multilinear and alternating in $\mathbf{r}_1, \dots, \mathbf{r}_n$ (Definition 5.1.1), hence so is the left side.

(b) Each row operation on A is the corresponding column operation on A^{T}, and $\det A = \det A^{\mathsf{T}}$. By Proposition 5.1.10, linearity in a column, and Lemma 5.1.8, the three column operations multiply $\det A^{\mathsf{T}}$ by -1, by c, and by 1. (Alternatively: a row operation multiplies A on the left by an elementary matrix E (Proposition 3.2.7), and $\det(EA) = \det E \det A$ with $\det E$ given by Proposition 5.4.1.)

(c) The rows of A are linearly dependent if and only if the columns of A^{T} are, which by Corollary 5.4.8 happens if and only if $\det A^{\mathsf{T}} = \det A$ is 0. □

In particular, a matrix with a zero row or two equal rows has determinant 0, and $\det(EA) = \det E \det A$ for elementary E, which is Lemma 5.4.2 with the factors in the other order.

Exercise 5.4.16. ROUTINE Suppose that $\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix} = 7$, where the entries lie in a field F. Find

$$\text{(a) } \det\begin{pmatrix} d & e & f \\ a & b & c \\ g & h & k \end{pmatrix},$$

$$\text{(b) } \det\begin{pmatrix} a & b & c \\ 2d-3a & 2e-3b & 2f-3c \\ g & h & k \end{pmatrix},$$

$$\text{(c) } \det\begin{pmatrix} a & d & g+5a \\ b & e & h+5b \\ c & f & k+5c \end{pmatrix},$$

$$\text{(d) } \det\begin{pmatrix} 3a & 3b & 3c \\ 3d & 3e & 3f \\ 3g & 3h & 3k \end{pmatrix}.$$

Solution on page 1416

5.4.5 Similar matrices and the determinant of an operator

Recall (Definition 4.5.23) that $A, B \in F^{n\times n}$ are **similar** if $B = S^{-1}AS$ for some invertible $S \in F^{n\times n}$. Then also $A = SBS^{-1}$, so the relation is symmetric.

Proposition 5.4.17 (Similar matrices have equal determinants). If $A, B \in F^{n\times n}$ are similar, then $\det A = \det B$.

Proof. ROUTINE

> *Uses.* Definition 4.5.23 (Similar matrices), Theorem 5.4.10 (Multiplicativity)
> *Hint 1.* Write $B = S^{-1}AS$ and apply Theorem 5.4.10.
> *Hint 2.* Determinants are scalars, so they commute even though matrices need not. Bring $\det S^{-1}$ and $\det S$ together.
> *Sketch.* $\det B = \det S^{-1} \det A \det S = \det A \det\bigl(S^{-1}S\bigr) = \det A \det I_n = \det A$.

Write $B = S^{-1}AS$ with S invertible. By Theorem 5.4.10, applied twice, and commutativity of multiplication in F,

$$\begin{aligned}\det B &= \det S^{-1} \det A \det S = \det A\bigl(\det S^{-1} \det S\bigr)\\ &= \det A \det\bigl(S^{-1}S\bigr) = \det A \det I_n = \det A,\end{aligned}$$

as desired. □

The converse fails. $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ and I_2 both have determinant 1, but they are not similar, since $S^{-1}I_2S = I_2$ for every invertible S. So the determinant can prove that two matrices are not similar, but not that they are. It joins the rank and the trace (Proposition 2.4.28) as a similarity invariant. Together with the trace, it appears among the coefficients of $\det(A + tI_n)$ (Proposition 5.3.19), and in Section 7.3 this polynomial becomes a finer invariant.

The main use of Proposition 5.4.17 is to define the determinant of a linear operator. Let V be a vector space with $\dim V = n \geq 1$, let $T \in \mathrm{L}(V)$, and let $\mathcal{B}$ and $\mathring{\mathcal{B}}$ be bases of V. By Theorem 4.5.17,

$$[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = [\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}}[T]_{\mathcal{B}\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}.$$

By Theorem 2.4.2, $[\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathcal{B}}[\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}} = [\mathrm{id}_V]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = I_n$, and likewise in the other order. So with $S = [\mathrm{id}_V]_{\mathcal{B}\mathring{\mathcal{B}}}$, we have $[T]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = S^{-1}[T]_{\mathcal{B}\mathcal{B}}S$. The two matrices are similar and have the same determinant.

Definition 5.4.18 (Determinant of an operator). Let V be a finite-dimensional vector space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. The **determinant** of T is

$$\det T = \det [T]_{\mathcal{B}\mathcal{B}},$$

where $\mathcal{B}$ is any basis of V. By the computation above and Proposition 5.4.17, this does not depend on the choice of $\mathcal{B}$.

For $V = F^n$ and $T = T_A$, the standard basis gives $[T_A]_{\mathcal{E}\mathcal{E}} = A$ (Proposition 2.3.16), so $\det T_A = \det A$. Translating between maps and matrices (Technique 2.3.27), the results of this section carry over to operators.

Example 5.4.19 (Two operators on matrices and polynomials).

(a) Let $T \in \mathrm{L}\big(F^{2\times 2}\big)$ be the transpose, $T(X) = X^{\mathsf{T}}$. Use the basis

$$U_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad U_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$U_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad U_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then T fixes U_{11} and U_{22} and exchanges U_{12} and U_{21}, so its matrix in this basis is the 4×4 swap E_{23}, and $\det T = -1$. Over $\mathbb{R}$ this is the expected sign: T fixes the three-dimensional subspace of symmetric matrices and negates the skew-symmetric matrix $U_{12} - U_{21}$, so it is a reflection of the four-dimensional space $\mathbb{R}^{2\times 2}$. Over $\mathbb{F}_2$, $\det T = -1 = 1$.

(b) Let $D \in \mathrm{L}\big(F[x]_{\leq 2}\big)$ be differentiation. Its matrix in the basis $1, x, x^2$ is upper triangular with zero diagonal (Example 2.3.4), so $\det D = 0$. This agrees with Proposition 5.4.20 below, since D is not invertible: $D(1) = \mathbf{0}$. For $\mathrm{id} + D$, which sends $1, x, x^2$ to 1, $1 + x$, and $2x + x^2$,

$$[\mathrm{id} + D]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \det(\mathrm{id} + D) = 1$$

by Proposition 5.3.15, where $\mathcal{B} = 1, x, x^2$.

Proposition 5.4.20 (Properties of the determinant of an operator). Let V be a vector space with $n = \dim V \geq 1$, let $S, T \in \mathrm{L}(V)$, and let $c \in F$.
(a) $\det \mathrm{id}_V = 1$, and $\det(cT) = c^n \det T$.
(b) $\det(ST) = \det S \det T$.
(c) T is invertible if and only if $\det T \neq 0$. In that case $\det\big(T^{-1}\big) = (\det T)^{-1}$.

Proof. STANDARD

Uses. Definition 5.4.18 (Determinant of an operator), Example 2.3.6 (The zero map and the identity map), Proposition 1.6.19 (Coordinates respect linear combinations), Exercise 5.1.5, Theorem 2.4.2 (The matrix of a composite), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.7 (Invertibility and the determinant), Theorem 2.2.1 (Linear maps are determined by a basis), Definition 2.3.1 (Matrix of a linear map)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Fix one basis $\mathcal{B}$ and translate each statement into a statement about the matrices of the operators with respect to $\mathcal{B}$.
Hint 2. For (c), if T is invertible, apply (b) to $TT^{-1} = \mathrm{id}_V$. Conversely, if $\det T \neq 0$, the matrix of T has an inverse C. Find a linear map whose matrix is C.
Sketch. With one basis $\mathcal{B}$ throughout, $[\mathrm{id}_V] = I_n$, $[cT] = c[T]$, and $[ST] = [S][T]$, so (a) and (b) follow from the matrix statements. If T is invertible, (b) gives $\det T \det T^{-1} = \det \mathrm{id}_V = 1$. If $\det T \neq 0$, then $[T]$ has an inverse C, and the linear map R with $[R] = C$ has $[RT] = [TR] = I_n = [\mathrm{id}_V]$, so $RT = TR = \mathrm{id}_V$.

Fix a basis $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ of V, and write $[R]$ for $[R]_{\mathcal{B}\mathcal{B}}$ for every $R \in \mathrm{L}(V)$. A linear map is determined by its matrix: if $[R] = [R']$, then $R\big(\mathbf{v}_j\big)$ and $R'\big(\mathbf{v}_j\big)$ have the same coordinate vector for every j, so they are equal, and $R = R'$ by the uniqueness in Theorem 2.2.1.

(a) $[\mathrm{id}_V] = I_n$ (Example 2.3.6), so $\det \mathrm{id}_V = 1$. Column j of $[cT]$ is $\big[cT\big(\mathbf{v}_j\big)\big]_{\mathcal{B}} = c\big[T\big(\mathbf{v}_j\big)\big]_{\mathcal{B}}$ by Proposition 1.6.19, so $[cT] = c[T]$, and $\det(cT) = c^n \det T$ by Exercise 5.1.5.

(b) By Theorem 2.4.2, $[ST] = [S][T]$, so by Theorem 5.4.10,

$$\begin{aligned} \det(ST) &= \det([S][T]) \\ &= \det[S] \det[T] = \det S \det T. \end{aligned}$$

(c) If T is invertible, then by (b) and (a), $\det T \det\big(T^{-1}\big) = \det\big(TT^{-1}\big) = \det \mathrm{id}_V = 1$. So $\det T \neq 0$, and $\det\big(T^{-1}\big) = (\det T)^{-1}$. Conversely, suppose that $\det T \neq 0$. Then $[T]$ is invertible by Theorem 5.4.7; let C be

its inverse, with entries c_{ij}. By Theorem 2.2.1 there is a linear map $R \in \mathrm{L}(V)$ with $R(\mathbf{v}_j) = \sum_i c_{ij}\mathbf{v}_i$ for every j, and then $[R] = C$ by Definition 2.3.1. By Theorem 2.4.2, $[RT] = C[T] = I_n = [\mathrm{id}_V]$ and $[TR] = [T]C = I_n$. Since a linear map is determined by its matrix, $RT = TR = \mathrm{id}_V$, and T is invertible. □

In Section 13.4 the determinant of an operator is defined directly, without a basis, and Definition 5.4.18 becomes a theorem.

5.4.6 Computing determinants

The Leibniz formula has $n!$ terms. Corollary 5.4.15 gives a much faster method: row reduce to a triangular matrix and keep track of how each operation changes the determinant.

Proposition 5.4.21 (Determinants by row reduction). Let $A \in F^{n\times n}$, and suppose that $E_t \cdots E_2 E_1 A = U$ for elementary matrices $E_1, \ldots, E_t$. Suppose that s of the E_l are swaps, r of them are scalings by $c_1, \ldots, c_r$, and the rest are replacements. Then

$$c_1 \cdots c_r \det A = (-1)^s \det U.$$

In particular, if U is triangular, then $\det A = (-1)^s (c_1 \cdots c_r)^{-1} u_{11} \cdots u_{nn}$.

Proof. ROUTINE

Uses. Theorem 5.4.10 (Multiplicativity), Proposition 5.4.1 (Determinants of elementary matrices), Proposition 5.3.15 (Triangular matrices)
Hint 1. Take determinants of both sides of $E_t \cdots E_1 A = U$.
Hint 2. Proposition 5.4.1 gives each $\det E_l$.
Sketch. By Theorem 5.4.10, $\det U = \det E_t \cdots \det E_1 \det A$, and the product of the $\det E_l$ is $(-1)^s c_1 \cdots c_r$ by Proposition 5.4.1. If U is triangular, $\det U$ is the product of its diagonal entries.

By Theorem 5.4.10, applied t times, $\det U = \det E_t \cdots \det E_1 \det A$. By Proposition 5.4.1, $\det E_l$ is -1 for each swap, c_q for the scaling by c_q, and 1 for each replacement, so $\det U = (-1)^s c_1 \cdots c_r \det A$. Multiplying by $(-1)^s$, which is its own inverse, gives the claim. If U is triangular, then $\det U = u_{11} \cdots u_{nn}$ by Proposition 5.3.15, and we may divide by $c_1 \cdots c_r$, which is nonzero. □

Scalings are never needed to reach a triangular matrix, so the simplest version of the method uses only swaps and replacements.

Algorithm 5.4.22 (Determinant by elimination)

Input A square matrix $A \in F^{n\times n}$.
Output $\det A$.

1. Set $s = 0$. For $k = 1, 2, \ldots, n$ in turn: if column k of the current matrix has a nonzero entry in one of the rows $k, \ldots, n$, swap such a row with row k (adding 1 to s if the rows are different), and then add multiples of row k to rows $k+1, \ldots, n$ to make their entries in column k equal to 0. If there is no such entry, go on to the next k.
2. Let U be the final matrix. Output $(-1)^s u_{11} u_{22} \cdots u_{nn}$.

Proof. ROUTINE

Uses. Proposition 5.4.21 (Determinants by row reduction), Proposition 3.2.7 (Row operations are left multiplications), Definition 3.3.31 (Triangular matrices)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. By Proposition 5.4.21 it is enough to show that U is upper triangular. Which entries does stage k make 0?
Hint 2. Check that later stages never touch column k below the diagonal: which rows do they swap, and which rows do they add?
Sketch. The algorithm uses only swaps and replacements, so by Proposition 5.4.21 it outputs $\det A$ provided U is upper triangular. After stage k the entries of column k below the diagonal are 0. A later stage $k' > k$ only swaps rows numbered

> at least k' and adds multiples of row k' to rows below it, and all these rows have entry 0 in column k, so column k stays cleared.

Every operation is a swap or a replacement, so the final matrix is $U = E_t \cdots E_1 A$ with s swaps among the E_l and no scalings (Proposition 3.2.7). By Proposition 5.4.21, the output equals $\det A$ provided U is upper triangular.

We show that after stage k, the entries of columns $1, \ldots, k$ below the diagonal are 0. This holds after stage 1, since stage 1 either clears column 1 below row 1 or finds it already zero there. Suppose it holds after stage $k-1$. Stage k only swaps two rows with indices at least k, and adds multiples of row k to rows with indices greater than k. For $j < k$, all of these rows have entry 0 in column j, by the induction hypothesis, so the entries of column j below the diagonal stay 0. And stage k makes the entries of column k below the diagonal 0, or finds them 0 already. After stage n, U is upper triangular (Definition 3.3.31). □

If some stage finds no nonzero entry, say at column k, then $u_{kk} = 0$, so $\det A = 0$ and A is not invertible. In practice one also uses scalings to pull a common factor out of a row, and one may mix in column operations freely: each row or column operation multiplies the determinant by the determinant of its elementary matrix (Corollary 5.4.15, Lemma 5.4.2).

Example 5.4.23 (A 4×4 determinant by elimination). Let

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & 4 \\ -1 & -3 & 3 & -2 \\ 3 & 8 & -1 & 9 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

Stage 1 adds -2, 1, and -3 times row 1 to rows 2, 3, and 4. In column 2, the new row 2 has a 0, so stage 2 swaps rows 2 and 3 and then adds 2 times the new row 2 to row 4:

$$A \xrightarrow{\text{stage } 1} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -2 \\ 0 & -1 & 2 & 1 \\ 0 & 2 & 2 & 0 \end{pmatrix},$$

$$\xrightarrow{\text{stage } 2} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 6 & 2 \end{pmatrix}.$$

Stage 3 adds -2 times row 3 to row 4, and stage 4 does nothing:

$$\xrightarrow{\text{stage } 3} \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 6 \end{pmatrix} = U.$$

There was one swap, so

$$\det A = (-1)^1 \cdot 1 \cdot (-1) \cdot 3 \cdot 6 = 18.$$

The Leibniz formula would have needed 24 terms. Since $18 = 2 \cdot 3^2$, Proposition 5.3.20 shows that the reduction of A modulo a prime p is invertible over $\mathbb{F}_p$ for every p except 2 and 3.

$$\det A = (-1)^1 \cdot 1 \cdot (-1) \cdot 3 \cdot 6 = 18$$

Gauss–Jordan elimination on the matrix A of Example 5.4.23, one elementary matrix at a time. It also scales each pivot to 1, by the factor $1/p$ for a pivot p, so by Proposition 5.4.21, $\det A$ is $(-1)^s$ times the product of the pivots before scaling, where s is the number of swaps. Here this is $(-1) \cdot 1 \cdot (-1) \cdot 3 \cdot 6 = 18$. Edit the matrix to try others. When some column has no pivot, $\operatorname{rref}(A)$ has a zero row and $\det A = 0$. (Interactive in the web version.)

Structure in a matrix can often be exploited before any elimination. A common trick: if every row has the same sum, add all columns to the first, or all rows to the first if every column has the same sum.

Example 5.4.24 (A matrix with constant row sums). For $n \geq 1$ and $x \in F$, let $M_n(x) \in F^{n \times n}$ have every diagonal entry equal to x and every other entry equal to 1. Then

$$\det M_n(x) = (x + n - 1)(x - 1)^{n-1}.$$

Every column of $M_n(x)$ sums to $x + n - 1$. Adding rows $2, \ldots, n$ to row 1 does not change the determinant (Corollary 5.4.15), and makes every entry of row 1 equal to $x + n - 1$. Pulling this factor out of row 1 leaves the row $(1, 1, \ldots, 1)$. Now subtract the new row 1 from each of rows $2, \ldots, n$. Row i had x in position i and 1 elsewhere, so it becomes $(x - 1)\mathbf{e}_i^{\mathsf{T}}$. The result is upper triangular with diagonal entries $1, x - 1, \ldots, x - 1$, so by Proposition 5.3.15,

$$\det M_n(x) = (x + n - 1) \cdot 1 \cdot (x - 1)^{n-1}.$$

So for $n \geq 2$, $M_n(x)$ is invertible unless $x = 1$ or $x = 1 - n$. These two values coincide when $n = 0$ in F: over $\mathbb{F}_3$, $\det M_3(x) = (x + 2)(x - 1)^2 = (x - 1)^3$.

Remark 5.4.25 (How much faster). Each term of the Leibniz formula is a product of n entries, which takes $n - 1$ multiplications, so evaluating the formula term by term takes $(n - 1)n!$ multiplications. Algorithm 5.4.22 takes about $n^3/3$: stage k computes a multiplier for each of the $n - k$ rows below row k and updates the $n - k$ remaining entries of each, for $(n - k)(n - k + 1)$ multiplications and divisions, and summing over k gives $\left(n^3 - n\right)/3$ (Remark 3.3.35). Multiplying the diagonal entries adds $n - 1$ more.

n	Leibniz: $(n-1)n!$	Elimination: $\left(n^3 - n\right)/3$
3	12	8
5	480	40
10	32 659 200	330
20	$\approx 4.6 \cdot 10^{19}$	2660

Table 5.4.1. Multiplications (and divisions) needed to evaluate an $n \times n$ determinant.

At 10^9 multiplications per second, the Leibniz formula for a single 20×20 matrix would take about 1500 years, while elimination takes a few microseconds. Cofactor expansion (Section 5.5) also takes on the order of $n!$ operations for a general matrix, although it is convenient for matrices with many zeros. The Leibniz formula remains the tool for theory (Technique 5.3.12, Technique 5.3.18), and elimination the tool for computation.

5.4.7 Block-triangular matrices

Proposition 5.3.17 computed the determinant of $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$. Its second proof never used that the upper right block is 0, only the lower left one, which suggests the following theorem. Recall (Definition 2.4.33) that a

square block matrix with square diagonal blocks is block upper triangular if all blocks below the diagonal blocks are 0; it is **block lower triangular** if all blocks above them are 0.

Theorem 5.4.26 (Block-triangular matrices). Let $A \in F^{k\times k}$, $B \in F^{k\times m}$, $C \in F^{m\times k}$, and $D \in F^{m\times m}$. Then

$$\det\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det A \det D,$$

$$\det\begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det A \det D.$$

Proof (by factoring). STANDARD

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.17 (Block-diagonal matrices), Lemma 5.1.8 (Column replacement), Theorem 5.4.14 (The transpose), Definition 2.4.21 (Transpose)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Write $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ as a product of two block matrices, one of which is block diagonal.
Hint 2. Check that $\begin{pmatrix} I_k & B \\ 0 & D \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$.
Hint 3. To find $\det\begin{pmatrix} I_k & B \\ 0 & D \end{pmatrix}$, use the first k columns, which are standard basis vectors, to clear B by column replacements.
Hint 4. For the second identity, take transposes.
Sketch. Block multiplication gives $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} I_k & B \\ 0 & D \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & I_m \end{pmatrix}$. The second factor has determinant $\det A$ by Proposition 5.3.17. In the first, column $k+j$ is $\sum_i b_{ij}\mathbf{e}_i$ plus column j of D placed below, and subtracting b_{ij} times column i for each $i \le k$ clears B without changing the determinant, leaving $\begin{pmatrix} I_k & 0 \\ 0 & D \end{pmatrix}$ with determinant $\det D$. For the lower block-triangular case, transpose.

Let $n = k + m$ and $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$. By block multiplication (Theorem 2.4.34),

$$\begin{pmatrix} I_k & B \\ 0 & D \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & I_m \end{pmatrix} = \begin{pmatrix} I_k A + B0 & I_k 0 + BI_m \\ 0A + D0 & 00 + DI_m \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}.$$

So by Theorem 5.4.10 and Proposition 5.3.17,

$$\det M = \det N \det\begin{pmatrix} A & 0 \\ 0 & I_m \end{pmatrix} = \det N \det A,$$

where $N = \begin{pmatrix} I_k & B \\ 0 & D \end{pmatrix}$. For $i \le k$, column i of N is $\mathbf{e}_i \in F^n$. For $j \le m$, column $k+j$ of N is $\sum_{i=1}^{k} b_{ij}\mathbf{e}_i + \mathbf{d}'_j$, where $\mathbf{d}'_j \in F^n$ has k zeros followed by the entries of column j of D. Subtracting b_{ij} times column i from column $k+j$, for every $i \le k$ and $j \le m$, does not change $\det N$ (Lemma 5.1.8), does not change the first k columns, and turns column $k+j$ into $\mathbf{d}'_j$. The result is $\begin{pmatrix} I_k & 0 \\ 0 & D \end{pmatrix}$, so by Proposition 5.3.17, $\det N = \det I_k \det D = \det D$. Hence $\det M = \det A \det D$.

For the second identity, transposing exchanges the entries in positions (i,j) and (j,i) (Definition 2.4.21), so the transpose of $\begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$ is $\begin{pmatrix} A^{\mathsf{T}} & C^{\mathsf{T}} \\ 0 & D^{\mathsf{T}} \end{pmatrix}$. By Theorem 5.4.14 and the first identity,

$$\det\begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det\begin{pmatrix} A^{\mathsf{T}} & C^{\mathsf{T}} \\ 0 & D^{\mathsf{T}} \end{pmatrix} = \det A^{\mathsf{T}} \det D^{\mathsf{T}} = \det A \det D.$$ □

Proof (via the Leibniz formula). STANDARD

Uses. Definition 5.3.6 (Determinant), Proposition 5.3.17 (Block-diagonal matrices), Theorem 5.4.14 (The transpose)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. Compare the Leibniz terms of $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ with those of $M' = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$. For which σ can they differ?

Hint 2. If σ sends some $j \le k$ to a row below k, the term of σ contains an entry of the lower left block, in both matrices. Otherwise σ maps $[k]$ onto $[k]$, and its term never touches the upper right block.
Sketch. Let $M' = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$. If $\sigma(j) > k$ for some $j \le k$, the terms of σ in $\det M$ and $\det M'$ both contain a 0 from the lower left block. Otherwise σ maps $[k]$ onto $[k]$, hence the rest onto the rest, and every factor of its term lies in A or in D, where M and M' agree. So $\det M = \det M' = \det A \det D$.

Let $n = k + m$, $K = \{k+1, \dots, n\}$, $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$, and $M' = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$. Write c_{ij} and c'_{ij} for the entries of M and M'. They agree unless $i \le k < j$, that is, unless position (i, j) lies in the upper right block. We compare the terms of a fixed $\sigma \in S_n$ in the Leibniz formulas for $\det M$ and $\det M'$.

If $\sigma(j) \in K$ for some $j \le k$, then position $(\sigma(j), j)$ lies in the lower left block, so $c_{\sigma(j)j} = c'_{\sigma(j)j} = 0$, and both terms are 0. Otherwise σ maps $[k]$ into $[k]$, hence onto $[k]$ since it is injective, and then it maps K onto K since it is a bijection. Then every position $(\sigma(j), j)$ lies in one of the two diagonal blocks, where M and M' agree, so the two terms are equal. Hence $\det M = \det M' = \det A \det D$ by Proposition 5.3.17. The second identity follows by transposing, as in the first proof. □

The first proof shows how block matrices can be manipulated like matrices of numbers, as long as the order of factors is respected (Technique 2.4.35). The second proof is the shorter one here, because the work was done in Proposition 5.3.17. Grouping the diagonal blocks two at a time extends the theorem to any number of blocks.

Corollary 5.4.27 (Several diagonal blocks). Let M be block upper triangular (or block lower triangular) with square diagonal blocks $A_1, \dots, A_r$. Then

$$\det M = \det A_1 \det A_2 \cdots \det A_r.$$

Proof. ROUTINE

Uses. Theorem 5.4.26 (Block-triangular matrices), Definition 2.4.33 (Block matrix), induction on r
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Induct on r. Regard M as a 2×2 block matrix whose upper left block is A_1.
Hint 2. The lower right block is block triangular with diagonal blocks $A_2, \dots, A_r$.
Sketch. For $r \ge 2$, group the last $r-1$ block rows and columns: $M = \begin{pmatrix} A_1 & X \\ 0 & M' \end{pmatrix}$, where M' is block upper triangular with diagonal blocks $A_2, \dots, A_r$. Then $\det M = \det A_1 \det M'$ by Theorem 5.4.26, and induction finishes. The lower triangular case is the same with $\begin{pmatrix} A_1 & 0 \\ Y & M' \end{pmatrix}$.

We induct on r; for $r = 1$ there is nothing to prove. Let $r \ge 2$, and suppose M is block upper triangular. Grouping the block rows and block columns $2, \dots, r$ together, $M = \begin{pmatrix} A_1 & X \\ 0 & M' \end{pmatrix}$, where the lower left block is 0 because all blocks below the diagonal are, and M' is block upper triangular with diagonal blocks $A_2, \dots, A_r$. By Theorem 5.4.26 and the induction hypothesis,

$$\det M = \det A_1 \det M' = \det A_1 \det A_2 \cdots \det A_r.$$

The block lower triangular case is identical, with $M = \begin{pmatrix} A_1 & 0 \\ Y & M' \end{pmatrix}$. □

With 1×1 blocks this is Proposition 5.3.15. For example, the matrix

$$M = \begin{pmatrix} 1 & 2 & 5 & 7 \\ 3 & 4 & 6 & 8 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

is block upper triangular with diagonal blocks $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, so $\det M = (-2) \cdot 1 = -2$. A 2×2 block matrix does not behave like a 2×2 matrix of numbers, though: in general $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is neither $\det A \det D - \det B \det C$ (Exercise 5.4.36) nor $\det(AD - BC)$ (the example in Exercise 5.4.50).

Block elimination BEYOND 21-242

Schur complements are standard in numerical linear algebra and statistics, but they are not part of the course.

When the upper left block is invertible, a general 2×2 block matrix can be reduced to the block-triangular case by one step of elimination with blocks.

Technique 5.4.28 (Block elimination). Row reduction works with blocks in place of entries, as long as the order of the factors in each product is respected. By Theorem 2.4.34, multiplying $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ on the left by $\begin{pmatrix} I & 0 \\ X & I \end{pmatrix}$ adds X times the first block row to the second, giving $\begin{pmatrix} A & B \\ XA+C & XB+D \end{pmatrix}$. Multiplying on the right by $\begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}$ adds the first block column times Y to the second, giving $\begin{pmatrix} A & AY+B \\ C & CY+D \end{pmatrix}$. Both multipliers have determinant 1 by Theorem 5.4.26, so these operations do not change the determinant. Choose X or Y to create a zero block, and finish with Theorem 5.4.26.

Reach for it when you want the determinant or the inverse of a 2×2 block matrix and one of the blocks is invertible, or when two block matrices are related by such operations. The canonical example is Proposition 5.4.29.

Used in. Proposition 5.4.29, Exercise 5.4.50, Theorem 5.5.27.

Proposition 5.4.29 (Schur complement). Let $A \in F^{k\times k}$ be invertible, and let $B \in F^{k\times m}$, $C \in F^{m\times k}$, and $D \in F^{m\times m}$. Then

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det\big(D - CA^{-1}B\big).$$

Proof. STANDARD

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.26 (Block-triangular matrices)

Techniques. Block elimination (Technique 5.4.28)

Hint 1. Imitate the first stage of elimination, with the block A as the pivot: clear the block C below it.

Hint 2. Multiply on the left by $\begin{pmatrix} I_k & 0 \\ X & I_m \end{pmatrix}$, and choose X so that the lower left block becomes 0.

Sketch. With $X = -CA^{-1}$, block multiplication gives $\begin{pmatrix} I_k & 0 \\ X & I_m \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}$. The left factor has determinant 1, and the right side has determinant $\det A \det\big(D - CA^{-1}B\big)$, both by Theorem 5.4.26.

Let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and $X = -CA^{-1} \in F^{m\times k}$. By Theorem 2.4.34,

$$\begin{aligned} \begin{pmatrix} I_k & 0 \\ X & I_m \end{pmatrix} M &= \begin{pmatrix} A & B \\ XA + C & XB + D \end{pmatrix} \\ &= \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \end{aligned}$$

By Theorem 5.4.26, the left factor has determinant $\det I_k \det I_m = 1$, and the right side has determinant $\det A \det\big(D - CA^{-1}B\big)$. So Theorem 5.4.10 gives the claim. □

The matrix $D - CA^{-1}B$ is the **Schur complement** of A in M. For $k = m = 1$ the proposition reads $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\big(d - ca^{-1}b\big) = ad - bc$. If also $AC = CA$, then $A\big(D - CA^{-1}B\big) = AD - CB$, and the proposition becomes $\det M = \det(AD - CB)$ (Exercise 5.4.50).

Exercises

Exercise 5.4.30. ROUTINE Let

$$A = \begin{pmatrix} 1 & 3 & 0 & 2 \\ -2 & -6 & 2 & -3 \\ 3 & 10 & -1 & 9 \\ 1 & 1 & 4 & 12 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

(a) Compute $\det A$ with Algorithm 5.4.22.

(b) For which primes p is the reduction of A modulo p invertible over $\mathbb{F}_p$?

Solution on page 1416

Exercise 5.4.31. ROUTINE Let $A, B \in \mathbb{R}^{3\times 3}$ with $\det A = 2$ and $\det B = -3$. Compute

$$\det(AB), \quad \det\left(A^{-1}B^2\right), \quad \det\left(2A^{\mathsf{T}}\right),$$
$$\det\left(BAB^{-1}\right), \quad \det\left(A^{\mathsf{T}}A\right), \quad \det\left(-A^3\right).$$

Solution on page 1416

Exercise 5.4.32. STANDARD For which $n \geq 1$ is there a matrix $A \in F^{n\times n}$ with $A^2 = -I_n$, when (a) $F = \mathbb{R}$, (b) $F = \mathbb{C}$, (c) $F = \mathbb{F}_3$, (d) $F = \mathbb{F}_5$?

Solution on page 1417

Exercise 5.4.33. STANDARD A matrix $A \in F^{n\times n}$ is **skew-symmetric** if $A^{\mathsf{T}} = -A$.

(a) Suppose that $1 + 1 \neq 0$ in F and that n is odd. Show that every skew-symmetric $A \in F^{n\times n}$ has $\det A = 0$.

(b) Keep the assumption $1 + 1 \neq 0$. Show that the skew-symmetric 3×3 matrices are exactly the matrices $\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$ with $a, b, c \in F$, and for such a matrix $A \neq 0$, find a nonzero $\mathbf{x}$ with $A\mathbf{x} = \mathbf{0}$.

(c) Show that (a) fails over $\mathbb{F}_2$.

Solution on page 1417

Exercise 5.4.34. STANDARD Let $n \geq 1$. In each case, find all possible values of $\det A$, and show that each of them occurs.

(a) $A \in F^{n\times n}$ with $A^2 = A$.

(b) $A \in F^{n\times n}$ with $A^k = 0$ for some $k \geq 1$.

(c) $A \in F^{n\times n}$ with $A^2 = I_n$.

(d) $A \in \mathbb{R}^{n\times n}$ with $A^{\mathsf{T}}A = I_n$.

Solution on page 1418

Exercise 5.4.35. STANDARD Find the error in the following argument, and repair it.

Claim. $\det(AB) = \det A \det B$ for all $A, B \in F^{n\times n}$.

Proof. By Theorem 3.5.19, we can write $A = E_1 \cdots E_k$ and $B = E_1' \cdots E_l'$ with elementary matrices E_i and E_j'. By Corollary 5.4.4,

$$\begin{aligned} \det(AB) &= \det\left(E_1 \cdots E_k E_1' \cdots E_l'\right) \\ &= \det E_1 \cdots \det E_k \det E_1' \cdots \det E_l' \end{aligned}$$

$$= \det A \det B.$$

Solution on page 1418

Exercise 5.4.36. STANDARD Let $A, B, C, D \in F^{m\times m}$ and $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

(a) Find the error in the following argument.

Claim. $\det M = \det A \det D - \det B \det C$.

Proof. By Theorem 2.4.34, block matrices multiply like 2×2 matrices whose entries are the blocks. So M is a 2×2 matrix with entries A, B, C, D, and Theorem 5.1.13 gives $\det M = \det A \det D - \det B \det C$.

(b) Show that the claim is false for $m = 2$, over every field.

(c) Show that the claim is true when $C = 0$.

Solution on page 1418

Exercise 5.4.37. STANDARD For $n \geq 1$, let $P_n, Q_n \in \mathbb{Q}^{n\times n}$ have entries $\min(i,j)$ and $\max(i,j)$, respectively, in position (i,j). For example,

$$P_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}, \quad Q_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix}.$$

Show that $\det P_n = 1$ and $\det Q_n = (-1)^{n-1}n$.

Solution on page 1419

Exercise 5.4.38. STANDARD Let $m < n$, $A \in F^{n\times m}$, and $B \in F^{m\times n}$. Show that $\det(AB) = 0$. Does the same hold when $m \geq n$?

Solution on page 1419

Exercise 5.4.39. STANDARD Let $A \in F^{n\times n}$, and define $L_A, R_A \in \mathrm{L}(F^{n\times n})$ by $L_A(X) = AX$ and $R_A(X) = XA$. Show that $\det L_A = \det R_A = (\det A)^n$.

Solution on page 1420

Exercise 5.4.40. STANDARD Let $T \in \mathrm{L}(F^{n\times n})$ be the transpose, $T(X) = X^{\mathsf{T}}$. Show that $\det T = (-1)^{n(n-1)/2}$. For which n is $\det T = 1$, when $1 + 1 \neq 0$ in F?

Solution on page 1420

Exercise 5.4.41. STANDARD Prove or disprove each statement, where $A, B \in F^{n\times n}$ and $n \geq 2$.

(a) If $\det A = \det B$ and $\operatorname{tr} A = \operatorname{tr} B$, then A and B are similar.

(b) If A and B are invertible, then $\det\left(ABA^{-1}B^{-1}\right) = 1$.

(c) If $F = \mathbb{R}$, n is odd, and $AB = -BA$, then A or B is not invertible.

(d) If $\det A = \det B$, then $\det(A + C) = \det(B + C)$ for every $C \in F^{n\times n}$.

Solution on page 1421

Exercise 5.4.42. ROUTINE Let $A \in \mathbb{Q}^{n\times n}$ have integer entries. Show that if A is invertible and A^{-1} also has integer entries, then $\det A = \pm 1$. (The converse holds as well; see Section 5.5.)

Solution on page 1421

Exercise 5.4.43. STANDARD Let $A, B \in F^{m\times m}$. Show that

$$\det\begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det(A+B)\det(A-B).$$

Solution on page 1421

Exercise 5.4.44. STANDARD Let

$$C(z) = \begin{pmatrix} 1 & z & z^2 \\ z^2 & 1 & z \\ z & z^2 & 1 \end{pmatrix}.$$

(a) Show that $\det C(z) = \left(1-z^3\right)^2$, for z in any field.

(b) Find all z for which $C(z)$ is not invertible, for $z \in \mathbb{C}$, for $z \in \mathbb{R}$, and for $z \in \mathbb{F}_7$.

Solution on page 1422

Exercise 5.4.45. HARD Let $B, C \in \mathbb{R}^{n\times n}$, and let $A = B + iC \in \mathbb{C}^{n\times n}$. Show that

$$\det\begin{pmatrix} B & -C \\ C & B \end{pmatrix} = |\det A|^2.$$

In particular, the real $2n \times 2n$ matrix on the left has a nonnegative determinant. (For $n = 1$ and $A = (a + bi)$, the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ is the matrix of multiplication by $a + bi$ on $\mathbb{C}$, regarded as a real vector space with basis $1, i$.)

Solution on page 1423

Exercise 5.4.46. HARD Let $A, B \in \mathbb{R}^{n\times n}$ with $AB = BA$. Show that $\det\left(A^2 + B^2\right) \geq 0$. Show by an example that the hypothesis $AB = BA$ cannot be dropped.

Solution on page 1423

Exercise 5.4.47. HARD Let p be a prime, let G be the set of invertible matrices in $\mathbb{F}_p^{n\times n}$, and let $H = \{A \in G \mid \det A = 1\}$.

(a) For $c \in \mathbb{F}_p$ with $c \neq 0$, show that $X \mapsto XE_{1c}$ is a bijection from H onto $\{A \in G \mid \det A = c\}$. Deduce that $|H| = |G|/(p-1)$.

(b) For $n = 2$, show that $|G| = \left(p^2-1\right)\left(p^2-p\right)$, and conclude that $|H| = p\left(p^2-1\right)$.

(c) Check (b) for $p = 2$ by listing H.

Solution on page 1424

Exercise 5.4.48. HARD Let $n \geq 2$.

(a) Let $A \in F^{n\times n}$ be invertible. Show that there are elementary matrices $E_1, \ldots, E_t$ of type 3 (replacements) with $E_t \cdots E_1 A = \operatorname{diag}(1, \ldots, 1, \det A)$.

(b) Deduce that every $A \in F^{n\times n}$ with $\det A = 1$ is a product of elementary matrices of type 3, and that every invertible A is such a product times the single scaling $E_{n,\det A}$.

Solution on page 1424

Exercise 5.4.49. CHALLENGE Let $n \geq 2$, and let $f : F^{n\times n} \to F$ satisfy $f(AB) = f(A)f(B)$ for all A, B, and $f(\operatorname{diag}(t, 1, \ldots, 1)) = t$ for all $t \in F$. Show that $f = \det$. (This extends Exercise 5.1.42 to all n and all fields.)

Solution on page 1425

Exercise 5.4.50. HARD BEYOND 21-242 Let $A, B, C, D \in F^{m\times m}$ with A invertible and $AC = CA$. Show that

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB).$$

Then show by an example with $m = 2$ that the formula can fail when A is invertible but $AC \neq CA$.

Solution on page 1426

5.5 Cofactor expansion

Expanding the Leibniz formula takes $n!$ products, and elimination computes a determinant in about $n^3/3$ operations (Remark 5.4.25). This section gives a third description of det, which is recursive: a **cofactor expansion** writes an $n \times n$ determinant as a combination of n determinants of size $n - 1$, one for each entry of a chosen row or column. For a general matrix this still takes on the order of $n!$ operations. Its strength lies elsewhere. It is quick for matrices with many zeros, and it sets up induction on the size of a matrix, which computes the determinants of whole families of matrices at once: tridiagonal matrices, companion matrices, and the Vandermonde matrices of Section 3.4. It also has a theoretical payoff. The coefficients in the expansions, the cofactors, assemble into a matrix that inverts B up to the scalar factor $\det B$. This gives explicit formulas for B^{-1} and for the solution of $B\mathbf{x} = \mathbf{y}$ (Cramer's rule).

5.5.1 Expansion along a row or a column

We start where Intuition (What the proof did, and where it leads) and Lemma 5.3.2 started, with multilinearity, but this time we expand only one row. Row i of $B \in F^{n\times n}$ is $\sum_j b_{ij}\mathbf{e}_j^{\mathsf{T}}$, so by linearity of det in row i (Corollary 5.4.15),

$$\det B = \sum_{j=1}^{n} b_{ij} \det B^{(j)},$$

where $B^{(j)}$ is B with row i replaced by $\mathbf{e}_j^{\mathsf{T}}$. Each $B^{(j)}$ has a row with a single nonzero entry. The lecture computes the determinant of such a matrix by comparing it with a matrix of size one less. We first name that smaller matrix.

Notation 5.5.1 (Minors). Let $n \geq 2$, $B \in F^{n\times n}$, and $i, j \in [n]$. We write $B_{\hat{i}\hat{j}} \in F^{(n-1)\times(n-1)}$ for the matrix obtained from B by deleting row i and column j. Its determinant $\det B_{\hat{i}\hat{j}}$ is the (i, j) **minor** of B. The hat marks the row and the column that are left out.

For instance, if $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$, then $B_{\hat{2}\hat{3}} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}$, with minor $8 - 14 = -6$, and $B_{\hat{1}\hat{1}} = \begin{pmatrix} 5 & 6 \\ 8 & 10 \end{pmatrix}$, with minor $50 - 48 = 2$. Two facts are used again and again. First, $B_{\hat{i}\hat{j}}$ involves no entry of row i and no entry of column j, so changing B in row i or in column j does not change $B_{\hat{i}\hat{j}}$. Second, deleting row j and column i of B^{T} gives the transpose of $B_{\hat{i}\hat{j}}$, that is, $\left(B^{\mathsf{T}}\right)_{\hat{j}\hat{i}} = \left(B_{\hat{i}\hat{j}}\right)^{\mathsf{T}}$.

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix} \longrightarrow B_{\widehat{23}} = \begin{pmatrix} b_{11} & b_{12} & b_{14} \\ b_{31} & b_{32} & b_{34} \\ b_{41} & b_{42} & b_{44} \end{pmatrix}$$

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

Figure 5.5.1. Top: deleting row 2 and column 3 of a 4×4 matrix B (the entries in grey) leaves the 3×3 matrix $B_{\widehat{23}}$. The remaining rows and columns keep their order. Bottom: the signs $(-1)^{i+j}$ that accompany the minors in Theorem 5.5.3 form a checkerboard, with + on the diagonal.

Intuition (Moving an entry into the corner). The lecture's starting point is the case $B = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$, in block form, where $\det B = \det A$. More generally, if the first row of B is $(b_{11}, 0, \ldots, 0)$, then B is block lower triangular with diagonal blocks (b_{11}) and $B_{\widehat{11}}$, and Theorem 5.4.26 gives $\det B = b_{11} \det B_{\widehat{11}}$. The entries below b_{11} play no role.

If the single nonzero entry of row i sits in column j, we move it into the corner. This must be done so that the rest of the matrix becomes exactly $B_{\hat{ij}}$. Swapping rows 1 and i directly does not achieve this: old row 1 would land in position i, out of order. Instead, move row i up one step at a time, swapping it with the row just above it. After $i-1$ swaps it is on top, and the other rows are still in their original order. Do the same with column j. Each swap changes the sign, so the price is $(-1)^{(i-1)+(j-1)} = (-1)^{i+j}$.

Lemma 5.5.2 (A row or column with one nonzero entry). Let $n \geq 2$, $B \in F^{n \times n}$, and $i, j \in [n]$. If every entry of row i of B other than b_{ij} is 0, then

$$\det B = (-1)^{i+j} b_{ij} \det B_{\hat{ij}}.$$

The same holds if every entry of column j other than b_{ij} is 0.

Proof. STANDARD

Uses. Corollary 5.4.15 (Rows), Proposition 5.1.10 (Alternating implies antisymmetric), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.14 (The transpose)
Hint 1. Settle the case $i = j = 1$ first. Which block structure does B have then?
Hint 2. For general i and j, move row i to the top and column j to the front so that the remaining rows and columns stay in their original order. How many swaps does this take?
Hint 3. Use only swaps of adjacent rows (and adjacent columns). For the column statement, transpose.
Sketch. Swap row i successively with rows $i-1, \ldots, 1$, and then column j with columns $j-1, \ldots, 1$. The result B' has first row $(b_{ij}, 0, \ldots, 0)$ and lower right block $B_{\hat{ij}}$, because the other rows and columns keep their order. So B' is block lower triangular and $\det B' = b_{ij} \det B_{\hat{ij}}$, while $\det B' = (-1)^{(i-1)+(j-1)} \det B$. For a column with one nonzero entry, apply the row case to B^{T}.

Suppose first that row i has the given form. Let B'' be the matrix obtained from B by swapping row i with row $i-1$, then with row $i-2$, and so on, up to row 1. These are $i-1$ swaps, and each multiplies the determinant by -1 (Corollary 5.4.15), so $\det B'' = (-1)^{i-1} \det B$. The rows of B'' are row i of B, followed by rows $1, \ldots, i-1, i+1, \ldots, n$ of B in this order. Next, let B' be obtained from B'' by swapping column j with column $j-1$, then with column $j-2$, and so on, up to column 1. By Proposition 5.1.10, $\det B' = (-1)^{j-1} \det B''$. The columns of B' are column j of B'', followed by its other columns in their original order.

So the first row of B' is row i of B rearranged, namely $(b_{ij}, 0, \ldots, 0)$, and deleting the first row and the first column of B' leaves the rows of B other than row i, in their original order, restricted to the columns other than column j, in their original order. That is exactly $B_{\hat{ij}}$. Hence, in block form,

$$B' = \begin{pmatrix} b_{ij} & 0 \\ \mathbf{w} & B_{\hat{i}\hat{j}} \end{pmatrix}$$

for some $\mathbf{w} \in F^{n-1}$, which is block lower triangular. By Theorem 5.4.26, $\det B' = b_{ij} \det B_{\hat{i}\hat{j}}$. Combining,

$$\det B = (-1)^{i-1}(-1)^{j-1} \det B' = (-1)^{i+j} b_{ij} \det B_{\hat{i}\hat{j}},$$

since $(-1)^{i+j-2} = (-1)^{i+j}$.

Now suppose instead that every entry of column j other than b_{ij} is 0. Then row j of B^{T} has the single nonzero entry b_{ij}, in column i. By the case already proved, Theorem 5.4.14, and $\left(B^{\mathsf{T}}\right)_{\hat{j}\hat{i}} = \left(B_{\hat{i}\hat{j}}\right)^{\mathsf{T}}$,

$$\det B = \det B^{\mathsf{T}} = (-1)^{j+i} b_{ij} \det \left(B_{\hat{i}\hat{j}}\right)^{\mathsf{T}} = (-1)^{i+j} b_{ij} \det B_{\hat{i}\hat{j}}.$$

□

The lecture states the case $i = 1$ and proves the case $j = 1$ (with $b_{11} = 1$) by observing that Gauss–Jordan elimination on $\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$ performs the same steps as on A. We had Theorem 5.4.26 available, which already contains that case.

The sign $(-1)^{i+j}$ depends only on the position, and it alternates like the colors of a checkerboard (Figure 5.5.1). Now we expand.

Theorem 5.5.3 (Cofactor expansion). Let $n \geq 2$ and $B \in F^{n \times n}$. For every $i \in [n]$,

$$\det B = \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det B_{\hat{i}\hat{j}} \tag{5.5.1}$$

(**expansion along row** i), and for every $j \in [n]$,

$$\det B = \sum_{i=1}^{n} (-1)^{i+j} b_{ij} \det B_{\hat{i}\hat{j}} \tag{5.5.2}$$

(**expansion along column** j).

Proof (by linearity in one row). ROUTINE

Uses. Corollary 5.4.15 (Rows), Lemma 5.5.2 (A row or column with one nonzero entry), Theorem 5.4.14 (The transpose)
Techniques. Expand by multilinearity (Technique 5.1.7), Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. Write row i as a linear combination of $\mathbf{e}_1^{\mathsf{T}}, \ldots, \mathbf{e}_n^{\mathsf{T}}$, and use that det is linear in row i.
Hint 2. Each resulting matrix has a row with a single nonzero entry. Apply Lemma 5.5.2, and check that the minor is the same as for B.
Hint 3. For columns, expand B^{T} along row j.
Sketch. Row i of B is $\sum_j b_{ij} \mathbf{e}_j^{\mathsf{T}}$. By linearity in row i, $\det B = \sum_j b_{ij} \det B^{(j)}$, where $B^{(j)}$ is B with row i replaced by $\mathbf{e}_j^{\mathsf{T}}$. By Lemma 5.5.2, $\det B^{(j)} = (-1)^{i+j} \det B^{(j)}_{\hat{i}\hat{j}}$, and $B^{(j)}_{\hat{i}\hat{j}} = B_{\hat{i}\hat{j}}$ since the two matrices differ only in row i. The expansion of B along column j is the expansion of B^{T} along row j.

Fix i. For $j \in [n]$, let $B^{(j)}$ be the matrix obtained from B by replacing row i with $\mathbf{e}_j^{\mathsf{T}}$. Since row i of B is $\sum_j b_{ij} \mathbf{e}_j^{\mathsf{T}}$ and det is linear in row i (Corollary 5.4.15),

$$\det B = \sum_{j=1}^{n} b_{ij} \det B^{(j)}.$$

The only nonzero entry of row i of $B^{(j)}$ is the 1 in column j, so $\det B^{(j)} = (-1)^{i+j} \det B^{(j)}_{\hat{i}\hat{j}}$ by Lemma 5.5.2. The matrices $B^{(j)}$ and B differ only in row i, which is deleted, so $B^{(j)}_{\hat{i}\hat{j}} = B_{\hat{i}\hat{j}}$. This proves (5.5.1).

For (5.5.2), fix j. Row j of B^{T} has entries $b_{1j}, \ldots, b_{nj}$, and $\left(B^{\mathsf{T}}\right)_{\hat{j}\hat{i}} = \left(B_{\hat{i}\hat{j}}\right)^{\mathsf{T}}$. By Theorem 5.4.14 and (5.5.1) for B^{T},

$$\det B = \det B^{\mathsf{T}} = \sum_{i=1}^{n} (-1)^{j+i} b_{ij} \det\left(B_{\hat{ij}}\right)^{\mathsf{T}} = \sum_{i=1}^{n} (-1)^{i+j} b_{ij} \det B_{\hat{ij}}.$$

□

Proof (via the Leibniz formula). HARD

> *Uses.* Definition 5.3.6 (Determinant), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign), Theorem 5.2.31 (Alternating functions and permutations), Proposition 5.3.17 (Block-diagonal matrices), Theorem 5.4.14 (The transpose)
> *Techniques.* Read structure off the Leibniz expansion (Technique 5.3.12)
> *Hint 1.* Every term of the Leibniz formula contains exactly one entry of column j. Group the terms according to which entry b_{ij} it is, that is, according to $i = \sigma(j)$.
> *Hint 2.* The terms with $\sigma(j) = i$ use one entry from each row other than i and each column other than j, like the terms of $\det B_{\hat{ij}}$. Match them with the permutations of $[n-1]$.
> *Hint 3.* To compare signs, write $\sigma = \rho_i \sigma' \rho_j^{-1}$ with $\sigma'(1) = 1$, where $\rho_m(1) = m$ and $\rho_m(p) = p - 1$ for $2 \le p \le m$ (and ρ_m fixes the rest). This ρ_m is an m-cycle.
> *Sketch.* Grouping by $i = \sigma(j)$ gives $\det B = \sum_i b_{ij} s_i$, where s_i is the sum of $\operatorname{sgn}(\sigma) \prod_{k \ne j} b_{\sigma(k)k}$ over the σ with $\sigma(j) = i$. The cycle ρ_m lists $[n] \setminus \{m\}$ in order on $2, \dots, n$, so $\sigma \mapsto \rho_i^{-1} \sigma \rho_j$ is a bijection onto the permutations fixing 1, which correspond to S_{n-1}; it multiplies signs by $\operatorname{sgn}(\rho_i)\operatorname{sgn}(\rho_j) = (-1)^{i+j}$ and turns the product into the Leibniz term of $B_{\hat{ij}}$. So $s_i = (-1)^{i+j} \det B_{\hat{ij}}$. Rows follow by transposing.

We prove (5.5.2); then (5.5.1) follows by applying it to B^{T}, using Theorem 5.4.14 and $\left(B^{\mathsf{T}}\right)_{\hat{ji}} = \left(B_{\hat{ij}}\right)^{\mathsf{T}}$.

Fix j. Each term $\operatorname{sgn}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n}$ of the Leibniz formula (Definition 5.3.6) contains exactly one factor from column j, namely $b_{\sigma(j)j}$. Grouping the terms by the value $i = \sigma(j)$,

$$\det B = \sum_{i=1}^{n} b_{ij} s_i, \quad \text{where } s_i = \sum_{\sigma \in S_n,\, \sigma(j) = i} \operatorname{sgn}(\sigma) \prod_{k \ne j} b_{\sigma(k)k}.$$

We show that $s_i = (-1)^{i+j} \det B_{\hat{ij}}$.

For $m \in [n]$, let $\rho_m \in S_n$ be given by $\rho_m(1) = m$, $\rho_m(p) = p - 1$ for $2 \le p \le m$, and $\rho_m(p) = p$ for $p > m$. For $m \ge 2$ this is the m-cycle $(1\ m\ m-1 \cdots 2)$, and $\rho_1 = \mathrm{id}$, so $\operatorname{sgn}(\rho_m) = (-1)^{m-1}$ by Proposition 5.2.26. As q runs through $1, \dots, n-1$, the values $\rho_m(q+1)$ run through $[n] \setminus \{m\}$ in increasing order. So row q of $B_{\hat{ij}}$ is row $\rho_i(q+1)$ of B, column q of $B_{\hat{ij}}$ is column $\rho_j(q+1)$ of B, and

$$\left(B_{\hat{ij}}\right)_{pq} = b_{\rho_i(p+1),\, \rho_j(q+1)}.$$

If $\sigma(j) = i$, then $\sigma' = \rho_i^{-1} \sigma \rho_j$ satisfies $\sigma'(1) = \rho_i^{-1}(\sigma(j)) = 1$. Conversely, if $\sigma' \in S_n$ fixes 1, then $\sigma = \rho_i \sigma' \rho_j^{-1}$ satisfies $\sigma(j) = \rho_i(\sigma'(1)) = i$. So $\sigma \mapsto \sigma'$ is a bijection from $\{\sigma \mid \sigma(j) = i\}$ onto the permutations fixing 1. These in turn correspond to the $\tau \in S_{n-1}$ via $\sigma'(q+1) = \tau(q) + 1$. By Proposition 5.2.25,

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\rho_i)\operatorname{sgn}(\sigma')\operatorname{sgn}\left(\rho_j\right)^{-1} = (-1)^{i+j} \operatorname{sgn}(\sigma').$$

Also $\operatorname{sgn}(\sigma') = \operatorname{sgn}(\tau)$: if $\tau = \tau_1 \cdots \tau_r$ with transpositions $\tau_l = (a_l\ b_l)$ of $[n-1]$, then σ' is the product of the r transpositions $(a_l + 1\ b_l + 1)$ of $[n]$, since both sides fix 1 and send $q+1$ to $\tau(q) + 1$.

For the product, the index $k = \rho_j(q+1)$ runs through $[n] \setminus \{j\}$ as q runs through $[n-1]$, and $\sigma\left(\rho_j(q+1)\right) = \rho_i(\sigma'(q+1)) = \rho_i(\tau(q)+1)$. Hence

$$\prod_{k \ne j} b_{\sigma(k)k} = \prod_{q=1}^{n-1} b_{\rho_i(\tau(q)+1),\, \rho_j(q+1)} = \prod_{q=1}^{n-1} \left(B_{\hat{ij}}\right)_{\tau(q)q}.$$

Summing over $\tau \in S_{n-1}$,

$$s_i = (-1)^{i+j} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) \prod_{q=1}^{n-1} \left(B_{\hat{ij}}\right)_{\tau(q)q} = (-1)^{i+j} \det B_{\hat{ij}},$$

as desired. □

Intuition (Two proofs compared). The first proof is Technique 5.1.7 stopped after one row. Lemma 5.3.2 expanded all n columns at once and needed permutations to sort out the n^n terms; expanding a single row produces only n terms, and the lemma about one nonzero entry evaluates each of them. The second proof reads the expansion off the Leibniz formula (Technique 5.3.12): every term uses exactly one entry of column j, and the terms that use b_{ij} are, up to the sign $(-1)^{i+j}$, the Leibniz terms of $B_{\hat{ij}}$. All of the work is in the bookkeeping of signs, which the first proof gets for free from the adjacent swaps.

A third route is Exercise 5.3.39. It takes the expansion along the first column as a *definition*, by induction on n, and checks the axioms directly.

For $n = 3$, expanding along row 1 gives

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix} = a(ek - fh) - b(dk - fg) + c(dh - eg),$$

whose six terms are those of the rule of Sarrus (Corollary 5.3.9). Expansion pays off when a row or column has many zeros, since every zero entry removes a minor from the sum.

Example 5.5.4 (A sparse 4×4 determinant). Let

$$A = \begin{pmatrix} 2 & 0 & 1 & 3 \\ 1 & 0 & 0 & 2 \\ 5 & 4 & 2 & 1 \\ 3 & 0 & 1 & 1 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

Column 2 has a single nonzero entry, $a_{32} = 4$, so expanding along column 2 (or applying Lemma 5.5.2 directly) gives

$$\det A = (-1)^{3+2} \cdot 4 \det A_{\hat{32}} = -4 \det\begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix}.$$

In the 3×3 minor, row 2 contains a zero. Expanding along row 2,

$$\det\begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix} = -1 \cdot \det\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + 0 - 2 \det\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} = -(-2) - 2(-1) = 4.$$

Hence $\det A = -16$.

In practice, expansion is combined with elimination. Row and column replacements do not change the determinant (Corollary 5.4.15, Lemma 5.1.8), so one first creates zeros in some row or column, and then expands along it. This reduces the size by one at the cost of one round of replacements. For a general matrix with no zeros, though, a full cofactor expansion computes n minors of size $n - 1$, each of these $n - 1$ minors of size $n - 2$, and so on, about $n!$ products in all; elimination is the better algorithm (Remark 5.4.25).

Exercise 5.5.5. ROUTINE Compute the determinant of

$$M = \begin{pmatrix} i & 0 & 2 & 0 \\ 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 2 & 1 & i & 1 \end{pmatrix} \in \mathbb{C}^{4\times 4}$$

with two cofactor expansions. Along which rows or columns should you expand?

Solution on page 1426

5.5.2 The cofactor matrix and the inverse

The numbers $(-1)^{i+j} \det B_{\hat{i}\hat{j}}$ appear in both expansions, so we give them a name and arrange them in a matrix.

Definition 5.5.6 (Cofactors, cofactor matrix, adjugate). Let $n \geq 2$ and $B \in F^{n\times n}$. The (i,j) **cofactor** of B is

$$c_{ij} = (-1)^{i+j} \det B_{\hat{i}\hat{j}}.$$

The **cofactor matrix** of B is $C_B = (c_{ij}) \in F^{n\times n}$, and the **adjugate** of B is its transpose,

$$\operatorname{adj} B = C_B^{\mathsf{T}},$$

whose (i,j) entry is c_{ji}. For $n = 1$ we set $\operatorname{adj}(b) = (1)$.

In this notation, Theorem 5.5.3 says that $\det B = \sum_j b_{ij}c_{ij}$ for each row i, and $\det B = \sum_i b_{ij}c_{ij}$ for each column j: multiply each entry of a row or column by its own cofactor and add. The lecture writes the adjugate as C_B^{T}, and we use both. The adjugate is sometimes called the classical adjoint, but it is unrelated to the adjoints of Section 6.6, so we avoid that name. (The convention for $n = 1$ makes the theorems below true for $n = 1$ as well. It corresponds to declaring the determinant of the empty 0×0 matrix to be 1.)

Example 5.5.7 (The 2×2 case). Let $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$. Deleting a row and a column leaves a single entry, so

$$c_{11} = s, \quad c_{12} = -r, \quad c_{21} = -q, \quad c_{22} = p,$$

and

$$C_B = \begin{pmatrix} s & -r \\ -q & p \end{pmatrix}, \quad \operatorname{adj} B = \begin{pmatrix} s & -q \\ -r & p \end{pmatrix}.$$

Then $B \operatorname{adj} B = \begin{pmatrix} ps-qr & -pq+qp \\ rs-sr & -rq+sp \end{pmatrix} = (ps-qr)I_2$, and similarly $(\operatorname{adj} B)B = (ps-qr)I_2$. So when $ps - qr \neq 0$, the inverse of B is $(ps-qr)^{-1}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$: swap the diagonal entries, negate the others, and divide by the determinant. The theorem below says that the same pattern holds for every n.

The product $B \operatorname{adj} B$ in the example has the determinant on the diagonal, which is expected: its (i,i) entry is $\sum_j b_{ij}c_{ij}$, the expansion along row i. Its off-diagonal entries pair row i of B with the cofactors of a different row k. The cofactors of row k do not involve row k at all, so pairing them with a vector $\mathbf{x}$ computes the determinant of B with row k replaced by $\mathbf{x}^{\mathsf{T}}$. When $\mathbf{x}^{\mathsf{T}}$ is another row of B, that matrix has two equal rows.

Lemma 5.5.8 (Replacing a row or a column). Let $n \geq 2$, let $B \in F^{n\times n}$ have cofactors c_{ij}, let $k \in [n]$, and let $\mathbf{x} \in F^n$.

(a) If B' is obtained from B by replacing row k with $\mathbf{x}^{\mathsf{T}}$, then $\det B' = \sum_{j=1}^n x_j c_{kj}$.

(b) If $B_k(\mathbf{x})$ is obtained from B by replacing column k with $\mathbf{x}$, then $\det B_k(\mathbf{x}) = \sum_{i=1}^n x_i c_{ik}$.

In particular, for all $i, k \in [n]$,

$$\sum_{j=1}^{n} b_{ij}c_{kj} = \delta_{ik} \det B \quad \text{and} \quad \sum_{j=1}^{n} b_{ji}c_{jk} = \delta_{ik} \det B,$$

where δ_{ik} is 1 if $i = k$ and 0 otherwise.

Proof. ROUTINE

Uses. Theorem 5.5.3 (Cofactor expansion), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Corollary 5.4.15 (Rows), Definition 5.1.1 (Determinant function)
Hint 1. Expand $\det B'$ along row k. What are the cofactors of row k of B'?
Hint 2. B' and B differ only in row k, which is deleted when forming the minors of row k.
Hint 3. For the last statement, take $\mathbf{x}^{\mathsf{T}}$ to be row i of B. If $i \neq k$, what does B' look like?
Sketch. The minors $B'_{\widehat{kj}}$ do not involve row k, so they equal $B_{\widehat{kj}}$, and the row k cofactors of B' are $c_{k1}, \ldots, c_{kn}$. Expanding along row k gives (a); (b) is the same with columns. With $\mathbf{x}^{\mathsf{T}}$ equal to row i of B, $B' = B$ if $i = k$, and B' has two equal rows if $i \neq k$. Likewise for columns.

(a) For each j, the matrices B' and B agree outside row k, so $B'_{\widehat{kj}} = B_{\widehat{kj}}$, and the (k, j) cofactor of B' is c_{kj}. Expanding $\det B'$ along row k (Theorem 5.5.3), whose entries are $x_1, \ldots, x_n$, gives $\det B' = \sum_j x_j c_{kj}$.

(b) Similarly, $B_k(\mathbf{x})$ agrees with B outside column k, so its (i, k) cofactor is c_{ik}, and expanding along column k gives $\det B_k(\mathbf{x}) = \sum_i x_i c_{ik}$.

For the last statement, let $\mathbf{x}^{\mathsf{T}}$ be row i of B, so that $x_j = b_{ij}$. If $i = k$, then $B' = B$, and (a) gives $\sum_j b_{kj}c_{kj} = \det B$. If $i \neq k$, then rows i and k of B' are both equal to row i of B, so $\det B' = 0$ (Corollary 5.4.15), and (a) gives $\sum_j b_{ij}c_{kj} = 0$. For the column version, let $\mathbf{x}$ be column i of B, so $x_j = b_{ji}$. If $i = k$, then $B_k(\mathbf{x}) = B$; if $i \neq k$, then $B_k(\mathbf{x})$ has equal columns i and k, so its determinant is 0 (Definition 5.1.1). Now apply (b). □

The sums with $i \neq k$ are often called expansions by *alien cofactors*. Part (b) will give Cramer's rule, and the last statement gives the main theorem of this subsection.

Theorem 5.5.9 (The adjugate and the inverse). For every $B \in F^{n\times n}$,

$$B \operatorname{adj} B = (\operatorname{adj} B)B = (\det B)I_n.$$

Consequently, if $\det B \neq 0$, then B is invertible and

$$B^{-1} = \frac{1}{\det B} \operatorname{adj} B = \frac{1}{\det B} C_B^{\mathsf{T}}.$$

Proof. STANDARD

Uses. Lemma 5.5.8 (Replacing a row or a column), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Definition 2.4.1 (Matrix product), Definition 2.5.12 (Invertible)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Compute the (i, k) entry of $B \operatorname{adj} B$. The (j, k) entry of $\operatorname{adj} B$ is c_{kj}.
Hint 2. You get $\sum_j b_{ij}c_{kj}$, which Lemma 5.5.8 evaluates. For $(\operatorname{adj} B)B$, use the column version.
Sketch. The (i, k) entry of $B \operatorname{adj} B$ is $\sum_j b_{ij}c_{kj} = \delta_{ik} \det B$, and that of $(\operatorname{adj} B)B$ is $\sum_j c_{ji}b_{jk} = \delta_{ik} \det B$, both by Lemma 5.5.8. Dividing by $\det B \neq 0$ shows that $(\det B)^{-1} \operatorname{adj} B$ is a two-sided inverse.

For $n = 1$, $B = (b)$ and $\operatorname{adj} B = (1)$, so both products are $(b) = (\det B)I_1$. Let $n \geq 2$. The (j, k) entry of $\operatorname{adj} B = C_B^{\mathsf{T}}$ is c_{kj}. So by Definition 2.4.1 and Lemma 5.5.8, the (i, k) entry of $B \operatorname{adj} B$ is

$$\sum_{j=1}^{n} b_{ij}c_{kj} = \delta_{ik} \det B,$$

which is the (i, k) entry of $(\det B)I_n$. Similarly, the (i, k) entry of $(\operatorname{adj} B)B$ is

$$\sum_{j=1}^{n} c_{ji}b_{jk} = \sum_{j=1}^{n} b_{jk}c_{ji} = \delta_{ki} \det B,$$

by the column version in Lemma 5.5.8 (with the roles of i and k exchanged). Since $\delta_{ki} = \delta_{ik}$, this is again the (i,k) entry of $(\det B)I_n$.

If $\det B \neq 0$, multiplying both identities by $(\det B)^{-1}$ gives

$$B\big((\det B)^{-1}\operatorname{adj} B\big) = \big((\det B)^{-1}\operatorname{adj} B\big)B = I_n,$$

so $(\det B)^{-1}\operatorname{adj} B$ is a two-sided inverse of B (Definition 2.5.12). □

The theorem gives a second proof of (II) ⇒ (I) in Theorem 5.4.7 that uses no elimination, and it produces the inverse explicitly. The identity $B\operatorname{adj} B = (\det B)I_n$ holds for every B, invertible or not; for instance, if $\det B = 0$, it says that every column of $\operatorname{adj} B$ lies in $\ker B$.

The lecture calls the formula for B^{-1} more of a curiosity than a useful way to compute inverses. For computing, that is right: it needs n^2 determinants of size $n-1$, while row reducing $(B \mid I_n)$ (Algorithm 3.3.24) takes on the order of n^3 operations, the same order as a single determinant. Its value is that it describes B^{-1} by a formula. Every entry of B^{-1} is a polynomial in the entries of B, divided by $\det B$. Two consequences follow below: integer matrices with determinant ± 1 have integer inverses, and the solution of $B\mathbf{x} = \mathbf{y}$ is given by Cramer's rule. The formula also shows how $\det B$ depends on a single entry (Remark 5.5.11), which leads to formulas for derivatives of determinants (Section 5.5.6).

Example 5.5.10 (A 3×3 inverse). Let $B = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$. The nine cofactors are

$$c_{11} = +\det\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 3, \quad c_{12} = -\det\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} = -4, c_{13} = +\det\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} = -1,$$

$$c_{21} = -\det\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = -1, c_{22} = +\det\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} = 2, \quad c_{23} = -\det\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} = 0,$$

$$c_{31} = +\det\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = -1, c_{32} = -\det\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} = 1, \quad c_{33} = +\det\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} = 1.$$

Expanding along row 1, $\det B = 2 \cdot 3 + 1 \cdot (-4) + 1 \cdot (-1) = 1$. Transposing C_B,

$$B^{-1} = \operatorname{adj} B = \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

Multiplying out confirms $B\operatorname{adj} B = I_3$. The inverse has integer entries because $\det B = 1$.

Remark 5.5.11 (Cofactors measure sensitivity). Let $B \in F^{n\times n}$ with $n \geq 2$, and let $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ be the matrix unit (Lemma 2.4.24). Adding tU_{ij} to B changes only the entry b_{ij}, which becomes $b_{ij} + t$. Expanding along row i, whose cofactors do not change,

$$\det\big(B + tU_{ij}\big) = \det B + tc_{ij} \quad \text{for all } t \in F.$$

So $\det B$ is an affine function of each single entry, and the cofactor c_{ij} is its slope. Over $\mathbb{R}$, $c_{ij} = \partial \det B/\partial b_{ij}$. For instance, if $c_{ij} \neq 0$, then changing b_{ij} to $b_{ij} - \det B/c_{ij}$ makes B singular, while if $c_{ij} = 0$, no change of b_{ij} alone affects $\det B$.

> **Corollary 5.5.12** (Integer inverses). Let $A \in \mathbb{Q}^{n\times n}$ have integer entries. Then $\operatorname{adj} A$ has integer entries, and so does $(\det A)A^{-1}$ if A is invertible. In particular, A is invertible with A^{-1} having integer entries if and only if $\det A = \pm 1$.

Proof. ROUTINE

Uses. Theorem 5.5.9 (The adjugate and the inverse), Proposition 5.3.20 (Reduction modulo p), Exercise 5.4.42, Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Each cofactor is $\pm$ the determinant of a matrix with integer entries.
Hint 2. For the converse direction, take determinants of $AA^{-1} = I_n$, as in Exercise 5.4.42.
Sketch. Determinants of integer matrices are integers (Proposition 5.3.20), so all cofactors and hence $\operatorname{adj} A$ are integral, and $(\det A)A^{-1} = \operatorname{adj} A$. If $\det A = \pm 1$, then $A^{-1} = \pm \operatorname{adj} A$ is integral. Conversely, if A^{-1} is integral, $\det A \det A^{-1} = 1$ with both factors integers forces $\det A = \pm 1$.

Each cofactor of A is $\pm$ the determinant of a matrix with integer entries, hence an integer by Proposition 5.3.20. So $\operatorname{adj} A$ has integer entries. If A is invertible, then $\det A \neq 0$ (Theorem 5.4.7), and $(\det A)A^{-1} = \operatorname{adj} A$ by Theorem 5.5.9. If $\det A = \pm 1$, then A is invertible and $A^{-1} = (\det A)^{-1} \operatorname{adj} A = \pm \operatorname{adj} A$ has integer entries. The converse is Exercise 5.4.42. □

So the denominators in A^{-1} always divide $\det A$. To manufacture integer matrices with integer inverses, start from an upper triangular matrix with 1s on the diagonal and apply row and column replacements with integer multipliers. These keep the entries integral and the determinant equal to 1.

Exercise 5.5.13. ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 1 & 1 \\ 0 & 4 & 3 \end{pmatrix} \in \mathbb{F}_7^{3\times 3}$. Compute $\det A$ and $\operatorname{adj} A$ over $\mathbb{F}_7$, and use them to find A^{-1}. Check your answer by multiplying.

Solution on page 1427

5.5.3 Cramer's rule

Intuition (Coefficients as ratios of areas). Suppose $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$ are independent and $\mathbf{y} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$. Replace $\mathbf{a}_1$ by $\mathbf{y}$ in the parallelogram spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$. The part $x_2\mathbf{a}_2$ of $\mathbf{y}$ is parallel to the other side, and shearing along $\mathbf{a}_2$ does not change signed area (Lemma 5.1.8). What remains, $x_1\mathbf{a}_1$, scales the area by x_1. So

$$\det(\mathbf{y}, \mathbf{a}_2) = \det(x_1\mathbf{a}_1 + x_2\mathbf{a}_2, \mathbf{a}_2) = x_1 \det(\mathbf{a}_1, \mathbf{a}_2),$$

and x_1 is the ratio $\det(\mathbf{y}, \mathbf{a}_2)/\det(\mathbf{a}_1, \mathbf{a}_2)$ of two signed areas (Figure 5.5.2). The same computation works for every coordinate and in every dimension.

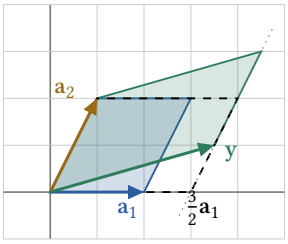


Figure 5.5.2. Here $\mathbf{a}_1 = (2, 0)$, $\mathbf{a}_2 = (1, 2)$, and $\mathbf{y} = \frac{3}{2}\mathbf{a}_1 + \frac{1}{2}\mathbf{a}_2 = \left(\frac{7}{2}, 1\right)$. The parallelogram on $\mathbf{a}_1, \mathbf{a}_2$ (blue) has signed area $\det(\mathbf{a}_1, \mathbf{a}_2) = 4$. The parallelogram on $\mathbf{y}, \mathbf{a}_2$ (green) has the same base $\mathbf{a}_2$ as the parallelogram on $\frac{3}{2}\mathbf{a}_1, \mathbf{a}_2$ (dashed), and its far side lies on the same dotted line, parallel to $\mathbf{a}_2$. So both have area $6 = \frac{3}{2} \cdot 4$, and $x_1 = \det(\mathbf{y}, \mathbf{a}_2)/\det(\mathbf{a}_1, \mathbf{a}_2) = \frac{3}{2}$.

Theorem 5.5.14 (Cramer's rule). Let $B \in F^{n\times n}$ with $\det B \neq 0$, and let $\mathbf{y} \in F^n$. Then the system $B\mathbf{x} = \mathbf{y}$ has exactly one solution $\mathbf{x} \in F^n$, and its entries are

$$x_i = \frac{\det B_i(\mathbf{y})}{\det B} \quad (i = 1, \ldots, n),$$

where $B_i(\mathbf{y})$ is the matrix obtained from B by replacing column i with $\mathbf{y}$.

Proof (by multilinearity). ROUTINE

Uses. Theorem 5.4.7 (Invertibility and the determinant), Definition 2.3.8 (Matrix–vector product), Definition 5.1.1 (Determinant function)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Existence and uniqueness come from invertibility. For the formula, write $\mathbf{y} = B\mathbf{x}$ as a combination of the columns of B.
Hint 2. Substitute $\mathbf{y} = \sum_k x_k \mathbf{b}_k$ into column i of $B_i(\mathbf{y})$ and expand by linearity in column i. Which terms survive?
Sketch. B is invertible, so $\mathbf{x} = B^{-1}\mathbf{y}$ is the only solution. Since $\mathbf{y} = \sum_k x_k \mathbf{b}_k$, linearity in column i gives $\det B_i(\mathbf{y}) = \sum_k x_k \det B_i(\mathbf{b}_k)$. For $k \neq i$, the matrix $B_i(\mathbf{b}_k)$ has $\mathbf{b}_k$ in columns i and k, so only $k = i$ survives, and $B_i(\mathbf{b}_i) = B$. So $\det B_i(\mathbf{y}) = x_i \det B$.

By Theorem 5.4.7, B is invertible. If $B\mathbf{x} = \mathbf{y}$, then $\mathbf{x} = B^{-1}B\mathbf{x} = B^{-1}\mathbf{y}$; conversely $B\big(B^{-1}\mathbf{y}\big) = \mathbf{y}$. So $\mathbf{x} = B^{-1}\mathbf{y}$ is the unique solution.

Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be the columns of B. By Definition 2.3.8, $\mathbf{y} = B\mathbf{x} = \sum_k x_k \mathbf{b}_k$. Fix i. By linearity in column i,

$$\det B_i(\mathbf{y}) = \sum_{k=1}^{n} x_k \det B_i(\mathbf{b}_k).$$

If $k \neq i$, then $B_i(\mathbf{b}_k)$ has $\mathbf{b}_k$ both in column i and in column k, so its determinant is 0 since det is alternating. If $k = i$, then $B_i(\mathbf{b}_i) = B$. So $\det B_i(\mathbf{y}) = x_i \det B$, and dividing by $\det B \neq 0$ gives the formula. □

Proof (via the adjugate). ROUTINE

Uses. Theorem 5.5.9 (The adjugate and the inverse), Lemma 5.5.8 (Replacing a row or a column), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate)
Hint 1. Write the solution as $\mathbf{x} = B^{-1}\mathbf{y}$ and use the formula for B^{-1}.
Hint 2. The i-th entry of $(\operatorname{adj} B)\mathbf{y}$ is $\sum_j c_{ji} y_j$. Compare with part (b) of Lemma 5.5.8.
Sketch. $\mathbf{x} = B^{-1}\mathbf{y} = (\det B)^{-1}(\operatorname{adj} B)\mathbf{y}$, and the i-th entry of $(\operatorname{adj} B)\mathbf{y}$ is $\sum_j c_{ji} y_j = \det B_i(\mathbf{y})$, the expansion of $\det B_i(\mathbf{y})$ along column i.

As in the first proof, the unique solution is $\mathbf{x} = B^{-1}\mathbf{y}$, which equals $(\det B)^{-1}(\operatorname{adj} B)\mathbf{y}$ by Theorem 5.5.9. For $n = 1$ this reads $x_1 = y_1/b_{11}$, which is the claim. For $n \geq 2$, the (i, j) entry of $\operatorname{adj} B$ is c_{ji}, so the i-th entry of $(\operatorname{adj} B)\mathbf{y}$ is $\sum_j y_j c_{ji}$, which is $\det B_i(\mathbf{y})$ by part (b) of Lemma 5.5.8. □

The two proofs reverse each other's logic. The second derives Cramer's rule from the inverse formula. The first needs nothing but the axioms, and running it backwards gives the inverse formula: column j of B^{-1} solves $B\mathbf{x} = \mathbf{e}_j$, so Cramer's rule computes its entries (Exercise 5.5.16). The first proof also shows more than the theorem states. Its computation never used that $\det B \neq 0$: for every $B \in F^{n\times n}$ and every solution $\mathbf{x}$ of $B\mathbf{x} = \mathbf{y}$,

$$x_i \det B = \det B_i(\mathbf{y}) \quad (i = 1, \dots, n). \tag{5.5.3}$$

So if $\det B = 0$ but $\det B_i(\mathbf{y}) \neq 0$ for some i, the system has no solution. The converse fails, as Exercise 5.5.35 shows.

Example 5.5.15 (One unknown at a time). Let B be the matrix of Example 5.5.10, with $\det B = 1$, and $\mathbf{y} = (1, 0, 4)$. To find only x_3 in the solution of $B\mathbf{x} = \mathbf{y}$, we need a single determinant:

$$x_3 = \det B_3(\mathbf{y}) = \det\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 0 \\ 2 & 1 & 4 \end{pmatrix} = 2(8-0) - 1(12-0) + 1(3-4) = 3.$$

The full solution, from $\mathbf{x} = (\operatorname{adj} B)\mathbf{y}$, is $\mathbf{x} = (-1, 0, 3)$, and indeed $B\mathbf{x} = (-2+3, -3+3, -2+6) = (1, 0, 4)$.

As a method for solving systems, Cramer's rule needs $n + 1$ determinants and is far slower than elimination. Its value is theoretical: it expresses each x_i as a ratio of two polynomials in the entries of B and $\mathbf{y}$, linear

in $\mathbf{y}$, so that, for instance, an integer system with $\det B = \pm 1$ has an integer solution. It also explains the geometry of Intuition (Coefficients as ratios of areas).

Exercise 5.5.16. ROUTINE Let $B \in F^{n\times n}$ with $\det B \neq 0$ and $n \geq 2$. Use Cramer's rule, applied to the systems $B\mathbf{x} = \mathbf{e}_j$, to show that the (i,j) entry of B^{-1} is $c_{ji}/\det B$. This proves the second half of Theorem 5.5.9 again.

Solution on page 1427

5.5.4 Determinants of structured matrices

Cofactor expansion reduces the size of a determinant by one. When the smaller determinants have the same shape as the original one, this sets up an induction.

Technique 5.5.17 (Induction on matrix size). Suppose a family of matrices $M_1, M_2, M_3, \ldots$, with $M_n \in F^{n\times n}$, has the property that expanding M_n along a well-chosen row or column produces minors that are earlier members of the family, or are triangular, or are otherwise known. Then the expansion gives a recurrence for $d_n = \det M_n$. Expand along the row or column with the fewest nonzero entries, identify each minor that appears, and write down the recurrence. Then either solve it, or compute d_1, d_2, d_3, guess a closed form, and prove it by induction on n.

Reach for it when a matrix has a pattern that repeats along the diagonal, so that deleting the last (or first) row and column gives the same kind of matrix. The canonical example is Proposition 5.5.18. Proposition 5.5.20 and the first proof of Theorem 5.5.21 induct in the same way, and Exercise 5.5.34 reproves Proposition 5.3.15.

Used in. Proposition 5.5.18, Proposition 5.5.20, Theorem 5.5.21, Exercise 5.5.34, Exercise 5.5.45, Exercise 5.5.53, Exercise 5.6.7, Problem 5.14, Theorem 8.3.23, Lemma 8.4.24.

Proposition 5.5.18 (Tridiagonal determinants). Let $a_1, \ldots, a_n$, $b_1, \ldots, b_{n-1}$, and $c_1, \ldots, c_{n-1}$ be elements of F. For $1 \leq m \leq n$ let

$$T_m = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ c_1 & a_2 & b_2 & \ddots & \vdots \\ 0 & c_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & a_{m-1} & b_{m-1} \\ 0 & \cdots & 0 & c_{m-1} & a_m \end{pmatrix} \in F^{m\times m},$$

with b_k in position $(k, k+1)$, c_k in position $(k+1, k)$, and zeros away from the three middle diagonals. Let $d_m = \det T_m$, and put $d_0 = 1$. Then $d_1 = a_1$, and

$$d_m = a_m d_{m-1} - b_{m-1}c_{m-1}d_{m-2} \quad (2 \leq m \leq n).$$

Proof. ROUTINE

Uses. Theorem 5.5.3 (Cofactor expansion), Lemma 5.5.2 (A row or column with one nonzero entry)
Techniques. Induction on matrix size (Technique 5.5.17)
Hint 1. Expand $\det T_m$ along its last row. It has at most two nonzero entries.
Hint 2. One minor is T_{m-1}. In the other, the last column has a single nonzero entry; apply Lemma 5.5.2 once more.
Sketch. The last row of T_m is $(0, \ldots, 0, c_{m-1}, a_m)$. Deleting row m and column m leaves T_{m-1}, with sign $+$. Deleting row m and column $m-1$ leaves a matrix whose last column is $(0, \ldots, 0, b_{m-1})$, with sign $-$; expanding along that column leaves T_{m-2}. So $d_m = a_m d_{m-1} - c_{m-1}b_{m-1}d_{m-2}$.

$T_1 = (a_1)$, so $d_1 = a_1$. Let $2 \leq m \leq n$. The last row of T_m is $(0, \ldots, 0, c_{m-1}, a_m)$, so expanding along it (Theorem 5.5.3),

$$d_m = (-1)^{m+m} a_m \det T_{\widehat{mm}} + (-1)^{m+m-1} c_{m-1} \det N = a_m d_{m-1} - c_{m-1} \det N,$$

where $T_{m\widehat{m}m} = T_{m-1}$ and $N = T_{m\widehat{mm-1}}$. The matrix N consists of rows $1, \dots, m-1$ and columns $1, \dots, m-2, m$ of T_m. Its last column is column m of T_m without the last entry, namely $(0, \dots, 0, b_{m-1})$. If $m = 2$, then $N = (b_1)$ and $\det N = b_1 = b_1 d_0$. If $m \geq 3$, the only nonzero entry of the last column of N is b_{m-1}, in position $(m-1, m-1)$, and deleting row $m-1$ and column $m-1$ of N leaves rows and columns $1, \dots, m-2$ of T_m, which form T_{m-2}. By Lemma 5.5.2, $\det N = b_{m-1} d_{m-2}$. In both cases $d_m = a_m d_{m-1} - b_{m-1} c_{m-1} d_{m-2}$. □

Only the products $b_k c_k$ enter, so the entries above and below the diagonal can be traded against each other without changing any d_m.

Example 5.5.19 (Two tridiagonal families).

(a) Let $a_k = 2$ and $b_k = c_k = -1$ for all k. This is the matrix of second differences, which approximates $-f''$ when f is sampled at equally spaced points. The recurrence reads $d_m = 2d_{m-1} - d_{m-2}$ with $d_0 = 1$ and $d_1 = 2$, so $d_2 = 3$, $d_3 = 4$, and by induction $d_m = m + 1$: if $d_{m-1} = m$ and $d_{m-2} = m - 1$, then $d_m = 2m - (m-1) = m + 1$. So the $n \times n$ matrix is invertible over $\mathbb{Q}$, and over $\mathbb{F}_p$ exactly when p does not divide $n + 1$ (Proposition 5.3.20).

(b) Over $\mathbb{C}$, let $a_k = 1$ and $b_k = c_k = i$. Then $-b_k c_k = -i^2 = 1$, so $d_m = d_{m-1} + d_{m-2}$ with $d_0 = d_1 = 1$. The determinants are the Fibonacci numbers $1, 1, 2, 3, 5, 8, 13, \dots$, although every off-diagonal entry is imaginary.

The next family has a single row with more than two nonzero entries. Its determinant is a polynomial chosen in advance.

Proposition 5.5.20 (Companion matrices). Let $n \geq 1$, let $a_0, \dots, a_{n-1} \in F$, let $p = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n \in F[x]$, and let

$$C = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in F^{n \times n},$$

whose column k is $\mathbf{e}_{k+1}$ for $k < n$, and whose last column is $-(a_0, \dots, a_{n-1})$. Then

$$\det(tI_n - C) = p(t) = t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0 \quad \text{for every } t \in F.$$

In particular, $\det C = (-1)^n a_0$.

Written out,

$$tI_n - C = \begin{pmatrix} t & 0 & \cdots & 0 & a_0 \\ -1 & t & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & t & a_{n-2} \\ 0 & \cdots & 0 & -1 & t + a_{n-1} \end{pmatrix}.$$

For $n = 1$ this is $(t + a_0)$, and for $n = 2$ it is $\begin{pmatrix} t & a_0 \\ -1 & t + a_1 \end{pmatrix}$, with determinant $t^2 + a_1 t + a_0$, as claimed. We give two proofs. The first is an induction in which the polynomial is built up by Horner's scheme $p(t) = a_0 + t(a_1 + t(a_2 + \cdots))$; the second performs the same scheme with row operations.

Proof (by expansion and induction). STANDARD

Uses. Theorem 5.5.3 (Cofactor expansion), Proposition 5.3.15 (Triangular matrices), Definition 1.2.17 (Evaluation), Exercise 5.1.5, induction on n

Techniques. Induction on matrix size (Technique 5.5.17)

Hint 1. Expand $\det(tI_n - C)$ along the first row, which has at most two nonzero entries.

Hint 2. Deleting row 1 and column 1 gives a matrix of the same kind, for the polynomial $a_1 + a_2x + \cdots + x^{n-1}$. Deleting row 1 and column n gives a triangular matrix.
Hint 3. You should find $\det(tI_n - C) = tq(t) + a_0$, where $q = a_1 + a_2x + \cdots + a_{n-1}x^{n-2} + x^{n-1}$.
Sketch. Induct on n; for $n = 1$ both sides are $t + a_0$. Let $M = tI_n - C$. Its first row is $(t, 0, \ldots, 0, a_0)$. The minor $M_{\widehat{11}}$ is $tI_{n-1} - C'$ for the companion matrix C' of $q = a_1 + a_2x + \cdots + x^{n-1}$, so its determinant is $q(t)$. The minor $M_{\widehat{1n}}$ is upper triangular with diagonal entries -1, so its determinant is $(-1)^{n-1}$. Hence $\det M = tq(t) + (-1)^{1+n}(-1)^{n-1}a_0 = tq(t) + a_0 = p(t)$.

We induct on n. For $n = 1$, $C = (-a_0)$ and $\det(tI_1 - C) = t + a_0 = p(t)$.

Let $n \geq 2$, assume the claim for $n - 1$, and let $M = tI_n - C$. Row k of M has -1 in column $k - 1$ (if $k \geq 2$), t in column k (if $k < n$), and a_{k-1} in column n, plus an extra t in position (n, n); all other entries are 0. The first row is $(t, 0, \ldots, 0, a_0)$, so expanding along it,

$$\det M = t \det M_{\widehat{11}} + (-1)^{1+n}a_0 \det M_{\widehat{1n}}.$$

The matrix $M_{\widehat{11}}$ consists of rows and columns $2, \ldots, n$ of M. Renumbering them as $1, \ldots, n - 1$, its row k has -1 in column $k - 1$ (if $k \geq 2$), t in column k (if $k < n - 1$), and a_k in its last column, plus an extra t in the last position. So $M_{\widehat{11}} = tI_{n-1} - C'$, where C' is the matrix of the proposition for the polynomial $q = a_1 + a_2x + \cdots + a_{n-1}x^{n-2} + x^{n-1}$. By the induction hypothesis, $\det M_{\widehat{11}} = q(t)$.

The matrix $M_{\widehat{1n}}$ consists of rows $2, \ldots, n$ and columns $1, \ldots, n - 1$ of M. Renumbered, its row k has -1 in column k and t in column $k + 1$ (if $k + 1 \leq n - 1$), and zeros elsewhere. It is upper triangular with every diagonal entry -1, so $\det M_{\widehat{1n}} = (-1)^{n-1}$ by Proposition 5.3.15.

Therefore, by Definition 1.2.17,

$$\begin{aligned}\det M &= tq(t) + (-1)^{1+n}(-1)^{n-1}a_0 = t\big(a_1 + a_2t + \cdots + a_{n-1}t^{n-2} + t^{n-1}\big) + a_0 \\ &= a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + t^n = p(t),\end{aligned}$$

since $(-1)^{1+n}(-1)^{n-1} = (-1)^{2n} = 1$. Finally, with $t = 0$, $\det(-C) = p(0) = a_0$, and $\det(-C) = (-1)^n \det C$ by Exercise 5.1.5, so $\det C = (-1)^n a_0$. □

Proof (by row operations). STANDARD

Uses. Corollary 5.4.15 (Rows), Lemma 5.5.2 (A row or column with one nonzero entry), Proposition 5.3.15 (Triangular matrices), Definition 1.2.17 (Evaluation)
Hint 1. Use the -1 below each t to clear the t: add t times row $k + 1$ to row k. In which order should the rows be processed?
Hint 2. Work from the bottom up. Track the last entry of each row: if q_{k+1} is the last entry of the new row $k + 1$, the new last entry of row k is $a_{k-1} + tq_{k+1}$.
Hint 3. At the end, row 1 is $(0, \ldots, 0, p(t))$. Expand along it.
Sketch. For $k = n - 1, n - 2, \ldots, 1$, add t times the current row $k + 1$ to row k. This clears the t in position (k, k) and makes the last entry of row k equal to $q_k = a_{k-1} + tq_{k+1}$, where $q_n = t + a_{n-1}$; so $q_1 = p(t)$ by Horner's scheme. The determinant is unchanged, row 1 becomes $(0, \ldots, 0, p(t))$, and the minor of its last entry is $-I_{n-1}$. So $\det(tI_n - C) = (-1)^{1+n}p(t)(-1)^{n-1} = p(t)$.

For $n = 1$ both sides are $t + a_0$, so let $n \geq 2$ and $M = tI_n - C$. Put $q_n = t + a_{n-1}$ and, for $k = n - 1, \ldots, 1$, $q_k = a_{k-1} + tq_{k+1}$. Then

$$q_1 = a_0 + t(a_1 + t(a_2 + \cdots + t(a_{n-1} + t))) = p(t)$$

by expanding the nested products (Horner's scheme).

For $k = n - 1, n - 2, \ldots, 1$ in this order, add t times the current row $k + 1$ to row k. We claim that afterwards, for each $k \geq 2$, row k has -1 in column $k - 1$, q_k in column n, and zeros elsewhere, and row 1 is $(0, \ldots, 0, q_1)$. Row n is never changed, and it has -1 in column $n - 1$ and $q_n = t + a_{n-1}$ in column n. Suppose row $k + 1$ already has its final form, with -1 in column k and q_{k+1} in column n. Row k of M has -1 in column $k - 1$ (if $k \geq 2$), t in column k, and a_{k-1} in column n. Adding t times row $k + 1$ changes only columns k and n: the entry in column k becomes $t + t(-1) = 0$, and the entry in column n becomes $a_{k-1} + tq_{k+1} = q_k$. This

proves the claim. Each operation adds a multiple of one row to another, so the determinant is unchanged (Corollary 5.4.15).

In the final matrix, row 1 has the single nonzero entry $q_1 = p(t)$, in column n. Deleting row 1 and column n leaves the matrix whose row $k-1$ is $-\mathbf{e}_{k-1}^{\mathsf{T}}$ for $k = 2, \ldots, n$, that is, $-I_{n-1}$, with determinant $(-1)^{n-1}$ (Proposition 5.3.15). By Lemma 5.5.2,

$$\det M = (-1)^{1+n} p(t) (-1)^{n-1} = p(t),$$

as desired. The value of $\det C$ follows as in the first proof. □

So every monic polynomial of degree n arises as $t \mapsto \det(tI_n - C)$ for some $C \in F^{n \times n}$. In Section 7.3, $\det(tI_n - A)$ becomes the characteristic polynomial of A, and C (or its transpose, which has the same determinant by Theorem 5.4.14; conventions differ) is called the companion matrix of p. There the identity of the proposition is upgraded to an identity in $F[x]$.

5.5.5 The Vandermonde determinant

Recall from Definition 3.4.22 that for $c_0, \ldots, c_n \in F$ the Vandermonde matrix $V(c_0, \ldots, c_n) \in F^{(n+1)\times(n+1)}$ has row $i+1$ equal to $\left(1, c_i, c_i^2, \ldots, c_i^n\right)$. In Section 3.4 we showed that it is invertible when the c_i are distinct (Theorem 3.4.24), by an argument about roots of polynomials. Now we compute its determinant.

Theorem 5.5.21 (The Vandermonde determinant). Let $n \geq 0$ and $c_0, c_1, \ldots, c_n \in F$. Then

$$\det V(c_0, \ldots, c_n) = \det \begin{pmatrix} 1 & c_0 & c_0^2 & \cdots & c_0^n \\ 1 & c_1 & c_1^2 & \cdots & c_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_n & c_n^2 & \cdots & c_n^n \end{pmatrix} = \prod_{0 \leq i < j \leq n} (c_j - c_i).$$

In particular, $V(c_0, \ldots, c_n)$ is invertible if and only if $c_0, \ldots, c_n$ are distinct.

For $n = 0$ both sides are 1 (the product is empty). For $n = 1$, $\det\begin{pmatrix} 1 & c_0 \\ 1 & c_1 \end{pmatrix} = c_1 - c_0$, and for $n = 2$ the product is $(c_1 - c_0)(c_2 - c_0)(c_2 - c_1)$, which matches the rule of Sarrus after expanding. The last sentence follows from the formula: by Proposition 1.1.17 the product is nonzero exactly when every factor is, and then Theorem 5.4.7 applies.

The first proof creates the factors $c_j - c_0$ with column operations. The following technique describes the move.

Technique 5.5.22 (Expose a factor with row and column operations). When the entries of a matrix are polynomial expressions in some parameters and the determinant is expected to have a factor such as $c_1 - c_0$, look for replacements that make this factor appear in every entry of one row (or column). Replacements do not change the determinant (Corollary 5.4.15, Lemma 5.1.8), and linearity in that row then pulls the factor out. Choose the replacements so that they also create zeros, so that a cofactor expansion reduces the size at the same step. Repeating this, usually by induction on the size, peels off the factors one at a time.

Reach for it when consecutive columns (or rows) are related by a common multiplier, as the powers c^k and c^{k+1} are related by c, or when all rows have a common sum. The canonical example is the first proof of Theorem 5.5.21. Example 5.4.24 and Exercise 5.4.44 used the same idea with row sums.

Used in. Theorem 5.5.21, Exercise 5.5.42, Exercise 5.5.53, Exercise 5.6.6, Exercise 5.6.7, Problem 5.10, Problem 5.19.

Proof (by column operations). STANDARD

Uses. Lemma 5.1.8 (Column replacement), Lemma 5.5.2 (A row or column with one nonzero entry), Corollary 5.4.15 (Rows), induction on n
Techniques. Expose a factor with row and column operations (Technique 5.5.22), Induction on matrix size (Technique 5.5.17)
Hint 1. Use the first node c_0: subtract c_0 times each column from the column to its right. In which order must this be done so that each step uses an unchanged column?
Hint 2. Process the columns from right to left. Row 1 becomes $(1, 0, \dots, 0)$, and in every other row $i+1$, each entry except the first acquires the factor $c_i - c_0$.
Hint 3. After expanding along row 1 and pulling out the factors, what remains is a smaller Vandermonde matrix.
Sketch. For $m = n+1, n, \dots, 2$, subtract c_0 times column $m-1$ from column m. Row $i+1$ becomes $\big(1, (c_i - c_0), (c_i - c_0)c_i, \dots, (c_i - c_0)c_i^{n-1}\big)$, and row 1 becomes $(1, 0, \dots, 0)$. Expanding along row 1 and pulling $c_i - c_0$ out of row i of the minor gives $\det V(c_0, \dots, c_n) = \prod_{i=1}^{n}(c_i - c_0) \det V(c_1, \dots, c_n)$, and induction finishes.

We induct on n. For $n = 0$, both sides equal 1. Let $n \geq 1$, assume the formula for n nodes, and let $V = V(c_0, \dots, c_n)$. Number the columns $1, \dots, n+1$, so that the entry of column m in row $i+1$ is c_i^{m-1}.

For $m = n+1, n, \dots, 2$, in this order, subtract c_0 times column $m-1$ from column m. When column m is changed, column $m-1$ has not been changed yet, so the new entry of column m in row $i+1$ is

$$c_i^{m-1} - c_0 c_i^{m-2} = (c_i - c_0)c_i^{m-2}.$$

Each step adds a multiple of one column to another, so the determinant does not change (Lemma 5.1.8). The new matrix V' has first row $(1, 0, \dots, 0)$, since $c_0 - c_0 = 0$, and for $i = 1, \dots, n$ its row $i+1$ is

$$\big(1, (c_i - c_0), (c_i - c_0)c_i, \dots, (c_i - c_0)c_i^{n-1}\big).$$

By Lemma 5.5.2 applied to row 1, $\det V = \det V'_{\widehat{11}}$. Row i of $V'_{\widehat{11}}$ is $(c_i - c_0)\big(1, c_i, \dots, c_i^{n-1}\big)$, so pulling the factor $c_i - c_0$ out of each row (Corollary 5.4.15) gives

$$\det V = (c_1 - c_0)(c_2 - c_0)\cdots(c_n - c_0) \det V(c_1, \dots, c_n).$$

By the induction hypothesis, $\det V(c_1, \dots, c_n) = \prod_{1 \leq i < j \leq n}\big(c_j - c_i\big)$. The factors $c_j - c_0$ are the factors of $\prod_{0 \leq i < j \leq n}\big(c_j - c_i\big)$ with $i = 0$, so $\det V$ is the full product, as desired. □

Intuition (Where the factors come from). Regard the last node c_n as a variable t, and the other nodes as fixed. Expanding $\det V(c_0, \dots, c_{n-1}, t)$ along the last row $\big(1, t, t^2, \dots, t^n\big)$ shows that it is a polynomial in t of degree at most n. If $t = c_i$ for some $i < n$, then two rows are equal and the determinant is 0. So when the nodes $c_0, \dots, c_{n-1}$ are distinct, the polynomial has these n roots, and a polynomial of degree at most n with n prescribed roots is pinned down by one more number, its coefficient of t^n. That coefficient is the cofactor of the entry t^n, a smaller Vandermonde determinant. The first proof found the factors by computing; this argument predicts them from the places where the matrix collapses.

To make this precise, we need the fact that a polynomial is determined by its roots and its leading coefficient. Polynomials cannot be multiplied until Section 7.1, so we state it for values.

Lemma 5.5.23 (A polynomial with prescribed roots). Let $n \geq 0$, let $P \in F[x]_{\leq n}$, and let a be its coefficient of x^n. If $c_1, \dots, c_n \in F$ are distinct and $P(c_1) = \cdots = P(c_n) = 0$, then

$$P(t) = a(t - c_1)(t - c_2)\cdots(t - c_n) \quad \text{for every } t \in F.$$

Proof. STANDARD

Uses. Lemma 3.4.23 (Too many roots), Proposition 1.2.18 (Degree and evaluation versus the operations), Definition 1.2.17 (Evaluation), induction on n
Hint 1. First show that $t \mapsto a(t - c_1)\cdots(t - c_n)$ is the evaluation of a polynomial $Q \in F[x]_{\leq n}$ whose coefficient of x^n is a.
Hint 2. Then $P - Q$ has degree at most $n - 1$. How many roots does it have?

Sketch. Multiplying out one factor at a time shows $(t-c_1)\cdots(t-c_n) = t^n + e_{n-1}t^{n-1} + \cdots + e_0$ for scalars e_k that do not depend on t. So $a\prod(t-c_k) = Q(t)$ for some $Q \in F[x]_{\le n}$ with coefficient a of x^n. Then $P - Q \in F[x]_{\le n-1}$ vanishes at the n distinct points c_k, so $P = Q$ by Lemma 3.4.23.

For $n = 0$, P is the constant a and the product is empty, so there is nothing to prove. Let $n \ge 1$.

First, there are $e_0, \ldots, e_{n-1} \in F$, depending only on $c_1, \ldots, c_n$, such that

$$(t-c_1)\cdots(t-c_n) = t^n + e_{n-1}t^{n-1} + \cdots + e_1 t + e_0 \quad \text{for every } t \in F.$$

We show this by induction on the number m of factors. For $m = 1$, take $e_0 = -c_1$. If $(t-c_1)\cdots(t-c_m) = t^m + \sum_{k<m} e_k t^k$, then multiplying by $t - c_{m+1}$ and distributing gives

$$t^{m+1} + \sum_{k<m} e_k t^{k+1} - c_{m+1}t^m - \sum_{k<m} c_{m+1}e_k t^k,$$

which is t^{m+1} plus a combination of $1, t, \ldots, t^m$ with coefficients that do not depend on t.

Let $Q = ae_0 + ae_1 x + \cdots + ae_{n-1}x^{n-1} + ax^n \in F[x]_{\le n}$. Then $Q(t) = a(t-c_1)\cdots(t-c_n)$ for every t. The coefficients of x^n in P and Q are both a, so $R = P - Q$ lies in $F[x]_{\le n-1}$, and for each k, $R(c_k) = P(c_k) - Q(c_k) = 0 - 0 = 0$ (Proposition 1.2.18), because the product $Q(c_k)$ contains the factor $c_k - c_k = 0$. So R has degree at most $n-1$ and vanishes at the n distinct points $c_1, \ldots, c_n$. By Lemma 3.4.23 (with $n-1$ in place of n), $R = \mathbf{0}$. Hence $P = Q$, and $P(t) = Q(t) = a(t-c_1)\cdots(t-c_n)$ for every t. □

Once polynomials can be multiplied, the lemma says $P = a(x-c_1)\cdots(x-c_n)$ in $F[x]$; it is the factor theorem of Section 7.1 (compare Proposition 7.1.22).

Technique 5.5.24 (Polynomial identity arguments). To prove that two quantities depending on a scalar t are equal, show that both are values $P(t)$ and $Q(t)$ of polynomials $P, Q \in F[x]_{\le n}$, and then show $P = Q$ by counting. Two polynomials of degree at most n that agree at $n+1$ distinct points are equal, since their difference has $n+1$ roots (Lemma 3.4.23). When zeros are easier to find than values, use Lemma 5.5.23 instead: a polynomial of degree at most n with n known distinct roots is determined by its coefficient of x^n. Then the proof splits into independent checks: a bound on the degree, the roots, and one coefficient.

Determinants are a rich source of such polynomials. If a parameter t occurs in only one row, the expansion along that row (Theorem 5.5.3) shows that the determinant is a polynomial in t; if the top power of t occurs in only one entry of that row, as in $(1, t, \ldots, t^n)$, its coefficient is a single cofactor. More generally, see Technique 5.3.18. Roots come from coincidences: a value of t that makes two rows equal, or the rows dependent, is a root.

Reach for it when the expected answer is a product of linear factors, or when an identity is easy to check for many special values of a parameter. Two pitfalls: forgetting the leading coefficient (roots determine a polynomial only up to a scalar factor; see Exercise 5.5.42), and counting roots that are not distinct. Over an infinite field, an identity between polynomials in t that holds for all but finitely many t holds for all t. This extends identities from invertible matrices to all matrices by way of $A + tI_n$ (Exercise 5.5.41). The canonical example is the second proof of Theorem 5.5.21. The technique returns in Section 7.3.

Used in. Theorem 5.5.21, Exercise 5.5.40, Exercise 5.5.41, Exercise 5.5.42, Exercise 5.5.45, Exercise 5.6.9, Exercise 5.6.10, Problem 5.18, Proposition 7.3.11, Proposition 7.3.22, Exercise 7.3.48, Exercise 7.3.49, Theorem 7.5.22, Lemma 7.6.21, Exercise 7.7.14, Proposition B.1.28, Exercise B.1.45.

Proof of Theorem 5.5.21 (by a polynomial identity). STANDARD

Uses. Theorem 5.5.3 (Cofactor expansion), Lemma 5.5.23 (A polynomial with prescribed roots), Corollary 5.4.15 (Rows), induction on n

Techniques. Polynomial identity arguments (Technique 5.5.24)

Hint 1. Induct on n. Fix $c_0, \ldots, c_{n-1}$ and consider $f(t) = \det V(c_0, \ldots, c_{n-1}, t)$ as a function of t.

Hint 2. Expand along the last row to see that f is a polynomial of degree at most n. What is its coefficient of t^n? Where does f vanish?

Hint 3. If $c_0, \dots, c_{n-1}$ are distinct, Lemma 5.5.23 applies. If not, both sides are 0.
Sketch. Let $f(t) = \det V(c_0, \dots, c_{n-1}, t)$. Expanding along the last row, f is a polynomial of degree at most n whose coefficient of t^n is the cofactor of the entry t^n, namely $D = \det V(c_0, \dots, c_{n-1})$. It vanishes at $c_0, \dots, c_{n-1}$ (equal rows). If these are distinct, Lemma 5.5.23 gives $f(t) = D\prod_{i<n}(t - c_i)$; put $t = c_n$ and use induction for D. If two of them coincide, both sides are 0.

We induct on n; for $n = 0$ both sides are 1. Let $n \geq 1$, and assume the formula for n nodes. Fix $c_0, \dots, c_{n-1}$, and for $t \in F$ let $f(t) = \det V(c_0, \dots, c_{n-1}, t)$.

f is a polynomial. The last row of $V(c_0, \dots, c_{n-1}, t)$ is $\left(1, t, t^2, \dots, t^n\right)$, with t^k in column $k+1$. Expanding along it (Theorem 5.5.3),

$$f(t) = \sum_{k=0}^{n} (-1)^{(n+1)+(k+1)} t^k m_k,$$

where m_k is the determinant of the matrix obtained by deleting the last row and column $k+1$. These matrices do not involve the last row, so the m_k do not depend on t. Hence $f(t) = P(t)$ for the polynomial $P = \sum_{k=0}^{n} (-1)^{n+k} m_k x^k \in F[x]_{\leq n}$. Its coefficient of x^n is m_n, and deleting the last row and the last column of $V(c_0, \dots, c_{n-1}, t)$ leaves $V(c_0, \dots, c_{n-1})$, so $m_n = D$, where $D = \det V(c_0, \dots, c_{n-1})$.

Roots. For $i < n$, the matrix $V(c_0, \dots, c_{n-1}, c_i)$ has rows $i+1$ and $n+1$ equal, so $P(c_i) = f(c_i) = 0$ (Corollary 5.4.15).

Conclusion. Suppose first that $c_0, \dots, c_{n-1}$ are distinct. By Lemma 5.5.23, $f(t) = D(t - c_0)\cdots(t - c_{n-1})$ for every $t \in F$. Taking $t = c_n$ and using the induction hypothesis for D,

$$\det V(c_0, \dots, c_n) = \prod_{0 \leq i < j \leq n-1} (c_j - c_i) \cdot \prod_{i=0}^{n-1} (c_n - c_i) = \prod_{0 \leq i < j \leq n} (c_j - c_i).$$

If instead $c_i = c_k$ for some $i < k < n$, then $V(c_0, \dots, c_n)$ has two equal rows, so its determinant is 0, and the product is 0 as well, since it contains the factor $c_k - c_i = 0$. □

Intuition (Two proofs compared). The first proof is a computation organized by induction: column operations expose all factors that involve c_0 at once, and what is left is a smaller Vandermonde matrix. It is the proof to imitate for determinants with a similar pattern of powers. The second proof hardly computes. It identifies the determinant as a polynomial in one node, locates its roots by asking when the matrix collapses, and pins down the one remaining constant with a cofactor. It explains why the answer is a product of differences: the factor $c_j - c_i$ is there because the matrix becomes singular when $c_j = c_i$. The second proof needs the nodes $c_0, \dots, c_{n-1}$ to be distinct and treats the degenerate case separately, while the first does not split into cases at all.

Remark 5.5.25 (Back to permutations and interpolation). The polynomial $\Delta(\mathbf{x}) = \prod_{1 \leq i < j \leq n}(x_j - x_i)$ of Approach 1: a product of differences is $\det V(x_1, \dots, x_n)$. Rearranging the entries of $\mathbf{x}$ by σ rearranges the rows of V: row k of $V\left(x_{\sigma(1)}, \dots, x_{\sigma(n)}\right)$ is row $\sigma(k)$ of $V(x_1, \dots, x_n)$, so $V(\mathbf{x}^\sigma) = E_\sigma^{\mathsf{T}} V(\mathbf{x})$, and $\Delta(\mathbf{x}^\sigma) = \operatorname{sgn}(\sigma)\Delta(\mathbf{x})$. For a transposition this is Lemma 5.2.20. This is not a new proof of Theorem 5.2.18, since the existence of det relied on it, but it explains why Δ was the right invariant.

The theorem also completes the second proof of invertibility promised in Section 3.4. For distinct nodes, $\det V(c_0, \dots, c_n) \neq 0$, so V is invertible and the interpolation problem has exactly one solution.

Exercise 5.5.26. ROUTINE

(a) Show that $\det V(1, 2, \dots, n) = 1! \cdot 2! \cdots (n-1)!$, and evaluate it for $n = 4$.

(b) Compute $\det V(0, 1, 2, 3, 4)$ in $\mathbb{F}_5$. Deduce that every function $g : \mathbb{F}_5 \to \mathbb{F}_5$ is of the form $t \mapsto P(t)$ for exactly one $P \in \mathbb{F}_5[x]_{\leq 4}$.

Solution on page 1427

5.5.6 Rank-one updates BEYOND 21-242

Not covered in lecture. The matrix determinant lemma is a standard application of the adjugate in statistics and numerical linear algebra.

Remark 5.5.11 computed $\det(B + tU_{ij})$, where $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ has rank one. The same computation works for every matrix of rank one, that is, for every $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}, \mathbf{v} \in F^n$.

Theorem 5.5.27 (Matrix determinant lemma). Let $A \in F^{n\times n}$ and $\mathbf{u}, \mathbf{v} \in F^n$. Then

$$\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det A + \mathbf{v}^{\mathsf{T}}(\operatorname{adj} A)\mathbf{u}.$$

In particular, if A is invertible, then $\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det A\left(1 + \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u}\right)$.

Proof (by multilinearity). STANDARD

Uses. Definition 2.4.1 (Matrix product), Definition 5.1.1 (Determinant function), Lemma 5.5.8 (Replacing a row or a column), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Theorem 5.5.9 (The adjugate and the inverse)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u}$. Expand $\det(\mathbf{a}_1 + v_1\mathbf{u}, \ldots, \mathbf{a}_n + v_n\mathbf{u})$ by multilinearity. Which terms vanish?
Hint 2. Every term that uses $\mathbf{u}$ in two columns vanishes. The terms that use $\mathbf{u}$ in exactly one column j are $v_j \det A_j(\mathbf{u})$; evaluate them with Lemma 5.5.8.
Sketch. Expanding every column $\mathbf{a}_j + v_j\mathbf{u}$ gives a sum over the sets S of columns in which $\mathbf{u}$ is chosen. Sets with two or more elements give a matrix with two columns equal to $\mathbf{u}$. So $\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det A + \sum_j v_j \det A_j(\mathbf{u})$, and $\det A_j(\mathbf{u}) = \sum_i u_i c_{ij}$. The double sum $\sum_{i,j} v_j c_{ij} u_i$ is $\mathbf{v}^{\mathsf{T}} C_A^{\mathsf{T}} \mathbf{u}$.

For $n = 1$ both sides are $a + vu$, since $\operatorname{adj}(a) = (1)$. Let $n \geq 2$. Column j of $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $v_j\mathbf{u}$ (Definition 2.4.1), so column j of $A + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $\mathbf{a}_j + v_j\mathbf{u}$. Expanding by linearity in each column, $\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)$ is the sum, over all subsets $S \subseteq [n]$, of $\left(\prod_{j\in S} v_j\right) \det A_S$, where A_S has column $\mathbf{u}$ in the positions $j \in S$ and column $\mathbf{a}_j$ in the other positions. If S has two or more elements, then A_S has two equal columns and $\det A_S = 0$. The set $S = \emptyset$ contributes $\det A$, and $S = \{j\}$ contributes $v_j \det A_j(\mathbf{u})$, where $A_j(\mathbf{u})$ is A with column j replaced by $\mathbf{u}$. By part (b) of Lemma 5.5.8, $\det A_j(\mathbf{u}) = \sum_i u_i c_{ij}$, where the c_{ij} are the cofactors of A. Hence

$$\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det A + \sum_{j=1}^{n}\sum_{i=1}^{n} v_j c_{ij} u_i = \det A + \mathbf{v}^{\mathsf{T}} C_A^{\mathsf{T}} \mathbf{u},$$

since the (j, i) entry of $C_A^{\mathsf{T}} = \operatorname{adj} A$ is c_{ij}. If A is invertible, then $\operatorname{adj} A = (\det A)A^{-1}$ by Theorem 5.5.9, and the right side is $\det A + \det A \cdot \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u}$. □

Proof (by block matrices, for invertible A). STANDARD

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.10 (Multiplicativity)
Techniques. Block elimination (Technique 5.4.28)
Hint 1. Consider the $(n + 1) \times (n + 1)$ matrix $M = \begin{pmatrix} A & -\mathbf{u} \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix}$, and compute $\det M$ in two ways.
Hint 2. Clear $\mathbf{v}^{\mathsf{T}}$ using the entry 1, or clear it using A. Each is a multiplication by a block triangular matrix with determinant 1.
Sketch. With $M = \begin{pmatrix} A & -\mathbf{u} \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix}$, block multiplication gives $\begin{pmatrix} I_n & \mathbf{u} \\ 0 & 1 \end{pmatrix} M = \begin{pmatrix} A+\mathbf{u}\mathbf{v}^{\mathsf{T}} & 0 \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix}$ and $\begin{pmatrix} I_n & 0 \\ -\mathbf{v}^{\mathsf{T}}A^{-1} & 1 \end{pmatrix} M = \begin{pmatrix} A & -\mathbf{u} \\ 0 & 1+\mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u} \end{pmatrix}$. Both multipliers have determinant 1, so $\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det M = \det A\left(1 + \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u}\right)$.

Let $M = \begin{pmatrix} A & -\mathbf{u} \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix} \in F^{(n+1)\times(n+1)}$. By Theorem 2.4.34,

$$\begin{pmatrix} I_n & \mathbf{u} \\ 0 & 1 \end{pmatrix} M = \begin{pmatrix} A + \mathbf{u}\mathbf{v}^{\mathsf{T}} & -\mathbf{u} + \mathbf{u} \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix} = \begin{pmatrix} A + \mathbf{u}\mathbf{v}^{\mathsf{T}} & 0 \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} I_n & 0 \\ -\mathbf{v}^{\mathsf{T}}A^{-1} & 1 \end{pmatrix} M = \begin{pmatrix} A & -\mathbf{u} \\ -\mathbf{v}^{\mathsf{T}} + \mathbf{v}^{\mathsf{T}} & \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u} + 1 \end{pmatrix} = \begin{pmatrix} A & -\mathbf{u} \\ 0 & 1 + \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u} \end{pmatrix}.$$

The two multipliers are block triangular with diagonal blocks I_n and (1), so they have determinant 1 by Theorem 5.4.26. The right sides are block triangular as well. Taking determinants with Theorem 5.4.10 and Theorem 5.4.26,

$$\det\left(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det M = \det A\left(1 + \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u}\right),$$

as desired. □

The first proof is shorter, works for every A, and shows that $\det\left(A + t\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det A + t\mathbf{v}^{\mathsf{T}}(\operatorname{adj} A)\mathbf{u}$ is affine in t. The second is an instance of Technique 5.4.28 and needs A^{-1}. For example, let $J \in F^{n\times n}$ be the matrix of all 1s, so $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ with $\mathbf{1} = (1, \ldots, 1)$. For $a, b \in F$, since $\operatorname{adj}(aI_n) = a^{n-1}I_n$,

$$\det(aI_n + bJ) = a^n + b\mathbf{1}^{\mathsf{T}}a^{n-1}I_n\mathbf{1} = a^{n-1}(a + nb).$$

With $a = x - 1$ and $b = 1$ this is Example 5.4.24 again.

Remark 5.5.28 (Derivatives of determinants). Let $A(s)$ be an $n \times n$ matrix whose entries are differentiable functions of $s \in \mathbb{R}$. Applying Remark 5.5.11 entry by entry with the chain rule (or Exercise 5.5.48) gives **Jacobi's formula**

$$\frac{d}{ds}\det A(s) = \sum_{i,j} c_{ij}(s)a'_{ij}(s) = \operatorname{tr}(\operatorname{adj}(A(s))A'(s)),$$

where $c_{ij}(s)$ are the cofactors of $A(s)$ and $A'(s)$ is the entrywise derivative. At an invertible $A(s)$ this reads $(\det A)' = \det A \cdot \operatorname{tr}\left(A^{-1}A'\right)$. For $A(s) = I_n + sH$ at $s = 0$ it gives $\operatorname{tr} H$, the coefficient of s in $\det(I_n + sH)$ found in Exercise 5.5.48 (with $A = I_n$ and $B = H$).

5.5.7 Minors and rank

BEYOND 21-242

Minors of arbitrary order appear in LADW 3.6 and in Cauchy–Binet, but not in the course.

The minors $\det B_{\hat{i}\hat{j}}$ come from square submatrices one size smaller than B. Submatrices of every size describe the rank.

Definition 5.5.29 (Minors of order k). Let $A \in F^{m\times n}$ and $1 \le k \le \min(m, n)$. For sets $I \subseteq [m]$ and $J \subseteq [n]$ with k elements each, the **submatrix** $A_{IJ} \in F^{k\times k}$ consists of the entries a_{ij} with $i \in I$ and $j \in J$, with rows and columns in increasing order. Its determinant $\det A_{IJ}$ is a **minor of order** k of A.

So $B_{\hat{i}\hat{j}} = B_{IJ}$ with $I = [n] \setminus \{i\}$ and $J = [n] \setminus \{j\}$, and the 1×1 minors are the entries.

Theorem 5.5.30 (Rank via minors). Let $A \in F^{m\times n}$ be nonzero. Then $\operatorname{rank} A$ is the largest k such that A has a nonzero minor of order k. In particular, every minor of order greater than $\operatorname{rank} A$ is 0.

Proof. HARD

Uses. Theorem 4.2.7 (The image of T_A), Theorem 4.1.10 (Independent lists are short), Theorem 1.6.10 (Every spanning list contains a basis), Theorem 4.3.8 (Row rank equals column rank), Corollary 5.4.8 (Vanishing determinants), Corollary 5.4.15 (Rows)

Hint 1. If A_{IJ} has nonzero determinant, its columns are independent. What does that say about the columns $j \in J$ of A?

Hint 2. Conversely, with $r = \operatorname{rank} A$, choose r independent columns of A. The $m \times r$ matrix they form has rank r, so its transpose does too. Choose r independent rows of it.

Sketch. If $\det A_{IJ} \neq 0$, the columns of A_{IJ} are independent, so the columns of A indexed by J are independent (a relation among them restricts to the rows in I). They lie in the column space, of dimension $\operatorname{rank} A$, so $k \le \operatorname{rank} A$.

Conversely, $r = \operatorname{rank} A$ columns of A form a basis of its column space; the $m \times r$ matrix A' they form has rank r, so by Theorem 4.3.8 some r rows of A' are independent, and the $r \times r$ submatrix they form has nonzero determinant.

The columns of A span $\operatorname{im} T_A$, which has dimension $\operatorname{rank} A$ by Theorem 4.2.7.

Suppose $\det A_{IJ} \neq 0$ for some I, J with k elements. By Corollary 5.4.8, the columns of A_{IJ} are linearly independent. For $j \in J$, the column of A_{IJ} that comes from column j of A is $\mathbf{a}'_j \in F^k$, the vector of entries a_{ij} with $i \in I$. We claim that the columns $\mathbf{a}_j$, $j \in J$, of A are linearly independent. If $\sum_{j\in J} x_j \mathbf{a}_j = \mathbf{0}$, then in particular the entries in the rows $i \in I$ vanish, that is, $\sum_{j\in J} x_j \mathbf{a}'_j = \mathbf{0}$; so all $x_j = 0$. Hence $\operatorname{im} T_A$ contains k independent vectors, and $k \leq \operatorname{rank} A$ by Theorem 4.1.10. This proves the last sentence of the theorem.

Now let $r = \operatorname{rank} A \geq 1$. By Theorem 1.6.10, the list of columns of A, which spans $\operatorname{im} T_A$, contains a basis of $\operatorname{im} T_A$; it has r elements, say the columns indexed by J. Let $A' \in F^{m\times r}$ be the matrix with these columns. Its columns are independent, so $\operatorname{rank} A' = r$ (Theorem 4.2.7), and $\operatorname{rank} A'^{\mathsf{T}} = r$ by Theorem 4.3.8. Applying the same argument to A'^{T}, some r columns of A'^{T} are independent, that is, some r rows of A', indexed by a set I, are independent. These rows form the $r \times r$ matrix A_{IJ}, which therefore has independent rows, so $\det A_{IJ} \neq 0$ by Corollary 5.4.15. So A has a nonzero minor of order r, and none of larger order. □

The theorem is not a practical way to find the rank, since row reduction is much faster than evaluating minors. Its value is that minors are polynomials in the entries, so statements about rank become statements about polynomials.

Corollary 5.5.31 (Rank does not depend on the field). Let K be a subfield of F (Definition 1.1.32), and let $A \in K^{m\times n}$. Then the rank of A is the same whether A is regarded as a matrix over K or over F.

Proof. ROUTINE

Uses. Theorem 5.5.30 (Rank via minors), Definition 5.3.6 (Determinant), Definition 1.1.32 (Subfield)
Hint 1. Is a minor of A computed over K the same element as the minor computed over F?
Hint 2. The Leibniz formula uses only sums and products, and K inherits these from F.
Sketch. Each minor is given by the Leibniz formula, which uses only the operations of K, and these agree with those of F. So A has the same minors over K and over F, and by Theorem 5.5.30 the same rank.

If $A = 0$, its rank is 0 over both fields. Otherwise, each minor $\det A_{IJ}$ is the sum in the Leibniz formula (Definition 5.3.6), a combination of sums, negatives, and products of entries of A. These operations in K are the restrictions of those in F (Definition 1.1.32), so the minor is the same element whether computed in K or in F. In particular, the minors that are nonzero are the same, and Theorem 5.5.30 gives the same rank. □

For example, a matrix with rational entries has the same rank over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. (Exercise 3.3.54 proves this with Gauss–Jordan elimination instead: the reduced row echelon form computed in K is also the one over F.) By contrast, reducing an integer matrix modulo a prime can lower its rank, because nonzero integer minors can become 0 in $\mathbb{F}_p$.

Remark 5.5.32 (The Cauchy–Binet formula). For $A \in F^{m\times n}$ and $B \in F^{n\times m}$, the determinant of the $m \times m$ matrix AB is a sum of products of minors of order m:

$$\det(AB) = \sum_{S\subseteq[n],\,|S|=m} \det A_{[m]S} \det B_{S[m]}.$$

If $m > n$, the sum is empty and $\det(AB) = 0$, in line with $\operatorname{rank}(AB) \leq n < m$. If $m = n$, there is one term, $\det A \det B$. For $A = B^{\mathsf{T}}$ over $\mathbb{R}$, it gives $\det\left(B^{\mathsf{T}}B\right) = \sum_S \left(\det B_{S[m]}\right)^2 \geq 0$. Exercise 5.5.52 proves the formula.

Exercises

Exercise 5.5.33. ROUTINE Let

$$A = \begin{pmatrix} 3 & 1 & 0 & 2 \\ 0 & 2 & 0 & 0 \\ 1 & 4 & 5 & 1 \\ 2 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

(a) Compute $\det A$ with cofactor expansions, choosing each row or column to keep the work small.

(b) For which primes p is the reduction of A modulo p invertible over $\mathbb{F}_p$?

Solution on page 1428

Exercise 5.5.34. ROUTINE Use cofactor expansion and induction on n to give another proof of Proposition 5.3.15: the determinant of an upper or lower triangular matrix is the product of its diagonal entries.

Solution on page 1428

Exercise 5.5.35. STANDARD For $t \in \mathbb{Q}$, consider the system $B(t)\mathbf{x} = \mathbf{y}$ with

$$B(t) = \begin{pmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{pmatrix}, \quad \mathbf{y} = (1, 2, 3).$$

(a) Show that $\det B(t) = -(t-1)^2(t+2)$. For which t does the system have exactly one solution?

(b) For those t, find x_2 with Cramer's rule.

(c) Show that for $t = 1$, every determinant $\det B_i(\mathbf{y})$ is 0, but the system has no solution. What happens for $t = -2$?

Solution on page 1429

Exercise 5.5.36. ROUTINE Construct a matrix $A \in \mathbb{Q}^{3\times 3}$ with integer entries, none of them 0, such that A^{-1} also has integer entries. Compute A^{-1} with the adjugate.

Solution on page 1429

Exercise 5.5.37. STANDARD Let $n \geq 2$ and let $B \in F^{n\times n}$ be upper triangular.

(a) Show that $\operatorname{adj} B$ is upper triangular, with diagonal entries $\prod_{k\neq i} b_{kk}$ for $i = 1, \ldots, n$.

(b) Deduce that if B is invertible, then B^{-1} is upper triangular with diagonal entries $b_{11}^{-1}, \ldots, b_{nn}^{-1}$.

Solution on page 1430

Exercise 5.5.38. STANDARD Let $n \geq 2$, $A, B \in F^{n\times n}$, and $c \in F$. Prove or disprove each statement. Where the answer depends on n, say how.

(a) $\operatorname{adj}\left(A^{\mathsf{T}}\right) = (\operatorname{adj} A)^{\mathsf{T}}$.

(b) $\operatorname{adj}(cA) = c^{n-1} \operatorname{adj} A$.

(c) $\operatorname{adj}(A + B) = \operatorname{adj} A + \operatorname{adj} B$.

(d) The map $\operatorname{adj} : F^{n\times n} \to F^{n\times n}$ is injective.

Solution on page 1431

Exercise 5.5.39. STANDARD Let $n \geq 2$ and $A \in F^{n\times n}$. Show that $\det(\operatorname{adj} A) = (\det A)^{n-1}$, including when A is not invertible.

Solution on page 1431

Exercise 5.5.40. HARD Let $n \geq 2$ and $A \in F^{n\times n}$. Show that

$$\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2} A,$$

with the convention $0^0 = 1$ when $n = 2$. Prove it first for invertible A, and then for all A when F is infinite.

Solution on page 1432

Solution on page 1432

Exercise 5.5.41. HARD Let F be an infinite field and $A, B \in F^{n\times n}$. Show that $\operatorname{adj}(AB) = (\operatorname{adj} B)(\operatorname{adj} A)$. (The identity holds over every field, but the proof suggested here needs infinitely many scalars.)

Solution on page 1433

Exercise 5.5.42. STANDARD Here is a "proof" that for all $a, b, c \in F$,

$$D(a,b,c) = \det\begin{pmatrix} 1 & a & a^3 \\ 1 & b & b^3 \\ 1 & c & c^3 \end{pmatrix} = (b-a)(c-a)(c-b).$$

Fix distinct a and b. Expanding along the last row, $D(a,b,c)$ is a polynomial in c of degree at most 3. It vanishes at $c = a$ and at $c = b$, since the matrix then has two equal rows. The right side also vanishes exactly at $c = a$ and $c = b$. Two polynomials with the same roots are equal, so $D(a,b,c) = (b-a)(c-a)(c-b)$. If $a = b$, both sides are 0.

(a) Find the errors, and show that the claimed formula is false for $F = \mathbb{Q}$.

(b) Find and prove the correct formula.

Solution on page 1433

Exercise 5.5.43. ROUTINE Use Theorem 5.5.21 to give short proofs of the following, which Section 1.5 proved without determinants.

(a) If $\lambda_1, \ldots, \lambda_k \in F$ are distinct, then the sequences $\mathbf{s}_{\lambda_i} = \left(1, \lambda_i, \lambda_i^2, \ldots\right)$ are linearly independent (Exercise 1.5.44).

(b) If $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ are distinct, then the functions $t \mapsto e^{\lambda_i t}$ are linearly independent in $\mathbb{R}^{\mathbb{R}}$ (Exercise 1.5.45).

Solution on page 1434

Exercise 5.5.44. HARD Let $c_0, \ldots, c_n \in F$ be distinct and $d_0, \ldots, d_n \in F$, and let $p \in F[x]_{\leq n}$ be the polynomial with $p(c_i) = d_i$ for all i (Theorem 3.4.24).

(a) Use Cramer's rule to show that the coefficient of x^n in p is

$$\sum_{i=0}^{n} \frac{d_i}{\prod_{j\neq i}(c_i - c_j)}.$$

(b) Deduce that $\sum_{i=0}^{n} c_i^m / \prod_{j\neq i}(c_i - c_j)$ is 0 for $0 \leq m < n$ and 1 for $m = n$.

Solution on page 1434

Exercise 5.5.45. HARD For $x \in \mathbb{R}$ and $n \geq 1$, let $M_n(x) \in \mathbb{R}^{n\times n}$ be tridiagonal with every diagonal entry equal to x and every entry directly above or below the diagonal equal to 1. Let $d_n(x) = \det M_n(x)$ and $d_0(x) = 1$.

(a) Show that $d_n(x) = xd_{n-1}(x) - d_{n-2}(x)$ for $n \geq 2$, and that $d_n(x) = P_n(x)$ for a polynomial $P_n \in \mathbb{R}[x]_{\leq n}$ with coefficient 1 of x^n.

(b) Show that if $x = 2\cos\theta$ with $0 < \theta < \pi$, then $d_n(x) = \sin((n+1)\theta)/\sin\theta$.

(c) Deduce that $M_n(x)$ is singular exactly for the n values $x = 2\cos(k\pi/(n+1))$, $k = 1, \ldots, n$, and that $d_n(x) = \prod_{k=1}^{n}(x - 2\cos(k\pi/(n+1)))$ for every $x \in \mathbb{R}$.

Solution on page 1435

Exercise 5.5.46. ROUTINE Find a matrix $A \in \mathbb{F}_3^{4\times 4}$ such that $\det(tI_4 - A) = t^4 + t + 2$ for every $t \in \mathbb{F}_3$. Then show that $tI_4 - A$ is invertible for every $t \in \mathbb{F}_3$.

Solution on page 1435

Exercise 5.5.47. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$.

(a) Show that $\det(I_m + AB) = \det(I_n + BA)$ (**Sylvester's determinant identity**). Hint: compute the determinant of $M = \begin{pmatrix} I_m & -A \\ B & I_n \end{pmatrix}$ in two ways.

(b) Deduce that $\det\left(I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = 1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}$ for $\mathbf{u}, \mathbf{v} \in F^n$, and compute $\det(I_n + J)$, where every entry of J is 1.

Solution on page 1436

Exercise 5.5.48. HARD Let $n \geq 2$ and $A, B \in F^{n\times n}$.

(a) Show that there are scalars $e_0, \ldots, e_n \in F$, depending only on A and B, such that $\det(A + tB) = e_0 + e_1 t + \cdots + e_n t^n$ for every $t \in F$, with $e_0 = \det A$, $e_n = \det B$, and

$$e_1 = \operatorname{tr}(\operatorname{adj}(A)B).$$

(b) Deduce that the coefficient of t in $\det(A + tI_n)$ (in the sense of Proposition 5.3.19) is $\sum_{i=1}^{n} \det A_{\hat{i}\hat{i}}$, the sum of the minors along the diagonal.

Solution on page 1436

Exercise 5.5.49. STANDARD Let $n \geq 2$, $B \in F^{n\times n}$, and $\mathbf{y} \in F^n$. Prove or disprove:

(a) If $\det B = 0$ and $\det B_i(\mathbf{y}) \neq 0$ for some i, then $B\mathbf{x} = \mathbf{y}$ has no solution.

(b) If $\det B = 0$ and $\det B_i(\mathbf{y}) = 0$ for every i, then $B\mathbf{x} = \mathbf{y}$ has a solution.

(c) If $F = \mathbb{Q}$, B and $\mathbf{y}$ have integer entries, and $\det B \neq 0$, then the solution of $B\mathbf{x} = \mathbf{y}$ has integer entries.

(d) If $\det B \neq 0$ and $\det B_i(\mathbf{y}) = 0$ for every i, then $\mathbf{y} = \mathbf{0}$.

Solution on page 1437

Exercise 5.5.50. STANDARD

(a) Compute $\det V(1, i, -1, -i)$ over $\mathbb{C}$ with Theorem 5.5.21.

(b) Let $V = V(1, i, -1, -i)$ and $\omega = i$, so V has $(j+1, k+1)$ entry ω^{jk} for $0 \leq j, k \leq 3$. Show that $V^2 = 4P$ for a permutation matrix P with $\det P = -1$, and use this to check (a) up to sign.

(c) In $\mathbb{F}_5$, the element 2 satisfies $2^4 = 1$, and its powers are $1, 2, 4, 3$. Compute $\det V(1, 2, 4, 3)$ in $\mathbb{F}_5$.

Solution on page 1437

Exercise 5.5.51. HARD BEYOND 21-242 Let $n \geq 2$ and $A \in F^{n\times n}$. Show that

$$\operatorname{rank}(\operatorname{adj} A) = \begin{cases} n \text{ if } \operatorname{rank} A = n \\ 1 \text{ if } \operatorname{rank} A = n-1 \\ 0 \text{ if } \operatorname{rank} A \leq n-2. \end{cases}$$

Solution on page 1438

Exercise 5.5.52. HARD BEYOND 21-242 Prove the Cauchy–Binet formula of Remark 5.5.32: for $A \in F^{m\times n}$ and $B \in F^{n\times m}$,

$$\det(AB) = \sum_{S\subseteq[n],\,|S|=m} \det A_{[m]S} \det B_{S[m]}.$$

Solution on page 1438

Exercise 5.5.53. CHALLENGE Let $x_1, \ldots, x_n, y_1, \ldots, y_n \in F$ with $x_i + y_j \neq 0$ for all i, j, and let $C_n \in F^{n\times n}$ have entries $1/(x_i + y_j)$.

(a) Show that

$$\det C_n = \frac{\prod_{1\leq i<j\leq n}(x_j - x_i)(y_j - y_i)}{\prod_{i,j=1}^{n}(x_i + y_j)}.$$

(b) Deduce that the Hilbert matrix $H_n \in \mathbb{Q}^{n\times n}$, with entries $1/(i+j-1)$, is invertible, and compute $\det H_3$.

Solution on page 1439

5.6 Proof patterns

The chapter described the determinant in four ways: as the only function satisfying three axioms (Definition 5.1.1, Theorem 5.3.4), as the Leibniz sum over permutations (Theorem 5.3.3), as the product of the factors collected while eliminating (Proposition 5.4.21), and as a cofactor expansion (Theorem 5.5.3). Most proofs in the chapter start by choosing one of these descriptions, and the choice decides how long the proof is. This section sorts the chapter's arguments by the question they answer, compares the descriptions on examples from the chapter, and says when each one is the wrong tool. The exercises at the end practice one move at a time, and the problem set (Problems for Chapter 5) mixes them.

5.6.1 Computing a determinant

For a matrix of numbers with no visible structure, use elimination (Algorithm 5.4.22). It takes about $n^3/3$ operations, while the Leibniz formula and a full cofactor expansion take on the order of $n!$ (Remark 5.4.25). Structure changes this, so look for it first. Roughly in order of how quickly they finish:

- *A known form.* Triangular and block-triangular matrices (Proposition 5.3.15, Corollary 5.4.27), permutation and monomial matrices (Theorem 5.2.31, Exercise 5.2.44), Vandermonde matrices (Theorem 5.5.21), and companion matrices (Proposition 5.5.20). Sometimes the form appears only after the rows or columns are reordered, which costs a sign (Exercise 5.3.37).
- *Few nonzero entries.* Ask which permutations avoid every zero (Technique 5.3.12). Often only a few terms survive: one in Example 5.6.2 below, and four of the 24 in Exercise 5.3.24. A row or column with one or

two nonzero entries calls for a cofactor expansion along it (Example 5.5.4), possibly after replacements that create the zeros.
- *A family.* If deleting the last row and column leaves a smaller member of the same family, expand to get a recurrence (Technique 5.5.17), as for tridiagonal and companion matrices (Proposition 5.5.18, Proposition 5.5.20).
- *Parameters.* When the entries are polynomials in some parameters, use replacements that expose a common factor (Technique 5.5.22), as for constant row sums (Example 5.4.24) and powers of a node (Theorem 5.5.21). If the answer should be a product of linear factors, its roots can be found before any computation (Technique 5.5.24).
- *A product.* If every entry is a sum of products, the matrix may factor, and then so does its determinant.

The last item has not been named yet, although several exercises of the chapter rely on it.

Technique 5.6.1 (Factor the matrix). To compute $\det M$, write M as a product of matrices whose determinants are known, and multiply (Theorem 5.4.10). There are two common sources of factorizations. If every entry is a sum of products, $m_{ij} = \sum_k x_{ik} d_k y_{kj}$, then $M = XDY$ with $D = \operatorname{diag}(d_1, \ldots, d_r)$, and X and Y are often triangular or Vandermonde matrices. If M is a block matrix, block row and column operations (Theorem 2.4.34) multiply it by matrices of determinant 1 until it is block triangular. When the inner size r is smaller than the size of M, the same factorization shows $\det M = 0$ instead (Exercise 5.4.38). For instance, the matrix with entries $1 + x_i y_j$ is XY with rows $(1, x_i)$ in X and columns $(1, y_j)$ in Y. For two rows this gives $(x_2 - x_1)(y_2 - y_1)$, and for three or more rows it gives 0.

Reach for it when the entries are power sums, binomial expansions, or sums over divisors, or when one block of a block matrix can be cleared. The chapter used it in Exercise 5.4.43, Exercise 5.4.46, and Exercise 5.5.47. The canonical example is the Hankel matrix of power sums, which is $V^{\mathsf{T}}V$ for a Vandermonde matrix V (Problem 5.9).

Used in. Exercise 5.6.7, Problem 5.9, Problem 5.12, Problem 5.13, Problem 5.19, Problem 5.20, Problem 7.18.

Example 5.6.2 (One determinant, three ways). Let

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 4 & 0 & 5 \\ 0 & 0 & 6 & 0 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

Counting the surviving terms. A term $\operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(4)4}$ is nonzero only if all four entries are. The only nonzero entry of column 1 is in row 2, so $\sigma(1) = 2$, and column 4 forces $\sigma(4) = 3$. Then column 2, with nonzero entries in rows 1 and 3, must use row 1, and column 3 must use row 4. Exactly one permutation survives, $\sigma = (1\ 2)(3\ 4)$, which is even. So

$$\det A = a_{21} a_{12} a_{43} a_{34} = 2 \cdot 1 \cdot 6 \cdot 5 = 60.$$

A recurrence. A is the tridiagonal matrix of Proposition 5.5.18 with every $a_k = 0$, $(b_1, b_2, b_3) = (1, 3, 5)$, and $(c_1, c_2, c_3) = (2, 4, 6)$. The recurrence becomes $d_m = -b_{m-1} c_{m-1} d_{m-2}$, so $d_1 = 0$, $d_2 = -2$, $d_3 = 0$, and $d_4 = -30 \cdot (-2) = 60$.

Elimination. Swap rows 1 and 2. Subtracting 4 times the new row 2 from row 3 turns row 3 into $(0, 0, 0, 5)$, and swapping rows 3 and 4 leaves an upper triangular matrix with diagonal $2, 1, 6, 5$. There were two swaps, so $\det A = 60$.

Elimination always works, but here it spends its steps on swaps and does not explain why the answer is a product of entries. The recurrence treats the whole family at once, and shows for example that $d_n = 0$ for every odd n. The count explains why: since the only nonzero entries are next to the diagonal, a surviving σ sends every odd index to an even one and every even index to an odd one, and when n is odd there are

more odd indices than even ones. On a dense matrix the same count lists all $n!$ permutations, and there it is the wrong tool.

Practice: Exercise 5.6.6 and Exercise 5.6.7.

5.6.2 Proving an identity for every matrix

Four arguments recur in the chapter. They differ in what they ask of the identity, and in how they deal with matrices that are not invertible.

- *Recognize the determinant by its axioms* (Technique 5.4.11). If one side, as a function of the columns of a matrix B, is multilinear and alternating, then it is its value at I_n times $\det B$. There are no cases and no formulas. This proved Theorem 5.4.10 (second proof) and Proposition 5.3.17 (first proof). It needs an expression that is visibly linear in each column of something.
- *Reduce to elementary matrices* (Technique 5.4.6). Check the identity on the three types and show that it survives multiplication. It handles transposes, which the axioms do not reach, but it only reaches invertible matrices, and the others need a separate argument (Exercise 5.4.35).
- *Manipulate the Leibniz formula.* Reindex the sum over S_n (Proposition 5.2.8), or apply an operation to every term: transposing (second proof of Theorem 5.4.14), conjugating (Exercise 5.3.28), reducing modulo p (Proposition 5.3.20). It works when the identity matches the terms of one sum with the terms of another.
- *Polynomial identities* (Technique 5.5.24). Prove the identity for invertible matrices, replace A by $A + tI_n$, and observe that both sides are polynomial in t and agree for all but finitely many t. This needs more scalars than the degree plus the number of bad values of t, which is automatic only over an infinite field (Exercise 5.6.10).

Example 5.6.3 (Multiplicativity, three ways). The text proved $\det(AB) = \det A \det B$ twice. The first proof (Technique 5.4.6) factors B into elementary matrices and needs a separate argument when B is not invertible: then AB is not invertible either, and both sides are 0. The second (Technique 5.4.11) notices that $B \mapsto \det(AB)$ is multilinear and alternating in the columns of B, so it equals $\det(AI_n)\det B$.

A third proof makes the second one explicit. Column j of AB is $A\mathbf{b}_j = \sum_k b_{kj}\mathbf{a}_k$ (Proposition 2.4.8). Expanding every column as in Lemma 5.3.2, but with $\mathbf{a}_k$ in place of $\mathbf{e}_k$,

$$\det(AB) = \sum_{f\in[n]^{[n]}} b_{f(1)1}\cdots b_{f(n)n}\det\bigl(\mathbf{a}_{f(1)},\dots,\mathbf{a}_{f(n)}\bigr).$$

The terms with f not injective have a repeated column and vanish. For a permutation σ, $\det\bigl(\mathbf{a}_{\sigma(1)},\dots,\mathbf{a}_{\sigma(n)}\bigr) = \operatorname{sgn}(\sigma)\det A$ by Theorem 5.2.31. So

$$\begin{aligned}\det(AB) &= \det A \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1)1}\cdots b_{\sigma(n)n}\\ &= \det A \det B.\end{aligned}$$

The same computation, with $\mathbf{e}_k$ in place of $\mathbf{a}_k$, proved Theorem 5.3.3; the second proof quotes its outcome, Theorem 5.3.4, instead of repeating it. Run on a rectangular A and B, the same expansion gives the Cauchy–Binet formula (Exercise 5.5.52), which neither of the first two proofs reaches.

The wrong tool here is the Leibniz formula applied to the entries of AB: each entry is a sum of n products, each of the $n!$ terms multiplies out into n^n products, and the cancellation is invisible. For $\det A^{\mathsf{T}} = \det A$ the roles change. There the axioms are the wrong tool, since the multilinearity in rows that Technique 5.4.11 would need is what the theorem is supposed to establish, while elementary matrices and the Leibniz formula both work (Intuition (What the two proofs see)).

Matrices that are not invertible were handled in three ways. In Theorem 5.4.10, both sides are 0 on them. In Exercise 5.5.40, a polynomial in t carries the identity from $A + tI_n$ to A, provided the field is infinite, and a separate rank argument covers every field. Proofs by the axioms never meet the case at all, which is one reason to prefer them when they apply. Practice: Exercise 5.6.8, Exercise 5.6.9, and Exercise 5.6.11.

5.6.3 Showing that a determinant is nonzero

By Theorem 5.4.7, "$\det A \neq 0$" and "A is invertible" are the same statement, so either can be proved by proving the other.

- *A product of nonzero factors.* A triangular matrix with nonzero diagonal, a Vandermonde matrix with distinct nodes, or any product of such matrices has nonzero determinant, since a field has no zero divisors (Proposition 1.1.17).
- *Reduce modulo a prime* (Technique 5.3.21). For an integer matrix, one prime p with $\det \overline{A} \neq 0$ in $\mathbb{F}_p$ suffices (Example 5.3.22). The converse inference fails: $\det \overline{A} = 0$ says only that p divides $\det A$.
- *A polynomial in a parameter* (Technique 5.3.18). A polynomial of degree n has at most n roots, so $A + tI_n$ is invertible for all but at most n values of t (Proposition 5.3.19). To show that such a polynomial is not the zero polynomial, one nonzero coefficient or one nonzero value is enough.
- *The kernel instead of the determinant.* Show that $A\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = \mathbf{0}$ (Corollary 5.4.8). This is often the shortest route, and sometimes the only one available.
- *A single surviving term.* If exactly one permutation avoids the zeros of A, then $\det A$ is $\pm$ a product of nonzero entries (Exercise 5.3.30).

To show $\det A = 0$, exhibit a dependence among the rows or columns, possibly after replacements; find a block of zeros that is too large (Exercise 5.3.29); factor A through a smaller size (Exercise 5.4.38); or use a symmetry, such as $\det A = \det A^{\mathsf{T}} = -\det A$ for skew-symmetric A of odd size (Exercise 5.4.33). The symmetry argument divides by $1 + 1$, and Exercise 5.6.13 replaces it by a pairing that does not.

Example 5.6.4 (One matrix, four certificates). Let $n \geq 2$, let $J \in \mathbb{Q}^{n\times n}$ have every entry equal to 1, and let $A = J - I_n$, with 0 on the diagonal and 1 elsewhere.

The determinant. Example 5.4.24 with $x = 0$ gives $\det A = (n-1)(-1)^{n-1} \neq 0$.

The kernel. Suppose $A\mathbf{x} = \mathbf{0}$, and let $s = x_1 + \cdots + x_n$. Row i of the system says $s - x_i = 0$. So every x_i equals s, hence $s = ns$, that is, $(n-1)s = 0$, and $s = 0$. So $\mathbf{x} = \mathbf{0}$.

An inverse. Since $J^2 = nJ$,

$$\begin{aligned}(J - I_n)(J - (n-1)I_n) &= nJ - (n-1)J - J + (n-1)I_n \\ &= (n-1)I_n,\end{aligned}$$

so $A^{-1} = (n-1)^{-1}(J - (n-1)I_n)$ (Technique 2.5.16).

Modulo 2. For even n, the reduction of A is invertible over $\mathbb{F}_2$ (Exercise 5.3.40), so $\det A$ is odd. For odd n, every row of A has the even sum $n - 1$, so the reduction sends $(1, \ldots, 1)$ to $\mathbf{0}$, and this route gives no information.

The determinant contains the most information. The kernel argument is the shortest, needs no formula, and shows exactly where the hypothesis enters: over a field in which $n - 1 = 0$, the vector $(1, \ldots, 1)$ lies in the kernel. The inverse is the strongest certificate, and finding it used the algebra of J, not determinants. Reducing modulo a prime suits matrices whose pattern is simple modulo that prime, and this one is simple modulo 2 only for even n.

Practice: Exercise 5.6.12 and Exercise 5.6.13.

5.6.4 Signs of permutations

There are three ways to compute $\operatorname{sgn}(\sigma)$: count the factors of any factorization into transpositions, count inversions (Corollary 5.2.24), or read off the lengths of the cycles (Proposition 5.2.26). Cycles are the quickest when σ is given in cycle notation, and inversions when it is given by its bottom row. In determinant computations, signs come from reordering rows or columns, and there it usually pays to count adjacent interchanges. That is how the cofactor sign $(-1)^{i+j}$ arose (Lemma 5.5.2), and how a cyclic shift of n columns costs $(-1)^{n-1}$ (Exercise 5.2.34).

Example 5.6.5 (Moving a block, three ways). Let $n = k + m$, and let $\pi \in S_n$ have bottom row $m+1, \ldots, m+k, 1, \ldots, m$, the permutation that moves the block $\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$ into block-diagonal form in Exercise 5.3.37.

Inversions. A pair of positions is out of order exactly when the first lies among the first k positions and the second among the last m, so $\operatorname{inv}(\pi) = km$ and $\operatorname{sgn}(\pi) = (-1)^{km}$.

Adjacent interchanges. Each of the m entries $1, \ldots, m$ moves to the front past each of the k larger entries, one adjacent interchange at a time: km interchanges.

Cycles. π adds m modulo n, so its cycles all have length $n/\gcd(n,m)$, there are $\gcd(n,m)$ of them, and $\operatorname{sgn}(\pi) = (-1)^{n-\gcd(n,m)}$. This is correct, but it depends on number theory, and checking that it agrees with $(-1)^{km}$ takes a parity argument about greatest common divisors. Cycle notation is the wrong tool for this permutation.

The definition of sgn through transpositions makes multiplicativity a one-line proof (Proposition 5.2.25), while proving it from inversions takes real work (Exercise 5.2.49). In general, define an invariant through the description in which its main property is easy, and compute it through the description in which it is quick.

Sums over S_n with signs are evaluated by pairing terms of opposite sign (Technique 5.3.5). Pairing σ with $\sigma\tau$ for a fixed transposition τ proved that L is alternating (Theorem 5.3.4) and that half of all permutations are even (Proposition 5.2.27). Pairing σ with σ^{-1} handles alternating matrices (Exercise 5.6.13). Parity also proves that some configurations cannot be reached (Technique 5.2.19): the 15 puzzle (Exercise 5.2.47) and the rotation puzzle (Exercise 5.2.48). Practice: Exercise 5.6.14.

5.6.5 Reading the chapter's proofs

Uniqueness first, then existence. The existence proof was found backwards. We assumed that a determinant function exists and asked what it would have to be. Expanding every column (Lemma 5.3.2), letting alternation kill the terms with a repeated column, and sorting the survivors (Theorem 5.2.31) produced Theorem 5.3.3: at most one candidate, given by a formula. Only then did Section 5.3.2 verify that the candidate satisfies the axioms. The 2×2 case went the same way (Theorem 5.1.13). The derivation also showed what existence requires: sorting the columns must produce a well-defined sign, which is the parity theorem (Remark 5.2.32). This is the "derive the formula" move of Section 1.7.2 on a larger scale. The verification also shows where the difficulty lay. Normalization and multilinearity can be checked one term at a time, but alternation cannot, because equal columns do not force any single term to vanish.

Pairing instead of reindexing. Reindexing the Leibniz sum by $\sigma \mapsto \sigma\tau$ shows $L(A) = -L(A)$, which gives $L(A) = 0$ only after dividing by $1 + 1$. Pairing each σ with $\sigma\tau$ writes $L(A)$ as a sum of pairs that cancel, and never divides (Intuition (Why alternation needs care)). The same distinction separates antisymmetry from alternation (Proposition 5.1.11, Example 5.1.12), and it is the reason the definition asks for alternation. An argument that ends with "$x = -x$, so $x = 0$" has used $1 + 1 \neq 0$ and should say so.

Tracking which axiom is used. Theorem 5.1.13 kept $D(I_2)$ as an unknown and obtained a statement about every multilinear alternating function at no extra cost. The same bookkeeping showed that dependent columns force $D(A) = 0$ without normalization (Proposition 5.1.17), while the converse fails for $D = 0$. So

every proof of Theorem 5.4.7 must use normalization somewhere, and the proof in the text uses it through Proposition 5.4.1.

Columns first. The axioms speak about columns, so every statement about rows waited for Theorem 5.4.14 (Corollary 5.4.15, Technique 3.2.27). Statements that mix rows and columns, such as cofactor expansion, were proved for rows and transposed, or proved with whichever kind of operation was convenient.

5.6.6 Common errors

Most of these appear in the chapter as find-the-error exercises.

- *det is not linear.* In general $\det(A+B) \neq \det A + \det B$, and $\det(cA) = c^n \det A$ (Exercise 5.1.21, Exercise 5.3.26, Remark 5.1.4). Multilinear means linear in one column at a time. In the same way, a 2×2 block matrix is not a 2×2 matrix of numbers (Exercise 5.4.36).
- *Sarrus for larger matrices.* The rule of Sarrus is a coincidence of size 3 (Remark 5.3.11, Exercise 5.3.25).
- *The sign of a cofactor.* The sign $(-1)^{i+j}$ comes from $i-1$ and $j-1$ adjacent interchanges. Interchanging row i with row 1 directly puts the old row 1 out of order, and for $i \geq 3$ what remains is not the minor $B_{\hat{i}\hat{j}}$ (Exercise 5.6.15). Forgetting the checkerboard altogether is the commoner version.
- *Row operations that scale.* Replacing row i by 2 times row i minus 3 times row 1 doubles the determinant (Exercise 5.4.16).
- *Factoring a singular matrix.* Only invertible matrices are products of elementary matrices (Exercise 5.4.35).
- *A converse for arbitrary D.* A multilinear alternating D can vanish on invertible matrices (Exercise 5.1.22).
- *Circularity.* The existence proof in Section 5.3.2 rests on the parity theorem, so quoting Theorem 5.3.4 to prove that theorem is circular (Exercise 5.2.38).
- *Inversions under a transposition.* They change by an odd number, not always by 1 (Exercise 5.2.35).
- *Polynomial arguments.* Roots determine a polynomial only together with its leading coefficient (Exercise 5.5.42), and agreement at "all but finitely many points" proves nothing over a small finite field (Exercise 5.6.10).
- *Reducing modulo p in the wrong direction.* $\det \overline{A} = 0$ in $\mathbb{F}_p$ does not make A singular over $\mathbb{Q}$ (Technique 5.3.21, Exercise 5.3.36).
- *Hidden division by* $1+1$. Antisymmetry does not imply alternation over $\mathbb{F}_2$ (Example 5.1.12), and neither does $\det A = -\det A$ imply $\det A = 0$ (Exercise 5.4.33).

Exercises

Exercise 5.6.6. ROUTINE For each matrix, choose a method, say why it fits, and compute the determinant.

(a) $A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 7 & 3 & 1 & 2 & 0 \\ -4 & 5 & 0 & 1 & 3 \\ 9 & 9 & 4 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{5\times 5}$.

(b) $B(x) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{pmatrix}$, for $x \in F$.

(c) $C = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 0 & 0 & 7 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$.

Solution on page 1440

Exercise 5.6.7. STANDARD For $n \geq 1$, let $P_n \in \mathbb{Q}^{n\times n}$ have entries $p_{ij} = \binom{i+j-2}{i-1}$, so that

$$P_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & 20 \end{pmatrix}.$$

(a) Show that $\det P_n = 1$ for every n. You may use Pascal's rule $\binom{m}{k} = \binom{m-1}{k} + \binom{m-1}{k-1}$ for $m, k \geq 1$, and $\binom{m}{k} = 0$ for $0 \leq m < k$.

(b) Deduce that P_n^{-1} has integer entries, and compute P_3^{-1}.

Solution on page 1440

Solution on page 1441

Exercise 5.6.8. STANDARD Let $A \in F^{n\times n}$. For $X \in F^{n\times n}$ with columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$, let

$$D(X) = \sum_{j=1}^{n} \det\bigl(\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}, A\mathbf{x}_j, \mathbf{x}_{j+1}, \ldots, \mathbf{x}_n\bigr),$$

the sum of the determinants obtained by applying A to one column at a time.

(a) Show that $D(X) = (\operatorname{tr} A)\det X$ for every $X \in F^{n\times n}$.

(b) Deduce that for every multilinear alternating $E : F^{n\times n} \to F$,

$$\sum_{j=1}^{n} E\bigl(\mathbf{x}_1, \ldots, A\mathbf{x}_j, \ldots, \mathbf{x}_n\bigr) = (\operatorname{tr} A)E(\mathbf{x}_1, \ldots, \mathbf{x}_n).$$

Solution on page 1442

Exercise 5.6.9. HARD Let $A, B \in F^{n\times n}$ and $t \in F$. We show that

$$\det(tI_n - AB) = \det(tI_n - BA).$$

(a) Prove this when A is invertible.

(b) Let F be infinite. Fix t, and show that $g(s) = \det(tI_n - (A + sI_n)B) - \det(tI_n - B(A + sI_n))$ satisfies $g(s) = e_0 + e_1 s + \cdots + e_n s^n$ for some $e_0, \ldots, e_n \in F$ and all $s \in F$, and that $g(s) = 0$ for all but at most n values of s. Conclude that the identity holds for every A.

(c) Prove the identity over every field, using Exercise 5.5.47 when $t \neq 0$.

Solution on page 1443

Exercise 5.6.10. STANDARD Find the error in the following argument, and give a matrix $A \in \mathbb{F}_2^{2\times 2}$ for which its key step fails. (The claim itself is true over every field; see the second approach to Exercise 5.5.40.)

Claim. For every field F, every $n \geq 2$, and every $A \in F^{n\times n}$, $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2}A$, with the convention $0^0 = 1$ when $n = 2$.

Proof. For invertible A this is a short computation (Exercise 5.5.40). For general A, the entries of $\operatorname{adj}(\operatorname{adj}(A + tI_n))$ and of $\det(A + tI_n)^{n-2}(A + tI_n)$ are polynomial functions of t. They agree whenever $A + tI_n$ is invertible, that is, for all but at most n values of t. Polynomial functions that agree at all but finitely many points are equal, so they also agree at $t = 0$.

Solution on page 1443

Exercise 5.6.11. STANDARD

(a) Let $A \in \mathbb{R}^{n\times n}$ be invertible. Show that $\det A > 0$ if and only if A is a product of replacement matrices E_{icj} and scaling matrices E_{ic} with $c > 0$.

(b) Write the swap $E_{12} \in \mathbb{R}^{2\times 2}$ as a product of three replacements and one scaling, and explain why some factor must be a scaling by a negative number.

Solution on page 1444

Exercise 5.6.12. STANDARD Prove each statement, choosing a suitable certificate of invertibility.

(a) The matrix $A \in \mathbb{Q}^{4\times 4}$ with entries $a_{ij} = i^j$ is invertible.

(b) For $x \in \mathbb{R}$, let $M(x) = \begin{pmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x+1 \end{pmatrix}$. Then $M(x)$ is invertible for every integer x, but not for every real x.

(c) For every $A \in \mathbb{R}^{m\times n}$, the matrix $I_n + A^{\mathsf{T}}A$ is invertible. (For $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^{\mathsf{T}}\mathbf{v} = v_1^2 + \cdots + v_n^2$.)

(d) For every $A \in \mathbb{Q}^{n\times n}$, the matrix $A + tI_n$ is invertible for some integer t with $0 \le t \le n$.

Solution on page 1444

Exercise 5.6.13. HARD Call $A \in F^{n\times n}$ *alternating* if $a_{ii} = 0$ and $a_{ji} = -a_{ij}$ for all i, j. (When $1 + 1 \neq 0$ this means skew-symmetric; over $\mathbb{F}_2$ it means symmetric with zero diagonal.) Let n be odd and A alternating, and write $t_\sigma(A) = \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ for the terms of the Leibniz formula.

(a) Show that $t_{\sigma^{-1}}(A) = -t_\sigma(A)$ for every $\sigma \in S_n$.

(b) Show that if $\sigma^2 = \mathrm{id}$, then σ has a fixed point, and $t_\sigma(A) = 0$.

(c) Deduce that $\det A = 0$, over every field.

(d) Exercise 5.4.33 proved (c) when $1 + 1 \neq 0$ by showing $\det A = -\det A$. Explain why that argument fails over $\mathbb{F}_2$, why this one does not, and where the hypothesis $a_{ii} = 0$ is used.

Solution on page 1445

Exercise 5.6.14. STANDARD Let $n \ge 1$, and let $\pi \in S_{2n}$ have $\pi(2i-1) = i$ and $\pi(2i) = n + i$ for $i = 1, \ldots, n$. Its bottom row is $1, n+1, 2, n+2, \ldots, n, 2n$, so the columns of ME_π interleave the first n columns of M with the last n (Proposition 5.2.30).

(a) Show that $\operatorname{inv}(\pi) = n(n-1)/2$.

(b) Write π in cycle notation for $n = 3$ and $n = 4$, and check (a) against Proposition 5.2.26.

(c) Let $A, B, C, D \in F^{n\times n}$ be diagonal, with diagonal entries a_i, b_i, c_i, d_i. Show that

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \prod_{i=1}^{n} (a_i d_i - b_i c_i).$$

Does the sign of π enter?

Solution on page 1446

Exercise 5.6.15. ROUTINE Find the error in the following argument, and show that its conclusion fails for

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad i = 3, \quad j = 1.$$

Claim. Let $n \ge 2$ and $B \in F^{n\times n}$. If every entry of row i of B other than b_{ij} is 0, then $\det B = \varepsilon b_{ij} \det B_{\hat{i}\hat{j}}$, where $\varepsilon = -1$ if exactly one of i and j equals 1, and $\varepsilon = 1$ otherwise.

Proof. Interchange rows 1 and i (if $i \neq 1$), and then columns 1 and j (if $j \neq 1$). Each interchange negates the determinant. The resulting matrix has b_{ij} in position $(1,1)$ and zeros in the rest of row 1, so it is block lower triangular, and by Theorem 5.4.26 its determinant is b_{ij} times the determinant of its lower right block, which is $B_{\hat{i}\hat{j}}$.

Solution on page 1447

Problems for Chapter 5

These problems combine the sections of Chapter 5 with each other and with earlier chapters. The first three are review prompts: state the result, prove it without looking, and then answer a follow-up question that the text did not ask. The rest are ordered roughly by difficulty. Problems of kind "competition" are harder than anything the course expects. Section 5.6 compares the methods these problems call for.

Review: state and prove

Problem 5.1. STANDARD

(a) State the theorem on the existence and uniqueness of determinant functions, including its description of all multilinear alternating functions on $F^{n\times n}$.

(b) Prove that the function L of (5.3.1) is alternating, by an argument that works over every field.

(c) Let $n \geq 2$ and $F = \mathbb{F}_2$, and define $D(A) = \det A + a_{11}a_{12}\cdots a_{1n}$ for $A \in \mathbb{F}_2^{n\times n}$. Show that D is normalized, multilinear, and antisymmetric (Definition 5.1.9), but not alternating. Explain why no such D exists over a field in which $1 + 1 \neq 0$.

Solution on page 1447

Problem 5.2. STANDARD

(a) State Theorem 5.4.7 and Theorem 5.4.10.

(b) Prove $\det(AB) = \det A \det B$ by recognizing the determinant by its axioms. Then deduce the implication "A invertible $\Rightarrow \det A \neq 0$" from it, without elementary matrices.

(c) For which $n \geq 1$ is there a matrix $A \in \mathbb{Q}^{n\times n}$ with $A^2 = 2I_n$? What changes over $\mathbb{R}$?

Solution on page 1448

Problem 5.3. STANDARD

(a) State Cramer's rule.

(b) Let $B \in F^{n\times n}$ and $\mathbf{y} \in F^n$, with no assumption on $\det B$. Prove that every solution of $B\mathbf{x} = \mathbf{y}$ satisfies $x_i \det B = \det B_i(\mathbf{y})$ for $i = 1, \ldots, n$, where $B_i(\mathbf{y})$ is B with column i replaced by $\mathbf{y}$. Deduce Cramer's rule.

(c) Let $B \in \mathbb{Q}^{n\times n}$ have integer entries and $d = \det B \neq 0$. Show that for every $\mathbf{y}$ with integer entries, the solution of $B\mathbf{x} = \mathbf{y}$ has $d\mathbf{x}$ integral. Show also that if $d \neq \pm 1$, then some $\mathbf{y}$ with integer entries has a solution that is not integral, and find one for $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$.

Solution on page 1449

Synthesis problems

Problem 5.4. STANDARD Prove or disprove each statement.

(a) If $A, B \in \mathbb{R}^{n\times n}$, with $n \geq 2$, have positive determinants, then so does $A + B$.

(b) If $A \in F^{n\times n}$ satisfies $A^3 = A$, then $\det A \in \{-1, 0, 1\}$.

(c) If $A \in F^{m\times n}$, then $\det\left(A^{\mathsf{T}}A\right) = \det\left(AA^{\mathsf{T}}\right)$.

(d) If every row of $A \in F^{n\times n}$ sums to 0, then $\det A = 0$.

(e) If every row of $A \in F^{n\times n}$ sums to 1, then $\det A = 1$.

Solution on page 1450

Problem 5.5. STANDARD Let p be a prime.

(a) Count the symmetric matrices $\begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathbb{F}_p^{2\times 2}$ with determinant 0.

(b) Count the matrices $A \in \mathbb{F}_p^{2\times 2}$ with $\operatorname{tr} A = 0$ and $\det A = 0$. Explain why the answers to (a) and (b) agree.

(c) Show that $A^2 - (\operatorname{tr} A)A + (\det A)I_2 = 0$ for every $A \in F^{2\times 2}$. Deduce that $A^2 = 0$ if and only if $\operatorname{tr} A = \det A = 0$, and count the $A \in \mathbb{F}_p^{2\times 2}$ with $A^2 = 0$.

Solution on page 1450

Problem 5.6. STANDARD For points $\mathbf{p}_1, \ldots, \mathbf{p}_m \in \mathbb{R}^2$ with $m \geq 3$, put $\mathbf{p}_{m+1} = \mathbf{p}_1$ and

$$S(\mathbf{p}_1, \ldots, \mathbf{p}_m) = \frac{1}{2}\sum_{k=1}^{m} \det(\mathbf{p}_k, \mathbf{p}_{k+1}),$$

where $\det(\mathbf{u}, \mathbf{v})$ is the determinant of the matrix with columns $\mathbf{u}$ and $\mathbf{v}$.

(a) Show that S does not change when every $\mathbf{p}_k$ is replaced by $\mathbf{p}_k + \mathbf{c}$, for a fixed $\mathbf{c} \in \mathbb{R}^2$.

(b) Show that $S(\mathbf{p}, \mathbf{q}, \mathbf{r}) = \frac{1}{2}\det(\mathbf{q} - \mathbf{p}, \mathbf{r} - \mathbf{p})$. This is the signed area of the triangle $\mathbf{pqr}$: its area, with a plus sign when $\mathbf{p}, \mathbf{q}, \mathbf{r}$ run counterclockwise (Example 5.1.15).

(c) Show that $S(\mathbf{p}_1, \ldots, \mathbf{p}_m) = \sum_{k=1}^{m} S(\mathbf{c}, \mathbf{p}_k, \mathbf{p}_{k+1})$ for every $\mathbf{c} \in \mathbb{R}^2$. Explain why S is the area of a convex polygon whose vertices $\mathbf{p}_1, \ldots, \mathbf{p}_m$ are listed counterclockwise. (Take for granted that the area of a region cut into triangles is the sum of their areas.)

(d) Find the area of the pentagon with vertices $(0, 0), (4, 0), (5, 3), (2, 5), (-1, 3)$.

Solution on page 1451

Problem 5.7. STANDARD Let $\mathbf{p}_0, \ldots, \mathbf{p}_n \in F^n$, and let $M \in F^{(n+1)\times(n+1)}$ have column $i + 1$ equal to $(1, \mathbf{p}_i)$, the entry 1 followed by the entries of $\mathbf{p}_i$.

(a) Show that $\det M = \det(\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_n - \mathbf{p}_0)$.

(b) Deduce that $\det M \neq 0$ if and only if $\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_n - \mathbf{p}_0$ are linearly independent.

(c) Suppose $\det M \neq 0$. Show that for each $\mathbf{x} \in F^n$ there are unique $\lambda_0, \ldots, \lambda_n \in F$ with $\lambda_0 + \cdots + \lambda_n = 1$ and $\lambda_0\mathbf{p}_0 + \cdots + \lambda_n\mathbf{p}_n = \mathbf{x}$, namely $\lambda_i = \det M_i / \det M$, where M_i is M with column $i + 1$ replaced by $(1, \mathbf{x})$.

(d) Let $F = \mathbb{R}$ and $n = 2$. Using Problem 5.6 (b), express each λ_i as a ratio of signed areas of triangles. Deduce that $\lambda_0, \lambda_1, \lambda_2 > 0$ exactly when the triangles $\mathbf{x}\mathbf{p}_1\mathbf{p}_2$, $\mathbf{p}_0\mathbf{x}\mathbf{p}_2$, and $\mathbf{p}_0\mathbf{p}_1\mathbf{x}$ have the same orientation as $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$, and compute the λ_i for $\mathbf{x} = (1, 2)$ and the triangle with vertices $(0, 0), (4, 0), (0, 4)$.

Solution on page 1452

Problem 5.8. STANDARD Let $\alpha, \beta \in F$, and let $T \in \mathrm{L}(F^{n\times n})$ be $T(X) = \alpha X + \beta X^{\mathsf{T}}$. Write U_{ij} for the matrix units (Lemma 2.4.24).

(a) Suppose $1 + 1 \neq 0$ in F. Show that the matrices U_{ii}, $U_{ij} + U_{ji}$, and $U_{ij} - U_{ji}$ (for $i < j$) form a basis of $F^{n\times n}$ in which the matrix of T is diagonal, and deduce that

$$\det T = (\alpha + \beta)^{n(n+1)/2}(\alpha - \beta)^{n(n-1)/2}.$$

(b) Check (a) against Exercise 5.4.40.

(c) Show that the formula holds for $n = 2$ over every field, including $\mathbb{F}_2$, by using the basis $U_{11}, U_{12}, U_{21}, U_{22}$.

(d) Keep $1 + 1 \neq 0$. For which α, β is T invertible? Describe $\ker T$ when $\alpha = \beta \neq 0$.

Solution on page 1453

Problem 5.9. STANDARD Let $x_1, \ldots, x_n \in F$, let $s_k = x_1^k + \cdots + x_n^k$ for $k \geq 0$ (so $s_0 = 1 + \cdots + 1$, with n terms), and let $H \in F^{n\times n}$ be the *Hankel matrix* with (i, j) entry s_{i+j-2}:

$$H = \begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{pmatrix}.$$

(a) Show that $H = V^{\mathsf{T}}V$ for the Vandermonde matrix $V = V(x_1, \ldots, x_n)$, and deduce that $\det H = \prod_{1\leq i<j\leq n} (x_j - x_i)^2$.

(b) Check (a) directly for $n = 2$.

(c) Let $F = \mathbb{R}$. Show that $\det H \geq 0$, with equality exactly when two of the x_i coincide. Show by an example that $\det H$ can be negative when $F = \mathbb{C}$.

(d) Real numbers x_1, x_2, x_3 have $s_1 = 0$, $s_2 = 6$, $s_3 = 0$, and $s_4 = 18$. Are they distinct?

Solution on page 1454

Problem 5.10. STANDARD For $a \in F$ and $n \geq 1$, let $K_n \in F^{n\times n}$ have (i, j) entry $a^{|i-j|}$, where $a^0 = 1$. For example,

$$K_3 = \begin{pmatrix} 1 & a & a^2 \\ a & 1 & a \\ a^2 & a & 1 \end{pmatrix}.$$

(a) Show that $\det K_n = \left(1 - a^2\right)^{n-1}$.

(b) Compute $\operatorname{adj} K_3$, and find K_3^{-1} when $a^2 \neq 1$.

(c) Let $n \geq 2$ and $a^2 \neq 1$. Guess K_n^{-1} from (b), and verify your guess.

Solution on page 1454

Problem 5.11. HARD Call $N \in F^{n\times n}$ *nilpotent* if $N^k = 0$ for some $k \geq 1$.

(a) Find the error in the following argument, and show that its claim is false.

Claim. If $A \in F^{n\times n}$ is invertible and $N \in F^{n\times n}$ is nilpotent, then $A+N$ is invertible.

Proof. Write $A+N = A\big(I_n + A^{-1}N\big)$. Since N is nilpotent, so is $M = A^{-1}N$, say $M^k = 0$. Then

$$(I_n + M)\Big(I_n - M + M^2 - \cdots + (-1)^{k-1}M^{k-1}\Big)$$
$$= I_n + (-1)^{k-1}M^k = I_n,$$

so $I_n + M$ is invertible, and $A+N$ is a product of invertible matrices.

(b) Show that the claim is true when $AN = NA$.

(c) Let $N \in \mathbb{C}^{n\times n}$ be nilpotent. Show that $\det(I_n + N) = 1$. (Hint: $t \mapsto \det(I_n + tN)$ is a polynomial function of t with no zeros.)

(d) Deduce that if $A \in \mathbb{C}^{n\times n}$ is invertible, N is nilpotent, and $AN = NA$, then $\det(A+N) = \det A$.

Solution on page 1455

Problem 5.12. HARD Let $n \geq 2$ and $c_0, \ldots, c_{n-1} \in \mathbb{C}$, and let $C \in \mathbb{C}^{n\times n}$ be the *circulant* matrix whose rows are the successive cyclic shifts of $(c_0, c_1, \ldots, c_{n-1})$ to the right. With rows and columns indexed by $0, \ldots, n-1$, its (j,k) entry is $c_{(k-j) \bmod n}$; for $n = 3$,

$$C = \begin{pmatrix} c_0 & c_1 & c_2 \\ c_2 & c_0 & c_1 \\ c_1 & c_2 & c_0 \end{pmatrix}.$$

Let $f(z) = c_0 + c_1 z + \cdots + c_{n-1}z^{n-1}$ and $\omega = \cos(2\pi/n) + i\sin(2\pi/n)$. You may use that $1, \omega, \ldots, \omega^{n-1}$ are distinct, that each of them satisfies $z^n = 1$, and that $\omega^{n/2} = -1$ when n is even (Section B.1).

(a) For $0 \leq m \leq n-1$, let $\mathbf{v}_m = \big(1, \omega^m, \omega^{2m}, \ldots, \omega^{(n-1)m}\big)$. Show that $C\mathbf{v}_m = f(\omega^m)\mathbf{v}_m$.

(b) Deduce that $\det C = f(1)f(\omega)\cdots f\big(\omega^{n-1}\big)$.

(c) Deduce that $\det\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix} = (a+b+c)\big(a+\omega b+\omega^2 c\big)\big(a+\omega^2 b+\omega c\big)$, where $\omega = -\frac{1}{2} + \big(\sqrt{3}/2\big)i$.

(d) Let C have first row $(1, 1, 0, \ldots, 0)$. Show that C is singular if and only if n is even, and that $\det C = 2$ when n is odd.

Solution on page 1456

Problem 5.13. HARD Let $n \geq 1$, let $x_1, \ldots, x_n, y_1, \ldots, y_n \in F$, and let $M \in F^{n\times n}$ have entries $m_{ij} = \big(x_i + y_j\big)^{n-1}$.

(a) Show that

$$\det M = \prod_{k=0}^{n-1}\binom{n-1}{k} \cdot \prod_{1\leq i<j\leq n}(x_j - x_i)(y_i - y_j),$$

where the binomial coefficients are read as elements of F.

(b) Check (a) directly for $n = 2$.

(c) For $n = 3$, decide when M is invertible, over $\mathbb{Q}$ and over $\mathbb{F}_2$.

Solution on page 1457

Problem 5.14. HARD Let $T \in F^{n\times n}$ be tridiagonal, with diagonal entries $a_1, \ldots, a_n$, entries b_k in positions $(k, k+1)$, and entries c_k in positions $(k+1, k)$, as in Proposition 5.5.18.

(a) Show that the permutations $\sigma \in S_n$ with $|\sigma(j) - j| \le 1$ for every j are exactly the products of disjoint adjacent transpositions $(k\ k+1)$ (the empty product included).

(b) Show that there are F_{n+1} such permutations, where $F_1 = F_2 = 1$ and $F_{m+1} = F_m + F_{m-1}$ are the Fibonacci numbers.

(c) Call a set P of pairwise disjoint pairs $\{k, k+1\}$ a *matching* of $[n]$, and call i *unmatched* if it lies in no pair of P. Show that

$$\det T = \sum_{P} \prod_{\{k,k+1\}\in P} (-b_k c_k) \prod_{i \text{ unmatched}} a_i,$$

the sum over all matchings P of $[n]$.

(d) Derive the recurrence of Proposition 5.5.18 from (c). Explain the Fibonacci numbers in Example 5.5.19 (b) without computing.

Solution on page 1458

Problem 5.15. HARD Let $A \in \mathbb{Q}^{n\times n}$ have integer entries, with a_{ii} even for every i and $a_{ij} + a_{ji}$ even for all i, j. (For example, A could be symmetric with even diagonal, or skew-symmetric.)

(a) Show that if n is odd, then $\det A$ is even.

(b) Show that the conclusion of (a) fails for every even n.

(c) Show that the hypothesis on the diagonal cannot be dropped, and check (a) on $J + I_n$, where every entry of J is 1.

Solution on page 1458

Problem 5.16. HARD Let $A \in \mathbb{R}^{n\times n}$ be skew-symmetric, $A^{\mathsf{T}} = -A$.

(a) Show that $\mathbf{x}^{\mathsf{T}} A\mathbf{x} = 0$ for every $\mathbf{x} \in \mathbb{R}^n$, and deduce that $I_n + tA$ is invertible for every $t \in \mathbb{R}$.

(b) Show that $\det(I_n + A) > 0$. (Consider $t \mapsto \det(I_n + tA)$ on $[0, 1]$.)

(c) Compute $\det(I_n + A)$ for $n = 2$ and $n = 3$ in terms of the entries above the diagonal.

(d) Show that the conclusion of (b) can fail if A is only assumed to have $\operatorname{tr} A = 0$.

Solution on page 1459

Problem 5.17. HARD Let W be a subspace of $F^{n\times n}$ in which every nonzero matrix is invertible.

(a) Show that $\dim W \le n$. (Consider the vectors $X\mathbf{e}_1$ for X in a basis of W.)

(b) Show that $W = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^{2\times 2}$ with this property and $\dim W = 2$.

(c) Let $F = \mathbb{C}$ and $n \ge 2$. Show that $\dim W \le 1$. (Take independent $A, B \in W$ and look at $\det(A + tB)$.)

(d) Let $F = \mathbb{R}$ and let n be odd. Show that $\dim W \le 1$.

Solution on page 1459

Problem 5.18. HARD Let $x, a, b \in F$ with $a \ne b$, and let $T_n \in F^{n\times n}$ have every diagonal entry equal to x, every entry above the diagonal equal to a, and every entry below the diagonal equal to b. Show that

$$\det T_n = \frac{b(x-a)^n - a(x-b)^n}{b-a},$$

and check the formula for $n = 2$. (Hint: let J be the matrix with every entry 1, and show that $t \mapsto \det(T_n + tJ)$ is affine in t.)

Solution on page 1460

Problem 5.19. CHALLENGE

(a) Let $x_1, \dots, x_n \in F$, and for $k = 0, \dots, n-1$ let $p_k \in F[x]$ have degree k and coefficient 1 of x^k. Show that the $n \times n$ matrix with $(i, k+1)$ entry $p_k(x_i)$ has determinant $\prod_{1 \le i < j \le n} (x_j - x_i)$.

(b) Let $x_1, \dots, x_n$ be integers. Show that $\prod_{1 \le i < j \le n} (x_j - x_i)$ is divisible by $0! \cdot 1! \cdots (n-1)!$.

(c) Deduce that $\prod_{1 \le i < j \le n} \frac{x_j - x_i}{j - i}$ is an integer for all integers $x_1, \dots, x_n$.

Solution on page 1461

Problem 5.20. CHALLENGE Let $G_n \in \mathbb{Q}^{n \times n}$ have entries $g_{ij} = \gcd(i, j)$. For $d \ge 1$, let $\varphi(d)$ be the number of $k \in \{1, \dots, d\}$ with $\gcd(k, d) = 1$. You may use that $\sum_{d \mid m} \varphi(d) = m$ for every integer $m \ge 1$.

(a) Show that $G_n = L \Phi L^{\mathsf{T}}$, where $\Phi = \operatorname{diag}(\varphi(1), \dots, \varphi(n))$ and $L \in \mathbb{Q}^{n \times n}$ has (i, d) entry 1 if $d \mid i$ and 0 otherwise.

(b) Deduce that $\det G_n = \varphi(1)\varphi(2) \cdots \varphi(n)$, and evaluate $\det G_6$.

(c) Show that G_n is invertible for every n, and that G_n^{-1} has integer entries only for $n \le 2$.

Solution on page 1462

Chapter 6

Inner product spaces

Inner products and norms, Cauchy–Schwarz, orthonormal bases, projections, least squares, and adjoints.

6.1 Inner products and norms

In $\mathbb{R}^2$ the dot product $\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2$ knows the length of every vector and the angle between any two. This section takes the algebraic properties of the dot product that are responsible for this and makes them axioms. A vector space over $\mathbb{R}$ or $\mathbb{C}$ equipped with a product satisfying them, an inner product, acquires lengths and angles, even when its vectors are functions or matrices and there is no picture to look at.

6.1.1 Length and angle in $\mathbb{R}^2$

For $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ in $\mathbb{R}^2$, the **dot product** is

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 \in \mathbb{R}.$$

Draw $\mathbf{v}$ as an arrow from the origin. Its length $\|\mathbf{v}\|$ is the hypotenuse of a right triangle with legs $|v_1|$ and $|v_2|$, so by Pythagoras

$$\|\mathbf{v}\|^2 = v_1^2 + v_2^2 = \mathbf{v} \cdot \mathbf{v}.$$

Now let $\mathbf{v}$ and $\mathbf{w}$ be nonzero, and let $\theta \in [0, \pi]$ be the angle between them. The arrows $\mathbf{v}$ and $\mathbf{w}$, together with the arrow $\mathbf{v} - \mathbf{w}$ from the tip of $\mathbf{w}$ to the tip of $\mathbf{v}$, form a triangle (Figure 6.1.1), and the law of cosines says

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\| \cos\theta.$$

We can also expand the left side algebraically. The dot product is symmetric, $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$, and it is linear in each slot, so

$$\|\mathbf{v} - \mathbf{w}\|^2 = (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} - 2\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w}.$$

Comparing the two expressions, and using $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$ and $\|\mathbf{w}\|^2 = \mathbf{w} \cdot \mathbf{w}$, we get

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta. \tag{6.1.1}$$

So lengths and angles can both be read off from the dot product. In particular, $\mathbf{v}$ and $\mathbf{w}$ are perpendicular exactly when $\mathbf{v} \cdot \mathbf{w} = 0$, which is where Section 6.3 starts.

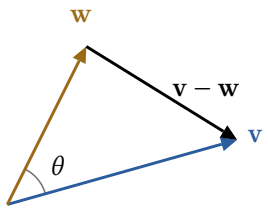


Figure 6.1.1. The triangle behind (6.1.1). Its sides have lengths $\|\mathbf{v}\|$, $\|\mathbf{w}\|$, and $\|\mathbf{v}-\mathbf{w}\|$, and the angle between the first two is θ.

Look at what the computation used: symmetry, linearity in each slot, and the fact that $\mathbf{v}\cdot\mathbf{v}$ is a nonnegative real number that vanishes only for $\mathbf{v}=\mathbf{0}$, so that $\|\mathbf{v}\|=\sqrt{\mathbf{v}\cdot\mathbf{v}}$ makes sense and is positive when $\mathbf{v}\neq\mathbf{0}$. Turned around, any product with these properties, on any vector space, lets us *define* the length of a vector as $\sqrt{\mathbf{v}\cdot\mathbf{v}}$, and (6.1.1) suggests a definition of the angle between two vectors. (Section 6.2 shows that the resulting value of $\cos\theta$ lies in $[-1,1]$.)

Which scalars?

The axioms will compare $\mathbf{v}\cdot\mathbf{v}$ with 0, and lengths are square roots. So the scalars need an order and square roots of their nonnegative elements. $\mathbb{R}$ has both, and the lack of either causes real trouble.

- In $\mathbb{Q}^2$, the vector $(1,1)$ would have length $\sqrt{2}\notin\mathbb{Q}$.
- In $\mathbb{F}_2^2$, the same formula gives $(1,1)\cdot(1,1)=1+1=0$, so a nonzero vector would have length 0. Ordering $\mathbb{F}_p$ cannot help: an order compatible with the field operations would force $0<1<1+1<\cdots<p\cdot 1=0$.
- In $\mathbb{C}^2$, $(1,i)\cdot(1,i)=1+i^2=0$, and $(1,1+i)\cdot(1,1+i)=1+2i$ is not even real.

$\mathbb{C}$ has no useful order either, but it contains $\mathbb{R}$, and the case $n=1$ shows how to land in $\mathbb{R}$. Viewing $z=a+bi$ as the point (a,b) of the plane, its length is $\sqrt{a^2+b^2}$, and $a^2+b^2=z\overline{z}$, not z^2. So on $\mathbb{C}^n$ we conjugate the second factor and use $\sum_k v_k\overline{w_k}$. Then $\mathbf{v}\cdot\mathbf{v}=\sum_k|v_k|^2$ is real and nonnegative, and the price is that symmetry becomes "conjugate symmetry". In this chapter the scalars are always real or complex.

Notation 6.1.1 (Real and complex scalars). Throughout Chapter 6, $\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$. For $z=a+bi\in\mathbb{C}$ with $a,b\in\mathbb{R}$, the **conjugate** of z is $\overline{z}=a-bi$, its **modulus** is $|z|=\sqrt{a^2+b^2}$, and its **real** and **imaginary parts** are $\operatorname{Re}z=a$ and $\operatorname{Im}z=b$. For all $z,w\in\mathbb{C}$,

$$\overline{z+w}=\overline{z}+\overline{w},\quad \overline{zw}=\overline{z}\,\overline{w},\quad \overline{\overline{z}}=z,$$
$$z\overline{z}=|z|^2,\quad z+\overline{z}=2\operatorname{Re}z,\quad \operatorname{Re}(-iz)=\operatorname{Im}z,$$

$|\operatorname{Re}z|\le|z|$, and $\overline{z}=z$ if and only if $z\in\mathbb{R}$. Each of these is checked by writing z and w in coordinates; Section B.1 develops conjugation and the modulus further. For real z, $\overline{z}=z$ and $|z|$ is the absolute value, so every formula in this chapter specializes to $\mathbb{R}$ by erasing the bars.

6.1.2 Inner products

Definition 6.1.2 (Inner product). Let V be a vector space over $\mathbb{F}$. An **inner product** on V is a function $V\times V\to\mathbb{F}$, written $(\mathbf{v},\mathbf{w})\mapsto\langle\mathbf{v},\mathbf{w}\rangle$, such that for all $\mathbf{v},\mathbf{w},\mathbf{x}\in V$ and all $a,b\in\mathbb{F}$:

- *Conjugate symmetry.* $\langle\mathbf{v},\mathbf{w}\rangle=\overline{\langle\mathbf{w},\mathbf{v}\rangle}$.
- *Linearity in the first slot.* $\langle a\mathbf{v}+b\mathbf{w},\mathbf{x}\rangle=a\langle\mathbf{v},\mathbf{x}\rangle+b\langle\mathbf{w},\mathbf{x}\rangle$.
- *Positivity.* $\langle\mathbf{v},\mathbf{v}\rangle\ge 0$.
- *Definiteness* (or *non-degeneracy*). $\langle\mathbf{v},\mathbf{v}\rangle=0$ if and only if $\mathbf{v}=\mathbf{0}$.

An **inner product space** is a vector space over $\mathbb{F}$ together with a specified inner product on it.

Remark 6.1.3. Taking $\mathbf{w} = \mathbf{v}$ in conjugate symmetry gives $\langle \mathbf{v}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{v} \rangle}$, so $\langle \mathbf{v}, \mathbf{v} \rangle$ is real for every $\mathbf{v}$, and positivity makes sense even when $\mathbb{F} = \mathbb{C}$. Over $\mathbb{R}$ conjugation does nothing, and conjugate symmetry is plain symmetry.

Half of definiteness is automatic: linearity gives $\langle \mathbf{0}, \mathbf{0} \rangle = \langle 0\mathbf{0}, \mathbf{0} \rangle = 0\langle \mathbf{0}, \mathbf{0} \rangle = 0$. So to verify definiteness one only has to show that $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ forces $\mathbf{v} = \mathbf{0}$.

A vector space usually carries many inner products, and lengths and angles depend on the choice (Example 6.1.21). So we always say which inner product we use.

Linearity is required only in the first slot. The second slot inherits a twisted version of it from conjugate symmetry.

Proposition 6.1.4 (Sesquilinearity). Let V be an inner product space. For all $\mathbf{v}, \mathbf{w}, \mathbf{x} \in V$ and $a, b \in \mathbb{F}$:

(a) $\langle \mathbf{v}, a\mathbf{w} + b\mathbf{x} \rangle = \overline{a}\langle \mathbf{v}, \mathbf{w} \rangle + \overline{b}\langle \mathbf{v}, \mathbf{x} \rangle$;

(b) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$.

Proof. ROUTINE

Uses. Definition 6.1.2 (Inner product), Notation 6.1.1 (Real and complex scalars)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. The only axiom that says anything about the second slot is conjugate symmetry. Use it to move the second slot into the first.
Hint 2. Flip with conjugate symmetry, expand with linearity in the first slot, and flip back. Conjugation turns a sum into a sum and a product into a product.
Hint 3. For (b), write $\mathbf{0} = \mathbf{0} + \mathbf{0}$ in the first slot.
Sketch. Conjugate symmetry turns $\langle \mathbf{v}, a\mathbf{w} + b\mathbf{x} \rangle$ into the conjugate of $a\langle \mathbf{w}, \mathbf{v} \rangle + b\langle \mathbf{x}, \mathbf{v} \rangle$, which is $\overline{a}\langle \mathbf{v}, \mathbf{w} \rangle + \overline{b}\langle \mathbf{v}, \mathbf{x} \rangle$. For (b), $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0} + \mathbf{0}, \mathbf{v} \rangle = 2\langle \mathbf{0}, \mathbf{v} \rangle$, and conjugate symmetry handles $\langle \mathbf{v}, \mathbf{0} \rangle$.

(a) We compute

$$\begin{aligned}
\langle \mathbf{v}, a\mathbf{w} + b\mathbf{x} \rangle &= \overline{\langle a\mathbf{w} + b\mathbf{x}, \mathbf{v} \rangle} && \text{(conjugate symmetry)} \\
&= \overline{a\langle \mathbf{w}, \mathbf{v} \rangle + b\langle \mathbf{x}, \mathbf{v} \rangle} && \text{(linearity in the first slot)} \\
&= \overline{a}\,\overline{\langle \mathbf{w}, \mathbf{v} \rangle} + \overline{b}\,\overline{\langle \mathbf{x}, \mathbf{v} \rangle} && \text{(conjugation respects } + \text{ and } \cdot\text{)} \\
&= \overline{a}\langle \mathbf{v}, \mathbf{w} \rangle + \overline{b}\langle \mathbf{v}, \mathbf{x} \rangle && \text{(conjugate symmetry),}
\end{aligned}$$

as desired.

(b) By linearity in the first slot, $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0} + \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle + \langle \mathbf{0}, \mathbf{v} \rangle$. Subtracting $\langle \mathbf{0}, \mathbf{v} \rangle$ from both sides gives $\langle \mathbf{0}, \mathbf{v} \rangle = 0$. Then $\langle \mathbf{v}, \mathbf{0} \rangle = \overline{\langle \mathbf{0}, \mathbf{v} \rangle} = \overline{0} = 0$. □

A function of two vectors that is linear in the first slot and satisfies (a) in the second is called **sesquilinear** ("one and a half times linear"). Over $\mathbb{R}$ the bars disappear: a real inner product is linear in each slot, that is, **bilinear**, and symmetric. Over $\mathbb{C}$, scalars leave the second slot conjugated, as in $\langle \mathbf{v}, i\mathbf{w} \rangle = -i\langle \mathbf{v}, \mathbf{w} \rangle$. Forgetting the bar is the most common error in computations with complex inner products.

Example 6.1.5 (Expanding an inner product). Let $\mathbf{x}, \mathbf{y}$ be vectors in a complex inner product space with $\langle \mathbf{x}, \mathbf{x} \rangle = 1$, $\langle \mathbf{y}, \mathbf{y} \rangle = 2$, and $\langle \mathbf{x}, \mathbf{y} \rangle = i$, so that $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{i} = -i$. To compute the inner product below, expand each slot, pulling scalars out of the first slot as they are and out of the second slot conjugated:

$$\begin{aligned}
&\langle (1+i)\mathbf{x} + \mathbf{y}, 2\mathbf{x} - i\mathbf{y} \rangle \\
&= (1+i)\cdot 2\langle \mathbf{x}, \mathbf{x} \rangle + (1+i)\overline{-i}\langle \mathbf{x}, \mathbf{y} \rangle \\
&\quad + 2\langle \mathbf{y}, \mathbf{x} \rangle + \overline{-i}\langle \mathbf{y}, \mathbf{y} \rangle \\
&= 2(1+i) + (1+i)i \cdot i + 2(-i) + i \cdot 2
\end{aligned}$$

$$= (2+2i) + (-1-i) - 2i + 2i = 1 + i.$$

Here $(1+i)i \cdot i = (1+i)(-1)$. This is Technique 5.1.7 for a sesquilinear function: each slot is split into its terms, and the four resulting values are known.

Over $\mathbb{C}$ the conjugation is forced on us, because positivity and definiteness together are incompatible with linearity in both slots.

Proposition 6.1.6 (Bilinear forms over $\mathbb{C}$ are not positive definite). Let $V \neq \{\mathbf{0}\}$ be a complex vector space, and let $B : V \times V \to \mathbb{C}$ be linear in each slot when the other is fixed. If $B(\mathbf{v}, \mathbf{v})$ is real and nonnegative for every $\mathbf{v} \in V$, then $B(\mathbf{v}, \mathbf{v}) = 0$ for every $\mathbf{v} \in V$. In particular, B cannot satisfy both positivity and definiteness.

Proof. ROUTINE

Uses. Definition 6.1.2 (Inner product)
Hint 1. Compare $B(\mathbf{v}, \mathbf{v})$ with $B(c\mathbf{v}, c\mathbf{v})$ for a scalar c.
Hint 2. Take $c = i$. What is $i \cdot i$?
Sketch. $B(i\mathbf{v}, i\mathbf{v}) = i^2 B(\mathbf{v}, \mathbf{v}) = -B(\mathbf{v}, \mathbf{v})$, and a real number whose negative is also nonnegative is 0. A nonzero $\mathbf{v}$ then violates definiteness.

Let $\mathbf{v} \in V$. Linearity in each slot gives $B(i\mathbf{v}, i\mathbf{v}) = i \cdot iB(\mathbf{v}, \mathbf{v}) = -B(\mathbf{v}, \mathbf{v})$. Both $B(\mathbf{v}, \mathbf{v})$ and $B(i\mathbf{v}, i\mathbf{v}) = -B(\mathbf{v}, \mathbf{v})$ are nonnegative real numbers, so $B(\mathbf{v}, \mathbf{v}) = 0$. Since $V \neq \{\mathbf{0}\}$, there is some $\mathbf{v} \neq \mathbf{0}$, and $B(\mathbf{v}, \mathbf{v}) = 0$ violates definiteness. □

With the conjugate in the second slot the problem disappears: $\langle i\mathbf{v}, i\mathbf{v}\rangle = \bar{i}i\langle \mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle$. This is what it should be, since multiplying by i rotates each coordinate of $\mathbf{v} \in \mathbb{C}^n$ without changing its modulus, so it should not change the length of $\mathbf{v}$.

6.1.3 Examples and non-examples

The dot product

The dot product on $\mathbb{C}^n$ is most compactly written with the conjugate transpose, which Section 6.3 studies in detail.

Notation 6.1.7 (Conjugate transpose). For $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, the **conjugate transpose** of A is the matrix $A^* \in \mathbb{C}^{n \times m}$ whose (i, j) entry is $\overline{a_{ji}}$. That is, $A^* = \overline{A}^{\mathsf{T}}$, where $\overline{A} = (\overline{a_{ij}})$ conjugates every entry (Definition 2.4.21). For a real matrix, $A^* = A^{\mathsf{T}}$. For $\mathbf{w} \in \mathbb{F}^n$, $\mathbf{w}^*$ is the row $(\overline{w_1}, \ldots, \overline{w_n})$, and for $\mathbf{v} \in \mathbb{F}^n$ the product $\mathbf{w}^*\mathbf{v}$ is a 1×1 matrix, which we identify with its entry $\sum_k \overline{w_k} v_k$.

Definition 6.1.8 (Dot product). The **dot product** on $\mathbb{F}^n$ is

$$\mathbf{v} \cdot \mathbf{w} = \sum_{k=1}^{n} v_k \overline{w_k} = \mathbf{w}^* \mathbf{v}$$

for $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ in $\mathbb{F}^n$. Over $\mathbb{R}$ it is $\sum_k v_k w_k = \mathbf{w}^{\mathsf{T}}\mathbf{v}$.

Checking the axioms for the dot product comes down to one elementary fact about real numbers, which we name because it is used constantly in this chapter.

Technique 6.1.9 (A sum of nonnegative terms). A sum of nonnegative real numbers is 0 only if every term is 0: if $t_1, \ldots, t_n \geq 0$ and some $t_j > 0$, then $t_1 + \cdots + t_n \geq t_j > 0$. So to show that many quantities vanish, it suffices to exhibit them as squares, or squared moduli $|z_k|^2$, in a single sum, and to show that the sum is 0.

Reach for it when a positivity or definiteness axiom must be verified for a concrete formula, and whenever one equation should force many. The canonical example is definiteness of the dot product (Proposition 6.1.10), where $\sum_k |v_k|^2 = 0$ forces every $v_k = 0$. Integrals of nonnegative continuous functions behave the same way (Lemma 6.1.14), and Technique 6.1.18 applies the idea to the entries of a matrix.
Used in. Exercise 1.5.48, Problem 3.15, Exercise 4.3.33, Proposition 6.1.10, Lemma 6.1.14, Proposition 6.1.13, Exercise 6.1.46, Lemma 6.2.6, Exercise 6.2.43, Proposition 6.3.13, Exercise 6.3.52, Theorem 6.4.33, Exercise 6.7.12, Problem 6.5, Problem 6.12, Problem 6.15, Problem 6.16, Theorem 8.1.3, Lemma 8.2.4, Lemma 8.3.8, Notation 8.3.10, Exercise 8.3.40, Lemma 8.4.24, Exercise 6.1.42, Exercise 6.2.47, Exercise 6.4.60, Exercise 8.4.35.

Proposition 6.1.10 (The dot product is an inner product). The dot product is an inner product on $\mathbb{F}^n$, and $\mathbf{v} \cdot \mathbf{v} = \sum_{k=1}^{n} |v_k|^2$ for all $\mathbf{v} \in \mathbb{F}^n$.

Proof. ROUTINE

Uses. Definition 6.1.2 (Inner product), Definition 6.1.8 (Dot product), Notation 6.1.1 (Real and complex scalars)
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Check the axioms one at a time with the coordinate formula $\sum_k v_k \overline{w_k}$.
Hint 2. Conjugate symmetry is $\overline{w_k \overline{v_k}} = \overline{w_k} v_k$ for each term. For positivity, $v_k \overline{v_k} = |v_k|^2$.
Hint 3. For definiteness, a sum of nonnegative real numbers is 0 only if each term is.
Sketch. Conjugating $\mathbf{w} \cdot \mathbf{v}$ term by term gives $\mathbf{v} \cdot \mathbf{w}$, and linearity in $\mathbf{v}$ is visible from the formula. Then $\mathbf{v} \cdot \mathbf{v} = \sum_k |v_k|^2 \geq 0$, and it vanishes only if every $|v_k| = 0$, that is, $\mathbf{v} = \mathbf{0}$.

Let $\mathbf{v}, \mathbf{w}, \mathbf{x} \in \mathbb{F}^n$ and $a, b \in \mathbb{F}$.

Conjugate symmetry. $\overline{\mathbf{w} \cdot \mathbf{v}} = \overline{\sum_k w_k \overline{v_k}} = \sum_k \overline{w_k} v_k = \mathbf{v} \cdot \mathbf{w}$.

Linearity in the first slot. The kth coordinate of $a\mathbf{v} + b\mathbf{w}$ is $av_k + bw_k$, so

$$(a\mathbf{v} + b\mathbf{w}) \cdot \mathbf{x} = \sum_k (av_k + bw_k)\overline{x_k} = a \sum_k v_k \overline{x_k} + b \sum_k w_k \overline{x_k} = a(\mathbf{v} \cdot \mathbf{x}) + b(\mathbf{w} \cdot \mathbf{x}).$$

Positivity. $\mathbf{v} \cdot \mathbf{v} = \sum_k v_k \overline{v_k} = \sum_k |v_k|^2 \geq 0$.

Definiteness. Suppose $\mathbf{v} \cdot \mathbf{v} = 0$. Then $\sum_k |v_k|^2 = 0$ is a sum of nonnegative real numbers, so $|v_k|^2 = 0$ for every k. Writing $v_k = a_k + b_k i$, this says $a_k^2 + b_k^2 = 0$, so $a_k = b_k = 0$ and $v_k = 0$. Hence $\mathbf{v} = \mathbf{0}$. By Remark 6.1.3, this is all that definiteness requires. □

Unless another inner product is specified, $\mathbb{F}^n$ carries the dot product. The lecture calls it "the dot product" precisely to distinguish it from the other inner products on $\mathbb{F}^n$ below.

Example 6.1.11 (Computing in $\mathbb{C}^2$). Let $\mathbf{v} = (1 + i, 2)$ and $\mathbf{w} = (i, 1 - i)$. Then

$$\mathbf{v} \cdot \mathbf{w} = (1+i)\bar{i} + 2\overline{(1-i)} = (1+i)(-i) + 2(1+i) = (1-i) + (2+2i) = 3 + i,$$
$$\mathbf{w} \cdot \mathbf{v} = i\overline{(1+i)} + (1-i)\overline{2} = i(1-i) + 2 - 2i = 3 - i,$$

as conjugate symmetry predicts. Also $\mathbf{v} \cdot \mathbf{v} = |1+i|^2 + |2|^2 = 2 + 4 = 6$ and $\mathbf{w} \cdot \mathbf{w} = 1 + 2 = 3$. Without the conjugate, $\mathbf{v}$ would have "squared length" $(1+i)^2 + 4 = 4 + 2i$.

Exercise 6.1.12. ROUTINE In $\mathbb{C}^3$, let $\mathbf{v} = (1, i, 2)$ and $\mathbf{w} = (i, 1 + i, 1 - i)$. Compute $\mathbf{v} \cdot \mathbf{w}$, $\mathbf{w} \cdot \mathbf{v}$, $\mathbf{v} \cdot \mathbf{v}$, and $\mathbf{w} \cdot \mathbf{w}$, and find $c \in \mathbb{C}$ with $(\mathbf{v} - c\mathbf{w}) \cdot \mathbf{w} = 0$.

Solution on page 1462

Continuous functions

Let $C([0,1], \mathbb{C})$ be the set of continuous functions $f : [0,1] \to \mathbb{C}$, that is, of functions $f = u + iv$ with $u, v \in C[0,1]$ (Example 1.2.12). It is a complex vector space under the pointwise operations, by the argument of Example 1.2.12. Such functions are integrated part by part, $\int_0^1 f = \int_0^1 u + i \int_0^1 v$. This integral is $\mathbb{C}$-linear and commutes with conjugation:

$$\int_0^1 (af+bg) = a\int_0^1 f + b\int_0^1 g \quad (a,b\in\mathbb{C}), \qquad \int_0^1 \overline{f} = \overline{\int_0^1 f}.$$

Both follow from the real case by splitting into real and imaginary parts. For instance, for $a = p+qi$ we have $af = (pu-qv)+i(pv+qu)$, whose integral is $(p+qi)\big(\int u + i\int v\big)$.

Proposition 6.1.13 (Integral inner products). The formula

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,\mathrm{d}t$$

defines an inner product on $C([0,1],\mathbb{C})$. On the real vector space $C[0,1]$ it reads $\langle f,g\rangle = \int_0^1 f(t)g(t)\,\mathrm{d}t$, and it is an inner product there as well.

Intuition. Conjugate symmetry and linearity are properties of the integral, and positivity holds because $f\overline{f} = |f|^2 \geq 0$. Definiteness is the one place where continuity matters. A function that is nonzero at only a single point has integral 0, so $\int_0^1 |f|^2 = 0$ does not force $f = 0$ for arbitrary integrable f. A continuous function that is positive at one point, however, is positive on a whole interval around it, and so has positive integral. This is Technique 6.1.9 with an integral in place of a sum.

Lemma 6.1.14 (Nonnegative functions with zero integral). Let $h \in C[0,1]$ with $h(t) \geq 0$ for all $t \in [0,1]$. If $\int_0^1 h = 0$, then $h = 0$.

Proof. STANDARD

Uses. continuity of h, monotonicity of the integral: if $g_1 \leq g_2$ on an interval, then their integrals compare the same way
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Prove the contrapositive: if $h(c) > 0$ for some c, then $\int_0^1 h > 0$.
Hint 2. By continuity, $h > h(c)/2$ on all points of $[0,1]$ close enough to c. Find an interval $I \subseteq [0,1]$ of positive length on which this holds.
Hint 3. Bound $\int_0^1 h$ below by the integral over I, using $h \geq 0$ off I.
Sketch. If $h(c) > 0$, continuity gives an interval $I \subseteq [0,1]$ of positive length ℓ on which $h > h(c)/2$. Since $h \geq 0$ everywhere, $\int_0^1 h \geq \int_I h \geq \ell h(c)/2 > 0$.

We prove the contrapositive. Suppose $h(c) > 0$ for some $c \in [0,1]$, and let $\varepsilon = h(c)/2 > 0$. By continuity of h at c, there is $\delta > 0$ such that $h(t) > \varepsilon$ for all $t \in [0,1]$ with $|t-c| < \delta$. Let $\ell = \min(\delta/2, 1/2)$. Either $c+\ell \leq 1$ or $c-\ell \geq 0$, since otherwise $1 < c+\ell$ and $c < \ell$, which give $1 < 2\ell \leq 1$. So one of the intervals $I = [c, c+\ell]$ or $I = [c-\ell, c]$ lies in $[0,1]$. Every $t \in I$ satisfies $|t-c| \leq \ell < \delta$, so $h > \varepsilon$ on I. Since $h \geq 0$ on all of $[0,1]$, monotonicity of the integral gives

$$\int_0^1 h \geq \int_I h \geq \ell\varepsilon > 0.$$

So if $\int_0^1 h = 0$, then $h(c) > 0$ for no c, and as $h \geq 0$, $h = 0$. □

Proof of Proposition 6.1.13. STANDARD

Uses. Definition 6.1.2 (Inner product), Lemma 6.1.14 (Nonnegative functions with zero integral), Notation 6.1.1 (Real and complex scalars), $\mathbb{C}$-linearity of the integral, and $\int \overline{f} = \overline{\int f}$
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. The products $f\overline{g}$ are continuous, so all the integrals exist. Check conjugate symmetry and linearity with the two properties of the integral above.
Hint 2. For positivity, $f\overline{f} = |f|^2$ is a nonnegative continuous real function.
Hint 3. For definiteness, apply Lemma 6.1.14 to $h = |f|^2$.
Sketch. $\langle g, f\rangle = \int g\overline{f} = \int \overline{f\overline{g}} = \overline{\langle f, g\rangle}$, and linearity in f is linearity of the integral. Then $\langle f, f\rangle = \int |f|^2 \geq 0$, and if it is 0, Lemma 6.1.14 gives $|f|^2 = 0$, so $f = 0$.

For $f, g \in C([0,1], \mathbb{C})$, the function $f\overline{g}$ is continuous, so $\langle f, g\rangle$ is defined. Let $f, g, h \in C([0,1], \mathbb{C})$ and $a, b \in \mathbb{C}$.

Conjugate symmetry. $\langle g, f\rangle = \int_0^1 g\overline{f} = \int_0^1 \overline{f\overline{g}} = \overline{\int_0^1 f\overline{g}} = \overline{\langle f, g\rangle}$.

Linearity in the first slot. $\langle af + bg, h\rangle = \int_0^1 \left(af\overline{h} + bg\overline{h}\right) = a\langle f, h\rangle + b\langle g, h\rangle$.

Positivity. Writing $f = u + iv$, we have $f\overline{f} = |f|^2 = u^2 + v^2$, a continuous function with values in $[0, \infty)$. So $\langle f, f\rangle = \int_0^1 |f|^2 \geq 0$.

Definiteness. If $\langle f, f\rangle = 0$, then Lemma 6.1.14 applied to $h = |f|^2$ gives $|f(t)|^2 = 0$, hence $f(t) = 0$, for every t. So $f = \mathbf{0}$.

For $C[0,1]$, the same argument applies with every bar erased. □

Nothing here depends on the interval being $[0,1]$. For any $a < b$, the proofs of Lemma 6.1.14 and Proposition 6.1.13 go through on $[a,b]$, with $(b-a)/2$ in place of $1/2$ in the choice of ℓ, so $\int_a^b f\overline{g}$ is an inner product on $C([a,b], \mathbb{C})$, and on $C[a,b]$ with the bars erased.

> *Intuition (Infinitely many coordinates).* Riemann sums explain why this integral behaves like a dot product. For large N,
>
> $$\langle f, g\rangle \approx \frac{1}{N}\sum_{k=1}^{N} f\left(\frac{k}{N}\right)\overline{g\left(\frac{k}{N}\right)} = \frac{1}{N}(\mathbf{f}_N \cdot \mathbf{g}_N),$$
>
> where $\mathbf{f}_N = (f(1/N), f(2/N), \ldots, f(1)) \in \mathbb{C}^N$ records the values of f at N sample points. So $\langle f, g\rangle$ is a limit of rescaled dot products of longer and longer vectors, with one coordinate for each point of $[0,1]$, and $\langle f, f\rangle$ is the average of $|f|^2$.

Example 6.1.15 (Computing with functions). In $C[0,1]$, $\langle t, t^2\rangle = \int_0^1 t^3\,dt = 1/4$ and $\langle 1, 2t-1\rangle = \int_0^1 (2t-1)\,dt = 0$. In $C([0,1], \mathbb{C})$, let $f(t) = t + i$ and $g(t) = 1 + it$. Then $f\overline{g} = (t+i)(1-it)$, which expands to $2t + i(1-t^2)$, so

$$\langle f, g\rangle = \int_0^1 2t\,dt + i\int_0^1 (1-t^2)\,dt = 1 + \frac{2i}{3},$$

and $\langle f, f\rangle = \int_0^1 (t^2+1)\,dt = 4/3$.

Matrices

The space $\mathbb{F}^{m\times n}$ is $\mathbb{F}^{mn}$ with the coordinates arranged in a rectangle, so it inherits a dot product: multiply corresponding entries, conjugating those of the second matrix, and add. The trace expresses this sum compactly.

Proposition 6.1.16 (Frobenius inner product). For $A = (a_{ij})$ and $B = (b_{ij})$ in $\mathbb{F}^{m\times n}$,

$$\operatorname{tr}(B^*A) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{b_{ij}}.$$

Consequently $\langle A, B\rangle = \operatorname{tr}(B^*A)$ is an inner product on $\mathbb{F}^{m\times n}$, the **Frobenius inner product**, and $\langle A, A\rangle = \sum_{i,j} |a_{ij}|^2$.

Proof. ROUTINE

Uses. Definition 2.4.26 (Trace), Definition 2.4.1 (Matrix product), Notation 6.1.7 (Conjugate transpose), Proposition 6.1.10 (The dot product is an inner product)
Hint 1. Compute the diagonal entries of B^*A from the definition of the matrix product.
Hint 2. The (j,j) entry of B^*A is $\sum_i \overline{b_{ij}}a_{ij}$: the dot product of column j of A with column j of B.
Hint 3. For the axioms, notice that the double sum is the formula of Definition 6.1.8 applied to mn coordinates.

Sketch. The jth diagonal entry of B^*A is the dot product of the jth columns of A and B; summing over j gives $\sum_{i,j} a_{ij}\overline{b_{ij}}$. That is the dot product of A and B regarded as vectors with mn coordinates, so the axioms hold by Proposition 6.1.10.

The (j,i) entry of $B^* \in \mathbb{F}^{n\times m}$ is $\overline{b_{ij}}$. By Definition 2.4.1, the (j,j) entry of $B^*A \in \mathbb{F}^{n\times n}$ is

$$(B^*A)_{jj} = \sum_{i=1}^{m} (B^*)_{ji} a_{ij} = \sum_{i=1}^{m} a_{ij}\overline{b_{ij}},$$

and summing over j (Definition 2.4.26) gives the formula.

The right side of the formula is the dot product formula of Definition 6.1.8, with the mn pairs (i,j) in place of the indices k. The proof of Proposition 6.1.10 checks each axiom from that formula alone, term by term, so it applies verbatim: $\langle A, B\rangle = \operatorname{tr}(B^*A)$ is an inner product on $\mathbb{F}^{m\times n}$, and $\langle A, A\rangle = \sum_{i,j} |a_{ij}|^2$. □

Example 6.1.17 (A Frobenius inner product). For $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix}$ in $\mathbb{R}^{2\times 2}$, the entrywise formula gives $\langle A, B\rangle = 1\cdot 3 + 2\cdot 0 + 0\cdot 1 + (-1)\cdot 1 = 2$. With the trace,

$$B^{\mathsf{T}}A = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 0 & -1 \end{pmatrix}, \quad \operatorname{tr}\left(B^{\mathsf{T}}A\right) = 3 - 1 = 2.$$

The identity $\langle A, A\rangle = \operatorname{tr}(A^*A) = \sum_{i,j} |a_{ij}|^2$ is useful well beyond computing lengths.

Technique 6.1.18 (The trace of A^*A). For every $A \in \mathbb{F}^{m\times n}$, $\operatorname{tr}(A^*A) = \sum_{i,j} |a_{ij}|^2$ (Proposition 6.1.16). So $\operatorname{tr}(A^*A) \geq 0$, with equality only for $A = 0$, and one scalar equation controls every entry of A. To show that a matrix A is zero, it suffices to show $\operatorname{tr}(A^*A) = 0$; in particular, it suffices to show $A^*A = 0$.

Reach for it when a matrix identity involves A^*A or AA^* (or $A^{\mathsf{T}}A$ over $\mathbb{R}$) and the goal is a conclusion about A itself. It is Technique 6.1.9 applied to the entries of a matrix. The canonical example is Corollary 6.1.19.

Used in. Corollary 6.1.19, Exercise 6.1.20, Exercise 6.1.54, Exercise 6.1.55, Exercise 6.6.50, Exercise 6.7.10, Exercise 7.5.41, Exercise 8.1.37, Exercise 8.1.38.

Corollary 6.1.19 ($A^*A = 0$ forces $A = 0$). Let $A \in \mathbb{F}^{m\times n}$. If $A^*A = 0$ or $AA^* = 0$, then $A = 0$.

Proof. ROUTINE

Uses. Proposition 6.1.16 (Frobenius inner product), Notation 6.1.7 (Conjugate transpose)
Techniques. The trace of A^*A (Technique 6.1.18)
Hint 1. Take the trace of $A^*A = 0$.
Hint 2. For $AA^* = 0$, apply the first case to A^*, using $(A^*)^* = A$.
Sketch. If $A^*A = 0$, then $\sum_{i,j} |a_{ij}|^2 = \operatorname{tr}(A^*A) = 0$, so every $a_{ij} = 0$. Since $(A^*)^* = A$, the case $AA^* = 0$ is the first case for A^*.

If $A^*A = 0$, then by Proposition 6.1.16, $\sum_{i,j} |a_{ij}|^2 = \operatorname{tr}(A^*A) = \operatorname{tr}(0) = 0$. This is a sum of nonnegative real numbers, so every $|a_{ij}|^2 = 0$, and $A = 0$.

The (i,j) entry of $(A^*)^*$ is the conjugate of the (j,i) entry of A^*, which is $\overline{\overline{a_{ij}}} = a_{ij}$, so $(A^*)^* = A$. If $AA^* = 0$, then $(A^*)^*A^* = 0$, and the first case applied to A^* gives $A^* = 0$. Hence $A = (A^*)^* = 0$. □

The conjugate transpose is essential here. The nilpotent matrix $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is nonzero with $N^2 = 0$, so $A^2 = 0$ does not force $A = 0$. Over $\mathbb{C}$, the transpose alone does not suffice either: the column $\mathbf{a} = (1, i)$ is nonzero, but $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + i^2 = 0$. This is the failure of $(1,i)\cdot(1,i)$ without conjugation that we met at the start of the section.

Exercise 6.1.20. ROUTINE Let $A \in \mathbb{C}^{n\times n}$ satisfy $A^* = A$ and $A^2 = 0$. Show that $A = 0$. Does the conclusion still hold if the hypothesis $A^* = A$ is dropped?

Solution on page 1462

Inner products on $\mathbb{F}^n$ from matrices

The dot product is not the only inner product on $\mathbb{F}^n$. Weighting the coordinates already gives others.

Example 6.1.21 (Weighted dot products). Let $c_1, \dots, c_n$ be positive real numbers. Then

$$\langle \mathbf{v}, \mathbf{w} \rangle_c = \sum_{k=1}^{n} c_k v_k \overline{w_k}$$

is an inner product on $\mathbb{F}^n$. The proof of Proposition 6.1.10 goes through with $c_k|v_k|^2$ in place of $|v_k|^2$, since $c_k|v_k|^2 \geq 0$, with equality only if $v_k = 0$. The weights must be positive: if $c_k \leq 0$, then $\langle \mathbf{e}_k, \mathbf{e}_k \rangle_c = c_k \leq 0$ while $\mathbf{e}_k \neq \mathbf{0}$.

Different weights give different geometry on the same set. With weights $(1, 4)$ on $\mathbb{R}^2$, the vector $\mathbf{e}_2$ has $\langle \mathbf{e}_2, \mathbf{e}_2 \rangle_c = 4$, so it is "twice as long" as $\mathbf{e}_1$, and the vectors with $\langle \mathbf{x}, \mathbf{x} \rangle_c = 1$ form the ellipse $x_1^2 + 4x_2^2 = 1$ rather than the unit circle (Figure 6.1.2).

In fact every inner product on $\mathbb{F}^n$ is given by a matrix. By sesquilinearity, $\langle \mathbf{x}, \mathbf{y} \rangle$ is determined by the n^2 numbers $\langle \mathbf{e}_j, \mathbf{e}_i \rangle$, and the question is which arrays of n^2 numbers occur.

Proposition 6.1.22 (Inner products on $\mathbb{F}^n$).

(a) Let $\langle \cdot, \cdot \rangle$ be an inner product on $\mathbb{F}^n$, and let $M \in \mathbb{F}^{n \times n}$ have entries $m_{ij} = \langle \mathbf{e}_j, \mathbf{e}_i \rangle$. Then $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* M \mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$.

(b) Conversely, let $M \in \mathbb{F}^{n \times n}$. The formula $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* M \mathbf{x}$ defines an inner product on $\mathbb{F}^n$ if and only if $M^* = M$ and $\mathbf{x}^* M \mathbf{x} > 0$ for every nonzero $\mathbf{x} \in \mathbb{F}^n$.

Proof. STANDARD

Uses. Proposition 6.1.4 (Sesquilinearity), Definition 1.6.8 (Standard bases), Definition 2.3.8 (Matrix–vector product), Notation 6.1.7 (Conjugate transpose), Remark 6.1.3
Techniques. Expand by multilinearity (Technique 5.1.7), Check it on a basis (Technique 2.2.7)
Hint 1. For (a), write $\mathbf{x} = \sum_j x_j \mathbf{e}_j$ and $\mathbf{y} = \sum_i y_i \mathbf{e}_i$, and expand using linearity in the first slot and Proposition 6.1.4 in the second.
Hint 2. For (b), the formula is linear in $\mathbf{x}$ for every M, and positivity together with definiteness says exactly that $\mathbf{x}^* M \mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$. So only conjugate symmetry needs thought.
Hint 3. Conjugate symmetry at the pair $(\mathbf{e}_j, \mathbf{e}_i)$ reads $m_{ij} = \overline{m_{ji}}$. For the converse, write $\mathbf{y}^* M \mathbf{x} = \sum_{i,j} \overline{y_i} m_{ij} x_j$, conjugate, and rename the indices.
Sketch. Expanding by sesquilinearity, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i,j} \overline{y_i} \langle \mathbf{e}_j, \mathbf{e}_i \rangle x_j = \mathbf{y}^* M \mathbf{x}$. For (b), linearity in $\mathbf{x}$ is automatic, and positivity plus definiteness is the condition $\mathbf{x}^* M \mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$. Conjugate symmetry tested at $(\mathbf{e}_j, \mathbf{e}_i)$ gives $M = M^*$; conversely, if $M = M^*$, conjugating the double sum for $\langle \mathbf{y}, \mathbf{x} \rangle$ and swapping i and j gives $\langle \mathbf{x}, \mathbf{y} \rangle$.

(a) Write $\mathbf{x} = \sum_j x_j \mathbf{e}_j$ and $\mathbf{y} = \sum_i y_i \mathbf{e}_i$. Linearity in the first slot and Proposition 6.1.4 give

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j \langle \mathbf{e}_j, \mathbf{y} \rangle = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \overline{y_i} \langle \mathbf{e}_j, \mathbf{e}_i \rangle = \sum_{i=1}^{n} \overline{y_i} \left(\sum_{j=1}^{n} m_{ij} x_j \right).$$

The inner sum is the ith coordinate of $M\mathbf{x}$ (Definition 2.3.8), so the whole expression is $\mathbf{y}^*(M\mathbf{x})$.

(b) Write $\langle \mathbf{x}, \mathbf{y} \rangle_M = \mathbf{y}^* M \mathbf{x} = \sum_{i,j} \overline{y_i} m_{ij} x_j$. For fixed $\mathbf{y}$ this is linear in $\mathbf{x}$, for every M, so linearity in the first slot always holds. Positivity and definiteness together say that $\langle \mathbf{x}, \mathbf{x} \rangle_M \geq 0$ for all $\mathbf{x}$, with equality only for $\mathbf{x} = \mathbf{0}$ (Remark 6.1.3); since $\langle \mathbf{0}, \mathbf{0} \rangle_M = 0$ in any case, this is the condition $\mathbf{x}^* M \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. It remains to show that conjugate symmetry holds if and only if $M^* = M$.

Suppose conjugate symmetry holds. For all i, j,

$$m_{ij} = \mathbf{e}_i^* M \mathbf{e}_j = \langle \mathbf{e}_j, \mathbf{e}_i \rangle_M = \overline{\langle \mathbf{e}_i, \mathbf{e}_j \rangle_M} = \overline{\mathbf{e}_j^* M \mathbf{e}_i} = \overline{m_{ji}},$$

which says $M^* = M$. Conversely, suppose $m_{ij} = \overline{m_{ji}}$ for all i, j. Then for all $\mathbf{x}, \mathbf{y}$,

$$\overline{\langle \mathbf{y}, \mathbf{x} \rangle_M} = \overline{\sum_{i,j} \overline{x_i} m_{ij} y_j} = \sum_{i,j} x_i \overline{m_{ij} y_j} = \sum_{i,j} \overline{y_j} m_{ji} x_i = \langle \mathbf{x}, \mathbf{y} \rangle_M,$$

where the last step renames the summation indices i and j as j and i. So conjugate symmetry holds. □

A matrix M with $M^* = M$ and $\mathbf{x}^* M \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$ is called **positive definite**, and Section 12.3 gives practical tests for it. Parts (a) and (b) say that every inner product on $\mathbb{F}^n$ is $\mathbf{y}^* M \mathbf{x}$ for a positive definite M. The dot product is the case $M = I_n$, and the weighted dot products are the diagonal M with positive diagonal entries.

Example 6.1.23 (A tilted inner product on $\mathbb{R}^2$). Let $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Then $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^\mathsf{T} M \mathbf{x} = 2x_1 y_1 + x_1 y_2 + x_2 y_1 + 2x_2 y_2$. Here $M^\mathsf{T} = M$, and for $\mathbf{x} \neq \mathbf{0}$,

$$\mathbf{x}^\mathsf{T} M \mathbf{x} = 2x_1^2 + 2x_1 x_2 + 2x_2^2 = x_1^2 + x_2^2 + (x_1 + x_2)^2 > 0,$$

a sum of squares that is not all zero. So Proposition 6.1.22 makes this an inner product. Measured with it, $(1, -1)$ has $\langle \mathbf{x}, \mathbf{x} \rangle = 2 - 2 + 2 = 2$ while $(1, 1)$ has $\langle \mathbf{x}, \mathbf{x} \rangle = 6$, and the vectors with $\langle \mathbf{x}, \mathbf{x} \rangle = 1$ form a tilted ellipse (Figure 6.1.2).

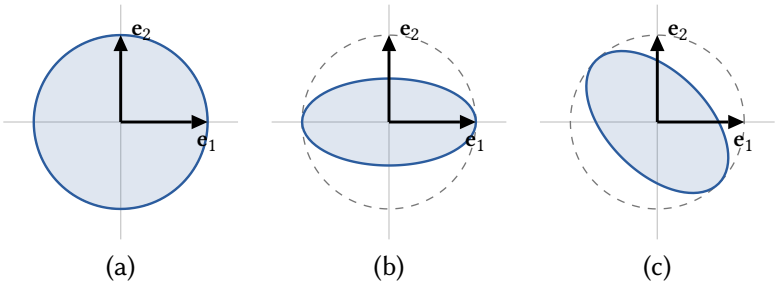


Figure 6.1.2. The vectors $\mathbf{x} \in \mathbb{R}^2$ with $\langle \mathbf{x}, \mathbf{x} \rangle = 1$ (solid) for three inner products on $\mathbb{R}^2$: (a) the dot product; (b) the weighted dot product $x_1 y_1 + 4x_2 y_2$ of Example 6.1.21, in which $\mathbf{e}_2$ lies outside the curve because it has length 2; (c) the inner product of Example 6.1.23. The dashed circle is the unit circle of the dot product.

Remark 6.1.24. Over $\mathbb{R}$, positivity of $\mathbf{x}^\mathsf{T} M \mathbf{x}$ does not force M to be symmetric. For $M = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $\mathbf{x}^\mathsf{T} M \mathbf{x} = x_1^2 + x_1 x_2 - x_2 x_1 + x_2^2 = x_1^2 + x_2^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$, but $\mathbf{y}^\mathsf{T} M \mathbf{x}$ is not an inner product, since $\langle \mathbf{e}_2, \mathbf{e}_1 \rangle = m_{12} = 1$ while $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = m_{21} = -1$. Over $\mathbb{C}$ this cannot happen: if $\mathbf{x}^* M \mathbf{x}$ is real for every $\mathbf{x} \in \mathbb{C}^n$, then $M^* = M$ (Exercise 6.1.47).

Non-examples

Non-example 6.1.25 (Formulas that are not inner products).

(a) On $\mathbb{C}^2$, $\langle \mathbf{x}, \mathbf{y}\rangle = x_1 y_1 + x_2 y_2$ is bilinear and symmetric, but $\langle (i,0),(i,0)\rangle = i^2 = -1 < 0$. This is Proposition 6.1.6 in action.

(b) On $\mathbb{C}^n$, $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_k \overline{x_k} y_k = \mathbf{x}^* \mathbf{y}$ is conjugate symmetric, positive, and definite, and it is linear in the *second* slot. But $\langle i\mathbf{x}, \mathbf{y}\rangle = -i\langle \mathbf{x}, \mathbf{y}\rangle$, so it is not linear in the first slot, and by our definition it is not an inner product. Many physics texts use this convention instead. It is the dot product with its arguments exchanged, $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y} \cdot \mathbf{x}$, so results translate by swapping the slots.

(c) On $\mathbb{R}^{2\times 2}$, $\langle A, B\rangle = \operatorname{tr}(AB)$ is bilinear and symmetric (as $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, Proposition 2.4.28). But $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $J^2 = -I_2$, so $\langle J, J\rangle = -2 < 0$. In the Frobenius inner product $\operatorname{tr}\left(B^{\mathsf{T}} A\right)$, the transpose is what turns $\langle A, A\rangle$ into a sum of squares.

(d) On $\mathbb{R}[x]_{\leq 2}$, $\langle p, q\rangle = p(0)q(0) + p(1)q(1)$ is symmetric, bilinear, and positive, but the nonzero polynomial $p = x^2 - x$ has $\langle p, p\rangle = 0$. On $\mathbb{R}[x]_{\leq 1}$, the same formula *is* an inner product: if $p = a + bx$ and $p(0)^2 + p(1)^2 = 0$, then $p(0) = a = 0$ and $p(1) = a + b = 0$, so $p = \mathbf{0}$. Whether a formula defines an inner product depends on the space it is defined on.

6.1.4 The norm

Definition 6.1.26 (Norm). Let V be an inner product space. The **norm** of $\mathbf{v} \in V$ is

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle},$$

the nonnegative square root of the nonnegative real number $\langle \mathbf{v}, \mathbf{v}\rangle$. A **unit vector** is a vector $\mathbf{u}$ with $\|\mathbf{u}\| = 1$.

The norm is the length of $\mathbf{v}$ as measured by the chosen inner product, and $\|\mathbf{v} - \mathbf{w}\|$ is the distance between $\mathbf{v}$ and $\mathbf{w}$. For the dot product on $\mathbb{F}^n$, $\|\mathbf{v}\|^2 = \sum_k |v_k|^2$, so $\|\mathbf{v}\|$ is the Euclidean length. For the Frobenius inner product, $\|A\|^2 = \sum_{i,j} |a_{ij}|^2$, and $\|A\|$ is often written $\|A\|_F$. For the integral inner product, $\|f\|^2 = \int_0^1 |f|^2$. A different inner product on the same space gives different lengths: with the weights $(1, 4)$ of Example 6.1.21, $\|\mathbf{e}_2\| = 2$.

Proposition 6.1.27 (Basic properties of the norm). Let V be an inner product space, $\mathbf{v} \in V$, and $a \in \mathbb{F}$.

(a) $\|\mathbf{v}\| \geq 0$, and $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

(b) $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$.

Proof. ROUTINE

Uses. Definition 6.1.26 (Norm), Definition 6.1.2 (Inner product), Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)
Hint 1. Part (a) is a restatement of two axioms.
Hint 2. For (b), compare squares: compute $\|a\mathbf{v}\|^2 = \langle a\mathbf{v}, a\mathbf{v}\rangle$ by taking a out of each slot.
Hint 3. The second slot contributes $\bar{a}$, and $a\bar{a} = |a|^2$. Two nonnegative reals with the same square are equal.
Sketch. (a) is positivity and definiteness. For (b), $\|a\mathbf{v}\|^2 = a\bar{a}\langle \mathbf{v}, \mathbf{v}\rangle = |a|^2\|\mathbf{v}\|^2$; take nonnegative square roots.

(a) The square root of a nonnegative real number is nonnegative. Also $\|\mathbf{v}\| = 0$ if and only if $\langle \mathbf{v}, \mathbf{v}\rangle = 0$, which by definiteness happens if and only if $\mathbf{v} = \mathbf{0}$.

(b) By linearity in the first slot and Proposition 6.1.4,

$$\|a\mathbf{v}\|^2 = \langle a\mathbf{v}, a\mathbf{v}\rangle = a\bar{a}\langle \mathbf{v}, \mathbf{v}\rangle = |a|^2\|\mathbf{v}\|^2.$$

Both $\|a\mathbf{v}\|$ and $|a|\|\mathbf{v}\|$ are nonnegative real numbers, and their squares are equal, so they are equal. □

In particular, if $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{v}/\|\mathbf{v}\|$ (that is, $(1/\|\mathbf{v}\|)\mathbf{v}$) is a unit vector, since by (b) its norm is $\|\mathbf{v}\|/\|\mathbf{v}\| = 1$. Replacing $\mathbf{v}$ by $\mathbf{v}/\|\mathbf{v}\|$ is called **normalizing** $\mathbf{v}$.

Exercise 6.1.28. ROUTINE Let $\mathbf{v} \neq \mathbf{0}$ be a vector in an inner product space. Find all scalars $a \in \mathbb{F}$ with $\|a\mathbf{v}\| = 1$, separately for $\mathbb{F} = \mathbb{R}$ and for $\mathbb{F} = \mathbb{C}$.

Solution on page 1463

Expanding norms

The computation that produced (6.1.1) expanded $\|\mathbf{v} - \mathbf{w}\|^2$. In a general inner product space the same expansion works, with one change over $\mathbb{C}$.

Proposition 6.1.29 (Expanding a norm). Let V be an inner product space, $\mathbf{x}, \mathbf{y} \in V$, and $a \in \mathbb{F}$. Then

$$\|\mathbf{x} + a\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\mathrm{Re}(\overline{a}\langle \mathbf{x}, \mathbf{y}\rangle) + |a|^2\|\mathbf{y}\|^2.$$

In particular,

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \pm 2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2.$$

Proof. ROUTINE

Uses. Definition 6.1.26 (Norm), Definition 6.1.2 (Inner product), Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Write $\|\mathbf{x} + a\mathbf{y}\|^2 = \langle \mathbf{x} + a\mathbf{y}, \mathbf{x} + a\mathbf{y}\rangle$ and expand both slots, the second with Proposition 6.1.4.
Hint 2. Of the four terms, the two cross terms are complex conjugates of each other. Use $z + \overline{z} = 2\,\mathrm{Re}\,z$.
Sketch. Expanding gives $\|\mathbf{x}\|^2 + \overline{a}\langle \mathbf{x}, \mathbf{y}\rangle + a\langle \mathbf{y}, \mathbf{x}\rangle + |a|^2\|\mathbf{y}\|^2$, and by conjugate symmetry the cross terms are z and $\overline{z}$ for $z = \overline{a}\langle \mathbf{x}, \mathbf{y}\rangle$. They add to $2\,\mathrm{Re}\,z$. Take $a = \pm 1$ for the second formula.

By linearity in the first slot and Proposition 6.1.4,

$$\begin{aligned}\|\mathbf{x} + a\mathbf{y}\|^2 &= \langle \mathbf{x} + a\mathbf{y}, \mathbf{x} + a\mathbf{y}\rangle \\ &= \langle \mathbf{x}, \mathbf{x}\rangle + \overline{a}\langle \mathbf{x}, \mathbf{y}\rangle + a\langle \mathbf{y}, \mathbf{x}\rangle + a\overline{a}\langle \mathbf{y}, \mathbf{y}\rangle.\end{aligned}$$

Let $z = \overline{a}\langle \mathbf{x}, \mathbf{y}\rangle$. By conjugate symmetry, $a\langle \mathbf{y}, \mathbf{x}\rangle = a\overline{\langle \mathbf{x}, \mathbf{y}\rangle} = \overline{\overline{a}\langle \mathbf{x}, \mathbf{y}\rangle} = \overline{z}$, so the two middle terms add to $z + \overline{z} = 2\,\mathrm{Re}\,z$. Since $a\overline{a} = |a|^2$, this is the first formula. For $a = 1$ it reads $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2$, and for $a = -1$ the middle term is $2\,\mathrm{Re}(-\langle \mathbf{x}, \mathbf{y}\rangle) = -2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$. □

Intuition (The law of cosines, everywhere). Over $\mathbb{R}$ the cross term is $2\langle \mathbf{x}, \mathbf{y}\rangle$, and in $\mathbb{R}^2$ with the dot product, (6.1.1) turns $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2$ back into the law of cosines. So Proposition 6.1.29 is the law of cosines of an arbitrary inner product space; the computation at the start of the section, run backwards.

Over $\mathbb{C}$, only the real part of $\langle \mathbf{x}, \mathbf{y}\rangle$ affects $\|\mathbf{x} \pm \mathbf{y}\|$. The imaginary part appears once $\mathbf{y}$ is multiplied by i: with $a = i$, $\mathrm{Re}\big(\overline{i}\langle \mathbf{x}, \mathbf{y}\rangle\big) = \mathrm{Re}(-i\langle \mathbf{x}, \mathbf{y}\rangle) = \mathrm{Im}\langle \mathbf{x}, \mathbf{y}\rangle$, so

$$\|\mathbf{x} + i\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\mathrm{Im}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2.$$

When $\langle \mathbf{x}, \mathbf{y}\rangle = 0$, the expansion gives $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, the Pythagorean theorem of Section 6.3.

The parallelogram law

Vectors $\mathbf{x}$ and $\mathbf{y}$ span a parallelogram whose diagonals are $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$ (Figure 6.1.3). In a rectangle the two diagonals are equally long, and Pythagoras says that each has squared length $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$. Skewing the rectangle lengthens one diagonal and shortens the other. Adding the two expansions in Proposition 6.1.29 shows that the changes in their squares cancel exactly.

Proposition 6.1.30 (Parallelogram law). For all $\mathbf{x}, \mathbf{y}$ in an inner product space,

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2.$$

Proof. ROUTINE

Uses. Proposition 6.1.29 (Expanding a norm)
Hint 1. Expand each term on the left with Proposition 6.1.29.
Hint 2. The cross terms $\pm 2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$ have opposite signs.
Sketch. Add the expansions of $\|\mathbf{x}+\mathbf{y}\|^2$ and $\|\mathbf{x}-\mathbf{y}\|^2$; the cross terms cancel.

By Proposition 6.1.29,

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= \left(\|\mathbf{x}\|^2 + 2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2\right) + \left(\|\mathbf{x}\|^2 - 2\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2\right) \\ &= 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2,\end{aligned}$$

as desired. □

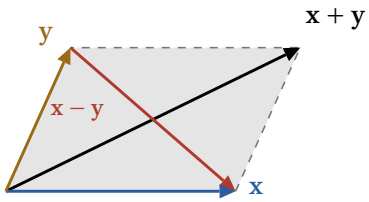


Figure 6.1.3. The parallelogram law: the squares of the diagonals $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ add up to the squares of the four sides, $2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$.

Unlike Proposition 6.1.29, the parallelogram law mentions only norms, not the inner product. That makes it a test: a notion of length that violates it cannot be the norm of any inner product (Aside (Norms that do not come from inner products)).

Exercise 6.1.31. ROUTINE In an inner product space, $\|\mathbf{x}\| = 3$, $\|\mathbf{y}\| = 4$, and $\|\mathbf{x}+\mathbf{y}\| = 6$. Find $\|\mathbf{x}-\mathbf{y}\|$ and $\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$.

Solution on page 1463

Polarization

The norm is defined from the inner product. Polarization goes the other way. Over $\mathbb{R}$, subtracting the two expansions in Proposition 6.1.29 isolates the cross term: $\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 = 4\langle \mathbf{x}, \mathbf{y}\rangle$. Over $\mathbb{C}$ the same subtraction yields only $4\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$, but the imaginary part is also a real part, since $\mathrm{Im}\langle \mathbf{x}, \mathbf{y}\rangle = \mathrm{Re}(-i\langle \mathbf{x}, \mathbf{y}\rangle) = \mathrm{Re}\langle \mathbf{x}, i\mathbf{y}\rangle$.

Proposition 6.1.32 (Polarization identities). Let V be an inner product space and $\mathbf{x}, \mathbf{y} \in V$.

(a) If $\mathbb{F} = \mathbb{R}$, then

$$\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right).$$

(b) If $\mathbb{F} = \mathbb{C}$, then

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y}\rangle &= \frac{1}{4}\sum_{k=0}^{3} i^k \left\|\mathbf{x} + i^k \mathbf{y}\right\|^2 \\ &= \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + i\|\mathbf{x}+i\mathbf{y}\|^2 - i\|\mathbf{x}-i\mathbf{y}\|^2\right).\end{aligned}$$

Proof (via real and imaginary parts). STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)

Hint 1. Subtract the two identities of Proposition 6.1.29.
Hint 2. Over $\mathbb{C}$ this gives only $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$. Write $\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle$ as the real part of another inner product.
Hint 3. $\langle \mathbf{x}, i\mathbf{y}\rangle = -i\langle \mathbf{x}, \mathbf{y}\rangle$ and $\operatorname{Re}(-iz) = \operatorname{Im} z$. Apply the real-part formula with $i\mathbf{y}$ in place of $\mathbf{y}$.
Sketch. Subtracting the expansions of $\|\mathbf{x} \pm \mathbf{y}\|^2$ gives $4\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$, which is $4\langle \mathbf{x}, \mathbf{y}\rangle$ over $\mathbb{R}$. Over $\mathbb{C}$, applying this to the pair $\mathbf{x}, i\mathbf{y}$ gives $4\operatorname{Re}\langle \mathbf{x}, i\mathbf{y}\rangle = 4\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2$. Then combine $\langle \mathbf{x}, \mathbf{y}\rangle = \operatorname{Re} + i \operatorname{Im}$.

Subtracting the two identities of Proposition 6.1.29 gives, for all $\mathbf{x}, \mathbf{y}$ in any inner product space,

$$\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\Big(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\Big). \tag{6.1.2}$$

(a) If $\mathbb{F} = \mathbb{R}$, then $\langle \mathbf{x}, \mathbf{y}\rangle$ is real, so it equals its real part, and (6.1.2) is (a).

(b) Let $\mathbb{F} = \mathbb{C}$. By Proposition 6.1.4, $\langle \mathbf{x}, i\mathbf{y}\rangle = \bar{i}\langle \mathbf{x}, \mathbf{y}\rangle = -i\langle \mathbf{x}, \mathbf{y}\rangle$, and $\operatorname{Re}(-iz) = \operatorname{Im} z$ for every $z \in \mathbb{C}$. So (6.1.2) applied to $\mathbf{x}$ and $i\mathbf{y}$ gives

$$\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle = \operatorname{Re}\langle \mathbf{x}, i\mathbf{y}\rangle = \frac{1}{4}\Big(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2\Big).$$

Since $\langle \mathbf{x}, \mathbf{y}\rangle = \operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle + i \operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle$, adding (6.1.2) to i times this identity gives the second line of (b). The terms of $\sum_{k=0}^{3} i^k \|\mathbf{x} + i^k\mathbf{y}\|^2$ for $k = 0, 1, 2, 3$ are $\|\mathbf{x} + \mathbf{y}\|^2$, $i\|\mathbf{x} + i\mathbf{y}\|^2$, $-\|\mathbf{x} - \mathbf{y}\|^2$, and $-i\|\mathbf{x} - i\mathbf{y}\|^2$, so the two lines agree. □

Proof (by averaging over powers of i). STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Notation 6.1.1 (Real and complex scalars)
Hint 1. Expand each $\|\mathbf{x} + i^k\mathbf{y}\|^2$ with Proposition 6.1.29, writing $2\operatorname{Re} w$ as $w + \overline{w}$.
Hint 2. After multiplying by i^k, the terms carry the factors i^k, 1, and i^{2k}. Which of these survive summing over $k = 0, 1, 2, 3$?
Sketch. With $z = \langle \mathbf{x}, \mathbf{y}\rangle$, $i^k\|\mathbf{x} + i^k\mathbf{y}\|^2 = i^k\big(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\big) + z + i^{2k}\bar{z}$. Summing over k kills the first and last terms, because $1 + i + i^2 + i^3 = 0$ and $1 - 1 + 1 - 1 = 0$, and leaves $4z$.

Let $\mathbb{F} = \mathbb{C}$, let $z = \langle \mathbf{x}, \mathbf{y}\rangle$, and let $c = i^k$ for some $k \in \{0, 1, 2, 3\}$, so that $|c| = 1$ and $\bar{c} = c^{-1}$. By Proposition 6.1.29, and since $2\operatorname{Re} w = w + \overline{w}$,

$$\|\mathbf{x} + c\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \bar{c}z + c\bar{z}.$$

Multiplying by c and using $c\bar{c} = 1$,

$$c\|\mathbf{x} + c\mathbf{y}\|^2 = c\Big(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\Big) + z + c^2\bar{z}.$$

Now sum over $c = 1, i, -1, -i$. The coefficients c add to $1 + i - 1 - i = 0$, and the coefficients c^2 add to $1 - 1 + 1 - 1 = 0$, so

$$\sum_{k=0}^{3} i^k\|\mathbf{x} + i^k\mathbf{y}\|^2 = 4z = 4\langle \mathbf{x}, \mathbf{y}\rangle,$$

which is (b). Part (a) is the same computation with $c = \pm 1$ and z real: then $c\|\mathbf{x} + c\mathbf{y}\|^2 = c\Big(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\Big) + 2z$, and summing over $c = \pm 1$ gives $\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4z$. □

The first proof explains why there are four terms: two give the real part, and two more give the imaginary part. The second explains the coefficients. Multiplying by i^k and summing over k is a filter: it keeps the term that carries no power of i and discards the terms that carry i^k and i^{2k}, because the powers of i add to 0. The same filter works with any $m \geq 3$ equally spaced points on the unit circle in place of $1, i, -1, -i$, and it is the idea behind the discrete Fourier transform.

Technique 6.1.33 (Polarization). An inner product is determined by its values $\langle \mathbf{v}, \mathbf{v}\rangle$ on single vectors: by Proposition 6.1.32, $\langle \mathbf{x}, \mathbf{y}\rangle$ is a combination of the numbers $\|\mathbf{x} + c\mathbf{y}\|^2$ with $c \in \{\pm 1, \pm i\}$. So a statement

about all pairs $(\mathbf{x}, \mathbf{y})$ can often be reduced to a statement about single vectors, applied to $\mathbf{x} + c\mathbf{y}$ for a few scalars c.

Reach for it when the hypothesis or the conclusion involves only lengths, or only "quadratic" quantities such as $\langle T\mathbf{v}, \mathbf{v}\rangle$, and you need information about two different vectors at once. The canonical example is Proposition 6.1.32 itself. Over $\mathbb{R}$ there is a caveat: a *symmetric* function $B(\mathbf{x}, \mathbf{y})$ that is linear in each slot is determined by its diagonal values $B(\mathbf{v}, \mathbf{v})$, but without symmetry this fails. For instance, $B(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1$ on $\mathbb{R}^2$ satisfies $B(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v}$ without being 0. Over $\mathbb{C}$ no symmetry is needed (Proposition 6.1.41).

Used in. Corollary 6.1.34, Exercise 6.1.35, Proposition 6.1.41, Exercise 6.1.45, Exercise 6.1.51, Exercise 6.1.53, Exercise 6.1.59, Exercise 6.7.8, Problem 6.8, Proposition 8.1.4, Proposition 8.1.13, Theorem 8.4.4.

Corollary 6.1.34 (The norm determines the inner product). Let $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ be inner products on V with $\langle \mathbf{v}, \mathbf{v}\rangle_1 = \langle \mathbf{v}, \mathbf{v}\rangle_2$ for all $\mathbf{v} \in V$. Then $\langle \mathbf{x}, \mathbf{y}\rangle_1 = \langle \mathbf{x}, \mathbf{y}\rangle_2$ for all $\mathbf{x}, \mathbf{y} \in V$.

Proof. ROUTINE

Uses. Proposition 6.1.32 (Polarization identities)
Techniques. Polarization (Technique 6.1.33)
Hint 1. Write both $\langle \mathbf{x}, \mathbf{y}\rangle_1$ and $\langle \mathbf{x}, \mathbf{y}\rangle_2$ with Proposition 6.1.32.
Hint 2. The right sides of the polarization identities only involve norms, and the two norms agree.
Sketch. Both inner products equal the same combination of the numbers $\|\mathbf{x} + c\mathbf{y}\|^2$, $c \in \{\pm 1, \pm i\}$, and these do not depend on which inner product defines the norm.

The two inner products define the same norm, $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle_1} = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle_2}$. By Proposition 6.1.32, applied once to each inner product, both $\langle \mathbf{x}, \mathbf{y}\rangle_1$ and $\langle \mathbf{x}, \mathbf{y}\rangle_2$ equal $\frac{1}{4}\left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\right)$ if $\mathbb{F} = \mathbb{R}$, and $\frac{1}{4}\sum_{k=0}^{3} i^k \|\mathbf{x} + i^k \mathbf{y}\|^2$ if $\mathbb{F} = \mathbb{C}$. So they are equal. □

Combined with Proposition 6.1.27(b), the inner product is even determined by its unit sphere $\{\mathbf{v} \in V \mid \|\mathbf{v}\| = 1\}$: for $\mathbf{v} \neq \mathbf{0}$, $\|\mathbf{v}\|$ is the unique $t > 0$ for which $\mathbf{v}/t$ lies on the sphere. So each curve in Figure 6.1.2 determines its inner product completely.

Exercise 6.1.35. ROUTINE Vectors $\mathbf{x}, \mathbf{y}$ in a complex inner product space satisfy

$$\|\mathbf{x} + \mathbf{y}\|^2 = 10, \quad \|\mathbf{x} - \mathbf{y}\|^2 = 2, \quad \|\mathbf{x} + i\mathbf{y}\|^2 = 4, \quad \|\mathbf{x} - i\mathbf{y}\|^2 = 8.$$

Find $\langle \mathbf{x}, \mathbf{y}\rangle$ and $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

Solution on page 1463

6.1.5 Inner products distinguish vectors

The lecture's next observation is that a vector is determined by its inner products with all vectors. The proof is a one-line instance of a move used throughout this chapter, so we name it first.

Technique 6.1.36 (Show a vector is zero via its norm). By definiteness, $\mathbf{v} = \mathbf{0}$ as soon as $\langle \mathbf{v}, \mathbf{v}\rangle = 0$. So to prove that a vector is zero, compute $\langle \mathbf{v}, \mathbf{v}\rangle$, or $\|\mathbf{v}\|^2$, and show that it vanishes. Often one knows that $\langle \mathbf{v}, \mathbf{x}\rangle = 0$ for many $\mathbf{x}$, and the move is to take $\mathbf{x} = \mathbf{v}$. To prove $\mathbf{v} = \mathbf{w}$, apply this to $\mathbf{v} - \mathbf{w}$, expanding $\|\mathbf{v} - \mathbf{w}\|^2$ with Proposition 6.1.29 if necessary.

Reach for it when the hypotheses are statements about inner products and the conclusion is an equation between vectors. The canonical example is Proposition 6.1.37. In Section 6.5 the same move shows that $A^*A\mathbf{x} = \mathbf{0}$ implies $A\mathbf{x} = \mathbf{0}$.

Used in. Proposition 6.1.37, Corollary 6.1.38, Exercise 6.1.50, Proposition 6.3.30, Theorem 6.5.5, Proposition 6.5.9, Theorem 6.6.21, Proposition 6.6.23, Exercise 6.7.9, Problem 6.3, Problem 6.8, Problem 6.9, Problem 6.11, Proposition 8.1.4, Exercise 8.1.6, Proposition 8.1.14, Theorem 8.2.5, Lemma 8.3.8, Exercise 6.6.51.

Proposition 6.1.37 (Inner products distinguish vectors). Let V be an inner product space and $\mathbf{v}, \mathbf{w} \in V$. If $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{w}, \mathbf{x} \rangle$ for all $\mathbf{x} \in V$, then $\mathbf{v} = \mathbf{w}$.

Proof. ROUTINE

Uses. Definition 6.1.2 (Inner product), Proposition 6.1.4 (Sesquilinearity)
Techniques. Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Reduce to the case $\mathbf{w} = \mathbf{0}$ by looking at $\mathbf{v} - \mathbf{w}$.
Hint 2. If $\langle \mathbf{u}, \mathbf{x} \rangle = 0$ for every $\mathbf{x}$, which choice of $\mathbf{x}$ tells you the most about $\mathbf{u}$?
Hint 3. Take $\mathbf{x} = \mathbf{u}$ and use definiteness.
Sketch. By linearity, $\langle \mathbf{v} - \mathbf{w}, \mathbf{x} \rangle = 0$ for every $\mathbf{x}$. Choosing $\mathbf{x} = \mathbf{v} - \mathbf{w}$ gives $\|\mathbf{v} - \mathbf{w}\|^2 = 0$, so $\mathbf{v} = \mathbf{w}$.

First suppose $\mathbf{w} = \mathbf{0}$. By Proposition 6.1.4, $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{0}, \mathbf{x} \rangle = 0$ for all $\mathbf{x} \in V$. In particular, for $\mathbf{x} = \mathbf{v}$ we get $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, and $\mathbf{v} = \mathbf{0}$ by definiteness.

For the general case, linearity in the first slot gives $\langle \mathbf{v} - \mathbf{w}, \mathbf{x} \rangle = \langle \mathbf{v}, \mathbf{x} \rangle - \langle \mathbf{w}, \mathbf{x} \rangle = 0 = \langle \mathbf{0}, \mathbf{x} \rangle$ for all $\mathbf{x} \in V$. By the first case, $\mathbf{v} - \mathbf{w} = \mathbf{0}$, that is, $\mathbf{v} = \mathbf{w}$, as desired. □

The same holds with the vectors in the second slot: if $\langle \mathbf{x}, \mathbf{v} \rangle = \langle \mathbf{x}, \mathbf{w} \rangle$ for all $\mathbf{x}$, conjugating gives $\langle \mathbf{v}, \mathbf{x} \rangle = \langle \mathbf{w}, \mathbf{x} \rangle$ for all $\mathbf{x}$. Nor do we need every $\mathbf{x}$: the vectors of a spanning list suffice (Exercise 6.1.50).

Corollary 6.1.38 (Linear maps are determined by inner products). Let V be a vector space over $\mathbb{F}$, let W be an inner product space, and let $S, T : V \to W$ be linear. If $\langle S\mathbf{v}, \mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{v} \in V$ and all $\mathbf{w} \in W$, then $S = T$.

Proof. ROUTINE

Uses. Proposition 6.1.37 (Inner products distinguish vectors)
Techniques. Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Fix $\mathbf{v} \in V$. What does the hypothesis say about the two vectors $S\mathbf{v}$ and $T\mathbf{v}$ of W?
Hint 2. Apply Proposition 6.1.37 in W.
Sketch. For each fixed $\mathbf{v}$, $\langle S\mathbf{v}, \mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{w} \in W$, so $S\mathbf{v} = T\mathbf{v}$ by Proposition 6.1.37.

Here and below we write $T\mathbf{v}$ for $T(\mathbf{v})$. Fix $\mathbf{v} \in V$. By hypothesis, the vectors $S\mathbf{v}$ and $T\mathbf{v}$ of W satisfy $\langle S\mathbf{v}, \mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{w} \in W$, so $S\mathbf{v} = T\mathbf{v}$ by Proposition 6.1.37. Since $\mathbf{v}$ was arbitrary, $S = T$. □

The proof never used that S and T are linear, so the corollary holds for arbitrary functions $V \to W$. It is how adjoints will be shown to be unique (Section 6.6).

Testing on the diagonal

Corollary 6.1.38 tests $\langle T\mathbf{v}, \mathbf{w} \rangle$ on all pairs of vectors. For an operator $T \in \mathrm{L}(V)$, it is natural to ask whether the values $\langle T\mathbf{v}, \mathbf{v} \rangle$, with the same vector in both slots, are already enough. Over $\mathbb{R}$ they are not.

Example 6.1.39 (A rotation). Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be $T\mathbf{x} = J\mathbf{x} = (-x_2, x_1)$, where $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$; this is the rotation by $\pi/2$. With the dot product,

$$\langle T\mathbf{x}, \mathbf{x} \rangle = (-x_2)x_1 + x_1 x_2 = 0 \quad \text{for every } \mathbf{x} \in \mathbb{R}^2,$$

as a rotation by a right angle moves every vector to a perpendicular one. But $T \neq 0$.

Over $\mathbb{C}$, however, the values $\langle T\mathbf{v}, \mathbf{v} \rangle$ do determine T. The proof feeds the hypothesis well-chosen vectors, a move that polarization has already used.

Technique 6.1.40 (Plug in clever test vectors). A hypothesis that holds for all vectors holds in particular for sums and multiples of the vectors you care about. Applying it to $\mathbf{v} = \mathbf{x} + \mathbf{y}$, $\mathbf{x} - \mathbf{y}$, $\mathbf{x} + i\mathbf{y}$, or to the very vector you want to control, turns one identity into a system of equations, which can then be solved for the quantity you want.

Reach for it when a hypothesis is quantified over all vectors but involves each vector only quadratically, like $\|\mathbf{v}\|$ or $\langle T\mathbf{v},\mathbf{v}\rangle$, and you need information about pairs of vectors. Over $\mathbb{C}$, the pair $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}+i\mathbf{y}$ is usually the right choice, because together they separate real and imaginary parts. The canonical example is Proposition 6.1.41. The polarization identities (Proposition 6.1.32) and the choice $\mathbf{x}=\mathbf{v}-\mathbf{w}$ in Proposition 6.1.37 are instances of the same move.
Used in. Proposition 6.1.41, Exercise 6.1.46, Exercise 6.1.47, Exercise 6.1.52, Proposition 6.3.5, Exercise 6.3.52, Exercise 6.7.7, Exercise 6.7.8, Exercise 6.7.15, Problem 6.6, Problem 7.8, Proposition 8.1.4, Theorem 8.3.23, Exercise 8.4.34.

Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v},\mathbf{v}\rangle$ determine T). Let V be a complex inner product space and $T\in \mathrm{L}(V)$. If $\langle T\mathbf{v},\mathbf{v}\rangle=0$ for all $\mathbf{v}\in V$, then $T=0$. Consequently, if $S,T\in\mathrm{L}(V)$ satisfy $\langle S\mathbf{v},\mathbf{v}\rangle=\langle T\mathbf{v},\mathbf{v}\rangle$ for all $\mathbf{v}\in V$, then $S=T$.

Intuition. By Corollary 6.1.38 it suffices to show $\langle T\mathbf{x},\mathbf{y}\rangle=0$ for all $\mathbf{x},\mathbf{y}$, but the hypothesis only speaks about one vector at a time. Plugging $\mathbf{v}=\mathbf{x}+\mathbf{y}$ into it isolates the sum $\langle T\mathbf{x},\mathbf{y}\rangle+\langle T\mathbf{y},\mathbf{x}\rangle$. Over $\mathbb{R}$ that is all one can get, and for the rotation of Example 6.1.39 this sum is 0 for all $\mathbf{x},\mathbf{y}$ although its terms are not. Over $\mathbb{C}$, plugging in $\mathbf{v}=\mathbf{x}+i\mathbf{y}$ produces a second combination, $-i$ times the difference $\langle T\mathbf{x},\mathbf{y}\rangle-\langle T\mathbf{y},\mathbf{x}\rangle$, and two independent equations in the two unknowns can be solved.

Proof (with test vectors). STANDARD

Uses. Corollary 6.1.38 (Linear maps are determined by inner products), Proposition 6.1.4 (Sesquilinearity), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space)
Techniques. Plug in clever test vectors (Technique 6.1.40)
Hint 1. It suffices to show $\langle T\mathbf{x},\mathbf{y}\rangle=0$ for all $\mathbf{x},\mathbf{y}\in V$. Apply the hypothesis to a vector built from $\mathbf{x}$ and $\mathbf{y}$.
Hint 2. $\mathbf{v}=\mathbf{x}+\mathbf{y}$ gives $\langle T\mathbf{x},\mathbf{y}\rangle+\langle T\mathbf{y},\mathbf{x}\rangle=0$. That alone cannot suffice, since the rotation satisfies it. Find a second equation.
Hint 3. Try $\mathbf{v}=\mathbf{x}+i\mathbf{y}$, and remember that i leaves the second slot as $-i$.
Sketch. Expanding $\langle T\mathbf{v},\mathbf{v}\rangle=0$ at $\mathbf{v}=\mathbf{x}+\mathbf{y}$ and at $\mathbf{v}=\mathbf{x}+i\mathbf{y}$ gives $\langle T\mathbf{x},\mathbf{y}\rangle+\langle T\mathbf{y},\mathbf{x}\rangle=0$ and $-i\langle T\mathbf{x},\mathbf{y}\rangle+i\langle T\mathbf{y},\mathbf{x}\rangle=0$. Dividing the second by $-i$ and adding, $\langle T\mathbf{x},\mathbf{y}\rangle=0$ for all $\mathbf{x},\mathbf{y}$, so $T=0$. Apply this to $S-T$ for the last claim.

Let $\mathbf{x},\mathbf{y}\in V$. Since T is linear and the inner product is sesquilinear,

$$\langle T(\mathbf{x}+\mathbf{y}),\mathbf{x}+\mathbf{y}\rangle=\langle T\mathbf{x},\mathbf{x}\rangle+\langle T\mathbf{x},\mathbf{y}\rangle+\langle T\mathbf{y},\mathbf{x}\rangle+\langle T\mathbf{y},\mathbf{y}\rangle.$$

The left side and the first and last terms on the right are 0 by hypothesis, so

$$\langle T\mathbf{x},\mathbf{y}\rangle+\langle T\mathbf{y},\mathbf{x}\rangle=0. \tag{6.1.3}$$

Similarly, $T(\mathbf{x}+i\mathbf{y})=T\mathbf{x}+iT\mathbf{y}$, and Proposition 6.1.4 gives

$$\langle T\mathbf{x}+iT\mathbf{y},\mathbf{x}+i\mathbf{y}\rangle=\langle T\mathbf{x},\mathbf{x}\rangle+\bar{i}\langle T\mathbf{x},\mathbf{y}\rangle+i\langle T\mathbf{y},\mathbf{x}\rangle+i\bar{i}\langle T\mathbf{y},\mathbf{y}\rangle.$$

Again the left side and the first and last terms vanish, so $-i\langle T\mathbf{x},\mathbf{y}\rangle+i\langle T\mathbf{y},\mathbf{x}\rangle=0$. Dividing by $-i$,

$$\langle T\mathbf{x},\mathbf{y}\rangle-\langle T\mathbf{y},\mathbf{x}\rangle=0.$$

Adding this to (6.1.3) gives $2\langle T\mathbf{x},\mathbf{y}\rangle=0$, so $\langle T\mathbf{x},\mathbf{y}\rangle=0=\langle 0\mathbf{x},\mathbf{y}\rangle$ for all $\mathbf{x},\mathbf{y}\in V$. By Corollary 6.1.38, $T=0$.

For the last statement, $S-T\in\mathrm{L}(V)$ (Proposition 2.1.23), and $\langle (S-T)\mathbf{v},\mathbf{v}\rangle=\langle S\mathbf{v},\mathbf{v}\rangle-\langle T\mathbf{v},\mathbf{v}\rangle=0$ for all $\mathbf{v}$. So $S-T=0$. □

Proof (by polarizing a sesquilinear function). STANDARD

Uses. Corollary 6.1.38 (Linear maps are determined by inner products), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.32 (Polarization identities)
Techniques. Polarization (Technique 6.1.33)
Hint 1. The function $B(\mathbf{x},\mathbf{y})=\langle T\mathbf{x},\mathbf{y}\rangle$ is linear in $\mathbf{x}$ and conjugate-linear in $\mathbf{y}$, like an inner product, but it need not be conjugate symmetric or positive. Try to redo the second proof of Proposition 6.1.32 for B; where that proof wrote $\bar{z}$ for $\langle\mathbf{y},\mathbf{x}\rangle$, keep $B(\mathbf{y},\mathbf{x})$ instead.
Hint 2. Expand $B\big(\mathbf{x}+i^k\mathbf{y},\mathbf{x}+i^k\mathbf{y}\big)$, multiply by i^k, and sum over $k=0,\dots,3$.

Sketch. Redo the roots-of-unity computation of Proposition 6.1.32 with $B(\mathbf{y},\mathbf{x})$ in place of $\overline{z}$; that term still carries the factor c^2, which sums to 0. So any B that is linear in the first slot and conjugate-linear in the second satisfies $B(\mathbf{x},\mathbf{y}) = \frac{1}{4}\sum_k i^k B\big(\mathbf{x}+i^k\mathbf{y},\mathbf{x}+i^k\mathbf{y}\big)$. For $B(\mathbf{x},\mathbf{y}) = \langle T\mathbf{x},\mathbf{y}\rangle$, every term on the right is 0.

Let $B(\mathbf{x},\mathbf{y}) = \langle T\mathbf{x},\mathbf{y}\rangle$. It is linear in $\mathbf{x}$, as T is linear and the inner product is linear in the first slot, and it satisfies $B(\mathbf{x},a\mathbf{y}+b\mathbf{z}) = \overline{a}B(\mathbf{x},\mathbf{y}) + \overline{b}B(\mathbf{x},\mathbf{z})$ by Proposition 6.1.4. For $c\in\mathbb{C}$ with $|c|=1$, expanding gives

$$B(\mathbf{x}+c\mathbf{y},\mathbf{x}+c\mathbf{y}) = B(\mathbf{x},\mathbf{x}) + \overline{c}B(\mathbf{x},\mathbf{y}) + cB(\mathbf{y},\mathbf{x}) + B(\mathbf{y},\mathbf{y}),$$

using $c\overline{c}=1$. Multiply by c and sum over $c=i^k$, $k=0,1,2,3$. Since $\sum_k i^k = 0$, $\sum_k i^k\overline{i^k} = 4$, and $\sum_k i^{2k} = 0$,

$$B(\mathbf{x},\mathbf{y}) = \frac{1}{4}\sum_{k=0}^{3} i^k B\big(\mathbf{x}+i^k\mathbf{y},\mathbf{x}+i^k\mathbf{y}\big). \tag{6.1.4}$$

We never used conjugate symmetry or positivity, only how B treats scalars in each slot. For our B, every term on the right is of the form $\langle T\mathbf{v},\mathbf{v}\rangle = 0$. So $\langle T\mathbf{x},\mathbf{y}\rangle = 0$ for all $\mathbf{x},\mathbf{y}$, and $T=0$ by Corollary 6.1.38. The last statement follows as in the first proof. □

The second proof shows more. Over $\mathbb{C}$, every function $B: V\times V\to\mathbb{C}$ that is linear in the first slot and conjugate-linear in the second is determined by its diagonal values $B(\mathbf{v},\mathbf{v})$, through (6.1.4), with no symmetry assumption. Over $\mathbb{R}$ the analogous computation gives

$$\frac{1}{4}(B(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) - B(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y})) = \frac{1}{2}(B(\mathbf{x},\mathbf{y}) + B(\mathbf{y},\mathbf{x})),$$

which recovers only the symmetric part of B. For the rotation of Example 6.1.39, $B(\mathbf{x},\mathbf{y}) = (J\mathbf{x})\cdot\mathbf{y} = x_1y_2 - x_2y_1$ is antisymmetric, so its symmetric part is 0 although B is not.

ASIDE: *Norms that do not come from inner products* BEYOND 21-242

The properties of $\|\cdot\|$ that matter in analysis are (a) and (b) of Proposition 6.1.27, together with the triangle inequality $\|\mathbf{v}+\mathbf{w}\| \le \|\mathbf{v}\| + \|\mathbf{w}\|$, which Section 6.2 proves for the norm of an inner product. Any function $V\to\mathbb{R}$ with these three properties is called a norm on V, and a vector space with a norm is a **normed space**. Every inner product space is a normed space, but $\mathbb{F}^n$ carries many norms that do not come from inner products, for instance

$$\|\mathbf{x}\|_1 = \sum_{k=1}^{n}|x_k|, \quad \|\mathbf{x}\|_\infty = \max_k |x_k|.$$

The first is the *taxicab norm*, the distance driven along a grid of streets; the second is the *maximum norm*. Checking the three properties for each is a short exercise. Their unit spheres in $\mathbb{R}^2$ are a diamond and a square (Figure 6.1.4), while those of inner products look like the ellipses of Figure 6.1.2.

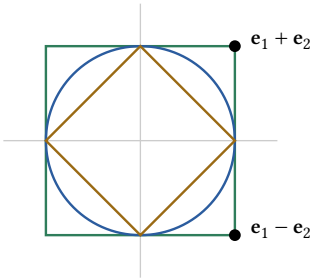


Figure 6.1.4. The unit spheres of $\|\cdot\|_1$ (diamond), of the Euclidean norm (circle), and of $\|\cdot\|_\infty$ (square) in $\mathbb{R}^2$. The marked points $\mathbf{e}_1\pm\mathbf{e}_2$ test the parallelogram law.

Corners suggest that no inner product is involved, and the parallelogram law (Proposition 6.1.30) makes this precise. In $\mathbb{R}^2$,

$$\|\mathbf{e}_1 + \mathbf{e}_2\|_1^2 + \|\mathbf{e}_1 - \mathbf{e}_2\|_1^2 = 4 + 4 = 8,$$
$$\|\mathbf{e}_1 + \mathbf{e}_2\|_\infty^2 + \|\mathbf{e}_1 - \mathbf{e}_2\|_\infty^2 = 1 + 1 = 2,$$

while $2\|\mathbf{e}_1\|^2 + 2\|\mathbf{e}_2\|^2 = 4$ for both norms. So neither satisfies the parallelogram law, and neither is the norm of any inner product on $\mathbb{R}^2$.

The parallelogram law is the only obstruction. The **Jordan–von Neumann theorem** says that a norm on a real or complex vector space is the norm of an inner product if and only if it satisfies the parallelogram law. The inner product is then given by the polarization identities, as it must be by Corollary 6.1.34. The real case is Exercise 6.1.61.

Exercises

Exercise 6.1.42. STANDARD Decide which of the following define inner products.

(a) On $\mathbb{R}^3$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_3 + x_3y_1$.

(b) On $\mathbb{R}^{2\times 2}$: $\langle A, B \rangle = \operatorname{tr}(A)\operatorname{tr}(B)$.

(c) On $\mathbb{C}^2$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1\overline{y_1} + ix_2\overline{y_2}$.

(d) On $\mathbb{R}[x]_{\le 2}$: $\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$.

(e) On $\mathbb{C}^2$: $\langle \mathbf{x}, \mathbf{y} \rangle = 2x_1\overline{y_1} - ix_2\overline{y_1} + ix_1\overline{y_2} + x_2\overline{y_2}$.

Solution on page 1463

Exercise 6.1.43. STANDARD The proof of Proposition 6.1.13 works on any closed interval $[a, b]$ in place of $[0, 1]$. Decide which of the following are inner products.

(a) $\langle f, g \rangle = \int_{-1}^{1} t^2 f(t)g(t)\,dt$ on $C[-1, 1]$.

(b) $\langle f, g \rangle = \int_{-1}^{1} t f(t)g(t)\,dt$ on $C[-1, 1]$.

(c) $\langle f, g \rangle = \int_{0}^{1/2} f(t)g(t)\,dt$ on $C[0, 1]$.

(d) $\langle f, g \rangle = f(0)g(0) + \int_0^1 f'(t)g'(t)\,dt$ on the space $C^1[0, 1]$ of continuously differentiable functions $[0, 1] \to \mathbb{R}$.

Solution on page 1464

Exercise 6.1.44. ROUTINE In a complex inner product space, $\|\mathbf{x}\| = 1$, $\|\mathbf{y}\| = 2$, and $\langle \mathbf{x}, \mathbf{y} \rangle = 1 - i$. Compute

(a) $\langle 2\mathbf{x} + i\mathbf{y}, \mathbf{x} - \mathbf{y} \rangle$;

(b) $\|\mathbf{x} - i\mathbf{y}\|^2$;

(c) $\|(1 + i)\mathbf{x} + \mathbf{y}\|^2$.

(d) Find $c \in \mathbb{C}$ with $\langle \mathbf{x} - c\mathbf{y}, \mathbf{y} \rangle = 0$, and compute $\|\mathbf{x} - c\mathbf{y}\|^2$.

Solution on page 1465

Exercise 6.1.45. STANDARD

(a) An inner product on $\mathbb{R}^2$ satisfies $\|\mathbf{e}_1\| = 1$, $\|\mathbf{e}_2\| = 2$, and $\|\mathbf{e}_1 + \mathbf{e}_2\| = 2$. Show that there is exactly one such inner product, and find a formula for $\langle \mathbf{x}, \mathbf{y} \rangle$.

(b) Compute $\|\mathbf{e}_1 - \mathbf{e}_2\|$ for the inner product of (a) in two ways: from your formula, and from the parallelogram law.

(c) Show that no inner product on $\mathbb{R}^2$ satisfies $\|\mathbf{e}_1\| = 1$, $\|\mathbf{e}_2\| = 2$, and $\|\mathbf{e}_1 + \mathbf{e}_2\| = 4$.

Solution on page 1465

Exercise 6.1.46. STANDARD For which $a, b, c \in \mathbb{R}$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2$$

an inner product on $\mathbb{R}^2$?

Solution on page 1466

Exercise 6.1.47. STANDARD Let $M \in \mathbb{C}^{n\times n}$, and suppose $\mathbf{x}^*M\mathbf{x} \in \mathbb{R}$ for every $\mathbf{x} \in \mathbb{C}^n$. Show that $M^* = M$. Conclude that over $\mathbb{C}$, the condition $M^* = M$ in Proposition 6.1.22(b) is implied by the condition $\mathbf{x}^*M\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$. (By Remark 6.1.24, this fails over $\mathbb{R}$.)

Solution on page 1466

Exercise 6.1.48. STANDARD Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be vectors in an inner product space, and let $\mathbf{m} = (\mathbf{x} + \mathbf{y})/2$. Prove **Apollonius's identity**

$$\|\mathbf{z} - \mathbf{x}\|^2 + \|\mathbf{z} - \mathbf{y}\|^2 = \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2 + 2\|\mathbf{z} - \mathbf{m}\|^2.$$

In $\mathbb{R}^2$, what does it say about the triangle with vertices $\mathbf{x}, \mathbf{y}, \mathbf{z}$?

Solution on page 1467

Exercise 6.1.49. STANDARD Let $\mathbf{x}, \mathbf{y}$ be vectors in an inner product space.

(a) Show that $\|\mathbf{x}\| = \|\mathbf{y}\|$ if and only if $\operatorname{Re}\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0$. Interpret the case $\mathbb{F} = \mathbb{R}$, $V = \mathbb{R}^2$ geometrically.

(b) Over $\mathbb{C}$, show that $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0$ if and only if $\|\mathbf{x}\| = \|\mathbf{y}\|$ and $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{R}$.

Solution on page 1467

Exercise 6.1.50. ROUTINE Let V be an inner product space, and let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span V. Show that if $\mathbf{x}, \mathbf{y} \in V$ satisfy $\langle \mathbf{x}, \mathbf{v}_j \rangle = \langle \mathbf{y}, \mathbf{v}_j \rangle$ for $j = 1, \ldots, k$, then $\mathbf{x} = \mathbf{y}$. Give an example showing that the spanning hypothesis cannot be dropped.

Solution on page 1467

Exercise 6.1.51. STANDARD Let V, W be inner product spaces over the same $\mathbb{F}$, and let $T : V \to W$ be linear with $\|T\mathbf{v}\| = \|\mathbf{v}\|$ for all $\mathbf{v} \in V$. Show that $\langle T\mathbf{v}, T\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$, and that T is injective. (Maps of this kind are studied in Section 8.4.)

Solution on page 1468

Exercise 6.1.52. STANDARD Let V be a real inner product space and $T \in \mathrm{L}(V)$.

(a) Prove or disprove: if $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for all $\mathbf{v} \in V$, then $T = 0$.

(b) Show that $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for all $\mathbf{v} \in V$ if and only if $\langle T\mathbf{v}, \mathbf{w} \rangle = -\langle \mathbf{v}, T\mathbf{w} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$.

(c) Suppose $\langle T\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, T\mathbf{w} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$. Show that if $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for all $\mathbf{v}$, then $T = 0$.

Solution on page 1468

Exercise 6.1.53. HARD Let V be a complex inner product space and $T \in \mathrm{L}(V)$. Show that $\langle T\mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$ for all $\mathbf{v} \in V$ if and only if $\langle T\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, T\mathbf{w} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$. (Such operators are called self-adjoint; see Section 6.6.)

Solution on page 1468

Exercise 6.1.54. HARD Let $A \in \mathbb{C}^{n\times n}$ satisfy $A^* = A$ and $A^k = 0$ for some $k \geq 1$. Show that $A = 0$. (Start by proving $(BC)^* = C^*B^*$ for square matrices B, C.)

Solution on page 1469

Exercise 6.1.55. ROUTINE Let $A \in \mathbb{R}^{n\times n}$ be skew-symmetric, $A^{\mathsf{T}} = -A$.

(a) Show that $\operatorname{tr}(A^2) \leq 0$, with equality only if $A = 0$.

(b) Deduce that no real skew-symmetric matrix satisfies $A^2 = I_n$. Show that $A^2 = -I_2$ is possible.

Solution on page 1469

Exercise 6.1.56. STANDARD For $\mathbf{u}, \mathbf{x} \in \mathbb{F}^m$ and $\mathbf{v}, \mathbf{y} \in \mathbb{F}^n$, the products $\mathbf{u}\mathbf{v}^*$ and $\mathbf{x}\mathbf{y}^*$ are $m \times n$ matrices. Show that, with the Frobenius inner product on $\mathbb{F}^{m\times n}$ and dot products on $\mathbb{F}^m$ and $\mathbb{F}^n$,

$$\langle \mathbf{u}\mathbf{v}^*, \mathbf{x}\mathbf{y}^* \rangle = (\mathbf{u} \cdot \mathbf{x})(\mathbf{y} \cdot \mathbf{v}),$$

and deduce that $\|\mathbf{u}\mathbf{v}^*\| = \|\mathbf{u}\|\|\mathbf{v}\|$.

Solution on page 1470

Exercise 6.1.57. HARD Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be unit vectors in an inner product space.

(a) Show that

$$\sum_{1 \leq i < j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|^2 = n^2 - \|\mathbf{x}_1 + \cdots + \mathbf{x}_n\|^2.$$

(b) Deduce that $\sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\|^2 \leq n^2$, with equality if and only if $\mathbf{x}_1 + \cdots + \mathbf{x}_n = \mathbf{0}$.

(c) For three unit vectors in $\mathbb{R}^2$, find the largest possible value of $\sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\|^2$ and a configuration attaining it.

Solution on page 1470

Exercise 6.1.58. HARD This exercise shows that for $n \geq 3$, the formula $\sum_k x_k y_k$ on $\mathbb{F}_p^n$ is never definite. (For $n = 1$ it is, since $x^2 = 0$ forces $x = 0$.)

(a) Find a nonzero $\mathbf{x} \in \mathbb{F}_5^2$ with $x_1^2 + x_2^2 = 0$.

(b) Show that for every prime p there is a nonzero $\mathbf{x} \in \mathbb{F}_p^3$ with $x_1^2 + x_2^2 + x_3^2 = 0$, and deduce the same in $\mathbb{F}_p^n$ for every $n \geq 3$.

Solution on page 1470

Exercise 6.1.59. STANDARD Find the error in the following argument, and decide whether the claim is true.

Claim. If $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is linear in each slot and $B(\mathbf{v}, \mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{R}^n$, then $B = 0$.

Proof. By the polarization identity, for all $\mathbf{x}, \mathbf{y}$,

$$B(\mathbf{x}, \mathbf{y}) = \frac{1}{4}(B(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - B(\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y})) = \frac{1}{4}(0 - 0) = 0.$$

Solution on page 1471

Exercise 6.1.60. STANDARD Let $R[0,1]$ be the real vector space of Riemann-integrable functions $[0,1] \to \mathbb{R}$. Find the error in the following argument, and decide whether the claim is true.

Claim. $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$ is an inner product on $R[0,1]$.

Proof. Products of integrable functions are integrable, so $\langle f, g\rangle$ is defined. Symmetry and linearity follow from the properties of the integral, and $\langle f, f\rangle = \int_0^1 f^2 \geq 0$ since $f^2 \geq 0$. Finally, if $\int_0^1 f^2 = 0$, then since $f^2 \geq 0$ we must have $f^2 = 0$, so $f = 0$.

Solution on page 1471

Exercise 6.1.61. CHALLENGE BEYOND 21-242 (*Jordan–von Neumann, real case.*) Let V be a real vector space with a function $\|\cdot\| : V \to \mathbb{R}$ satisfying, for all $\mathbf{v}, \mathbf{w} \in V$ and $a \in \mathbb{R}$: $\|\mathbf{v}\| \geq 0$ with equality only for $\mathbf{v} = \mathbf{0}$; $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$; $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$; and the parallelogram law $\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2\|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2$. Define

$$\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}\Big(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\Big).$$

(a) Show that $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{x}\rangle$, $\langle \mathbf{x}, \mathbf{x}\rangle = \|\mathbf{x}\|^2$, and $\langle -\mathbf{x}, \mathbf{y}\rangle = -\langle \mathbf{x}, \mathbf{y}\rangle$.

(b) Show that $\langle \mathbf{x} + \mathbf{z}, \mathbf{y}\rangle + \langle \mathbf{x} - \mathbf{z}, \mathbf{y}\rangle = 2\langle \mathbf{x}, \mathbf{y}\rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z}$.

(c) Deduce that $\langle \mathbf{u} + \mathbf{w}, \mathbf{y}\rangle = \langle \mathbf{u}, \mathbf{y}\rangle + \langle \mathbf{w}, \mathbf{y}\rangle$ for all $\mathbf{u}, \mathbf{w}, \mathbf{y}$.

(d) Show that $\langle t\mathbf{x}, \mathbf{y}\rangle = t\langle \mathbf{x}, \mathbf{y}\rangle$ for all $t \in \mathbb{R}$, and conclude that $\langle \cdot, \cdot\rangle$ is an inner product on V whose norm is $\|\cdot\|$.

Solution on page 1471

6.2 Cauchy–Schwarz and the triangle inequality

In Section 6.1.1 we found that the dot product on $\mathbb{R}^2$ records angles, $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta$, and proposed to turn this around: in any real inner product space, define the angle θ between nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ by

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

This defines an angle only if the right side lies in $[-1, 1]$. In $\mathbb{R}^2$ that is a property of the cosine, but for functions or matrices it is a claim about the inner product, and it needs a proof. The claim is the Cauchy–Schwarz inequality. It also gives the triangle inequality, which is what makes $\|\mathbf{x} - \mathbf{y}\|$ behave like a distance, and applied to the inner products of Section 6.1.3 it produces inequalities between sums and between integrals that have no visible connection to geometry.

Throughout, V is an inner product space over $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ (Notation 6.1.1).

6.2.1 The Cauchy–Schwarz inequality

Intuition (The shadow is shorter than the vector). Take $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ with $\mathbf{y} \neq \mathbf{0}$, and drop a perpendicular from the tip of $\mathbf{x}$ to the line through $\mathbf{y}$ (Figure 6.2.1). Its foot is a multiple $c\mathbf{y}$, the *shadow* of $\mathbf{x}$ on the line, and the rest of $\mathbf{x}$, namely $\mathbf{w} = \mathbf{x} - c\mathbf{y}$, is perpendicular to $\mathbf{y}$. By (6.1.1), the shadow has length

$$\|c\mathbf{y}\| = \|\mathbf{x}\||\cos\theta| = \frac{|\mathbf{x} \cdot \mathbf{y}|}{\|\mathbf{y}\|}.$$

The vectors $c\mathbf{y}$, $\mathbf{w}$, and $\mathbf{x}$ form a right triangle with hypotenuse $\mathbf{x}$, and a leg is never longer than the hypotenuse. So

$$\frac{|\mathbf{x}\cdot\mathbf{y}|}{\|\mathbf{y}\|}\le\|\mathbf{x}\|.$$

Multiplying by $\|\mathbf{y}\|$ gives $|\mathbf{x}\cdot\mathbf{y}|\le\|\mathbf{x}\|\|\mathbf{y}\|$, with equality exactly when the other leg $\mathbf{w}$ is $\mathbf{0}$, that is, when $\mathbf{x}$ lies on the line through $\mathbf{y}$.

In $\mathbb{R}^2$ this is just $|\cos\theta|\le 1$. But the argument never used cosines. It used the decomposition $\mathbf{x}=c\mathbf{y}+\mathbf{w}$ and Pythagoras, and both make sense in every inner product space. All we need is the right c. The requirement that $\mathbf{w}$ be perpendicular to $\mathbf{y}$ is $\langle\mathbf{x}-c\mathbf{y},\mathbf{y}\rangle=0$, which by linearity in the first slot reads $\langle\mathbf{x},\mathbf{y}\rangle=c\|\mathbf{y}\|^2$. So

$$c=\frac{\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{y}\|^2}.$$

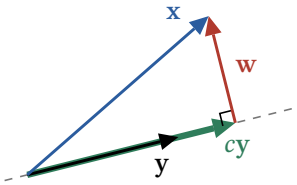


Figure 6.2.1. The shadow of $\mathbf{x}$ on the line through $\mathbf{y}$ is $c\mathbf{y}$, where $c=\langle\mathbf{x},\mathbf{y}\rangle/\|\mathbf{y}\|^2$, and $\mathbf{w}=\mathbf{x}-c\mathbf{y}$ is perpendicular to $\mathbf{y}$. By Pythagoras, $\|\mathbf{x}\|^2=\|c\mathbf{y}\|^2+\|\mathbf{w}\|^2$, so the shadow is no longer than $\mathbf{x}$.

Theorem 6.2.1 (Cauchy–Schwarz inequality). Let V be an inner product space and $\mathbf{x},\mathbf{y}\in V$.

(a) $|\langle\mathbf{x},\mathbf{y}\rangle|\le\|\mathbf{x}\|\|\mathbf{y}\|$.

(b) Equality holds in (a) if and only if one of $\mathbf{x},\mathbf{y}$ is a scalar multiple of the other.

Both sides of (a) are nonnegative, so (a) is equivalent to its square, $|\langle\mathbf{x},\mathbf{y}\rangle|^2\le\langle\mathbf{x},\mathbf{x}\rangle\langle\mathbf{y},\mathbf{y}\rangle$, which avoids square roots. For the dot product on $\mathbb{F}^n$ it says that

$$\left|\sum_{k=1}^{n}x_k\overline{y_k}\right|^2\le\left(\sum_{k=1}^{n}|x_k|^2\right)\left(\sum_{k=1}^{n}|y_k|^2\right)$$

for all complex numbers $x_1,\dots,x_n,y_1,\dots,y_n$, which is not obvious even for $n=2$.

Remark 6.2.2. In (b), "one of $\mathbf{x},\mathbf{y}$" cannot be shortened to "$\mathbf{x}$": if $\mathbf{x}\neq\mathbf{0}$ and $\mathbf{y}=\mathbf{0}$, both sides of (a) are 0, but $\mathbf{x}$ is not a multiple of $\mathbf{0}$. The condition in (b) says exactly that the list $\mathbf{x},\mathbf{y}$ is linearly dependent (Definition 1.5.1). Indeed, $\mathbf{x}=a\mathbf{y}$ gives the nontrivial relation $1\mathbf{x}-a\mathbf{y}=\mathbf{0}$, and $\mathbf{y}=b\mathbf{x}$ gives $b\mathbf{x}-1\mathbf{y}=\mathbf{0}$. Conversely, suppose $a\mathbf{x}+b\mathbf{y}=\mathbf{0}$ with a,b not both 0. If $a\neq 0$, then $\mathbf{x}=-(b/a)\mathbf{y}$. If $a=0$, then $b\neq 0$ and $b\mathbf{y}=\mathbf{0}$, so $\mathbf{y}=\mathbf{0}=0\mathbf{x}$ by Proposition 1.2.35.

Over $\mathbb{C}$ the scalar may be complex. For $\mathbf{x}\neq\mathbf{0}$, $|\langle\mathbf{x},i\mathbf{x}\rangle|=|-i|\|\mathbf{x}\|^2=\|\mathbf{x}\|\|i\mathbf{x}\|$, so $\mathbf{x}$ and $i\mathbf{x}$ give equality, even though in $\mathbb{C}=\mathbb{R}^2$ the numbers 1 and i are perpendicular arrows. We return to this in Section 6.2.3.

One direction of (b) is a direct computation, which we record first.

Lemma 6.2.3 (The easy half of (b)). If one of $\mathbf{x},\mathbf{y}\in V$ is a scalar multiple of the other, then $|\langle\mathbf{x},\mathbf{y}\rangle|=\|\mathbf{x}\|\|\mathbf{y}\|$.

Proof. ROUTINE

Uses. Proposition 6.1.27 (Basic properties of the norm), Proposition 6.1.4 (Sesquilinearity), Definition 6.1.26 (Norm)

Hint 1. Substitute $\mathbf{x} = a\mathbf{y}$ and take a out of the inner product.
Hint 2. Compare $|a|\|\mathbf{y}\|^2$ with $\|a\mathbf{y}\|\|\mathbf{y}\|$, using Proposition 6.1.27. For $\mathbf{y} = b\mathbf{x}$, the scalar leaves the second slot conjugated, and $|\overline{b}| = |b|$.
Sketch. If $\mathbf{x} = a\mathbf{y}$, both sides equal $|a|\|\mathbf{y}\|^2$. If $\mathbf{y} = b\mathbf{x}$, both sides equal $|b|\|\mathbf{x}\|^2$.

If $\mathbf{x} = a\mathbf{y}$, then by linearity in the first slot and Proposition 6.1.27,

$$|\langle a\mathbf{y}, \mathbf{y}\rangle| = |a|\|\mathbf{y}\|^2 = \|a\mathbf{y}\|\|\mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\|.$$

If $\mathbf{y} = b\mathbf{x}$, then by Proposition 6.1.4 and Proposition 6.1.27, $|\langle \mathbf{x}, b\mathbf{x}\rangle| = |\overline{b}|\|\mathbf{x}\|^2 = \|\mathbf{x}\|\|b\mathbf{x}\|$, since $|\overline{b}| = |b|$. □

So the content of the theorem is (a), together with the converse of Lemma 6.2.3: equality forces one vector to be a multiple of the other. We give three proofs. Each one finds a squared norm that is nonnegative for an obvious reason, expands it, and reads off (a). Each one gets the equality case by asking when that squared norm is 0.

First proof: the shadow

The first proof is the picture of Figure 6.2.1, carried out in V. The move it makes is used again and again in this chapter, so we name it.

Technique 6.2.4 (Orthogonal decomposition). To compare a vector $\mathbf{x}$ with a nonzero vector $\mathbf{y}$, split $\mathbf{x}$ into a part along $\mathbf{y}$ and a part perpendicular to $\mathbf{y}$:

$$\mathbf{x} = c\mathbf{y} + \mathbf{w}, \quad c = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{y}\|^2}, \quad \mathbf{w} = \mathbf{x} - c\mathbf{y}.$$

This choice of c is exactly the one that makes $\langle \mathbf{w}, \mathbf{y}\rangle = 0$. Then the cross term in Proposition 6.1.29 vanishes, and $\|\mathbf{x}\|^2 = |c|^2\|\mathbf{y}\|^2 + \|\mathbf{w}\|^2$. So the length of $\mathbf{x}$ splits into a contribution from the direction $\mathbf{y}$ and a contribution from everything else. Dropping $\|\mathbf{w}\|^2 \geq 0$ gives a lower bound for $\|\mathbf{x}\|$ that is sharp exactly when $\mathbf{w} = \mathbf{0}$.

Reach for it when a statement compares $\mathbf{x}$ with one fixed direction, or asks how close $\mathbf{x}$ comes to the line through $\mathbf{y}$. The canonical example is the first proof of Theorem 6.2.1. In Section 6.4 the line is replaced by an arbitrary subspace, the part along $\mathbf{y}$ becomes the orthogonal projection, and the same computation solves best-approximation problems (Theorem 6.4.33) and least squares (Section 6.5).
Used in. Theorem 6.2.1, Exercise 6.2.10, Exercise 6.2.27, Exercise 6.2.45, Theorem 6.3.24, Theorem 6.4.33, Proposition 6.5.27, Theorem 6.6.3, Exercise 6.7.12, Problem 6.1, Problem 6.13, Proposition 8.4.21, Exercise 6.4.57, Exercise 6.4.62.

Proof of Theorem 6.2.1 (by orthogonal decomposition). STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.1.4 (Sesquilinearity), Lemma 6.2.3 (The easy half of (b))
Techniques. Orthogonal decomposition (Technique 6.2.4)
Hint 1. Dispose of $\mathbf{y} = \mathbf{0}$ first. For $\mathbf{y} \neq \mathbf{0}$, write $\mathbf{x}$ as a multiple of $\mathbf{y}$ plus a vector $\mathbf{w}$ with $\langle \mathbf{w}, \mathbf{y}\rangle = 0$.
Hint 2. The multiple is $c\mathbf{y}$ with $c = \langle \mathbf{x}, \mathbf{y}\rangle/\|\mathbf{y}\|^2$. Expand $\|\mathbf{x}\|^2 = \|\mathbf{w} + c\mathbf{y}\|^2$ with Proposition 6.1.29.
Hint 3. You should get $\|\mathbf{x}\|^2 = \|\mathbf{w}\|^2 + |\langle \mathbf{x}, \mathbf{y}\rangle|^2/\|\mathbf{y}\|^2$. Drop $\|\mathbf{w}\|^2$. For the equality case, ask when that term is 0.
Sketch. If $\mathbf{y} = \mathbf{0}$, both sides vanish. Otherwise put $c = \langle \mathbf{x}, \mathbf{y}\rangle/\|\mathbf{y}\|^2$ and $\mathbf{w} = \mathbf{x} - c\mathbf{y}$, so that $\langle \mathbf{w}, \mathbf{y}\rangle = 0$. Expanding $\|\mathbf{w} + c\mathbf{y}\|^2$, the cross term vanishes, and $\|\mathbf{x}\|^2 = \|\mathbf{w}\|^2 + |\langle \mathbf{x}, \mathbf{y}\rangle|^2/\|\mathbf{y}\|^2$. This gives (a), with equality if and only if $\mathbf{w} = \mathbf{0}$, that is, $\mathbf{x} = c\mathbf{y}$.

If $\mathbf{y} = \mathbf{0}$, then $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ by Proposition 6.1.4 and $\|\mathbf{y}\| = 0$, so both sides of (a) are 0. Equality holds, and $\mathbf{y} = 0\mathbf{x}$ is a multiple of $\mathbf{x}$, so (b) holds as well.

Suppose $\mathbf{y} \neq \mathbf{0}$, so that $\|\mathbf{y}\| > 0$ by Proposition 6.1.27. Let

$$c = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{y}\|^2} \quad \text{and} \quad \mathbf{w} = \mathbf{x} - c\mathbf{y}.$$

By linearity in the first slot, $\langle \mathbf{w}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle - c\langle \mathbf{y}, \mathbf{y} \rangle = 0$. Since $\mathbf{x} = \mathbf{w} + c\mathbf{y}$, Proposition 6.1.29 with $a = c$ gives

$$\begin{aligned} \|\mathbf{x}\|^2 &= \|\mathbf{w}\|^2 + 2\,\mathrm{Re}(\bar{c}\langle \mathbf{w}, \mathbf{y} \rangle) + |c|^2\|\mathbf{y}\|^2 \\ &= \|\mathbf{w}\|^2 + \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{y}\|^4}\|\mathbf{y}\|^2 \\ &= \|\mathbf{w}\|^2 + \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{y}\|^2}. \end{aligned} \tag{6.2.1}$$

Since $\|\mathbf{w}\|^2 \geq 0$, it follows that $|\langle \mathbf{x}, \mathbf{y} \rangle|^2 / \|\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2$. Multiplying by $\|\mathbf{y}\|^2 > 0$ gives $|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2$, and taking nonnegative square roots gives (a).

For (b), (6.2.1) shows that equality holds in (a) if and only if $\|\mathbf{w}\|^2 = 0$, that is (by Proposition 6.1.27), if and only if $\mathbf{x} = c\mathbf{y}$. So if equality holds, then $\mathbf{x}$ is a multiple of $\mathbf{y}$. The converse is Lemma 6.2.3. □

Remark 6.2.5 (The size of the gap). Multiplying (6.2.1) by $\|\mathbf{y}\|^2$ gives, for $\mathbf{y} \neq \mathbf{0}$,

$$\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 = \|\mathbf{y}\|^2\|\mathbf{x} - c\mathbf{y}\|^2. \tag{6.2.2}$$

So the first proof yields more than an inequality: it computes the gap, and the gap measures how far $\mathbf{x}$ is from the line through $\mathbf{y}$. In $\mathbb{R}^2$ with the dot product, $\|\mathbf{y}\|$ is the base and $\|\mathbf{x} - c\mathbf{y}\|$ the height of the parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$, so the right side is the square of its area. A direct expansion shows that the left side is $(x_1y_2 - x_2y_1)^2$, and Exercise 6.2.43 extends this to $\mathbb{R}^n$.

Second proof: a quadratic in t

The second proof needs no picture and no clever choice. It needs one fact about real quadratics.

Lemma 6.2.6 (Nonnegative quadratics). Let $a, b, c \in \mathbb{R}$ with $a > 0$, and let $q(t) = at^2 - 2bt + c$ for $t \in \mathbb{R}$.

(a) $q(t) \geq 0$ for all $t \in \mathbb{R}$ if and only if $b^2 \leq ac$.

(b) If $b^2 \leq ac$, then q has a real root if and only if $b^2 = ac$, and in that case its only root is $t = b/a$.

Proof. ROUTINE

Uses. completing the square
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Complete the square. Multiplying by a first avoids fractions.
Hint 2. Check that $aq(t) = (at - b)^2 + (ac - b^2)$ for all t. For the converse in (a), evaluate at the t that kills the square.
Sketch. $aq(t) = (at - b)^2 + (ac - b^2)$. If $b^2 \leq ac$, both terms are nonnegative, so $q \geq 0$; conversely $aq(b/a) = ac - b^2$. A sum of two nonnegative terms vanishes only if both do, which happens only for $t = b/a$ and $b^2 = ac$.

For every $t \in \mathbb{R}$,

$$(at - b)^2 + (ac - b^2) = a^2t^2 - 2abt + ac = aq(t).$$

(a) If $b^2 \leq ac$, then $aq(t)$ is a sum of two nonnegative numbers, so $aq(t) \geq 0$, and $q(t) \geq 0$ since $a > 0$. Conversely, if $q(t) \geq 0$ for all t, then in particular $aq(b/a) = 0 + (ac - b^2) \geq 0$.

(b) Suppose $b^2 \leq ac$. Then $aq(t)$ is the sum of the nonnegative numbers $(at - b)^2$ and $ac - b^2$, so $q(t) = 0$ if and only if both are 0, that is, if and only if $t = b/a$ and $b^2 = ac$. So if $b^2 = ac$, then b/a is a root and the only one, and if $b^2 < ac$, there is no root. □

Technique 6.2.7 (A nonnegative quadratic has nonpositive discriminant). If a real quadratic $at^2 - 2bt + c$ with $a > 0$ is nonnegative for every real t, then $b^2 \leq ac$, with equality exactly when the quadratic has a real root (Lemma 6.2.6). So to prove an inequality of the shape $b^2 \leq ac$, look for an expression that is nonnegative for every real t for an obvious reason, usually a squared norm $\|\mathbf{x} - t\mathbf{y}\|^2$ or a sum of squares,

and that expands to $at^2 - 2bt + c$. The equality case comes for free: it says that the nonnegative expression vanishes for some t.

Reach for it when the target has the shape $b^2 \le ac$ with $a, c \ge 0$, especially when no geometric picture suggests a decomposition. The parameter t must be real; over $\mathbb{C}$, first make the middle coefficient real (Technique 6.2.8). The canonical example is the second proof of Theorem 6.2.1.

Used in. Theorem 6.2.1, Exercise 6.2.35, Problem 6.1, Problem 6.17, Lemma 8.2.10.

Over $\mathbb{C}$ there is a complication. For real t, Proposition 6.1.29 gives $\|\mathbf{x} - t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2t\,\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle + t^2\|\mathbf{y}\|^2$, in which only the real part of $\langle \mathbf{x}, \mathbf{y}\rangle$ appears. The cure is to rotate $\mathbf{x}$ first.

Technique 6.2.8 (Rotate the phase). Every $z \in \mathbb{C}$ can be written $z = |z|\mu$ with $|\mu| = 1$: take $\mu = z/|z|$ if $z \neq 0$, and $\mu = 1$ if $z = 0$. Then $\overline{\mu} z = |z|$, so multiplying by $\overline{\mu}$ rotates z onto the nonnegative real axis without changing its modulus. In an inner product space, replacing $\mathbf{x}$ by $\overline{\mu}\mathbf{x}$ changes no norm, since $|\overline{\mu}| = 1$, and it turns $\langle \mathbf{x}, \mathbf{y}\rangle$ into $\overline{\mu}\langle \mathbf{x}, \mathbf{y}\rangle$. So in a statement that involves only $\|\mathbf{x}\|$, $\|\mathbf{y}\|$, and $|\langle \mathbf{x}, \mathbf{y}\rangle|$, we may assume that $\langle \mathbf{x}, \mathbf{y}\rangle$ is a nonnegative real number, and real-variable arguments apply. The same observation gives

$$|z| = \max_{|\lambda|=1} \mathrm{Re}(\lambda z),$$

since $\mathrm{Re}(\lambda z) \le |\lambda z| = |z|$ for $|\lambda| = 1$, with equality at $\lambda = \overline{\mu}$.

Reach for it when a real-variable argument, such as a discriminant argument, must be run over $\mathbb{C}$, or when $\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$ and $|\langle \mathbf{x}, \mathbf{y}\rangle|$ must be compared. The canonical example is the second proof of Theorem 6.2.1.

Used in. Theorem 6.2.1, Exercise 6.2.29, Exercise 6.7.6, Theorem B.1.11, Exercise B.1.44, Exercise B.1.49.

Proof of Theorem 6.2.1 (by a quadratic in t). STANDARD

Uses. Lemma 6.2.6 (Nonnegative quadratics), Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm), Lemma 6.2.3 (The easy half of (b)), Notation 6.1.1 (Real and complex scalars)
Techniques. A nonnegative quadratic has nonpositive discriminant (Technique 6.2.7), Rotate the phase (Technique 6.2.8)
Hint 1. For real t, the number $\|\mathbf{x} - t\mathbf{y}\|^2$ is nonnegative. Expand it as a function of t.
Hint 2. Over $\mathbb{R}$ you get $\|\mathbf{x}\|^2 - 2t\langle \mathbf{x}, \mathbf{y}\rangle + t^2\|\mathbf{y}\|^2 \ge 0$ for all t, and Lemma 6.2.6 applies when $\mathbf{y} \neq \mathbf{0}$. Over $\mathbb{C}$ the middle coefficient is only $\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$.
Hint 3. Over $\mathbb{C}$, first replace $\mathbf{x}$ by $\lambda\mathbf{x}$, with $|\lambda| = 1$ chosen so that $\langle \lambda\mathbf{x}, \mathbf{y}\rangle = |\langle \mathbf{x}, \mathbf{y}\rangle|$. This changes neither side of (a).
Sketch. Assume $\mathbf{y} \neq \mathbf{0}$. Choose $|\lambda| = 1$ with $\lambda\langle \mathbf{x}, \mathbf{y}\rangle = |\langle \mathbf{x}, \mathbf{y}\rangle|$. For real t, $0 \le \|\lambda\mathbf{x} - t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2t|\langle \mathbf{x}, \mathbf{y}\rangle| + t^2\|\mathbf{y}\|^2$. A quadratic with positive leading coefficient that is never negative has $b^2 \le ac$, which is (a). Equality means that the quadratic has a real root t_0, so $\lambda\mathbf{x} = t_0\mathbf{y}$.

If $\mathbf{y} = \mathbf{0}$, both sides of (a) are 0, and $\mathbf{y} = 0\mathbf{x}$, so (a) and (b) hold. Suppose $\mathbf{y} \neq \mathbf{0}$. By Lemma 6.2.3, it remains to prove (a) and to show that equality forces one vector to be a multiple of the other.

Write $\langle \mathbf{x}, \mathbf{y}\rangle = |\langle \mathbf{x}, \mathbf{y}\rangle|\mu$ with $|\mu| = 1$, and let $\lambda = \overline{\mu}$ and $\mathbf{x}' = \lambda\mathbf{x}$. Since $\lambda\mu = |\mu|^2 = 1$,

$$\langle \mathbf{x}', \mathbf{y}\rangle = \lambda\langle \mathbf{x}, \mathbf{y}\rangle = |\langle \mathbf{x}, \mathbf{y}\rangle|,$$

a nonnegative real number, and $\|\mathbf{x}'\| = |\lambda|\|\mathbf{x}\| = \|\mathbf{x}\|$ by Proposition 6.1.27. Let $t \in \mathbb{R}$. By Proposition 6.1.29 with the scalar $-t$, using $\overline{-t} = -t$ and that $\langle \mathbf{x}', \mathbf{y}\rangle$ is real,

$$0 \le \|\mathbf{x}' - t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2t|\langle \mathbf{x}, \mathbf{y}\rangle| + t^2\|\mathbf{y}\|^2.$$

The right side is $q(t) = at^2 - 2bt + c$ with

$$a = \|\mathbf{y}\|^2 > 0, \quad b = |\langle \mathbf{x}, \mathbf{y}\rangle|, \quad c = \|\mathbf{x}\|^2,$$

and it is nonnegative for every real t. By Lemma 6.2.6, $b^2 \le ac$, that is, $|\langle \mathbf{x}, \mathbf{y}\rangle|^2 \le \|\mathbf{x}\|^2\|\mathbf{y}\|^2$. Taking nonnegative square roots gives (a).

If equality holds in (a), then $b^2 = ac$, so q has a real root t_0 by Lemma 6.2.6. Then $\|\mathbf{x}' - t_0\mathbf{y}\| = 0$, so $\lambda\mathbf{x} = t_0\mathbf{y}$ by Proposition 6.1.27. Multiplying by $\overline{\lambda}$ and using $\overline{\lambda}\lambda = 1$ gives $\mathbf{x} = \overline{\lambda}t_0\mathbf{y}$, a multiple of $\mathbf{y}$. □

Over $\mathbb{R}$, $\lambda = \pm 1$, and the rotation only flips the sign of $\mathbf{x}$ when $\langle \mathbf{x}, \mathbf{y} \rangle < 0$; one could instead apply Lemma 6.2.6 with $b = \langle \mathbf{x}, \mathbf{y} \rangle$ directly. Over $\mathbb{C}$ the rotation cannot be skipped. Run on $\mathbf{x}$ itself, the argument proves only $(\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle)^2 \le \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$, which is weaker than (a) whenever $\langle \mathbf{x}, \mathbf{y} \rangle$ is not real. Rotating moves all of $|\langle \mathbf{x}, \mathbf{y} \rangle|$ into the real part. (One can also let t be complex. The best t is then c from the first proof, and the computation becomes the first proof.)

Third proof: normalizing

> *Intuition (Balancing the two sides).* The cheapest inequality available is $\|\mathbf{x} - \mathbf{y}\|^2 \ge 0$, which by Proposition 6.1.29 says
>
> $$2 \operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle \le \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$
>
> It bounds a product by a sum, and when $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ are very different it is far from sharp. But both sides of (a) are *homogeneous*: replacing $\mathbf{x}$ by $s\mathbf{x}$ with $s > 0$ multiplies both $|\langle \mathbf{x}, \mathbf{y} \rangle|$ and $\|\mathbf{x}\|\|\mathbf{y}\|$ by s, and similarly for $\mathbf{y}$. So it suffices to prove (a) after rescaling $\mathbf{x}$ and $\mathbf{y}$, and we can rescale them to unit vectors. For unit vectors the right side above is $1 + 1$, which is twice the product $\|\mathbf{x}\|\|\mathbf{y}\| = 1$, and the cheap inequality becomes sharp enough.
>
> Equivalently, apply the cheap inequality to $s\mathbf{x}$ and $\mathbf{y}/s$, and choose $s > 0$ to make the bound $s^2\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2/s^2$ as small as possible. The choice $s^2 = \|\mathbf{y}\|/\|\mathbf{x}\|$ makes the two terms equal, each $\|\mathbf{x}\|\|\mathbf{y}\|$, and the bound becomes $2\|\mathbf{x}\|\|\mathbf{y}\|$.

Proof of Theorem 6.2.1 (by normalizing). STANDARD

> *Uses.* Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.1.4 (Sesquilinearity), Lemma 6.2.3 (The easy half of (b))
> *Techniques.* Rotate the phase (Technique 6.2.8)
> *Hint 1.* Dispose of the case where $\mathbf{x}$ or $\mathbf{y}$ is $\mathbf{0}$. Otherwise, compare $\mathbf{x}/\|\mathbf{x}\|$ and $\mathbf{y}/\|\mathbf{y}\|$.
> *Hint 2.* For unit vectors $\mathbf{u}, \mathbf{v}$, expand $0 \le \|\mathbf{u} - \mathbf{v}\|^2$. You get $\operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle \le 1$.
> *Hint 3.* To get $|\langle \mathbf{u}, \mathbf{v} \rangle| \le 1$, first rotate $\mathbf{u}$ by a scalar of modulus 1 so that $\langle \mathbf{u}, \mathbf{v} \rangle$ is a nonnegative real number.
> *Sketch.* For nonzero $\mathbf{x}, \mathbf{y}$, let $\mathbf{u} = \lambda\mathbf{x}/\|\mathbf{x}\|$ and $\mathbf{v} = \mathbf{y}/\|\mathbf{y}\|$, where $|\lambda| = 1$ makes $\langle \mathbf{u}, \mathbf{v} \rangle = |\langle \mathbf{x}, \mathbf{y} \rangle|/(\|\mathbf{x}\|\|\mathbf{y}\|)$ real and nonnegative. Then $0 \le \|\mathbf{u} - \mathbf{v}\|^2 = 2 - 2\langle \mathbf{u}, \mathbf{v} \rangle$ gives (a), and equality forces $\mathbf{u} = \mathbf{v}$, so $\mathbf{y}$ is a multiple of $\mathbf{x}$.

If $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$, both sides of (a) are 0 by Proposition 6.1.4, and one vector is 0 times the other, so (a) and (b) hold. Suppose $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. By Lemma 6.2.3, it remains to prove (a) and to show that equality forces one vector to be a multiple of the other.

Choose $\lambda \in \mathbb{F}$ with $|\lambda| = 1$ and $\lambda\langle \mathbf{x}, \mathbf{y} \rangle = |\langle \mathbf{x}, \mathbf{y} \rangle|$, as in Technique 6.2.8, and let

$$\mathbf{u} = \frac{\lambda}{\|\mathbf{x}\|}\mathbf{x}, \quad \mathbf{v} = \frac{1}{\|\mathbf{y}\|}\mathbf{y}.$$

By Proposition 6.1.27, $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. Since $1/\|\mathbf{y}\|$ is real, it leaves the second slot unconjugated, and

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{\lambda}{\|\mathbf{x}\|\|\mathbf{y}\|}\langle \mathbf{x}, \mathbf{y} \rangle = \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|},$$

a nonnegative real number. By Proposition 6.1.29,

$$0 \le \|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 = 2 - 2\frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

Rearranging gives (a).

If equality holds in (a), the right side is 0, so $\|\mathbf{u} - \mathbf{v}\| = 0$ and $\mathbf{u} = \mathbf{v}$ by Proposition 6.1.27. That is, $\lambda\mathbf{x}/\|\mathbf{x}\| = \mathbf{y}/\|\mathbf{y}\|$, so $\mathbf{y} = (\lambda\|\mathbf{y}\|/\|\mathbf{x}\|)\mathbf{x}$ is a multiple of $\mathbf{x}$. □

Intuition (Comparing the three proofs). All three proofs come from one fact: $\|\mathbf{x}-t\mathbf{y}\|^2 \geq 0$ for a scalar t (in the second and third proofs, after rotating the phase of $\mathbf{x}$). They differ in how t is chosen.

The first proof takes the best t, namely $t = c = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$. It makes $\mathbf{x} - t\mathbf{y}$ perpendicular to $\mathbf{y}$, so $c\mathbf{y}$ is the point of the line through $\mathbf{y}$ closest to $\mathbf{x}$. Finding c takes the picture or a computation, but once it is found the proof is one expansion, and it produces the identity (6.2.2) rather than just an inequality.

The second proof does not choose t at all. It uses every real t at once and lets the discriminant do the optimizing: $b^2 \leq ac$ says that the minimum value of the quadratic, attained at $t = b/a$, is nonnegative. For $\mathbf{x}$ rotated so that $\langle \mathbf{x}, \mathbf{y}\rangle \geq 0$, that minimizer $b/a = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$ is exactly the coefficient of the first proof. This is the proof to remember when an inequality has the shape $b^2 \leq ac$ and there is no picture: find a quantity that is nonnegative and quadratic in a real parameter.

The third proof rescales so that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$, after which the crude choice $t = 1$ already suffices. It needs a separate zero case and a normalization step, but its idea, that a homogeneous inequality can be checked on normalized vectors, applies far beyond inner products.

The equality case works the same way in all three: equality means that the nonnegative quantity is 0 for the chosen t, and that says one vector is a multiple of the other.

Example 6.2.9 (Instances of Cauchy–Schwarz). Each inner product of Section 6.1.3 gives an inequality.

(a) For the dot product on $\mathbb{R}^n$: $\left(\sum_k a_k b_k\right)^2 \leq \left(\sum_k a_k^2\right)\left(\sum_k b_k^2\right)$ for all real a_k, b_k. For $(1,2,3)$ and $(4,5,6)$ this reads $32^2 = 1024 \leq 14 \cdot 77 = 1078$.

(b) For the integral inner product on $C[a,b]$ (Proposition 6.1.13 and the paragraph after it):

$$\left(\int_a^b f(t)g(t)\,dt\right)^2 \leq \left(\int_a^b f(t)^2\,dt\right)\left(\int_a^b g(t)^2\,dt\right).$$

For $f(t) = 1$ and $g(t) = t$ on $[0,1]$, this reads $1/4 \leq 1/3$.

(c) For the Frobenius inner product on $\mathbb{F}^{m\times n}$ (Proposition 6.1.16): $|\operatorname{tr}(B^*A)|^2 \leq \operatorname{tr}(A^*A)\operatorname{tr}(B^*B)$.

In each case equality holds exactly for proportional data. In (b), that means $g = cf$ or $f = cg$ as functions on $[a,b]$.

Exercise 6.2.10. ROUTINE In $\mathbb{C}^2$ with the dot product, let $\mathbf{x} = (1+i, 2)$ and $\mathbf{y} = (i, 1-i)$, as in Example 6.1.11, where we found $\mathbf{x}\cdot\mathbf{y} = 3+i$, $\|\mathbf{x}\|^2 = 6$, and $\|\mathbf{y}\|^2 = 3$. Find $c \in \mathbb{C}$ and $\mathbf{w} \in \mathbb{C}^2$ with $\mathbf{x} = c\mathbf{y} + \mathbf{w}$ and $\mathbf{w}\cdot\mathbf{y} = 0$, and check (6.2.1) numerically.

Solution on page 1472

6.2.2 The triangle inequality

In a triangle, each side is at most as long as the other two together. The sides of the triangle with vertices $\mathbf{0}$, $\mathbf{x}$, and $\mathbf{x}+\mathbf{y}$ are $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x}+\mathbf{y}$ (Figure 6.2.2), so the statement to prove is $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. Both sides are nonnegative, so we may compare squares:

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2,$$
$$(\|\mathbf{x}\| + \|\mathbf{y}\|)^2 = \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2.$$

So the triangle inequality is the statement $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle \leq \|\mathbf{x}\|\|\mathbf{y}\|$, and Cauchy–Schwarz gives this with room to spare, since $\operatorname{Re} z \leq |z|$.



(a) (b)

Figure 6.2.2. (a) The triangle inequality: the side $\mathbf{x}+\mathbf{y}$ is no longer than the path $\mathbf{x}$ followed by $\mathbf{y}$. (b) Equality: when $\mathbf{y}$ is a positive multiple of $\mathbf{x}$, the triangle is flat. (The arrow $\mathbf{x}+\mathbf{y}$ is drawn slightly offset.)

Theorem 6.2.11 (Triangle inequality). Let V be an inner product space and $\mathbf{x}, \mathbf{y} \in V$.

(a) $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$.

(b) The following are equivalent:
  (i) $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$;
  (ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\|$;
  (iii) one of $\mathbf{x}, \mathbf{y}$ is a nonnegative real multiple of the other.

Proof. STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)
Hint 1. Both sides of (a) are nonnegative, so compare their squares. Expand $\|\mathbf{x}+\mathbf{y}\|^2$ with Proposition 6.1.29.
Hint 2. The squared inequality reduces to $\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle \le \|\mathbf{x}\|\|\mathbf{y}\|$. Pass through $|\langle \mathbf{x}, \mathbf{y} \rangle|$.
Hint 3. For (b), equality in (a) forces both steps $\operatorname{Re} z \le |z|$ and Cauchy–Schwarz to be equalities. When is $\operatorname{Re} z = |z|$? For (ii) $\Rightarrow$ (iii), use Theorem 6.2.1(b) and then compare $\langle a\mathbf{y}, \mathbf{y} \rangle$ with $\|a\mathbf{y}\|\|\mathbf{y}\|$.
Sketch. In the expansion of $\|\mathbf{x}+\mathbf{y}\|^2$, the cross term $2\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle$ is at most $2|\langle \mathbf{x}, \mathbf{y} \rangle|$, which is at most $2\|\mathbf{x}\|\|\mathbf{y}\|$ by Cauchy–Schwarz; so $\|\mathbf{x}+\mathbf{y}\|^2 \le (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$. Equality holds if and only if both steps are equalities, that is, $\langle \mathbf{x}, \mathbf{y} \rangle$ is a nonnegative real number equal to $\|\mathbf{x}\|\|\mathbf{y}\|$. By Cauchy–Schwarz, this forces $\mathbf{x} = a\mathbf{y}$ (when $\mathbf{y} \neq \mathbf{0}$), and then $a\|\mathbf{y}\|^2 = |a|\|\mathbf{y}\|^2$ gives $a \ge 0$. The converse is a direct computation.

(a) By Proposition 6.1.29, $\operatorname{Re} z \le |z|$ (Notation 6.1.1), and Theorem 6.2.1,

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\ &\le \|\mathbf{x}\|^2 + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|^2 \\ &\le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned} \tag{6.2.3}$$

Both $\|\mathbf{x}+\mathbf{y}\|$ and $\|\mathbf{x}\| + \|\mathbf{y}\|$ are nonnegative, so taking square roots gives (a).

(b) We prove (i) $\Leftrightarrow$ (ii) and (ii) $\Leftrightarrow$ (iii).

(i) $\Leftrightarrow$ (ii). Since both sides of (a) are nonnegative, (i) holds if and only if the first and last expressions in (6.2.3) are equal. Their difference is the sum of the two nonnegative gaps in the chain, so this happens if and only if both inequalities in the chain are equalities:

$$\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle = |\langle \mathbf{x}, \mathbf{y} \rangle| \quad \text{and} \quad |\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\|\|\mathbf{y}\|.$$

For $z \in \mathbb{C}$, $\operatorname{Re} z = |z|$ if and only if z is a nonnegative real number. Indeed, writing $z = a + bi$ with $a, b \in \mathbb{R}$, we have $\operatorname{Re} z = a \le |a| \le \sqrt{a^2 + b^2} = |z|$, with equality in the first step if and only if $a \ge 0$ and in the second if and only if $b = 0$. So the two equalities together say that $\langle \mathbf{x}, \mathbf{y} \rangle$ is a nonnegative real number equal to $\|\mathbf{x}\|\|\mathbf{y}\|$, which is (ii). Conversely, (ii) gives both equalities.

(iii) $\Rightarrow$ (ii). If $\mathbf{x} = a\mathbf{y}$ with $a \ge 0$, then $|a| = a$, and by Proposition 6.1.27, $\langle \mathbf{x}, \mathbf{y} \rangle = a\|\mathbf{y}\|^2 = \|a\mathbf{y}\|\|\mathbf{y}\| = \|\mathbf{x}\|\|\mathbf{y}\|$. If $\mathbf{y} = b\mathbf{x}$ with $b \ge 0$, then $\overline{b} = b$, and by Proposition 6.1.4, $\langle \mathbf{x}, \mathbf{y} \rangle = b\|\mathbf{x}\|^2 = \|\mathbf{x}\|\|\mathbf{y}\|$ in the same way.

(ii) ⇒ (iii). Assume (ii). If $\mathbf{y} = \mathbf{0}$, then $\mathbf{y} = 0\mathbf{x}$ and (iii) holds. Suppose $\mathbf{y} \neq \mathbf{0}$. By (ii), $|\langle \mathbf{x}, \mathbf{y}\rangle| = \|\mathbf{x}\|\|\mathbf{y}\|$, so by Theorem 6.2.1(b) one of $\mathbf{x}, \mathbf{y}$ is a multiple of the other. In fact $\mathbf{x}$ is a multiple of $\mathbf{y}$: if $\mathbf{y} = b\mathbf{x}$, then $b \neq 0$ because $\mathbf{y} \neq \mathbf{0}$, and $\mathbf{x} = b^{-1}\mathbf{y}$. Write $\mathbf{x} = a\mathbf{y}$ with $a \in \mathbb{F}$. Then

$$a\|\mathbf{y}\|^2 = \langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x}\|\|\mathbf{y}\| = |a|\|\mathbf{y}\|^2,$$

and dividing by $\|\mathbf{y}\|^2 > 0$ gives $a = |a|$, a nonnegative real number. □

Intuition (Same line versus same direction). Equality in Cauchy–Schwarz asks that $\mathbf{x}$ and $\mathbf{y}$ lie on a common line $\mathbb{F}\mathbf{v}$ through $\mathbf{0}$. Equality in the triangle inequality asks for more, a common ray $\{t\mathbf{v} \mid t \geq 0\}$. Over $\mathbb{R}$, if $\mathbf{y} = -\mathbf{x} \neq \mathbf{0}$, then $|\langle \mathbf{x}, \mathbf{y}\rangle| = \|\mathbf{x}\|\|\mathbf{y}\|$ but $\|\mathbf{x} + \mathbf{y}\| = 0 < 2\|\mathbf{x}\|$. Over $\mathbb{C}$, $\mathbf{y} = i\mathbf{x}$ gives equality in Cauchy–Schwarz, but $\|\mathbf{x} + i\mathbf{x}\| = |1 + i|\|\mathbf{x}\| = \sqrt{2}\|\mathbf{x}\|$. In the proof, it is the first inequality of (6.2.3), $\mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle \leq |\langle \mathbf{x}, \mathbf{y}\rangle|$, that detects the direction.

Exercise 6.2.12. ROUTINE Let $\mathbf{x} \neq \mathbf{0}$ be a vector in a complex inner product space. Find all $a \in \mathbb{C}$ with $\|\mathbf{x} + a\mathbf{x}\| = \|\mathbf{x}\| + \|a\mathbf{x}\|$, and all $a \in \mathbb{C}$ with $\|\mathbf{x} + a\mathbf{x}\| = |\|\mathbf{x}\| - \|a\mathbf{x}\||$.

Solution on page 1473

The triangle inequality extends to any number of vectors, and it has a reversed form.

Corollary 6.2.13 (Sums of several vectors). For $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$,

$$\|\mathbf{x}_1 + \cdots + \mathbf{x}_n\| \leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_n\|.$$

Proof. ROUTINE

Uses. Theorem 6.2.11 (Triangle inequality)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Induct on n.
Hint 2. Write $\mathbf{x}_1 + \cdots + \mathbf{x}_n = (\mathbf{x}_1 + \cdots + \mathbf{x}_{n-1}) + \mathbf{x}_n$ and apply Theorem 6.2.11 to these two vectors.
Sketch. The case $n = 1$ is trivial. For $n \geq 2$, the triangle inequality splits off $\|\mathbf{x}_n\|$, and the induction hypothesis bounds the norm of the remaining sum.

We induct on n. For $n = 1$ there is nothing to prove. Let $n \geq 2$, and assume the claim for $n - 1$ vectors. With $\mathbf{s} = \mathbf{x}_1 + \cdots + \mathbf{x}_{n-1}$, Theorem 6.2.11 and the induction hypothesis give

$$\|\mathbf{s} + \mathbf{x}_n\| \leq \|\mathbf{s}\| + \|\mathbf{x}_n\| \leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_{n-1}\| + \|\mathbf{x}_n\|,$$

as desired. □

Corollary 6.2.14 (Distances and the reverse triangle inequality). For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$:

(a) $\|\mathbf{x} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|$;

(b) $|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$.

Proof. ROUTINE

Uses. Theorem 6.2.11 (Triangle inequality), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. For (a), write $\mathbf{x} - \mathbf{z}$ as a sum of two vectors.
Hint 2. For (b), write $\mathbf{x} = (\mathbf{x} - \mathbf{y}) + \mathbf{y}$ and apply the triangle inequality. This bounds $\|\mathbf{x}\| - \|\mathbf{y}\|$. Then exchange the roles of $\mathbf{x}$ and $\mathbf{y}$.
Sketch. (a) is the triangle inequality for $(\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})$. For (b), $\|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|$ and $\|\mathbf{y}\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x}\|$, and $\|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$.

(a) Apply Theorem 6.2.11 to $\mathbf{x} - \mathbf{z} = (\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})$.

(b) By Theorem 6.2.11, $\|\mathbf{x}\| = \|(\mathbf{x} - \mathbf{y}) + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|$, so $\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$. Exchanging $\mathbf{x}$ and $\mathbf{y}$, $\|\mathbf{y}\| - \|\mathbf{x}\| \leq \|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$, where the last step is Proposition 6.1.27 with $a = -1$. Since $|r|$ is the larger of r and $-r$, (b) follows. □

Part (a) is the triangle inequality for the triangle with vertices $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$: going from $\mathbf{x}$ to $\mathbf{z}$ directly is never longer than going through $\mathbf{y}$. Part (b) says that a side of a triangle is at least the difference of the other two. Read as a statement about the function $\mathbf{v} \mapsto \|\mathbf{v}\|$, it says that moving $\mathbf{v}$ by a vector of norm δ changes $\|\mathbf{v}\|$ by at most δ; in particular, the norm is a continuous function. (Exercise 6.1.61 used exactly this.)

Remark 6.2.15. By Proposition 6.1.27 and Theorem 6.2.11, the norm of an inner product satisfies the three properties of a norm in the sense of Aside (Norms that do not come from inner products), so every inner product space is a normed space. By Corollary 6.2.14(a), the distance $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ also behaves as a distance should: $d(\mathbf{x}, \mathbf{y}) \geq 0$ with equality only for $\mathbf{x} = \mathbf{y}$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, and $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$.

6.2.3 Angles

Let V be a real inner product space and $\mathbf{x}, \mathbf{y} \in V$ nonzero. By Theorem 6.2.1,

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1,$$

and the cosine is a strictly decreasing continuous function from $[0, \pi]$ onto $[-1, 1]$, hence a bijection. So the following definition makes sense, and the motivation at the start of the section is justified.

Definition 6.2.16 (Angle). Let V be a real inner product space, and let $\mathbf{x}, \mathbf{y} \in V$ be nonzero. The **angle** between $\mathbf{x}$ and $\mathbf{y}$ is the unique $\theta \in [0, \pi]$ with

$$\cos\theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

For the dot product on $\mathbb{R}^2$, this is the usual angle, by (6.1.1). The familiar properties of angles follow from what we have proved.

Proposition 6.2.17 (Properties of angles). Let V be a real inner product space, let $\mathbf{x}, \mathbf{y} \in V$ be nonzero, and let θ be the angle between them.

(a) (Law of cosines.)

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta.$$

(b) $\theta = 0$ if and only if $\mathbf{y} = a\mathbf{x}$ for some $a > 0$, and $\theta = \pi$ if and only if $\mathbf{y} = a\mathbf{x}$ for some $a < 0$.

(c) $\theta = \pi/2$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Proof. ROUTINE

Uses. Definition 6.2.16 (Angle), Proposition 6.1.29 (Expanding a norm), Theorem 6.2.11 (Triangle inequality), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. For (a), expand $\|\mathbf{x} - \mathbf{y}\|^2$ and substitute the definition of θ.
Hint 2. For (b), $\theta = 0$ means $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\|$. Which result characterizes that equation? For $\theta = \pi$, replace $\mathbf{y}$ by $-\mathbf{y}$.
Sketch. (a) is Proposition 6.1.29 with $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$. In (b), $\cos\theta = 1$ means $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\|$, which by Theorem 6.2.11(b) means one vector is a nonnegative multiple of the other; as both are nonzero, the multiple is positive. The case $\theta = \pi$ is the case $\theta = 0$ for $\mathbf{x}$ and $-\mathbf{y}$. (c) is $\cos\theta = 0$.

(a) By Proposition 6.1.29, $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2$ (over $\mathbb{R}$, $\mathrm{Re}\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$), and $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$ by Definition 6.2.16.

(b) Since cos is injective on $[0, \pi]$ with $\cos 0 = 1$, we have $\theta = 0$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\|$. By Theorem 6.2.11(b), this holds if and only if one of $\mathbf{x}, \mathbf{y}$ is a nonnegative multiple of the other. As both are nonzero, the multiple is nonzero, hence positive, and if $\mathbf{x} = b\mathbf{y}$ with $b > 0$, then $\mathbf{y} = b^{-1}\mathbf{x}$ with $b^{-1} > 0$. So $\theta = 0$ if and only if $\mathbf{y} = a\mathbf{x}$ with $a > 0$.

Similarly, $\theta = \pi$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = -\|\mathbf{x}\|\|\mathbf{y}\|$, that is, $\langle \mathbf{x}, -\mathbf{y} \rangle = \|\mathbf{x}\|\|-\mathbf{y}\|$, using Proposition 6.1.4 and Proposition 6.1.27. By the first case applied to $\mathbf{x}$ and $-\mathbf{y}$, this holds if and only if $-\mathbf{y} = a\mathbf{x}$ with $a > 0$, that is, $\mathbf{y} = (-a)\mathbf{x}$ with $-a < 0$.

(c) $\cos\theta = 0$ for $\theta \in [0, \pi]$ if and only if $\theta = \pi/2$. □

Part (c) is why perpendicularity is defined by $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ in Section 6.3. In spaces without a picture, the angle still measures how far two vectors are from being proportional, and its values can be surprising.

Example 6.2.18 (Angles in $\mathbb{R}^n$ and in $C[0,1]$).

(a) In $\mathbb{R}^n$ with the dot product, let $\mathbf{d} = (1, 1, \ldots, 1)$, the diagonal of the unit cube, and $\mathbf{e}_1$, one of its edges. Then $\langle \mathbf{d}, \mathbf{e}_1 \rangle = 1$, $\|\mathbf{d}\| = \sqrt{n}$, and $\|\mathbf{e}_1\| = 1$, so the angle θ_n between them satisfies $\cos\theta_n = 1/\sqrt{n}$. For $n = 2, 3, 4$ this gives $\theta_n = \pi/4$, about $54.7°$, and $\pi/3$. As $n \to \infty$, $\cos\theta_n \to 0$ and $\theta_n \to \pi/2$: in high dimension the diagonal of a cube is almost perpendicular to every edge.

(b) In $C[0,1]$ with the integral inner product, $\langle 1, t \rangle = \int_0^1 t\,dt = 1/2$, $\|1\| = 1$, and $\|t\| = \left(\int_0^1 t^2\,dt\right)^{1/2} = 1/\sqrt{3}$. So the angle between the functions 1 and t satisfies $\cos\theta = \sqrt{3}/2$, that is, $\theta = \pi/6$.

Example 6.2.19 (Correlation is a cosine). Given data points $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^2$, let m_x and m_y be the means of the x_k and of the y_k, and let

$$\tilde{\mathbf{x}} = (x_1 - m_x, \ldots, x_m - m_x), \quad \tilde{\mathbf{y}} = \left(y_1 - m_y, \ldots, y_m - m_y\right)$$

be the centered data in $\mathbb{R}^m$. If both are nonzero, the **correlation coefficient** of the data is

$$r = \frac{\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}}}{\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|},$$

the cosine of the angle between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$. Cauchy–Schwarz says $-1 \le r \le 1$, and Proposition 6.2.17(b) says that $r = \pm 1$ exactly when $\tilde{\mathbf{y}} = a\tilde{\mathbf{x}}$ with a of the same sign as r, that is, when all data points lie on the line $y = m_y + a(x - m_x)$. For the points $(0,1)$, $(1,1)$, $(2,4)$, we get $\tilde{\mathbf{x}} = (-1, 0, 1)$ and $\tilde{\mathbf{y}} = (-1, -1, 2)$, so $r = 3/\left(\sqrt{2}\sqrt{6}\right) = \sqrt{3}/2$. Least squares (Section 6.5) finds the line that comes closest to such data.

> **Exercise 6.2.20.** ROUTINE In $C[0,1]$ with the integral inner product, find $\cos\theta$ for the angle θ between t and t^2. Then find all $a \in \mathbb{R}$ for which the angle between the functions 1 and $t + a$ is $\pi/3$.

Solution on page 1473

Over $\mathbb{C}$

Let V be a complex inner product space and $\mathbf{x}, \mathbf{y} \in V$ nonzero. Now $\langle \mathbf{x}, \mathbf{y} \rangle$ is a complex number, and there are two natural real candidates for a cosine:

$$\operatorname{Re}\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \in [-1, 1] \quad \text{and} \quad \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|} \in [0, 1].$$

Both lie in the stated intervals by Theorem 6.2.1, since $|\operatorname{Re} z| \le |z|$. They measure different things.

The first is the cosine of an honest real angle. Regard V as a real vector space, forgetting multiplication by non-real scalars. Then $\operatorname{Re}\langle \cdot, \cdot \rangle$ is a real inner product on it: it is symmetric because $\operatorname{Re}\overline{z} = \operatorname{Re} z$, linear in the first slot for real scalars, and $\operatorname{Re}\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$. For $V = \mathbb{C}^n$ this is the dot product of $\mathbb{R}^{2n}$, and it is the inner product that appears in the expansion $\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle$. But it does not see complex scalars: $\operatorname{Re}\langle \mathbf{x}, i\mathbf{x} \rangle = \operatorname{Re}\left(-i\|\mathbf{x}\|^2\right) = 0$, so $\mathbf{x}$ and $i\mathbf{x}$ are perpendicular in this sense, although $i\mathbf{x}$ is a multiple of $\mathbf{x}$.

The second ignores phases. By Technique 6.2.8, $|\langle \mathbf{x}, \mathbf{y} \rangle| = \max_{|\lambda|=1} \operatorname{Re}\langle \lambda\mathbf{x}, \mathbf{y} \rangle$, so it is the cosine of the smallest real angle between $\mathbf{y}$ and a unit-modulus multiple $\lambda\mathbf{x}$. It does not change when $\mathbf{x}$ and $\mathbf{y}$ are

multiplied by arbitrary nonzero scalars (by Proposition 6.1.27, $|\langle a\mathbf{x}, b\mathbf{y}\rangle| = |a||b||\langle \mathbf{x}, \mathbf{y}\rangle|$ and the norms scale the same way), so it depends only on the complex lines $\mathbb{C}\mathbf{x}$ and $\mathbb{C}\mathbf{y}$, and it is really an angle between them. By Theorem 6.2.1(b) it equals 1 exactly when these lines coincide.

Neither number is "the" angle, and we do not define one over $\mathbb{C}$. What both notions agree on is the right angle in the strong sense: $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ if and only if $\mathrm{Re}\langle \lambda\mathbf{x}, \mathbf{y}\rangle = 0$ for every λ with $|\lambda| = 1$, that is, $\mathbf{y}$ is perpendicular, in the real sense, to every unit-modulus multiple of $\mathbf{x}$. (For the "if", take $\lambda = \overline{\mu}$ as in Technique 6.2.8.) This is the notion of orthogonality used from Section 6.3 on.

6.2.4 Choosing the inner product

Every inner product gives an instance of Theorem 6.2.1, and Section 6.1.3 has many, so the theorem is a machine for producing inequalities between numbers. Using it well is a skill of its own.

Technique 6.2.21 (Apply Cauchy–Schwarz to a well-chosen inner product). To prove an inequality between sums or integrals, recognize one side as $|\langle \mathbf{x}, \mathbf{y}\rangle|$ and the other as $\|\mathbf{x}\|\|\mathbf{y}\|$ for some inner product. The work is in the choices. Pick the space and the inner product (the dot product, a weighted dot product, an integral, the Frobenius inner product), and split each term of the quantity to be bounded as a product of two factors, one belonging to $\mathbf{x}$ and one to $\mathbf{y}$. A useful move is to insert weights: write $a_k b_k$ as $\left(\sqrt{w_k}a_k\right)\left(b_k/\sqrt{w_k}\right)$ with $w_k > 0$ chosen so that the resulting norms are quantities you control.

Then check the equality case. Equality holds exactly when one of $\mathbf{x}, \mathbf{y}$ is a multiple of the other, so it tells you whether the bound is sharp and which data attain it. If the data that would attain it violate your constraints, the bound is not sharp, and a different split may do better.

Reach for it when the target has the shape $\left(\sum_k a_k b_k\right)^2 \le \left(\sum_k a_k^2\right)\left(\sum_k b_k^2\right)$ after renaming, when a linear quantity such as $\sum_k a_k$ or $\int f$ must be bounded by a quadratic one such as $\sum_k a_k^2$ or $\int f^2$, or when a quadratic quantity must be minimized subject to a linear constraint. The canonical example is Proposition 6.2.22.

Used in. Proposition 6.2.22, Proposition 6.2.24, Corollary 6.2.26, Exercise 6.2.31, Exercise 6.2.32, Exercise 6.2.36, Exercise 6.2.37, Exercise 6.2.38, Exercise 6.2.39, Exercise 6.2.46, Exercise 6.2.47, Exercise 6.4.52, Exercise 6.5.50, Problem 6.15, Problem 6.16, Exercise 8.2.43, Exercise 8.3.32, Exercise 8.3.33, Exercise 8.4.36, Exercise 6.3.53, Exercise 6.6.49.

Proposition 6.2.22 (The square of a sum). For $a_1, \dots, a_n \in \mathbb{F}$,

$$|a_1 + \cdots + a_n|^2 \le n\left(|a_1|^2 + \cdots + |a_n|^2\right),$$

with equality if and only if $a_1 = a_2 = \cdots = a_n$.

Proof. ROUTINE

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.10 (The dot product is an inner product)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. A sum $a_1 + \cdots + a_n$ is an inner product of $(a_1, \dots, a_n)$ with a fixed vector. Which one?
Hint 2. Take $\mathbf{1} = (1, \dots, 1)$. Then $\mathbf{a} \cdot \mathbf{1} = \sum_k a_k$ and $\|\mathbf{1}\|^2 = n$. For the equality case, $\mathbf{1} \neq \mathbf{0}$.
Sketch. Apply Cauchy–Schwarz to $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{1} = (1, \dots, 1)$ in $\mathbb{F}^n$: $|\mathbf{a} \cdot \mathbf{1}|^2 \le \|\mathbf{a}\|^2\|\mathbf{1}\|^2$ is the claim. Equality holds if and only if $\mathbf{a}$ is a multiple of $\mathbf{1}$, that is, all a_k are equal.

Let $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{1} = (1, \dots, 1)$ in $\mathbb{F}^n$, with the dot product. Then $\mathbf{a} \cdot \mathbf{1} = \sum_k a_k \overline{1} = \sum_k a_k$, $\|\mathbf{1}\|^2 = n$, and $\|\mathbf{a}\|^2 = \sum_k |a_k|^2$ by Proposition 6.1.10. By Theorem 6.2.1,

$$\left|\sum_k a_k\right|^2 = |\mathbf{a} \cdot \mathbf{1}|^2 \le \|\mathbf{a}\|^2\|\mathbf{1}\|^2 = n\sum_k |a_k|^2.$$

By Theorem 6.2.1(b), equality holds if and only if one of $\mathbf{a}, \mathbf{1}$ is a multiple of the other. Since $\mathbf{1} \neq \mathbf{0}$, this means $\mathbf{a} = c\mathbf{1}$ for some $c \in \mathbb{F}$ (if $\mathbf{1} = b\mathbf{a}$, then $b \neq 0$ and $\mathbf{a} = b^{-1}\mathbf{1}$). And $\mathbf{a} = c\mathbf{1}$ says that every a_k equals c. □

Corollary 6.2.23 (Arithmetic mean and quadratic mean). For $a_1, \dots, a_n \in \mathbb{R}$,

$$\frac{a_1 + \cdots + a_n}{n} \le \sqrt{\frac{a_1^2 + \cdots + a_n^2}{n}},$$

with equality if and only if $a_1 = \cdots = a_n \ge 0$.

Proof. ROUTINE

Uses. Proposition 6.2.22 (The square of a sum)
Hint 1. Bound the left side by its absolute value, then use Proposition 6.2.22.
Hint 2. For equality, both steps must be equalities: when is $r = |r|$?
Sketch. $\sum_k a_k/n \le |\sum_k a_k|/n \le \sqrt{n \sum_k a_k^2}/n$, and the right side is the quadratic mean. Equality requires $\sum_k a_k \ge 0$ and all a_k equal, hence all a_k equal to a common value $c \ge 0$.

By Proposition 6.2.22,

$$\frac{\sum_k a_k}{n} \le \frac{|\sum_k a_k|}{n} \le \frac{\sqrt{n \sum_k a_k^2}}{n} = \sqrt{\frac{\sum_k a_k^2}{n}}.$$

Equality holds if and only if both steps are equalities, that is, $\sum_k a_k \ge 0$ and (by Proposition 6.2.22) all a_k are equal to some c. Together these say that all a_k equal c with $nc \ge 0$, that is, $c \ge 0$. □

Intuition (Mean plus deviation). The proof applied Cauchy–Schwarz with $\mathbf{y} = \mathbf{1}$, and the first proof of Theorem 6.2.1 computes the gap. Here $c = (\mathbf{a} \cdot \mathbf{1})/\|\mathbf{1}\|^2 = (a_1 + \cdots + a_n)/n$ is the mean m of the a_k, the shadow $m\mathbf{1}$ is the constant vector with that mean, and $\mathbf{w} = \mathbf{a} - m\mathbf{1}$ is the vector of deviations from the mean. So (6.2.1) reads

$$\sum_{k=1}^{n} |a_k|^2 = n|m|^2 + \sum_{k=1}^{n} |a_k - m|^2.$$

Dividing by n: the mean of the squares is the square of the mean plus the variance $(1/n)\sum_k |a_k - m|^2$. So the gap in the squared inequality $|m|^2 \le (1/n)\sum_k |a_k|^2$ is exactly the variance, and it vanishes exactly for constant data.

The same choice of $\mathbf{y}$ works for integrals. The following inequality is used constantly in analysis.

Proposition 6.2.24 ($\int|f|$ versus $\int|f|^2$). Let $a < b$ and $f \in C([a,b], \mathbb{C})$. Then

$$\left(\int_a^b |f(t)|\,dt\right)^2 \le (b-a)\int_a^b |f(t)|^2\,dt,$$

with equality if and only if $|f|$ is constant. In particular, $\left(\int_0^1 |f|\right)^2 \le \int_0^1 |f|^2$.

Proof. ROUTINE

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.13 (Integral inner products)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Work in the real space $C[a,b]$ with $\langle g, h\rangle = \int_a^b gh$, applied to the real function $g = |f|$.
Hint 2. $\int_a^b |f| = \langle |f|, 1\rangle$, and $\|1\|^2 = b - a$.
Sketch. Apply Cauchy–Schwarz in $C[a,b]$ to $g = |f|$ and the constant function 1: $\langle g, 1\rangle^2 \le \|g\|^2\|1\|^2$ is the claim. Equality holds if and only if g is a multiple of 1, that is, $|f|$ is constant.

Write $f = u + iv$ with $u, v \in C[a,b]$. Then $g = |f| = \sqrt{u^2 + v^2}$ is continuous and real-valued, so $g \in C[a,b]$. Give $C[a,b]$ the integral inner product $\langle g, h\rangle = \int_a^b gh$ (Proposition 6.1.13 and the paragraph after it), and let 1 denote the constant function. Then

$$\langle g, 1\rangle = \int_a^b |f|, \quad \|g\|^2 = \int_a^b |f|^2, \quad \|1\|^2 = b - a,$$

and Theorem 6.2.1 gives $\langle g, 1\rangle^2 \le \|g\|^2\|1\|^2$, which is the inequality. By Theorem 6.2.1(b), equality holds if and only if one of $g, 1$ is a multiple of the other; since $1 \neq \mathbf{0}$, this means $g = c \cdot 1$ for some $c \in \mathbb{R}$, that is, $|f|$ is constant. □

Weights let the same idea handle unequal coefficients. The next example also shows the equality case doing real work: it finds the minimizer.

Example 6.2.25 (A constrained minimum). We find the minimum of $a^2 + 2b^2 + 3c^2$ over all $(a, b, c) \in \mathbb{R}^3$ with $a + b + c = 1$. The quantity to be minimized is $\|\mathbf{v}\|^2$ for $\mathbf{v} = (a, b, c)$ in the weighted inner product $\langle \mathbf{x}, \mathbf{y}\rangle_w = x_1y_1 + 2x_2y_2 + 3x_3y_3$ (Example 6.1.21). So we write the constraint as an inner product in the same space: $a + b + c = \langle \mathbf{v}, \mathbf{z}\rangle_w$ for $\mathbf{z} = (1, 1/2, 1/3)$; the weights 2 and 3 cancel the entries $1/2$ and $1/3$. Since

$$\|\mathbf{z}\|_w^2 = 1 + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{9} = \frac{11}{6},$$

Cauchy–Schwarz gives $1 = \langle \mathbf{v}, \mathbf{z}\rangle_w^2 \le \|\mathbf{v}\|_w^2 \cdot 11/6$, that is,

$$a^2 + 2b^2 + 3c^2 \ge \frac{6}{11}.$$

Equality requires $\mathbf{v} = t\mathbf{z}$, and the constraint then gives $t \cdot 11/6 = 1$, so $t = 6/11$ and $(a, b, c) = (6/11, 3/11, 2/11)$. This point satisfies the constraint, and indeed

$$\left(\frac{6}{11}\right)^2 + 2\left(\frac{3}{11}\right)^2 + 3\left(\frac{2}{11}\right)^2 = \frac{36 + 18 + 12}{121} = \frac{6}{11}.$$

So the minimum is $6/11$. Geometrically, we found the point of the plane $a + b + c = 1$ closest to $\mathbf{0}$ in the geometry of $\langle \cdot, \cdot \rangle_w$.

A variant of Cauchy–Schwarz with weights in denominators is a standard tool in competition problems.

Corollary 6.2.26 (Engel's form). Let $a_1, \dots, a_n \in \mathbb{R}$ and $b_1, \dots, b_n > 0$. Then

$$\frac{a_1^2}{b_1} + \cdots + \frac{a_n^2}{b_n} \ge \frac{(a_1 + \cdots + a_n)^2}{b_1 + \cdots + b_n},$$

with equality if and only if $a_1/b_1 = a_2/b_2 = \cdots = a_n/b_n$.

Proof. ROUTINE

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.10 (The dot product is an inner product)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Write $a_k = \left(a_k/\sqrt{b_k}\right) \cdot \sqrt{b_k}$.
Hint 2. Apply Cauchy–Schwarz in $\mathbb{R}^n$ to $\mathbf{x} = \left(a_k/\sqrt{b_k}\right)_k$ and $\mathbf{y} = \left(\sqrt{b_k}\right)_k$.
Sketch. With $\mathbf{x} = \left(a_k/\sqrt{b_k}\right)_k$ and $\mathbf{y} = \left(\sqrt{b_k}\right)_k$, we have $\mathbf{x} \cdot \mathbf{y} = \sum_k a_k$, $\|\mathbf{x}\|^2 = \sum_k a_k^2/b_k$, and $\|\mathbf{y}\|^2 = \sum_k b_k$. Cauchy–Schwarz gives the claim after dividing by $\sum_k b_k$. Equality holds if and only if $\mathbf{x} = c\mathbf{y}$, that is, $a_k = cb_k$ for all k.

Let $\mathbf{x} = \left(a_1/\sqrt{b_1}, \dots, a_n/\sqrt{b_n}\right)$ and $\mathbf{y} = \left(\sqrt{b_1}, \dots, \sqrt{b_n}\right)$ in $\mathbb{R}^n$. Then

$$\mathbf{x} \cdot \mathbf{y} = \sum_k a_k, \quad \|\mathbf{x}\|^2 = \sum_k \frac{a_k^2}{b_k}, \quad \|\mathbf{y}\|^2 = \sum_k b_k > 0.$$

By Theorem 6.2.1, $\left(\sum_k a_k\right)^2 \le \left(\sum_k a_k^2/b_k\right)\left(\sum_k b_k\right)$, and dividing by $\sum_k b_k$ gives the inequality. Since $\mathbf{y} \neq \mathbf{0}$, equality holds if and only if $\mathbf{x} = c\mathbf{y}$ for some $c \in \mathbb{R}$ (as in the proof of Proposition 6.2.22), that is, $a_k/\sqrt{b_k} = c\sqrt{b_k}$, or $a_k/b_k = c$, for every k. □

For instance, for positive reals a, b, c, Corollary 6.2.26 with numerators a, b, c and denominators b, c, a gives

$$\frac{a^2}{b} + \frac{b^2}{c} + \frac{c^2}{a} \geq \frac{(a+b+c)^2}{b+c+a} = a + b + c.$$

Equality requires $a/b = b/c = c/a$. The product of these three ratios is 1, so the common value is 1, and $a = b = c$.

Exercises

Exercise 6.2.27. ROUTINE Give $\mathbb{R}^3$ the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + 2x_2 y_2 + 3x_3 y_3$ (Example 6.1.21), and let $\mathbf{x} = (1, 1, 1)$ and $\mathbf{y} = (1, -1, 2)$.

(a) Verify the Cauchy–Schwarz inequality for $\mathbf{x}$ and $\mathbf{y}$.

(b) Find $c \in \mathbb{R}$ and $\mathbf{w} \in \mathbb{R}^3$ with $\mathbf{x} = c\mathbf{y} + \mathbf{w}$ and $\langle \mathbf{w}, \mathbf{y} \rangle = 0$, and check (6.2.1).

(c) Find the cosine of the angle between $\mathbf{x}$ and $\mathbf{y}$ in this inner product, and compare it with the cosine for the dot product.

Solution on page 1473

Exercise 6.2.28. ROUTINE In $\mathbb{C}^2$ with the dot product, let $\mathbf{x} = (1, i)$ and $\mathbf{y} = (z, 1)$ with $z \in \mathbb{C}$.

(a) For which z does equality hold in the Cauchy–Schwarz inequality for $\mathbf{x}, \mathbf{y}$?

(b) For which z does $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ hold?

Solution on page 1474

Exercise 6.2.29. ROUTINE In $\mathbb{C}^2$ with the dot product, let $\mathbf{x} = (1, 1)$ and $\mathbf{y} = (1, i)$.

(a) Compute $\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle / (\|\mathbf{x}\|\|\mathbf{y}\|)$ and $|\langle \mathbf{x}, \mathbf{y} \rangle| / (\|\mathbf{x}\|\|\mathbf{y}\|)$, and the corresponding angles in $[0, \pi]$.

(b) Find $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ that maximizes $\operatorname{Re}\langle \lambda\mathbf{x}, \mathbf{y} \rangle$, and the maximum value.

Solution on page 1474

Exercise 6.2.30. ROUTINE Let $z_1, \ldots, z_n, w_1, \ldots, w_n \in \mathbb{C}$. Show that

$$\left| \sum_{k=1}^{n} z_k w_k \right|^2 \leq \left(\sum_{k=1}^{n} |z_k|^2 \right) \left(\sum_{k=1}^{n} |w_k|^2 \right),$$

and that equality holds if and only if $(z_1, \ldots, z_n)$ and $(\overline{w_1}, \ldots, \overline{w_n})$ are linearly dependent. Give an example where $(z_1, \ldots, z_n)$ and $(w_1, \ldots, w_n)$ are linearly dependent but the inequality is strict.

Solution on page 1474

Exercise 6.2.31. STANDARD Find the largest value of $x + y + z$ over all $(x, y, z) \in \mathbb{R}^3$ with $x^2 + 4y^2 + 9z^2 \leq 1$, and the point where it is attained.

Solution on page 1475

Exercise 6.2.32. STANDARD

(a) Find all $(x, y, z) \in \mathbb{R}^3$ with $x + 2y + 3z = 14$ and $x^2 + y^2 + z^2 = 14$.

(b) Show that there is no $(x, y, z) \in \mathbb{R}^3$ with $x + 2y + 3z = 14$ and $x^2 + y^2 + z^2 = 13$.

Solution on page 1475

Exercise 6.2.33. STANDARD Let $\mathbf{x}, \mathbf{y}$ be vectors in an inner product space.

(a) Show that $\left|\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right| \le \|\mathbf{x}+\mathbf{y}\|\|\mathbf{x}-\mathbf{y}\|$.

(b) Show that equality holds if and only if one of $\mathbf{x}, \mathbf{y}$ is a *real* multiple of the other. (Over $\mathbb{R}$ this just says that $\mathbf{x}, \mathbf{y}$ are linearly dependent.)

Solution on page 1475

Exercise 6.2.34. STANDARD Let $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be vectors in an inner product space such that equality holds in the Cauchy–Schwarz inequality for the pair $\mathbf{x}, \mathbf{y}$ and for the pair $\mathbf{y}, \mathbf{z}$. Prove or disprove: equality holds for the pair $\mathbf{x}, \mathbf{z}$. What changes if $\mathbf{y} \neq \mathbf{0}$?

Solution on page 1476

Exercise 6.2.35. STANDARD Find the errors in the following argument, and give a numerical example in $\mathbb{C}$ (with the inner product $\langle z, w\rangle = z\overline{w}$) showing that its first displayed formula is false.

Claim. For $\mathbf{x}, \mathbf{y}$ in a complex inner product space, $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\|\mathbf{y}\|$.

Proof. We may assume $\mathbf{y} \neq \mathbf{0}$. For every $t \in \mathbb{C}$,

$$0 \le \|\mathbf{x} - t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2t\langle \mathbf{x}, \mathbf{y}\rangle + t^2\|\mathbf{y}\|^2.$$

A quadratic in t that is never negative has nonpositive discriminant, so $4\langle \mathbf{x}, \mathbf{y}\rangle^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \le 0$. Taking square roots, $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\|\mathbf{y}\|$.

Solution on page 1476

Exercise 6.2.36. STANDARD Show that for every $f \in C[0,1]$,

$$\left(\int_0^1 f(t)\,dt\right)^2 \le \left(1 - e^{-1}\right)\int_0^1 e^t f(t)^2\,dt,$$

and find all f for which equality holds.

Solution on page 1477

Exercise 6.2.37. STANDARD Let $p = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$ and $t \in \mathbb{R}$.

(a) Show that

$$|p(t)| \le \left(a_0^2 + \cdots + a_n^2\right)^{1/2}\left(1 + t^2 + \cdots + t^{2n}\right)^{1/2},$$

and that for each t and n some nonzero p of degree at most n attains equality.

(b) Deduce that if $|t| < 1$, then $|p(t)| \le \left(a_0^2 + \cdots + a_n^2\right)^{1/2}/\sqrt{1-t^2}$ for every polynomial p, whatever its degree.

Solution on page 1477

Exercise 6.2.38. STANDARD

(a) Show that $\sqrt{1}+\sqrt{2}+\cdots+\sqrt{n} \le n\sqrt{(n+1)/2}$ for every $n \ge 1$.

(b) Show that for all real $a_1, \ldots, a_n$,

$$\frac{a_1}{1}+\frac{a_2}{2}+\cdots+\frac{a_n}{n} \le \sqrt{2}\left(a_1^2+\cdots+a_n^2\right)^{1/2}.$$

(You may use $1/k^2 < 1/(k-1) - 1/k$ for $k \ge 2$.)

Solution on page 1477

Exercise 6.2.39. STANDARD

(a) Show that for every $A \in \mathbb{F}^{n\times n}$, $|\operatorname{tr} A|^2 \le n \operatorname{tr}(A^*A)$, with equality if and only if A is a scalar multiple of I_n.

(b) Find all $A \in \mathbb{R}^{2\times 2}$ with $\operatorname{tr} A = 2$ and $\sum_{i,j} a_{ij}^2 = 2$.

Solution on page 1478

Exercise 6.2.40. STANDARD Write $\|A\|_F = \operatorname{tr}(A^*A)^{1/2}$ for the Frobenius norm (Proposition 6.1.16).

(a) Show that $\|AB\|_F \le \|A\|_F \|B\|_F$ for all $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times p}$.

(b) Deduce that every entry of A^k has modulus at most $\|A\|_F^k$, for $A \in \mathbb{F}^{n\times n}$ and $k \ge 1$.

(c) Give $A, B \in \mathbb{R}^{2\times 2}$, both nonzero, with $\|AB\|_F < \|A\|_F \|B\|_F$.

Solution on page 1478

Exercise 6.2.41. STANDARD

(a) Let $\mathbf{x}, \mathbf{y}$ be unit vectors in an inner product space with $\mathbf{x} \ne \mathbf{y}$, and let $0 < s < 1$. Show that $\|(1-s)\mathbf{x} + s\mathbf{y}\| < 1$. (So the segment between two points of the unit sphere lies strictly inside it, except at its endpoints.)

(b) Show that (a) fails for the taxicab norm $\|\mathbf{x}\|_1 = |x_1| + |x_2|$ on $\mathbb{R}^2$ (Aside (Norms that do not come from inner products)). Conclude again that the taxicab norm does not come from an inner product.

Solution on page 1479

Exercise 6.2.42. STANDARD Let $\mathbf{x}, \mathbf{y}$ be vectors in an inner product space V. Show that $\mathbf{z} \in V$ satisfies

$$\|\mathbf{x}-\mathbf{z}\| + \|\mathbf{z}-\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}\|$$

if and only if $\mathbf{z} = (1-s)\mathbf{x} + s\mathbf{y}$ for some $s \in [0,1]$. That is, the points for which the detour through $\mathbf{z}$ costs nothing are exactly the points of the segment from $\mathbf{x}$ to $\mathbf{y}$.

Solution on page 1479

Exercise 6.2.43. STANDARD

(a) Prove *Lagrange's identity*: for real numbers $a_1, \dots, a_n, b_1, \dots, b_n$,

$$\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right) - \left(\sum_{k=1}^{n} a_k b_k\right)^2 = \sum_{1 \le j < k \le n} (a_j b_k - a_k b_j)^2.$$

(b) Deduce the Cauchy–Schwarz inequality for the dot product on $\mathbb{R}^n$, including its equality case, without using Theorem 6.2.1.

(c) For $n = 2$, compare (a) with (6.2.2), and conclude that the parallelogram spanned by $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ has area $|a_1 b_2 - a_2 b_1|$.

Solution on page 1479

Exercise 6.2.44. HARD Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be vectors in an inner product space, not all $\mathbf{0}$. Show that

$$\|\mathbf{x}_1 + \cdots + \mathbf{x}_n\| = \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_n\|$$

if and only if there is a unit vector $\mathbf{u}$ and real numbers $r_1, \dots, r_n \ge 0$ with $\mathbf{x}_k = r_k \mathbf{u}$ for every k.

Solution on page 1480

Exercise 6.2.45. HARD Let V be a real inner product space, and write $\theta(\mathbf{x}, \mathbf{y})$ for the angle between nonzero vectors (Definition 6.2.16). Show that for all nonzero $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$,

$$\theta(\mathbf{x}, \mathbf{z}) \le \theta(\mathbf{x}, \mathbf{y}) + \theta(\mathbf{y}, \mathbf{z}).$$

(Angles measure distances on the unit sphere, and this is their triangle inequality.)

Solution on page 1480

Exercise 6.2.46. HARD Find the minimum of $\int_0^1 f(t)^2 \, dt$ over all $f \in C[0,1]$ with

$$\int_0^1 f(t)\, dt = 1 \quad \text{and} \quad \int_0^1 t f(t)\, dt = 1,$$

and all functions attaining it.

Solution on page 1481

Exercise 6.2.47. HARD (*Nesbitt's inequality.*) Show that for all positive real numbers a, b, c,

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \ge \frac{3}{2},$$

with equality if and only if $a = b = c$.

Solution on page 1482

Exercise 6.2.48. CHALLENGE (*Ptolemy's inequality.*)

(a) Show that for nonzero vectors $\mathbf{x}, \mathbf{y}$ in an inner product space,

$$\left\| \frac{\mathbf{x}}{\|\mathbf{x}\|^2} - \frac{\mathbf{y}}{\|\mathbf{y}\|^2} \right\| = \frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

(b) Deduce that for all $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}$ in an inner product space,

$$\|\mathbf{x}-\mathbf{y}\|\|\mathbf{z}-\mathbf{w}\| \le \|\mathbf{x}-\mathbf{z}\|\|\mathbf{y}-\mathbf{w}\| + \|\mathbf{x}-\mathbf{w}\|\|\mathbf{y}-\mathbf{z}\|.$$

(For four points in the plane: the product of the diagonals of a quadrilateral is at most the sum of the products of opposite sides.)

Solution on page 1482

6.3 Orthogonality and orthonormal bases

For nonzero $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$, $\mathbf{v} \cdot \mathbf{w} = 0$ exactly when $\mathbf{v}$ and $\mathbf{w}$ are perpendicular ((6.1.1)), and in every real inner product space the angle between nonzero vectors (Definition 6.2.16) is a right angle exactly when $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ (Proposition 6.2.17). This section takes the condition $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ as the definition of perpendicularity, now over $\mathbb{R}$ and $\mathbb{C}$ alike, and studies bases made of perpendicular unit vectors. For such a basis, finding the coordinates of a vector requires no linear system: each coordinate is a single inner product, computed without reference to the others. Once the coordinates are known, lengths and inner products are computed from them exactly as in $\mathbb{F}^n$ with the dot product. In $\mathbb{F}^n$ these bases are the columns of the matrices U with $U^*U = I$, and we end with the most important example after the standard basis, the Fourier basis of $\mathbb{C}^n$.

Throughout, V is an inner product space over $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ (Notation 6.1.1).

6.3.1 Orthogonality

Definition 6.3.1 (Orthogonality). Let V be an inner product space. Vectors $\mathbf{u}, \mathbf{v} \in V$ are **orthogonal**, written $\mathbf{u} \perp \mathbf{v}$, if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. A list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V is **orthogonal** if $\mathbf{v}_i \perp \mathbf{v}_j$ whenever $i \neq j$.

Orthogonality is symmetric: $\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}$, and a complex number vanishes exactly when its conjugate does. By Proposition 6.1.4(b), $\mathbf{0}$ is orthogonal to every vector, itself included, and by definiteness it is the only vector orthogonal to itself. It is also the only vector orthogonal to every vector of V: if $\mathbf{v} \perp \mathbf{x}$ for all $\mathbf{x}$, take $\mathbf{x} = \mathbf{v}$. (This is Proposition 6.1.37 with $\mathbf{w} = \mathbf{0}$.)

In the definition of an orthogonal list only entries in different positions are compared. So a list of length 0 or 1 is orthogonal, an orthogonal list may contain $\mathbf{0}$ any number of times, and a nonzero vector cannot occur twice in an orthogonal list, since $\langle \mathbf{v}, \mathbf{v} \rangle \neq 0$.

Example 6.3.2 (Orthogonal vectors).

(a) In $\mathbb{R}^2$, $(a, b) \perp (-b, a)$ for all a, b, since $a(-b) + ba = 0$. The second vector is the first rotated by $\pi/2$. For instance $(1, 2) \perp (-2, 1)$.

(b) In $\mathbb{C}^2$, $(1, i) \cdot (1, -i) = 1 \cdot 1 + i \cdot \overline{-i} = 1 + i \cdot i = 0$, so $(1, i) \perp (1, -i)$. Forgetting the conjugate gives $1 + i(-i) = 2$ instead.

(c) Give $C[-1, 1]$ the inner product $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$ (Proposition 6.1.13, on $[-1, 1]$). If f is even and g is odd, then fg is odd, so $\langle f, g \rangle = 0$. For instance $1 \perp t$ and $t \perp t^2$. On the other hand $\langle 1, t^2 \rangle = 2/3 \neq 0$.

Orthogonality depends on the inner product as well as on the vectors. For the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^{\mathsf{T}} M \mathbf{x}$ on $\mathbb{R}^2$ with $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (Example 6.1.23), $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = m_{21} = 1$, so $\mathbf{e}_1$ and $\mathbf{e}_2$ are not orthogonal. The vectors $(1, 1)$ and $(1, -1)$ are orthogonal for both inner products: $M(1, 1) = (3, 3)$, and $(1, -1) \cdot (3, 3) = 0$.

Over $\mathbb{C}$, orthogonality is more demanding than perpendicularity of arrows. In $\mathbb{C} = \mathbb{C}^1$ with $\langle z, w\rangle = z\overline{w}$, the numbers 1 and i are perpendicular as arrows in the plane, but $\langle 1, i\rangle = \bar{i} = -i \neq 0$. They are not orthogonal, and they should not be, since $i = i \cdot 1$ is a scalar multiple of 1. As we saw at the end of Section 6.2.3, $\langle \mathbf{u}, \mathbf{v}\rangle = 0$ says that $\mathbf{v}$ is perpendicular, in the real sense, to $\lambda\mathbf{u}$ for *every* λ with $|\lambda| = 1$.

The Pythagorean theorem

Theorem 6.3.3 (Pythagorean theorem). Let $\mathbf{u}, \mathbf{v} \in V$.

(a) If $\mathbf{u} \perp \mathbf{v}$, then $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

(b) If $\mathbb{F} = \mathbb{R}$, the converse holds: if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u} \perp \mathbf{v}$.

Proof. ROUTINE

Uses. Proposition 6.1.29 (Expanding a norm), Definition 6.3.1 (Orthogonality)
Hint 1. Expand $\|\mathbf{u} + \mathbf{v}\|^2$ with Proposition 6.1.29. Which term has to vanish?
Hint 2. The identity holds if and only if $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Over $\mathbb{R}$, what is $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle$?
Sketch. By Proposition 6.1.29, $\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 = 2\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle$. So the identity holds exactly when $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$. This follows from $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, and over $\mathbb{R}$, where $\langle \mathbf{u}, \mathbf{v}\rangle$ is real, it is equivalent to it.

By Proposition 6.1.29,

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2, \tag{6.3.1}$$

so the identity $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ holds if and only if $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

(a) If $\mathbf{u} \perp \mathbf{v}$, then $\langle \mathbf{u}, \mathbf{v}\rangle = 0$, so its real part is 0.

(b) If $\mathbb{F} = \mathbb{R}$, then $\langle \mathbf{u}, \mathbf{v}\rangle$ is real and equals its real part. So $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$ says that $\mathbf{u} \perp \mathbf{v}$. □

In $\mathbb{R}^2$ with the dot product, (a) is the theorem about the right triangle with legs $\mathbf{u}$, $\mathbf{v}$ and hypotenuse $\mathbf{u} + \mathbf{v}$, and (b) is its converse. In other spaces it says new things. In $C[-1, 1]$, $1 \perp t$ by Example 6.3.2(c), so

$$\int_{-1}^{1} (1 + t)^2 \, dt = \|1\|^2 + \|t\|^2 = 2 + \frac{2}{3} = \frac{8}{3},$$

which agrees with the direct computation $\int_{-1}^{1}\left(1 + 2t + t^2\right) dt = 2 + 0 + 2/3$.

Non-example 6.3.4 (The converse fails over $\mathbb{C}$). In $\mathbb{C}$ with $\langle z, w\rangle = z\overline{w}$, take $\mathbf{u} = 1$ and $\mathbf{v} = i$. Then $|1 + i|^2 = 2 = |1|^2 + |i|^2$, but $\langle 1, i\rangle = -i \neq 0$. More generally, let $\mathbf{x} \neq \mathbf{0}$ be a vector in any complex inner product space. By Proposition 6.1.27,

$$\|\mathbf{x} + i\mathbf{x}\|^2 = |1 + i|^2\|\mathbf{x}\|^2 = 2\|\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|i\mathbf{x}\|^2,$$

while $\langle \mathbf{x}, i\mathbf{x}\rangle = \bar{i}\|\mathbf{x}\|^2 = -i\|\mathbf{x}\|^2 \neq 0$. The proof of Theorem 6.3.3 shows what goes wrong: the Pythagorean identity only sees $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle$, and here $\langle \mathbf{u}, \mathbf{v}\rangle$ is purely imaginary.

The imaginary part becomes visible once $\mathbf{v}$ is multiplied by i. With $a = i$, Proposition 6.1.29 gives $\|\mathbf{u} + i\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\operatorname{Im}\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$, as computed after that proposition. So over $\mathbb{C}$ two Pythagorean identities together detect orthogonality.

Proposition 6.3.5 (Orthogonality over $\mathbb{C}$ through Pythagoras). Let V be a complex inner product space and $\mathbf{u}, \mathbf{v} \in V$. Then $\mathbf{u} \perp \mathbf{v}$ if and only if both

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \quad \text{and}$$
$$\|\mathbf{u} + i\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|i\mathbf{v}\|^2.$$

In that case $\|\mathbf{u} + a\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|a\mathbf{v}\|^2$ for every $a \in \mathbb{C}$.

Proof. ROUTINE

> *Uses.* Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm), Notation 6.1.1 (Real and complex scalars)
> *Techniques.* Plug in clever test vectors (Technique 6.1.40)
> *Hint 1.* Expand $\|\mathbf{u}+a\mathbf{v}\|^2$ with Proposition 6.1.29 for a general scalar a. The Pythagorean identity for $\mathbf{u}$ and $a\mathbf{v}$ says that one term vanishes.
> *Hint 2.* For $a=1$ that term is $2\,\mathrm{Re}\langle\mathbf{u},\mathbf{v}\rangle$. For $a=i$ it is $2\,\mathrm{Re}(-i\langle\mathbf{u},\mathbf{v}\rangle)$; which part of $\langle\mathbf{u},\mathbf{v}\rangle$ is that?
> *Sketch.* By Proposition 6.1.29 and $\|a\mathbf{v}\|^2=|a|^2\|\mathbf{v}\|^2$, the identity for $\mathbf{u}$ and $a\mathbf{v}$ holds exactly when $\mathrm{Re}(\overline{a}\langle\mathbf{u},\mathbf{v}\rangle)=0$. For $a=1$ and $a=i$ this says $\mathrm{Re}\langle\mathbf{u},\mathbf{v}\rangle=0$ and $\mathrm{Im}\langle\mathbf{u},\mathbf{v}\rangle=0$, which together say $\langle\mathbf{u},\mathbf{v}\rangle=0$. If $\langle\mathbf{u},\mathbf{v}\rangle=0$, then $\mathrm{Re}(\overline{a}\cdot 0)=0$ for every a.

Let $a\in\mathbb{C}$. By Proposition 6.1.29 and Proposition 6.1.27,

$$\|\mathbf{u}+a\mathbf{v}\|^2-\|\mathbf{u}\|^2-\|a\mathbf{v}\|^2=2\,\mathrm{Re}(\overline{a}\langle\mathbf{u},\mathbf{v}\rangle).$$

For $a=1$ the right side is $2\,\mathrm{Re}\langle\mathbf{u},\mathbf{v}\rangle$, and for $a=i$ it is $2\,\mathrm{Re}(-i\langle\mathbf{u},\mathbf{v}\rangle)=2\,\mathrm{Im}\langle\mathbf{u},\mathbf{v}\rangle$ (Notation 6.1.1). So the two displayed identities hold together if and only if the real and imaginary parts of $\langle\mathbf{u},\mathbf{v}\rangle$ both vanish, that is, if and only if $\mathbf{u}\perp\mathbf{v}$. Finally, if $\langle\mathbf{u},\mathbf{v}\rangle=0$, the right side is 0 for every a. □

The last statement also follows from Theorem 6.3.3(a): if $\mathbf{u}\perp\mathbf{v}$, then $\langle\mathbf{u},a\mathbf{v}\rangle=\overline{a}\langle\mathbf{u},\mathbf{v}\rangle=0$, so $\mathbf{u}\perp a\mathbf{v}$ for every scalar a. Over $\mathbb{C}$, then, orthogonality is "Pythagoras for $\mathbf{u}$ and every multiple of $\mathbf{v}$", which matches the description of orthogonality through unit-modulus multiples above.

The Pythagorean theorem extends to any number of vectors, and pairwise orthogonality is all it needs.

Corollary 6.3.6 (Pythagoras for orthogonal lists). Let $\mathbf{v}_1,\dots,\mathbf{v}_k$ be an orthogonal list in V, and let $a_1,\dots,a_k\in\mathbb{F}$. Then

$$\|a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k\|^2=|a_1|^2\|\mathbf{v}_1\|^2+\cdots+|a_k|^2\|\mathbf{v}_k\|^2.$$

Proof (by expanding). ROUTINE

> *Uses.* Proposition 6.1.4 (Sesquilinearity), Definition 6.1.26 (Norm), Definition 6.3.1 (Orthogonality)
> *Techniques.* Expand by multilinearity (Technique 5.1.7)
> *Hint 1.* Write the left side as an inner product of a sum with itself and expand both slots.
> *Hint 2.* You get $\sum_j\sum_l a_j\overline{a_l}\langle\mathbf{v}_j,\mathbf{v}_l\rangle$. Which terms survive?
> *Sketch.* Expanding both slots gives k^2 terms $a_j\overline{a_l}\langle\mathbf{v}_j,\mathbf{v}_l\rangle$. Those with $j\neq l$ vanish by orthogonality, and those with $j=l$ are $|a_j|^2\|\mathbf{v}_j\|^2$.

By linearity in the first slot and Proposition 6.1.4 in the second,

$$\left\|\sum_{j=1}^{k}a_j\mathbf{v}_j\right\|^2=\left\langle\sum_{j=1}^{k}a_j\mathbf{v}_j,\sum_{l=1}^{k}a_l\mathbf{v}_l\right\rangle=\sum_{j=1}^{k}\sum_{l=1}^{k}a_j\overline{a_l}\langle\mathbf{v}_j,\mathbf{v}_l\rangle.$$

For $j\neq l$ the term is 0 because $\mathbf{v}_j\perp\mathbf{v}_l$. For $j=l$ it is $a_j\overline{a_j}\langle\mathbf{v}_j,\mathbf{v}_j\rangle=|a_j|^2\|\mathbf{v}_j\|^2$. Adding these gives the claim. □

Proof (by induction on k). ROUTINE

> *Uses.* Theorem 6.3.3 (Pythagorean theorem), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.1.4 (Sesquilinearity)
> *Techniques.* Induction on the length of a list (Technique 1.6.9)
> *Hint 1.* Split off the last term: $\mathbf{w}=a_1\mathbf{v}_1+\cdots+a_{k-1}\mathbf{v}_{k-1}$ and $a_k\mathbf{v}_k$.
> *Hint 2.* Check that $\mathbf{w}\perp a_k\mathbf{v}_k$ by expanding $\langle\mathbf{w},a_k\mathbf{v}_k\rangle$, then apply Theorem 6.3.3.
> *Sketch.* The case $k=1$ is Proposition 6.1.27. For $k\geq 2$, the vector $\mathbf{w}=\sum_{j<k}a_j\mathbf{v}_j$ satisfies $\langle\mathbf{w},a_k\mathbf{v}_k\rangle=\sum_{j<k}a_j\overline{a_k}\langle\mathbf{v}_j,\mathbf{v}_k\rangle=0$, so Theorem 6.3.3 splits off $|a_k|^2\|\mathbf{v}_k\|^2$, and the induction hypothesis handles $\|\mathbf{w}\|^2$.

We induct on k. For $k=1$ the claim is $\|a_1\mathbf{v}_1\|^2=|a_1|^2\|\mathbf{v}_1\|^2$, which is Proposition 6.1.27. Let $k\geq 2$ and assume the claim for orthogonal lists of length $k-1$. Let $\mathbf{w}=a_1\mathbf{v}_1+\cdots+a_{k-1}\mathbf{v}_{k-1}$. Then

$$\langle \mathbf{w}, a_k \mathbf{v}_k \rangle = \sum_{j=1}^{k-1} a_j \overline{a_k} \langle \mathbf{v}_j, \mathbf{v}_k \rangle = 0,$$

so $\mathbf{w} \perp a_k \mathbf{v}_k$, and by Theorem 6.3.3(a), Proposition 6.1.27, and the induction hypothesis applied to the orthogonal list $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$,

$$\|\mathbf{w} + a_k \mathbf{v}_k\|^2 = \|\mathbf{w}\|^2 + |a_k|^2 \|\mathbf{v}_k\|^2 = \sum_{j=1}^{k} |a_j|^2 \|\mathbf{v}_j\|^2,$$

as desired. □

The second proof explains why pairwise orthogonality is enough: a vector orthogonal to each of $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ is orthogonal to every linear combination of them. We record this fact in Proposition 6.3.9.

Exercise 6.3.7. ROUTINE In $\mathbb{C}^3$ with the dot product, find all $a \in \mathbb{C}$ with $(1, i, a) \perp (2, 1+i, i)$. For that a, verify Theorem 6.3.3(a) for the two vectors by computing all three norms.

Solution on page 1483

Orthogonality to a set

We often need all vectors orthogonal to a given collection of vectors.

Definition 6.3.8 (Vectors orthogonal to a set). Let $A \subseteq V$ be any subset. A vector $\mathbf{v} \in V$ is **orthogonal to** A, written $\mathbf{v} \perp A$, if $\mathbf{v} \perp \mathbf{a}$ for every $\mathbf{a} \in A$. The set of all such vectors is

$$A^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{a} \rangle = 0 \text{ for all } \mathbf{a} \in A\},$$

read "A perp". For a list $\mathcal{L}$, $\mathcal{L}^\perp$ is the perp of the set of entries of $\mathcal{L}$.

When A is a subspace, $A^\perp$ is called the *orthogonal complement* of A, and Section 6.4 shows that in finite dimensions it is a complement of A in the sense of Section 4.4. For now A is arbitrary.

Proposition 6.3.9 (Properties of $A^\perp$). Let $A, B \subseteq V$.

(a) $A^\perp$ is a subspace of V.

(b) If $A \subseteq B$, then $B^\perp \subseteq A^\perp$.

(c) For a list $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ in V, $\operatorname{span}(\mathcal{L})^\perp = \mathcal{L}^\perp$. That is, a vector is orthogonal to every linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ if and only if it is orthogonal to each $\mathbf{v}_j$.

(d) $\{\mathbf{0}\}^\perp = V$, $V^\perp = \{\mathbf{0}\}$, and $A \cap A^\perp \subseteq \{\mathbf{0}\}$.

Proof. ROUTINE

Uses. Theorem 1.4.3 (Subspace criterion), Proposition 6.1.4 (Sesquilinearity), Proposition 1.3.17 (A list lies in its span), Definition 1.3.10 (Span), Definition 6.1.2 (Inner product)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For (a), check the conditions of Theorem 1.4.3 using linearity in the first slot.
Hint 2. For (c), one containment is (b). For the other, take $\mathbf{v} \perp \mathbf{v}_j$ for all j and expand $\left\langle \mathbf{v}, \sum_j c_j \mathbf{v}_j \right\rangle$ in the second slot.
Hint 3. For (d), a vector in $V^\perp$ or in $A \cap A^\perp$ is orthogonal to itself.
Sketch. (a) $\mathbf{0} \in A^\perp$, and $\langle c\mathbf{v} + \mathbf{w}, \mathbf{a} \rangle = c\langle \mathbf{v}, \mathbf{a} \rangle + \langle \mathbf{w}, \mathbf{a} \rangle$ shows closure. (b) Orthogonality to all of B includes orthogonality to all of A. (c) $\mathcal{L} \subseteq \operatorname{span}(\mathcal{L})$ gives one containment by (b); conversely $\left\langle \mathbf{v}, \sum_j c_j \mathbf{v}_j \right\rangle = \sum_j \overline{c_j} \langle \mathbf{v}, \mathbf{v}_j \rangle = 0$. (d) Everything is orthogonal to $\mathbf{0}$, and a vector orthogonal to itself is $\mathbf{0}$.

(a) By Proposition 6.1.4(b), $\langle \mathbf{0}, \mathbf{a} \rangle = 0$ for all $\mathbf{a} \in A$, so $\mathbf{0} \in A^\perp$. Let $\mathbf{v}, \mathbf{w} \in A^\perp$ and $c \in \mathbb{F}$. For every $\mathbf{a} \in A$, linearity in the first slot gives

$$\langle c\mathbf{v} + \mathbf{w}, \mathbf{a} \rangle = c\langle \mathbf{v}, \mathbf{a} \rangle + \langle \mathbf{w}, \mathbf{a} \rangle = c \cdot 0 + 0 = 0,$$

so $c\mathbf{v}+\mathbf{w}\in A^\perp$. By Theorem 1.4.3, $A^\perp$ is a subspace.

(b) If $\mathbf{v}\in B^\perp$, then $\mathbf{v}\perp\mathbf{b}$ for every $\mathbf{b}\in B$, in particular for every element of $A\subseteq B$. So $\mathbf{v}\in A^\perp$.

(c) Each $\mathbf{v}_j$ lies in $\operatorname{span}(\mathcal{L})$ (Proposition 1.3.17), so (b) gives $\operatorname{span}(\mathcal{L})^\perp\subseteq\mathcal{L}^\perp$. Conversely, let $\mathbf{v}\in\mathcal{L}^\perp$, and let $\mathbf{w}=c_1\mathbf{v}_1+\cdots+c_k\mathbf{v}_k$ be any element of $\operatorname{span}(\mathcal{L})$. By Proposition 6.1.4,

$$\langle\mathbf{v},\mathbf{w}\rangle=\overline{c_1}\langle\mathbf{v},\mathbf{v}_1\rangle+\cdots+\overline{c_k}\langle\mathbf{v},\mathbf{v}_k\rangle=0,$$

so $\langle\mathbf{w},\mathbf{v}\rangle=\overline{\langle\mathbf{v},\mathbf{w}\rangle}=0$ as well, and $\mathbf{v}\in\operatorname{span}(\mathcal{L})^\perp$.

(d) Every $\mathbf{v}\in V$ satisfies $\langle\mathbf{v},\mathbf{0}\rangle=0$ by Proposition 6.1.4(b), so $\{\mathbf{0}\}^\perp=V$. If $\mathbf{v}\in V^\perp$, then in particular $\langle\mathbf{v},\mathbf{v}\rangle=0$, so $\mathbf{v}=\mathbf{0}$ by definiteness; and $\mathbf{0}\in V^\perp$ by (a). Finally, if $\mathbf{v}\in A\cap A^\perp$, then $\mathbf{v}\perp\mathbf{v}$, so $\mathbf{v}=\mathbf{0}$. □

Part (c) is how $A^\perp$ is computed in practice: to test orthogonality to a subspace, test it on a spanning list. The same argument shows $A^\perp=\operatorname{span}(A)^\perp$ for an arbitrary subset A, with $\operatorname{span}(A)$ as in Aside (Spans of infinite sets). Part (a) holds even when A itself is not a subspace.

Example 6.3.10 (Perps are solution sets).

(a) In $\mathbb{R}^3$, $\{(1,1,1)\}^\perp=\{\mathbf{x}\in\mathbb{R}^3\mid x_1+x_2+x_3=0\}$, the plane through $\mathbf{0}$ with normal vector $(1,1,1)$.

(b) Let $A\in\mathbb{F}^{m\times n}$. The ith entry of $A\mathbf{x}$ is $\sum_j a_{ij}x_j=\mathbf{x}\cdot\mathbf{r}_i$, where $\mathbf{r}_i=(\overline{a_{i1}},\ldots,\overline{a_{in}})\in\mathbb{F}^n$ is the ith row of A with its entries conjugated. So

$$\ker A=\{\mathbf{r}_1,\ldots,\mathbf{r}_m\}^\perp.$$

Over $\mathbb{R}$, $\mathbf{r}_i$ is just the ith row of A, and solving a homogeneous system means finding all vectors orthogonal to the rows of its coefficient matrix. For instance, for $A=\begin{pmatrix}1&2&0\\0&1&-1\end{pmatrix}$, solving $A\mathbf{x}=\mathbf{0}$ gives $\ker A=\operatorname{span}((-2,1,1))$, and indeed $(-2,1,1)$ is orthogonal to $(1,2,0)$ and to $(0,1,-1)$. The vectors $\mathbf{r}_i$ are the columns of the matrix A^* of Definition 6.3.26 below, so by Proposition 6.3.9(c) this says $\ker A=(\operatorname{im}A^*)^\perp$; Section 6.6 proves the analogous statement for linear maps.

Exercise 6.3.11. ROUTINE Let $A,B\subseteq V$. Show that $(A\cup B)^\perp=A^\perp\cap B^\perp$ and that $A\subseteq\left(A^\perp\right)^\perp$. Give an example with $A\neq\left(A^\perp\right)^\perp$.

Solution on page 1483

6.3.2 Orthonormal lists and bases

Orthogonal lists are independent

Three mutually perpendicular nonzero arrows in $\mathbb{R}^3$ cannot lie in a common plane, so they are linearly independent. To prove the general version, start from a linear relation $a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k=\mathbf{0}$ among orthogonal vectors and try to show that one coefficient, say a_j, is 0. The inner product with $\mathbf{v}_j$ is the natural thing to apply: it is linear, so it can be applied term by term, and it annihilates every $\mathbf{v}_l$ with $l\neq j$. This move is used throughout the rest of the book.

Technique 6.3.12 (Take an inner product to isolate a coefficient). If a vector is written as a combination $\mathbf{w}=a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k$ of an orthogonal list, take the inner product of both sides with $\mathbf{v}_j$. Linearity in the first slot distributes it over the sum, and orthogonality removes every term except the jth:

$$\langle\mathbf{w},\mathbf{v}_j\rangle=\sum_{l=1}^{k}a_l\langle\mathbf{v}_l,\mathbf{v}_j\rangle=a_j\|\mathbf{v}_j\|^2.$$

So each coefficient is determined by one inner product, with no reference to the other coefficients. Put $\mathbf{v}_j$ in the *second* slot; in the first slot it produces $\langle\mathbf{v}_j,\mathbf{w}\rangle=\overline{a_j}\|\mathbf{v}_j\|^2$, the conjugate.

Reach for it when a vector or a linear relation is expressed through an orthogonal list and you need the coefficients, or need to show that they vanish. It is the inner-product form of Technique 4.1.25: the linear map $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{v}_j \rangle$ from V to $\mathbb{F}$ kills every $\mathbf{v}_l$ with $l \neq j$. The canonical example is Proposition 6.3.13. The coordinate formula Proposition 6.3.18, the Fourier inversion formula (Proposition 6.3.41), and Gram–Schmidt (Algorithm 6.4.14) use it again.

Used in. Proposition 6.3.13, Proposition 6.3.18, Corollary 6.3.21, Theorem 6.3.24, Proposition 6.4.4, Theorem 6.4.6, Corollary 6.4.25, Theorem 6.6.3, Theorem 6.6.15, Exercise 6.7.7, Problem 6.5, Lemma 8.4.23, Exercise 6.4.55.

Proposition 6.3.13 (Orthogonal lists are independent). An orthogonal list of nonzero vectors is linearly independent.

Proof (by isolating coefficients). ROUTINE

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 6.3.1 (Orthogonality), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Start from a linear relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. You must show that every a_j is 0.
Hint 2. Fix j and take the inner product of both sides with $\mathbf{v}_j$.
Hint 3. You get $a_j\|\mathbf{v}_j\|^2 = 0$. Why is $\|\mathbf{v}_j\| \neq 0$?
Sketch. If $\sum_l a_l\mathbf{v}_l = \mathbf{0}$, then for each j, $0 = \langle \sum_l a_l\mathbf{v}_l, \mathbf{v}_j \rangle = a_j\|\mathbf{v}_j\|^2$ by orthogonality. Since $\mathbf{v}_j \neq \mathbf{0}$, $\|\mathbf{v}_j\|^2 > 0$, so $a_j = 0$. Only the trivial relation exists.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be an orthogonal list of nonzero vectors, and suppose $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. Fix $j \in \{1, \ldots, k\}$. By Proposition 6.1.4(b) and linearity in the first slot,

$$0 = \langle \mathbf{0}, \mathbf{v}_j \rangle = \left\langle \sum_{l=1}^{k} a_l\mathbf{v}_l, \mathbf{v}_j \right\rangle = \sum_{l=1}^{k} a_l \langle \mathbf{v}_l, \mathbf{v}_j \rangle$$
$$= a_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = a_j\|\mathbf{v}_j\|^2,$$

where the fourth step uses $\langle \mathbf{v}_l, \mathbf{v}_j \rangle = 0$ for $l \neq j$. Since $\mathbf{v}_j \neq \mathbf{0}$, $\|\mathbf{v}_j\|^2 > 0$ by Proposition 6.1.27, so $a_j = 0$. As j was arbitrary, the only linear relation among $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is the trivial one, and the list is independent by Theorem 1.5.5. □

Proof (by Pythagoras). ROUTINE

Uses. Corollary 6.3.6 (Pythagoras for orthogonal lists), Theorem 1.5.5 (Independence is decided at zero), Proposition 6.1.27 (Basic properties of the norm)
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Take the squared norm of both sides of a relation $\sum_l a_l\mathbf{v}_l = \mathbf{0}$.
Hint 2. Corollary 6.3.6 turns the left side into a sum of nonnegative terms.
Sketch. If $\sum_l a_l\mathbf{v}_l = \mathbf{0}$, then $0 = \|\sum_l a_l\mathbf{v}_l\|^2 = \sum_l |a_l|^2\|\mathbf{v}_l\|^2$. A sum of nonnegative terms vanishes only if each does, and $\|\mathbf{v}_l\| > 0$, so every $a_l = 0$.

Suppose $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$. By Corollary 6.3.6,

$$0 = \|\mathbf{0}\|^2 = |a_1|^2\|\mathbf{v}_1\|^2 + \cdots + |a_k|^2\|\mathbf{v}_k\|^2.$$

Each term is a nonnegative real number, so each term is 0. Since $\|\mathbf{v}_l\|^2 > 0$ for every l, it follows that $|a_l|^2 = 0$, that is, $a_l = 0$, for every l. By Theorem 1.5.5 the list is independent. □

The first proof extracts the coefficients one at a time; the second shows with a single number that they all vanish. The second also says more: by Corollary 6.3.6, a combination of orthogonal unit vectors has length $\left(\sum_l |a_l|^2\right)^{1/2}$, so it is short only when all its coefficients are small. Both proofs use that $\|\mathbf{v}_j\|^2 \neq 0$ for $\mathbf{v}_j \neq \mathbf{0}$. For the "dot product" $\sum_k x_k y_k$ on $\mathbb{F}_2^n$ with $n \geq 2$, which is not an inner product, this fails, and so does the proposition (Exercise 6.3.47). The hypothesis that the vectors are nonzero is needed too, since the orthogonal list $\mathbf{v}, \mathbf{0}$ is dependent.

Corollary 6.3.14 (Counting orthogonal vectors). Let V be finite-dimensional with $\dim V = n$. An orthogonal list of nonzero vectors in V has length at most n, and if its length is exactly n, it is a basis of V.

Proof. ROUTINE

Uses. Proposition 6.3.13 (Orthogonal lists are independent), Theorem 4.1.10 (Independent lists are short), an independent list of length $\dim V$ is a basis (Section 4.1)
Hint 1. By Proposition 6.3.13, such a list is independent. What does Section 4.1 say about independent lists?
Hint 2. Independent lists have length at most $\dim V$, and an independent list of length $\dim V$ is a basis.
Sketch. The list is independent by Proposition 6.3.13. Independent lists in V have length at most n, and independent lists of length n are bases.

By Proposition 6.3.13 the list is linearly independent. By Theorem 4.1.10 its length is at most n, and an independent list of length $n = \dim V$ is a basis of V (Section 4.1). □

So $\mathbb{R}^n$ contains at most n mutually orthogonal nonzero vectors, and to check that n of them form a basis it suffices to check orthogonality.

Orthonormal lists

Definition 6.3.15 (Orthonormal list, orthonormal basis). A list $\mathbf{u}_1, \ldots, \mathbf{u}_k$ in V is **orthonormal** if it is orthogonal and every $\mathbf{u}_j$ is a unit vector. That is,

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij} \quad \text{for all } i, j \in \{1, \ldots, k\},$$

with the Kronecker delta of Notation 2.4.7. An **orthonormal basis** of V is an orthonormal list that is a basis of V.

Unit vectors are nonzero, so an orthonormal list is linearly independent by Proposition 6.3.13, and by Corollary 6.3.14 an orthonormal list in an n-dimensional space is an orthonormal basis exactly when its length is n. An orthogonal list of nonzero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ becomes orthonormal after normalizing (Section 6.1.4): the vectors $\mathbf{v}_j / \|\mathbf{v}_j\|$ are unit vectors, and scaling does not affect orthogonality, as $\langle a\mathbf{v}, b\mathbf{w} \rangle = a\bar{b}\langle \mathbf{v}, \mathbf{w} \rangle$. We write $\mathcal{O}$ for orthonormal lists, as the lecture does.

Example 6.3.16 (Orthonormal bases).

(a) The standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is an orthonormal basis of $\mathbb{F}^n$ with the dot product, since $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

(b) For every $\theta \in \mathbb{R}$, $\mathbf{u}_1 = (\cos\theta, \sin\theta)$, $\mathbf{u}_2 = (-\sin\theta, \cos\theta)$ is an orthonormal basis of $\mathbb{R}^2$: the standard basis rotated by θ.

(c) In $\mathbb{R}^3$, let

$$\mathbf{u}_1 = \frac{1}{3}(1, 2, 2), \quad \mathbf{u}_2 = \frac{1}{3}(2, 1, -2), \quad \mathbf{u}_3 = \frac{1}{3}(2, -2, 1).$$

Each has squared norm $(1 + 4 + 4)/9 = 1$, and $\mathbf{u}_1 \cdot \mathbf{u}_2 = (2 + 2 - 4)/9 = 0$, $\mathbf{u}_1 \cdot \mathbf{u}_3 = (2 - 4 + 2)/9 = 0$, $\mathbf{u}_2 \cdot \mathbf{u}_3 = (4 - 2 - 2)/9 = 0$. So this is an orthonormal list of length $3 = \dim \mathbb{R}^3$, hence an orthonormal basis.

(d) In $\mathbb{C}^2$, $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, i)$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, -i)$ form an orthonormal basis: $\|\mathbf{u}_j\|^2 = (1 + 1)/2 = 1$, and $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$ by Example 6.3.2(b).

(e) With the Frobenius inner product (Proposition 6.1.16), the matrix units E_{ij}, which have entry 1 in position (i, j) and 0 elsewhere, form an orthonormal basis of $\mathbb{F}^{m \times n}$: $\langle E_{ij}, E_{kl} \rangle$ is the sum of the products of corresponding entries, which is 1 if $(i, j) = (k, l)$ and 0 otherwise.

Example 6.3.17 (Orthogonal polynomials). Give $\mathbb{R}[x]_{\leq 2}$ the inner product $\langle p, q\rangle = \int_{-1}^{1} p(t)q(t)\,dt$. (Definiteness: if $\int_{-1}^{1} p(t)^2\,dt = 0$, then p vanishes on $[-1,1]$ by Lemma 6.1.14, so it has infinitely many roots and $p = \mathbf{0}$ by Proposition 7.1.22.) The monomials $1, x, x^2$ are not orthogonal, since $\langle 1, x^2\rangle = 2/3$. But $1, x, x^2 - 1/3$ is an orthogonal list: $\langle 1, x\rangle = 0$ and $\langle x, x^2 - 1/3\rangle = 0$ because the integrands are odd, and

$$\left\langle 1, x^2 - \frac{1}{3}\right\rangle = \int_{-1}^{1}\left(t^2 - \frac{1}{3}\right)dt = \frac{2}{3} - \frac{2}{3} = 0.$$

The squared norms are $\|1\|^2 = 2$, $\|x\|^2 = 2/3$, and

$$\left\|x^2 - \frac{1}{3}\right\|^2 = \int_{-1}^{1}\left(t^4 - \frac{2}{3}t^2 + \frac{1}{9}\right)dt = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45}.$$

So $\frac{1}{\sqrt{2}}, \sqrt{3/2}\,x, \sqrt{45/8}\,(x^2 - 1/3)$ is an orthonormal basis of $\mathbb{R}[x]_{\leq 2}$, by Corollary 6.3.14. Up to scaling these are the first three *Legendre polynomials*. Section 6.4 explains how to produce them from $1, x, x^2$ systematically.

Coordinates are inner products

Intuition (Why orthonormal coordinates come for free). For a general basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of V, finding $[\mathbf{v}]_{\mathcal{B}}$ means solving $\mathbf{v} = a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n$ for $a_1, \ldots, a_n$. In $\mathbb{F}^n$ this is a linear system of n equations in which, in general, every equation involves every unknown, solved by row reducing the augmented matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_n, \mathbf{v}$. Geometrically, in $\mathbb{R}^2$ the coordinate a_1 is found by drawing the line through the tip of $\mathbf{v}$ parallel to $\mathbf{b}_2$ and seeing where it meets the line through $\mathbf{b}_1$ (Figure 6.3.1(a)). So a_1 depends on $\mathbf{b}_2$ as well as on $\mathbf{b}_1$: tilting $\mathbf{b}_2$ moves the point of intersection.

For an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2$, the line through the tip of $\mathbf{v}$ parallel to $\mathbf{u}_2$ is perpendicular to $\mathbf{u}_1$, so $a_1\mathbf{u}_1$ is the foot of the perpendicular from $\mathbf{v}$ to the line through $\mathbf{u}_1$ (Figure 6.3.1(b)). That foot is the shadow of Technique 6.2.4, $c\mathbf{u}_1$ with $c = \langle \mathbf{v}, \mathbf{u}_1\rangle / \|\mathbf{u}_1\|^2 = \langle \mathbf{v}, \mathbf{u}_1\rangle$. So each coordinate is one inner product. It does not involve the other basis vectors, and computing it involves no system at all.

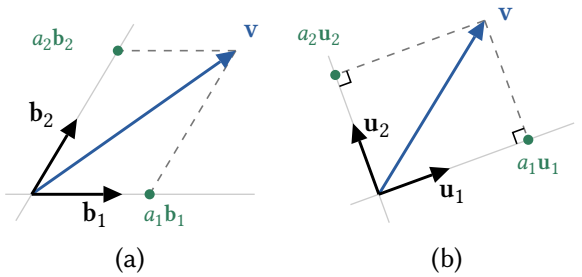


Figure 6.3.1. Reading off coordinates. (a) For a general basis, $a_1\mathbf{b}_1$ is found by moving from the tip of $\mathbf{v}$ parallel to $\mathbf{b}_2$, so it depends on both basis vectors. (b) For an orthonormal basis, $a_j\mathbf{u}_j$ is the foot of the perpendicular from $\mathbf{v}$ to the line through $\mathbf{u}_j$, and $a_j = \langle \mathbf{v}, \mathbf{u}_j\rangle$.

Proposition 6.3.18 (Coordinates in an orthonormal basis). Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal list in V.

(a) If $\mathbf{v} \in \operatorname{span}(\mathcal{O})$, then

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1\rangle\mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k\rangle\mathbf{u}_k.$$

(b) In particular, if $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ is an orthonormal basis of V, then every $\mathbf{v} \in V$ satisfies (a), that is,

$$[\mathbf{v}]_{\mathcal{O}} = (\langle \mathbf{v}, \mathbf{u}_1\rangle, \ldots, \langle \mathbf{v}, \mathbf{u}_n\rangle).$$

Proof. ROUTINE

> *Uses.* Definition 1.3.10 (Span), Definition 6.3.15 (Orthonormal list, orthonormal basis), Definition 1.6.14 (Coordinates)
> *Techniques.* Take an inner product to isolate a coefficient (Technique 6.3.12)
> *Hint 1.* Since $\mathbf{v} \in \operatorname{span}(\mathcal{O})$, write $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$. You need to identify a_j.
> *Hint 2.* Take the inner product of both sides with $\mathbf{u}_j$, with $\mathbf{u}_j$ in the second slot.
> *Sketch.* Write $\mathbf{v} = \sum_l a_l\mathbf{u}_l$. Then $\langle \mathbf{v}, \mathbf{u}_j\rangle = \sum_l a_l\langle \mathbf{u}_l, \mathbf{u}_j\rangle = \sum_l a_l\delta_{lj} = a_j$, so the coefficients are the inner products. For (b), $\operatorname{span}(\mathcal{O}) = V$ and coordinates are the unique coefficients.

(a) Since $\mathbf{v} \in \operatorname{span}(\mathcal{O})$, there are scalars with $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$. For each j, linearity in the first slot gives

$$\langle \mathbf{v}, \mathbf{u}_j\rangle = \sum_{l=1}^{k} a_l\langle \mathbf{u}_l, \mathbf{u}_j\rangle = \sum_{l=1}^{k} a_l\delta_{lj} = a_j.$$

Substituting $a_j = \langle \mathbf{v}, \mathbf{u}_j\rangle$ into $\mathbf{v} = \sum_j a_j\mathbf{u}_j$ gives (a).

(b) If $\mathcal{O}$ is a basis, then $\operatorname{span}(\mathcal{O}) = V$, so (a) applies to every $\mathbf{v} \in V$. The coordinates $[\mathbf{v}]_{\mathcal{O}}$ are the unique coefficients $a_1, \ldots, a_n$ with $\mathbf{v} = \sum_j a_j\mathbf{u}_j$ (Definition 1.6.14), and we have just seen that $a_j = \langle \mathbf{v}, \mathbf{u}_j\rangle$. □

Remark 6.3.19 (Which slot). The coefficient of $\mathbf{u}_j$ is $\langle \mathbf{v}, \mathbf{u}_j\rangle$, with $\mathbf{v}$ in the first slot, the slot in which the inner product is linear. The other order gives $\langle \mathbf{u}_j, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}_j\rangle}$. Over $\mathbb{R}$ the two agree, but over $\mathbb{C}$ the conjugated numbers are the wrong coordinates. Take the orthonormal basis $\mathbf{u}_1 = (1, i)/\sqrt{2}$, $\mathbf{u}_2 = (1, -i)/\sqrt{2}$ of Example 6.3.16(d) and $\mathbf{v} = (0, 1)$. Then

$$\langle \mathbf{v}, \mathbf{u}_1\rangle = \frac{1}{\sqrt{2}}\big(0 \cdot 1 + 1 \cdot \bar{i}\big) = -\frac{i}{\sqrt{2}},$$
$$\langle \mathbf{v}, \mathbf{u}_2\rangle = \frac{1}{\sqrt{2}}\big(0 \cdot 1 + 1 \cdot \overline{-i}\big) = \frac{i}{\sqrt{2}},$$

and indeed $(-i/2)(1, i) + (i/2)(1, -i) = (0, 1)$. The conjugated coefficients give $(i/2)(1, i) - (i/2)(1, -i) = (0, -1)$ instead. Texts that make inner products linear in the second slot, as in Non-example 6.1.25(b), write the coefficient as $\langle \mathbf{u}_j, \mathbf{v}\rangle$; it is the same number.

Example 6.3.20 (Coordinates without a linear system). Take the orthonormal basis $\mathcal{O} = \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ of $\mathbb{R}^3$ from Example 6.3.16(c) and $\mathbf{v} = (1, 2, 3)$. By Proposition 6.3.18,

$$\langle \mathbf{v}, \mathbf{u}_1\rangle = \frac{1 + 4 + 6}{3} = \frac{11}{3},$$
$$\langle \mathbf{v}, \mathbf{u}_2\rangle = \frac{2 + 2 - 6}{3} = -\frac{2}{3},$$
$$\langle \mathbf{v}, \mathbf{u}_3\rangle = \frac{2 - 4 + 3}{3} = \frac{1}{3},$$

so $[\mathbf{v}]_{\mathcal{O}} = (11/3, -2/3, 1/3)$. Check:

$$\frac{11}{9}(1, 2, 2) - \frac{2}{9}(2, 1, -2) + \frac{1}{9}(2, -2, 1)$$
$$= \frac{1}{9}(9, 18, 27) = (1, 2, 3).$$

Three dot products replaced the row reduction of a 3×4 augmented matrix.

The same isolation works for orthogonal bases that are not normalized; each coefficient just picks up the factor $1/\|\mathbf{v}_j\|^2$.

> **Corollary 6.3.21** (Coordinates in an orthogonal basis). Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be an orthogonal basis of V, that is, an orthogonal list of nonzero vectors that is a basis. Then every $\mathbf{v} \in V$ satisfies

$$\mathbf{v} = \sum_{j=1}^{n} \frac{\langle \mathbf{v}, \mathbf{v}_j \rangle}{\|\mathbf{v}_j\|^2} \mathbf{v}_j.$$

Proof. ROUTINE

Uses. Definition 1.6.1 (Basis), Definition 6.3.1 (Orthogonality)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Write $\mathbf{v} = \sum_l a_l \mathbf{v}_l$ and take the inner product with $\mathbf{v}_j$.
Hint 2. You get $\langle \mathbf{v}, \mathbf{v}_j \rangle = a_j \|\mathbf{v}_j\|^2$, and $\|\mathbf{v}_j\| \neq 0$.
Sketch. With $\mathbf{v} = \sum_l a_l \mathbf{v}_l$, isolating coefficients gives $\langle \mathbf{v}, \mathbf{v}_j \rangle = a_j \|\mathbf{v}_j\|^2$; divide by $\|\mathbf{v}_j\|^2 > 0$.

Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis, $\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n$ for some scalars. Taking the inner product with $\mathbf{v}_j$ and using orthogonality,

$$\langle \mathbf{v}, \mathbf{v}_j \rangle = \sum_{l=1}^{n} a_l \langle \mathbf{v}_l, \mathbf{v}_j \rangle = a_j \|\mathbf{v}_j\|^2.$$

As $\mathbf{v}_j \neq \mathbf{0}$, we may divide by $\|\mathbf{v}_j\|^2 > 0$, which gives $a_j = \langle \mathbf{v}, \mathbf{v}_j \rangle / \|\mathbf{v}_j\|^2$. □

Each term $\langle \mathbf{v}, \mathbf{v}_j \rangle / \|\mathbf{v}_j\|^2 \mathbf{v}_j$ is the shadow of $\mathbf{v}$ on the line through $\mathbf{v}_j$ (Technique 6.2.4). So in an orthogonal basis a vector is the sum of its shadows on the basis lines. For a basis that is not orthogonal this fails. In $\mathbb{R}^2$ with $\mathbf{v}_1 = (1, 0)$, $\mathbf{v}_2 = (1, 1)$, and $\mathbf{v} = (0, 1) = -\mathbf{v}_1 + \mathbf{v}_2$, the shadows are $0 \cdot \mathbf{v}_1$ and $(1/2)\mathbf{v}_2$, which add up to $(1/2, 1/2) \neq \mathbf{v}$.

Parseval's identity

Coordinates with respect to an orthonormal basis do more than locate a vector. They compute lengths and inner products, by the same formulas as in $\mathbb{F}^n$.

Theorem 6.3.22 (Parseval's identity). Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ be an orthonormal basis of V, and let $\mathbf{v}, \mathbf{w} \in V$.

(a) $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle \overline{\langle \mathbf{w}, \mathbf{u}_j \rangle}$. In coordinates,

$$\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{v}]_{\mathcal{O}} \cdot [\mathbf{w}]_{\mathcal{O}} = [\mathbf{w}]_{\mathcal{O}}^{*} [\mathbf{v}]_{\mathcal{O}}.$$

(b) $\|\mathbf{v}\|^2 = \sum_{j=1}^{n} |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2$, which is the squared norm of $[\mathbf{v}]_{\mathcal{O}}$ in $\mathbb{F}^n$.

Proof. ROUTINE

Uses. Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 6.1.4 (Sesquilinearity), Definition 6.1.8 (Dot product), Notation 6.1.7 (Conjugate transpose), Proposition 6.1.10 (The dot product is an inner product)
Hint 1. It suffices to expand one of the two vectors in the basis. Which one?
Hint 2. Write $\mathbf{w} = \sum_j \langle \mathbf{w}, \mathbf{u}_j \rangle \mathbf{u}_j$ by Proposition 6.3.18 and pull the coefficients out of the second slot, remembering the conjugates.
Sketch. Substituting $\mathbf{w} = \sum_j \langle \mathbf{w}, \mathbf{u}_j \rangle \mathbf{u}_j$ into the second slot gives $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_j \overline{\langle \mathbf{w}, \mathbf{u}_j \rangle} \langle \mathbf{v}, \mathbf{u}_j \rangle$, which is the dot product of the coordinate vectors. Take $\mathbf{w} = \mathbf{v}$ for (b).

(a) By Proposition 6.3.18, $\mathbf{w} = \sum_{j=1}^{n} \langle \mathbf{w}, \mathbf{u}_j \rangle \mathbf{u}_j$. By Proposition 6.1.4,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \left\langle \mathbf{v}, \sum_{j=1}^{n} \langle \mathbf{w}, \mathbf{u}_j \rangle \mathbf{u}_j \right\rangle = \sum_{j=1}^{n} \overline{\langle \mathbf{w}, \mathbf{u}_j \rangle} \langle \mathbf{v}, \mathbf{u}_j \rangle.$$

By Proposition 6.3.18(b), the jth entries of $[\mathbf{v}]_{\mathcal{O}}$ and $[\mathbf{w}]_{\mathcal{O}}$ are $\langle \mathbf{v}, \mathbf{u}_j \rangle$ and $\langle \mathbf{w}, \mathbf{u}_j \rangle$, so by Definition 6.1.8 and Notation 6.1.7 this sum is $[\mathbf{v}]_{\mathcal{O}} \cdot [\mathbf{w}]_{\mathcal{O}} = [\mathbf{w}]_{\mathcal{O}}^{*} [\mathbf{v}]_{\mathcal{O}}$.

(b) Take $\mathbf{w} = \mathbf{v}$ in (a) and use $z\bar{z} = |z|^2$. By Proposition 6.1.10 the result is the squared norm of $[\mathbf{v}]_{\mathcal{O}}$ for the dot product. □

Part (b) is also Corollary 6.3.6 applied to $\mathbf{v} = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$. For the basis of Example 6.3.20, it says $(121 + 4 + 1)/9 = 14 = \|(1, 2, 3)\|^2$, and since $[\mathbf{e}_3]_{\mathcal{O}} = (2/3, -2/3, 1/3)$, part (a) says

$$\mathbf{v} \cdot \mathbf{e}_3 = \frac{11}{3} \cdot \frac{2}{3} + \left(-\frac{2}{3}\right)\left(-\frac{2}{3}\right) + \frac{1}{3} \cdot \frac{1}{3} = \frac{27}{9} = 3,$$

which is indeed the third entry of $\mathbf{v}$.

Intuition (An orthonormal basis turns V into $\mathbb{F}^n$). The coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{O}}$ is a linear bijection $V \to \mathbb{F}^n$ (Proposition 1.6.19), and Parseval's identity says that it carries the inner product of V to the dot product of $\mathbb{F}^n$. So once an orthonormal basis is chosen, every question about lengths, angles, and orthogonality in V can be answered in coordinates, with the formulas of $\mathbb{F}^n$. Gram–Schmidt (Theorem 6.4.20) shows that every finite-dimensional inner product space has an orthonormal basis, so every n-dimensional inner product space looks like $\mathbb{F}^n$ with the dot product; Section 8.4 makes "looks like" precise.

This is special to orthonormal bases. For the basis $\mathcal{B} = (1,0),(1,1)$ of $\mathbb{R}^2$, the orthogonal vectors $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = -(1,0) + (1,1)$ have coordinate vectors $(1,0)$ and $(-1,1)$, whose dot product is -1. In fact $[\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}$ is a different inner product on $\mathbb{R}^2$, the unique one for which $\mathcal{B}$ is orthonormal (Exercise 6.3.51).

Exercise 6.3.23. ROUTINE Show that $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1,0)$, $\mathbf{u}_2 = \frac{1}{\sqrt{3}}(1,-1,1)$, $\mathbf{u}_3 = \frac{1}{\sqrt{6}}(-1,1,2)$ is an orthonormal basis of $\mathbb{R}^3$. Find the coordinates of $\mathbf{v} = (3,1,2)$ in this basis, and check your answer with Theorem 6.3.22(b).

Solution on page 1483

Bessel's inequality

Parseval's identity needs a basis. For an orthonormal list that does not span V, the sum $\sum_j |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2$ misses part of $\mathbf{v}$, and an inequality remains.

Intuition (What is left over). In $\mathbb{R}^3$ with $\mathbf{u}_1 = \mathbf{e}_1$ and $\mathbf{u}_2 = \mathbf{e}_2$, the sum is $v_1^2 + v_2^2$, which falls short of $\|\mathbf{v}\|^2$ by v_3^2, the squared length of the part of $\mathbf{v}$ perpendicular to the plane of $\mathbf{u}_1, \mathbf{u}_2$. In general, split $\mathbf{v} = \mathbf{p} + \mathbf{w}$, where $\mathbf{p} = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$ is what the formula of Proposition 6.3.18 produces and $\mathbf{w} = \mathbf{v} - \mathbf{p}$ is the remainder. Isolating coefficients shows that $\mathbf{w}$ is orthogonal to every $\mathbf{u}_j$, hence to $\mathbf{p}$, and Pythagoras gives $\|\mathbf{v}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{w}\|^2$. For a single unit vector $\mathbf{u}_1$, this is the decomposition of the first proof of Theorem 6.2.1, and the inequality below is the Cauchy–Schwarz inequality $|\langle \mathbf{v}, \mathbf{u}_1 \rangle| \le \|\mathbf{v}\|$.

Theorem 6.3.24 (Bessel's inequality). Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal list in V, and let $\mathbf{v} \in V$. Then

$$\sum_{j=1}^{k} |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2 \le \|\mathbf{v}\|^2,$$

with equality if and only if $\mathbf{v} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$.

Proof. STANDARD

Uses. Corollary 6.3.6 (Pythagoras for orthogonal lists), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.3.9 (Properties of $A^\perp$), Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Orthogonal decomposition (Technique 6.2.4), Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Let $\mathbf{p} = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$ and $\mathbf{w} = \mathbf{v} - \mathbf{p}$. Compute $\langle \mathbf{w}, \mathbf{u}_l \rangle$ for each l.
Hint 2. $\mathbf{w}$ is orthogonal to every $\mathbf{u}_l$, hence to $\mathbf{p}$. Apply Theorem 6.3.3 to $\mathbf{v} = \mathbf{p} + \mathbf{w}$, and compute $\|\mathbf{p}\|^2$ with Corollary 6.3.6.
Hint 3. Equality means $\mathbf{w} = \mathbf{0}$. For the converse, use Proposition 6.3.18(a).
Sketch. With $\mathbf{p} = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$ and $\mathbf{w} = \mathbf{v} - \mathbf{p}$, we get $\langle \mathbf{w}, \mathbf{u}_l \rangle = \langle \mathbf{v}, \mathbf{u}_l \rangle - \langle \mathbf{v}, \mathbf{u}_l \rangle = 0$, so $\mathbf{w} \perp \mathbf{p}$. Pythagoras gives $\|\mathbf{v}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{w}\|^2 \ge \|\mathbf{p}\|^2 = \sum_j |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2$. Equality holds iff $\mathbf{w} = \mathbf{0}$, that is, $\mathbf{v} = \mathbf{p}$. If $\mathbf{v} = \mathbf{p}$, then $\mathbf{v}$ lies in the span; conversely, if $\mathbf{v}$ lies in the span, then $\mathbf{v} = \mathbf{p}$ by Proposition 6.3.18.

Let $\mathbf{p} = \sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$ and $\mathbf{w} = \mathbf{v} - \mathbf{p}$. For each l,

$$\langle \mathbf{w}, \mathbf{u}_l \rangle = \langle \mathbf{v}, \mathbf{u}_l \rangle - \sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{u}_j \rangle \langle \mathbf{u}_j, \mathbf{u}_l \rangle$$
$$= \langle \mathbf{v}, \mathbf{u}_l \rangle - \langle \mathbf{v}, \mathbf{u}_l \rangle = 0.$$

So $\mathbf{w} \in \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}^\perp = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)^\perp$ by Proposition 6.3.9(c), and in particular $\mathbf{w} \perp \mathbf{p}$. By Theorem 6.3.3(a) applied to $\mathbf{v} = \mathbf{p} + \mathbf{w}$, and Corollary 6.3.6 applied to $\mathbf{p}$,

$$\|\mathbf{v}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{w}\|^2 = \sum_{j=1}^{k} \left| \langle \mathbf{v}, \mathbf{u}_j \rangle \right|^2 + \|\mathbf{w}\|^2. \tag{6.3.2}$$

Since $\|\mathbf{w}\|^2 \geq 0$, the inequality follows.

By (6.3.2), equality holds if and only if $\|\mathbf{w}\| = 0$, that is (Proposition 6.1.27), if and only if $\mathbf{v} = \mathbf{p}$. If $\mathbf{v} = \mathbf{p}$, then $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Conversely, if $\mathbf{v} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, then $\mathbf{v} = \mathbf{p}$ by Proposition 6.3.18(a). □

The space V need not be finite-dimensional here. The vector $\mathbf{p}$ is the *orthogonal projection* of $\mathbf{v}$ onto $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, the subject of Section 6.4, where we will also see that $\|\mathbf{w}\| = \|\mathbf{v} - \mathbf{p}\|$ is the distance from $\mathbf{v}$ to that span (Theorem 6.4.33). Bessel's inequality says that projecting onto a subspace cannot increase length, and its equality case gives a membership test: $\mathbf{v}$ lies in the span of an orthonormal list exactly when the list captures all of $\|\mathbf{v}\|^2$.

Example 6.3.25 (Bessel in $\mathbb{R}^3$). Keep $\mathbf{v} = (1, 2, 3)$ and the vectors $\mathbf{u}_1, \mathbf{u}_2$ of Example 6.3.16(c), but drop $\mathbf{u}_3$. Then

$$|\langle \mathbf{v}, \mathbf{u}_1 \rangle|^2 + |\langle \mathbf{v}, \mathbf{u}_2 \rangle|^2 = \frac{121}{9} + \frac{4}{9} = \frac{125}{9} < 14,$$

and the deficit $1/9 = \langle \mathbf{v}, \mathbf{u}_3 \rangle^2$ is the part that $\mathbf{u}_3$ would have contributed. So $(1, 2, 3) \notin \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$. On the other hand $\mathbf{x} = (1, 1, 0)$ has $\langle \mathbf{x}, \mathbf{u}_1 \rangle = \langle \mathbf{x}, \mathbf{u}_2 \rangle = 1$, so the sum is $2 = \|\mathbf{x}\|^2$, and therefore $\mathbf{x} \in \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$; indeed $\mathbf{x} = \mathbf{u}_1 + \mathbf{u}_2$.

ASIDE: *Fourier coefficients* BEYOND 21-242

Bessel's inequality matters most in infinite-dimensional spaces, where an orthonormal list can never be a basis. On $C([0, 2\pi], \mathbb{C})$, take the inner product $\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}\, dt$ (Proposition 6.1.13 on $[0, 2\pi]$, scaled by a positive constant), and let $e_k(t) = \cos kt + i \sin kt$ for $k \in \mathbb{Z}$. The angle subtraction formulas give $e_j \overline{e_k} = e_{j-k}$, and $\int_0^{2\pi} \cos mt\, dt = \int_0^{2\pi} \sin mt\, dt = 0$ for $m \neq 0$, so $\langle e_j, e_k \rangle = \delta_{jk}$. Bessel's inequality for the orthonormal list $e_{-N}, \ldots, e_N$ says that the *Fourier coefficients* $c_k = \langle f, e_k \rangle$ of f satisfy

$$\sum_{k=-N}^{N} |c_k|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2\, dt \quad \text{for every } N.$$

The bound does not depend on N, so the series $\sum_k |c_k|^2$ converges, and in particular $c_k \to 0$ as $|k| \to \infty$. Parseval's identity, with equality, also holds for the full infinite sum, but proving it takes analysis (Section 15.3). Both names come from this setting: the identity and the inequality were first found for Fourier series.

6.3.3 Matrices with orthonormal columns

In $\mathbb{F}^m$ with the dot product, a list of n vectors is the same thing as an $m \times n$ matrix, whose columns are the vectors. Orthonormality of the list is a statement about the dot products of the columns with each other, and these are the entries of a single matrix product.

Definition 6.3.26 (Conjugate transpose). Let $A = (a_{ij}) \in \mathbb{F}^{m\times n}$. The **conjugate transpose** of A is the matrix $A^* \in \mathbb{F}^{n\times m}$ with entries

$$(A^*)_{ij} = \overline{a_{ji}} \quad (1 \le i \le n, 1 \le j \le m).$$

That is, $A^* = \overline{A}^{\mathsf{T}}$, where $\overline{A} = (\overline{a_{ij}})$ conjugates every entry (Definition 2.4.21). If A is real, $A^* = A^{\mathsf{T}}$. A square matrix A with $A^* = A$ is called **Hermitian**; a real Hermitian matrix is a symmetric matrix.

This is the matrix of Notation 6.1.7. In particular, for $\mathbf{w} \in \mathbb{F}^n$, $\mathbf{w}^*$ is the row $(\overline{w_1}, \dots, \overline{w_n})$, and $\mathbf{w}^*\mathbf{v} = \mathbf{v} \cdot \mathbf{w}$ for $\mathbf{v} \in \mathbb{F}^n$ (Definition 6.1.8). For example,

$$\begin{pmatrix} 1 & i & 2 \\ 0 & 1-i & 3i \end{pmatrix}^* = \begin{pmatrix} 1 & 0 \\ -i & 1+i \\ 2 & -3i \end{pmatrix}.$$

Proposition 6.3.27 (Properties of the conjugate transpose). Let $A, A' \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times p}$, and $c \in \mathbb{F}$.

(a) $(A^*)^* = A$ and $(cA + A')^* = \bar{c}A^* + A'^*$.

(b) $(AB)^* = B^*A^*$.

(c) If $m = n$ and A is invertible, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.

(d) $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y})$ for all $\mathbf{x} \in \mathbb{F}^n$ and $\mathbf{y} \in \mathbb{F}^m$.

Proof. ROUTINE

Uses. Definition 6.3.26 (Conjugate transpose), Proposition 2.4.22 (Transpose is a linear involution), Proposition 2.4.23 (Transpose of a product), Proposition 2.4.14 (Algebra of matrices), Definition 2.5.12 (Invertible), Definition 6.1.8 (Dot product), Notation 6.1.1 (Real and complex scalars)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. (a) is entrywise. For (b), first show $\overline{AB} = \overline{A}\,\overline{B}$ entry by entry, then transpose with Proposition 2.4.23.
Hint 2. For (c), apply $*$ to $AA^{-1} = I$ and $A^{-1}A = I$ using (b).
Hint 3. For (d), write both sides as matrix products: $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{y}^*(A\mathbf{x})$. Regroup and use (a) and (b) to recognize $\mathbf{y}^*A$ as $(A^*\mathbf{y})^*$.
Sketch. (a) Conjugating twice and transposing twice undo themselves, and conjugation respects sums and products. (b) $\overline{AB} = \overline{A}\,\overline{B}$ entrywise, and $(\overline{A}\,\overline{B})^{\mathsf{T}} = \overline{B}^{\mathsf{T}}\overline{A}^{\mathsf{T}}$. (c) $A^*(A^{-1})^* = (A^{-1}A)^* = I^* = I$, and similarly on the other side. (d) $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{y}^*A\mathbf{x} = (A^*\mathbf{y})^*\mathbf{x} = \mathbf{x} \cdot (A^*\mathbf{y})$.

(a) The (i,j) entry of $(A^*)^*$ is $\overline{(A^*)_{ji}} = \overline{\overline{a_{ij}}} = a_{ij}$. The (i,j) entry of $(cA + A')^*$ is $\overline{ca_{ji} + a'_{ji}} = \bar{c}\,\overline{a_{ji}} + \overline{a'_{ji}}$, the (i,j) entry of $\bar{c}A^* + A'^*$.

(b) Since conjugation respects sums and products (Notation 6.1.1), the (i,j) entry of $\overline{AB}$ is $\overline{\sum_l a_{il}b_{lj}} = \sum_l \overline{a_{il}}\,\overline{b_{lj}}$, the (i,j) entry of $\overline{A}\,\overline{B}$. So $\overline{AB} = \overline{A}\,\overline{B}$, and by Proposition 2.4.23,

$$(AB)^* = (\overline{A}\,\overline{B})^{\mathsf{T}} = \overline{B}^{\mathsf{T}}\,\overline{A}^{\mathsf{T}} = B^*A^*.$$

(c) By (b), $A^*(A^{-1})^* = (A^{-1}A)^* = I^* = I$ and $(A^{-1})^*A^* = (AA^{-1})^* = I^* = I$, since I is real and symmetric. So $(A^{-1})^*$ is a two-sided inverse of A^* (Definition 2.5.12).

(d) Identifying 1×1 matrices with scalars, $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{y}^*(A\mathbf{x}) = (\mathbf{y}^*A)\mathbf{x}$ by associativity (Proposition 2.4.14). By (b) and (a), $(A^*\mathbf{y})^* = \mathbf{y}^*(A^*)^* = \mathbf{y}^*A$. Hence $(A\mathbf{x}) \cdot \mathbf{y} = (A^*\mathbf{y})^*\mathbf{x} = \mathbf{x} \cdot (A^*\mathbf{y})$. □

Part (d) is what makes A^* useful: it moves a matrix from one side of the dot product to the other, at the price of a star. Section 6.6 turns this property into the definition of the adjoint of a linear map.

Gram matrices

All the inner products among the vectors of a list fit into one matrix.

Definition 6.3.28 (Gram matrix). The **Gram matrix** of a list $\mathbf{v}_1, \dots, \mathbf{v}_k$ in V is the matrix $G \in \mathbb{F}^{k\times k}$ with entries

$$g_{ij} = \langle \mathbf{v}_j, \mathbf{v}_i \rangle.$$

The order of the indices, with j in the first slot, matches Proposition 6.1.22 and keeps conjugates out of the next results. By conjugate symmetry $G^* = G$, and the diagonal entries are the squared norms $\|\mathbf{v}_i\|^2$. The list is orthogonal if and only if G is diagonal, and orthonormal if and only if $G = I_k$.

Theorem 6.3.29 (Orthonormal columns). Let $A \in \mathbb{F}^{m\times n}$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$. Then

$$(A^*A)_{ij} = \mathbf{a}_j \cdot \mathbf{a}_i \quad (1 \le i, j \le n),$$

so A^*A is the Gram matrix of the columns of A. In particular, the columns of A are orthonormal if and only if $A^*A = I_n$, and orthogonal if and only if A^*A is diagonal.

Proof. ROUTINE

Uses. Proposition 2.4.8 (Four views of the product), Definition 6.3.26 (Conjugate transpose), Definition 6.1.8 (Dot product), Definition 6.3.28 (Gram matrix), Definition 6.3.15 (Orthonormal list, orthonormal basis), Definition 2.3.5 (Identity matrix)
Hint 1. Compute the (i, j) entry of A^*A as row i of A^* times column j of A.
Hint 2. Row i of A^* is $\mathbf{a}_i^*$. What is $\mathbf{a}_i^*\mathbf{a}_j$?
Sketch. Row i of A^* is $\mathbf{a}_i^*$ and column j of A is $\mathbf{a}_j$, so $(A^*A)_{ij} = \mathbf{a}_i^*\mathbf{a}_j = \mathbf{a}_j \cdot \mathbf{a}_i$. Orthonormality says $\mathbf{a}_j \cdot \mathbf{a}_i = \delta_{ij}$, which is the statement $A^*A = I_n$.

By Definition 6.3.26, row i of A^* is $(\overline{a_{1i}}, \dots, \overline{a_{mi}}) = \mathbf{a}_i^*$. By Proposition 2.4.8(a),

$$(A^*A)_{ij} = \mathbf{a}_i^*\mathbf{a}_j = \sum_{l=1}^{m} a_{lj}\overline{a_{li}} = \mathbf{a}_j \cdot \mathbf{a}_i,$$

which is the (i, j) entry of the Gram matrix of $\mathbf{a}_1, \dots, \mathbf{a}_n$ for the dot product. The columns are orthonormal if and only if $\mathbf{a}_j \cdot \mathbf{a}_i = \delta_{ij}$ for all i, j (Definition 6.3.15), that is, if and only if $A^*A = (\delta_{ij}) = I_n$ (Definition 2.3.5). Similarly, they are orthogonal if and only if every off-diagonal entry of A^*A is 0. □

For example, the matrix with the columns $\mathbf{u}_1 = (1, i)/\sqrt{2}$ and $\mathbf{u}_2 = (1, -i)/\sqrt{2}$ of Example 6.3.16(d) satisfies

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+1 & 1-1 \\ 1-1 & 1+1 \end{pmatrix} = I_2.$$

The condition $A^*A = I_n$ does not require A to be square. The 3×2 matrix with columns $\mathbf{u}_1, \mathbf{u}_2$ from Example 6.3.16(c) has $A^*A = I_2$. But $AA^* \neq I_3$, since AA^* has rank at most 2 (Theorem 4.3.3); Section 6.4 identifies AA^* as the projection onto the plane spanned by the columns.

Gram matrices are useful beyond orthonormality. They detect linear independence of any list.

Proposition 6.3.30 (Gram matrices detect independence). Let G be the Gram matrix of a list $\mathbf{v}_1, \dots, \mathbf{v}_k$ in V, and let $\mathbf{c} = (c_1, \dots, c_k) \in \mathbb{F}^k$. Then:

(a) $\mathbf{c}^*G\mathbf{c} = \|c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k\|^2$;

(b) $G\mathbf{c} = \mathbf{0}$ if and only if $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.

Consequently, G is invertible if and only if $\mathbf{v}_1, \dots, \mathbf{v}_k$ is linearly independent.

Proof. STANDARD

Uses. Definition 6.3.28 (Gram matrix), Definition 2.3.8 (Matrix–vector product), Proposition 6.1.4 (Sesquilinearity), Definition 6.1.2 (Inner product), Theorem 1.5.5 (Independence is decided at zero), a square matrix whose kernel is $\{\mathbf{0}\}$ is invertible (Section 3.5)

Techniques. Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Let $\mathbf{w} = \sum_j c_j \mathbf{v}_j$. Show that the ith entry of $G\mathbf{c}$ is $\langle \mathbf{w}, \mathbf{v}_i \rangle$.
Hint 2. For (a), $\mathbf{c}^* G\mathbf{c} = \sum_i \overline{c_i} (G\mathbf{c})_i$; recognize this as $\left\langle \mathbf{w}, \sum_i c_i \mathbf{v}_i \right\rangle$.
Hint 3. For (b), one direction is immediate from the formula for $(G\mathbf{c})_i$. For the other, use (a) and definiteness.
Sketch. With $\mathbf{w} = \sum_j c_j \mathbf{v}_j$, $(G\mathbf{c})_i = \sum_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle c_j = \langle \mathbf{w}, \mathbf{v}_i \rangle$, and then $\mathbf{c}^* G\mathbf{c} = \sum_i \overline{c_i} \langle \mathbf{w}, \mathbf{v}_i \rangle = \langle \mathbf{w}, \mathbf{w} \rangle$. If $\mathbf{w} = \mathbf{0}$, then $G\mathbf{c} = \mathbf{0}$; if $G\mathbf{c} = \mathbf{0}$, then $\|\mathbf{w}\|^2 = \mathbf{c}^* \mathbf{0} = 0$. So $\ker G$ is the set of coefficient vectors of linear relations, which is $\{\mathbf{0}\}$ exactly when the list is independent.

Let $\mathbf{w} = c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k$. By Definition 2.3.8 and linearity in the first slot,

$$(G\mathbf{c})_i = \sum_{j=1}^{k} g_{ij} c_j = \sum_{j=1}^{k} c_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \langle \mathbf{w}, \mathbf{v}_i \rangle.$$

(a) By the formula above and Proposition 6.1.4,

$$\begin{aligned} \mathbf{c}^* G\mathbf{c} &= \sum_{i=1}^{k} \overline{c_i} (G\mathbf{c})_i = \sum_{i=1}^{k} \overline{c_i} \langle \mathbf{w}, \mathbf{v}_i \rangle \\ &= \left\langle \mathbf{w}, \sum_{i=1}^{k} c_i \mathbf{v}_i \right\rangle = \|\mathbf{w}\|^2. \end{aligned}$$

(b) If $\mathbf{w} = \mathbf{0}$, every entry $\langle \mathbf{w}, \mathbf{v}_i \rangle$ of $G\mathbf{c}$ is 0. Conversely, if $G\mathbf{c} = \mathbf{0}$, then by (a), $\|\mathbf{w}\|^2 = \mathbf{c}^* \mathbf{0} = 0$, so $\mathbf{w} = \mathbf{0}$ by definiteness.

By (b), $\ker G$ consists of the coefficient vectors of linear relations among $\mathbf{v}_1, \ldots, \mathbf{v}_k$. So $\ker G = \{\mathbf{0}\}$ if and only if the list is linearly independent (Theorem 1.5.5). Since G is square, $\ker G = \{\mathbf{0}\}$ if and only if G is invertible (Section 3.5). □

Technique 6.3.31 (Gram matrices). To study a list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in an inner product space, record all the inner products $\langle \mathbf{v}_j, \mathbf{v}_i \rangle$ in its Gram matrix G. Questions about the list become questions about one matrix: the list is orthonormal if and only if $G = I$, orthogonal if and only if G is diagonal, and independent if and only if G is invertible, and the combination with coefficient vector $\mathbf{c}$ has squared length $\mathbf{c}^* G\mathbf{c}$. When the vectors are the columns of $A \in \mathbb{F}^{m \times n}$, $G = A^* A$, so these become statements about $A^* A$.

Reach for it when the hypotheses are about inner products among the vectors of a list (lengths, angles, orthogonality) and the conclusion is about the list itself, or when the product $A^* A$ appears. The canonical example is Proposition 6.3.30. In Section 6.5 the normal equations have the coefficient matrix $A^* A$, and Proposition 6.3.30 shows that they have a unique solution exactly when the columns of A are independent.
Used in. Proposition 6.3.41, Exercise 6.3.54, Exercise 6.3.56, Exercise 6.4.59, Proposition 6.5.23, Exercise 6.5.43, Exercise 6.7.11, Problem 6.4, Problem 6.9, Problem 6.12, Theorem 8.3.5, Proposition 8.3.15, Exercise 8.3.32, Exercise 8.3.35, Exercise 6.6.46.

Example 6.3.32 (A relation read off from a Gram matrix). In $\mathbb{R}^3$, let $\mathbf{v}_1 = (1, 0, 1)$, $\mathbf{v}_2 = (0, 1, 1)$, $\mathbf{v}_3 = (1, -1, 0)$. Their Gram matrix is

$$G = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix},$$

and $G(1, -1, -1) = (2 - 1 - 1, 1 - 2 + 1, 1 + 1 - 2) = \mathbf{0}$. By Proposition 6.3.30(b), $\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$, which is easily checked. So the list is dependent, and G is not invertible.

Unitary and orthogonal matrices

Definition 6.3.33 (Unitary and orthogonal matrices). A square matrix $U \in \mathbb{F}^{n\times n}$ is **unitary** if $U^*U = I_n$, that is (Theorem 6.3.29), if its columns form an orthonormal basis of $\mathbb{F}^n$. A real unitary matrix, that is, a matrix $Q \in \mathbb{R}^{n\times n}$ with $Q^{\mathsf{T}}Q = I_n$, is called **orthogonal**.

An $n \times n$ matrix has n columns, so by Corollary 6.3.14 its columns are an orthonormal basis of $\mathbb{F}^n$ as soon as they are orthonormal. The name "orthogonal matrix" is traditional but misleading: the columns of an orthogonal matrix are orthonormal, not merely orthogonal. The matrix $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ has orthogonal columns of length $\sqrt{2}$ and is not orthogonal; dividing it by $\sqrt{2}$ makes it orthogonal.

Proposition 6.3.34 (Characterizations of unitary matrices). For $U \in \mathbb{F}^{n\times n}$, the following are equivalent:

(i) $U^*U = I_n$, that is, the columns of U are orthonormal;
(ii) U is invertible and $U^{-1} = U^*$;
(iii) $UU^* = I_n$, that is, the rows of U, read as vectors of $\mathbb{F}^n$, are orthonormal;
(iv) U^* is unitary.

Proof. ROUTINE

Uses. Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 6.3.27 (Properties of the conjugate transpose), Theorem 6.3.29 (Orthonormal columns), Definition 2.5.12 (Invertible), Definition 6.1.8 (Dot product)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. (i) says that U^* is a left inverse of the square matrix U. What do you know about one-sided inverses of square matrices?
Hint 2. Prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i), then compare (iii) and (iv) using $(U^*)^* = U$.
Hint 3. For the reading of (iii), compute the (i, j) entry of UU^* as a dot product of two rows.
Sketch. For a square matrix a one-sided inverse is two-sided, so $U^*U = I$ and $UU^* = I$ each imply $U^{-1} = U^*$, which implies both. Statement (iv) says $(U^*)^*U^* = I$, which is $UU^* = I$. Finally $(UU^*)_{ij} = \sum_l u_{il}\overline{u_{jl}}$ is the dot product of rows i and j.

We prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i) and (iii) ⇔ (iv).

(i) ⇒ (ii). U^* is a left inverse of the square matrix U, so it is a two-sided inverse by Theorem 3.5.23. That is, U is invertible with $U^{-1} = U^*$.

(ii) ⇒ (iii). $UU^* = UU^{-1} = I_n$.

(iii) ⇒ (i). U^* is a right inverse of the square matrix U, so again it is a two-sided inverse by Theorem 3.5.23, and $U^*U = I_n$.

(iii) ⇔ (iv). By Proposition 6.3.27(a), $(U^*)^*U^* = UU^*$, so U^* is unitary exactly when $UU^* = I_n$.

Finally, the (i, j) entry of UU^* is $\sum_l u_{il}\overline{u_{jl}}$, which is the dot product of $\mathbf{r}_i = (u_{i1}, \ldots, u_{in})$ with $\mathbf{r}_j$ (Definition 6.1.8). So $UU^* = I_n$ says that $\mathbf{r}_i \cdot \mathbf{r}_j = \delta_{ij}$, that is, the rows are orthonormal. □

That the rows of a square matrix are orthonormal whenever its columns are is not visible from either list by itself; it comes from the fact that one-sided inverses of square matrices are two-sided. For non-square matrices it fails: the columns of $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ are orthonormal, but its rows $(1, 0)$, $(0, 1)$, $(0, 0)$ are not.

Example 6.3.35 (Unitary and orthogonal matrices).

(a) For every $\theta \in \mathbb{R}$, the rotation and the reflection

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

are orthogonal. The columns of R_θ are the orthonormal basis of Example 6.3.16(b), and those of S_θ are the same two vectors up to the sign of the second.

(b) Every permutation matrix (Definition 5.2.29) is orthogonal, since its columns are the standard basis vectors in some order.

(c) A diagonal matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is unitary if and only if $|\lambda_j| = 1$ for every j, since its columns $\lambda_j \mathbf{e}_j$ are orthogonal with norms $|\lambda_j|$.

(d) The matrices

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \quad \text{and} \quad \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$$

are unitary and orthogonal respectively, by Example 6.3.16(d) and (c).

The definition only asks for orthonormal columns. What unitary matrices do geometrically, namely preserve lengths and inner products, and what they look like (rotations and reflections, in the real case) is the subject of Section 8.4. Here we only record why they belong in this section: they are the matrices whose linear systems are solved by the method of Proposition 6.3.18.

Corollary 6.3.36 (Orthonormal coordinates in $\mathbb{F}^n$). Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ be an orthonormal basis of $\mathbb{F}^n$, and let U be the unitary matrix with columns $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then $[\mathbf{x}]_\mathcal{O} = U^*\mathbf{x}$ for every $\mathbf{x} \in \mathbb{F}^n$.

Proof (by inner products). ROUTINE

Uses. Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 2.4.8 (Four views of the product), Definition 6.3.26 (Conjugate transpose)
Hint 1. By Proposition 6.3.18, the jth coordinate is $\mathbf{x} \cdot \mathbf{u}_j$. Write it as a matrix product.
Hint 2. $\mathbf{x} \cdot \mathbf{u}_j = \mathbf{u}_j^*\mathbf{x}$, and $\mathbf{u}_j^*$ is row j of U^*.
Sketch. The jth coordinate is $\mathbf{x} \cdot \mathbf{u}_j = \mathbf{u}_j^*\mathbf{x}$, and row j of U^* is $\mathbf{u}_j^*$, so this is the jth entry of $U^*\mathbf{x}$.

By Proposition 6.3.18(b), the jth entry of $[\mathbf{x}]_\mathcal{O}$ is $\mathbf{x} \cdot \mathbf{u}_j = \mathbf{u}_j^*\mathbf{x}$. Row j of U^* is $\mathbf{u}_j^*$ (Definition 6.3.26), so by Proposition 2.4.8 this is the jth entry of $U^*\mathbf{x}$. □

Proof (by solving the system). ROUTINE

Uses. Definition 1.6.14 (Coordinates), Definition 2.3.8 (Matrix–vector product), Proposition 6.3.34 (Characterizations of unitary matrices)
Hint 1. The coordinate vector $\mathbf{c} = [\mathbf{x}]_\mathcal{O}$ is the unique solution of $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{x}$. Write this as a matrix equation.
Hint 2. The equation is $U\mathbf{c} = \mathbf{x}$. Solve it with U^{-1}.
Sketch. $[\mathbf{x}]_\mathcal{O}$ is the solution of $U\mathbf{c} = \mathbf{x}$, and $U^{-1} = U^*$ by Proposition 6.3.34, so $\mathbf{c} = U^*\mathbf{x}$.

By Definition 1.6.14 and Definition 2.3.8, $\mathbf{c} = [\mathbf{x}]_\mathcal{O}$ is the unique vector with $U\mathbf{c} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{x}$. By Proposition 6.3.34, U is invertible with $U^{-1} = U^*$, so $\mathbf{c} = U^{-1}\mathbf{x} = U^*\mathbf{x}$. □

The first proof computes the coordinates one at a time as inner products. The second says that the linear system $U\mathbf{c} = \mathbf{x}$, which computing coordinates always requires, is solved for unitary U by a single multiplication by U^*, with no elimination. In the notation of Section 4.5, $U = [\mathrm{id}]_{\mathcal{E}\mathcal{O}}$ and $U^* = [\mathrm{id}]_{\mathcal{O}\mathcal{E}}$, and the general fact that these two change-of-coordinates matrices are inverse to each other becomes $U^{-1} = U^*$.

Exercise 6.3.37. ROUTINE Let $U, W \in \mathbb{F}^{n\times n}$ be unitary. Show that UW and U^{-1} are unitary, and that $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ and $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$.

Solution on page 1484

Exercise 6.3.38. ROUTINE Find all $a, b \in \mathbb{C}$ for which $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & a \\ i & b \end{pmatrix}$ is unitary.

Solution on page 1484

6.3.4 The discrete Fourier basis

After the standard basis, the most useful orthonormal basis of $\mathbb{C}^n$ consists of sampled waves. Computing coordinates in it is the *discrete Fourier transform* of signal processing. Its orthogonality rests on one computation with roots of unity.

Fix $n \geq 1$ and let

$$\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}.$$

By de Moivre's formula (Section B.1), $\omega^m = \cos(2\pi m/n) + i\sin(2\pi m/n)$ for every integer m. So $|\omega| = 1$, which gives $\overline{\omega^m} = \omega^{-m}$, and $\omega^m = 1$ if and only if m/n is an integer, that is, $n \mid m$. In particular $\omega^n = 1$, so ω^m depends only on the remainder of m modulo n.

Lemma 6.3.39 (Sums of roots of unity). For every integer m,

$$\sum_{j=0}^{n-1} \omega^{jm} = \begin{cases} n \text{ if } n \mid m \\ 0 \text{ if } n \nmid m. \end{cases}$$

Intuition. The numbers ω^{jm}, $j = 0, \ldots, n-1$, go around the unit circle in equal steps, and when $n \nmid m$ they are spread evenly around it. Rotating them all by the step $z = \omega^m$ permutes them, so their sum S satisfies $zS = S$; as $z \neq 1$, this forces $S = 0$. When $n \mid m$ every step is a full turn, and all n terms are 1.

Proof. ROUTINE

Uses. $\omega^m = 1$ if and only if $n \mid m$, Proposition 1.1.17 (No zero divisors)
Hint 1. If $n \mid m$, what is each term? If not, let $z = \omega^m \neq 1$ and $S = \sum_{j=0}^{n-1} z^j$. Compare zS with S.
Hint 2. $zS = z + z^2 + \cdots + z^n$, and $z^n = (\omega^n)^m = 1$.
Sketch. If $n \mid m$, every term is 1. Otherwise $z = \omega^m \neq 1$, and $zS = \sum_{j=1}^{n} z^j = S$ because $z^n = 1 = z^0$. So $(z-1)S = 0$ with $z - 1 \neq 0$, and $S = 0$. (Equivalently, sum the geometric series: $S = (z^n - 1)/(z - 1) = 0$.)

If $n \mid m$, then $\omega^m = 1$, so $\omega^{jm} = (\omega^m)^j = 1$ for every j, and the sum is n. Suppose $n \nmid m$, and let $z = \omega^m$, so $z \neq 1$ and $z^n = (\omega^n)^m = 1$. Let $S = \sum_{j=0}^{n-1} z^j$. Then

$$\begin{aligned} zS &= z + z^2 + \cdots + z^{n-1} + z^n \\ &= z + z^2 + \cdots + z^{n-1} + 1 = S. \end{aligned}$$

So $(z-1)S = 0$, and since $z - 1 \neq 0$, Proposition 1.1.17 gives $S = 0$. □

Definition 6.3.40 (Fourier matrix). The $n \times n$ **Fourier matrix** is the matrix $\Phi_n \in \mathbb{C}^{n\times n}$ with entries

$$(\Phi_n)_{jk} = \omega^{jk},$$

where, unlike elsewhere in the book, rows and columns are indexed by $j, k = 0, 1, \ldots, n-1$. Its columns are $\mathbf{f}_0, \ldots, \mathbf{f}_{n-1} \in \mathbb{C}^n$, where

$$\mathbf{f}_k = \left(1, \omega^k, \omega^{2k}, \ldots, \omega^{(n-1)k}\right).$$

For example, $\Phi_1 = (1)$, $\Phi_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, and since $\omega = i$ for $n = 4$,

$$\Phi_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

Since $jk = kj$, Φ_n is symmetric, so its rows are also $\mathbf{f}_0, \ldots, \mathbf{f}_{n-1}$.

Intuition (The columns are waves). The jth entry of $\mathbf{f}_k$ is $\omega^{jk} = \cos(2\pi kj/n) + i\sin(2\pi kj/n)$. So $\mathbf{f}_k$ records the values of $t \mapsto \cos(2\pi kt) + i\sin(2\pi kt)$, which winds k times around the unit circle as t runs over $[0, 1]$, at the n sample points $t = j/n$. Its real part samples a cosine wave with k full periods (Figure 6.3.2). The coordinate of a vector $\mathbf{x} \in \mathbb{C}^n$ along $\mathbf{f}_k$ measures how much of the wave of frequency k the data $\mathbf{x}$ contains. Since $\omega^{j(n-k)} = \omega^{-jk}$, we have $\mathbf{f}_{n-k} = \overline{\mathbf{f}_k}$: at only n sample points, frequency $n-k$ looks exactly like frequency $-k$.

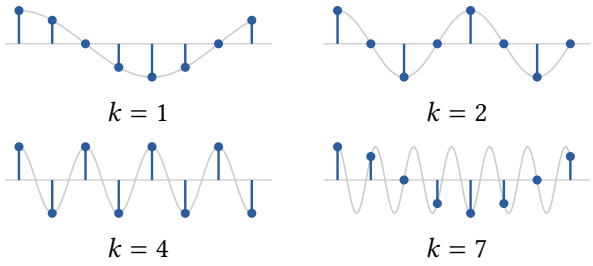


Figure 6.3.2. Real parts of four columns $\mathbf{f}_k$ of Φ_8: the samples $\cos(2\pi kj/8)$ at $j = 0, \ldots, 7$, joined by the curve $s \mapsto \cos(2\pi ks/8)$. Frequency 4 is the fastest oscillation that 8 samples can show. The samples for $k = 7$ coincide with those for $k = 1$, because $\mathbf{f}_7 = \overline{\mathbf{f}_1}$ has the same real part as $\mathbf{f}_1$.

Proposition 6.3.41 (The Fourier basis). The columns $\mathbf{f}_0, \ldots, \mathbf{f}_{n-1}$ of Φ_n are orthogonal, and $\|\mathbf{f}_k\|^2 = n$ for every k. Consequently:

(a) $\Phi_n^*\Phi_n = nI_n$, and $\Phi_n/\sqrt{n}$ is unitary;

(b) Φ_n is invertible, with $\Phi_n^{-1} = \frac{1}{n}\Phi_n^* = \frac{1}{n}\overline{\Phi_n}$;

(c) every $\mathbf{x} \in \mathbb{C}^n$ satisfies

$$\mathbf{x} = \frac{1}{n}\sum_{k=0}^{n-1}\langle \mathbf{x}, \mathbf{f}_k\rangle \mathbf{f}_k, \quad \text{where} \quad \langle \mathbf{x}, \mathbf{f}_k\rangle = \sum_{j=0}^{n-1} x_j \omega^{-jk};$$

(d) every $\mathbf{x} \in \mathbb{C}^n$ satisfies $\sum_{j=0}^{n-1} |x_j|^2 = \frac{1}{n}\sum_{k=0}^{n-1} |\langle \mathbf{x}, \mathbf{f}_k\rangle|^2$.

Proof. STANDARD

Uses. Lemma 6.3.39 (Sums of roots of unity), Theorem 6.3.29 (Orthonormal columns), Proposition 6.3.34 (Characterizations of unitary matrices), Corollary 6.3.21 (Coordinates in an orthogonal basis), Theorem 6.3.22 (Parseval's identity), Corollary 6.3.14 (Counting orthogonal vectors)
Techniques. Gram matrices (Technique 6.3.31)
Hint 1. Compute $\mathbf{f}_k \cdot \mathbf{f}_l$ using $\overline{\omega^{jl}} = \omega^{-jl}$. You get a sum of the kind in Lemma 6.3.39.
Hint 2. The sum is $\sum_j \omega^{j(k-l)}$. For $0 \le k, l \le n-1$, when does n divide $k-l$?
Hint 3. For (a), use Theorem 6.3.29. For (c) and (d), $\mathbf{f}_0, \ldots, \mathbf{f}_{n-1}$ is an orthogonal basis; apply Corollary 6.3.21 and Theorem 6.3.22 to $\mathbf{f}_k/\sqrt{n}$.
Sketch. $\mathbf{f}_k \cdot \mathbf{f}_l = \sum_j \omega^{jk}\omega^{-jl} = \sum_j \omega^{j(k-l)}$, which is n if $k = l$ and 0 otherwise by Lemma 6.3.39, as $|k-l| < n$. So $\Phi_n^*\Phi_n = nI_n$ by Theorem 6.3.29, $\Phi_n/\sqrt{n}$ is unitary, and its inverse is its conjugate transpose. The vectors $\mathbf{f}_k$ form an orthogonal basis with $\|\mathbf{f}_k\|^2 = n$, which gives (c) by Corollary 6.3.21, and (d) is Parseval for the orthonormal basis $\mathbf{f}_k/\sqrt{n}$.

For $0 \le k, l \le n-1$, since $\overline{\omega^{jl}} = \omega^{-jl}$,

$$\mathbf{f}_k \cdot \mathbf{f}_l = \sum_{j=0}^{n-1} \omega^{jk}\overline{\omega^{jl}} = \sum_{j=0}^{n-1} \omega^{j(k-l)}.$$

Since $|k-l| \le n-1$, $n \mid (k-l)$ only when $k = l$. By Lemma 6.3.39, $\mathbf{f}_k \cdot \mathbf{f}_l = n$ if $k = l$ and 0 otherwise. So the columns are orthogonal with $\|\mathbf{f}_k\|^2 = n$.

(a) By Theorem 6.3.29, $(\Phi_n^*\Phi_n)_{kl} = \mathbf{f}_l \cdot \mathbf{f}_k = n\delta_{kl}$, so $\Phi_n^*\Phi_n = nI_n$. Hence $\left(\Phi_n/\sqrt{n}\right)^*\left(\Phi_n/\sqrt{n}\right) = (1/n)\Phi_n^*\Phi_n = I_n$ (Proposition 6.3.27(a)), and $\Phi_n/\sqrt{n}$ is unitary.

(b) By Proposition 6.3.34, $\Phi_n/\sqrt{n}$ is invertible with inverse $\left(\Phi_n/\sqrt{n}\right)^* = \Phi_n^*/\sqrt{n}$. So Φ_n is invertible, and $\Phi_n^{-1} = (1/n)\Phi_n^*$. Since Φ_n is symmetric, $\Phi_n^* = \overline{\Phi_n}^{\mathsf{T}} = \overline{\Phi_n}$.

(c) The vectors $\mathbf{f}_0, \ldots, \mathbf{f}_{n-1}$ are n orthogonal nonzero vectors in $\mathbb{C}^n$, hence an orthogonal basis (Corollary 6.3.14). By Corollary 6.3.21, $\mathbf{x} = \sum_k \langle \mathbf{x}, \mathbf{f}_k\rangle / \|\mathbf{f}_k\|^2 \mathbf{f}_k$, and $\|\mathbf{f}_k\|^2 = n$. Finally $\langle \mathbf{x}, \mathbf{f}_k\rangle = \mathbf{x} \cdot \mathbf{f}_k = \sum_j x_j \overline{\omega^{jk}} = \sum_j x_j \omega^{-jk}$.

(d) The vectors $\mathbf{f}_k/\sqrt{n}$ form an orthonormal basis of $\mathbb{C}^n$, so by Theorem 6.3.22(b),

$$\|\mathbf{x}\|^2 = \sum_{k=0}^{n-1} \left|\left\langle \mathbf{x}, \frac{\mathbf{f}_k}{\sqrt{n}}\right\rangle\right|^2 = \frac{1}{n}\sum_{k=0}^{n-1} |\langle \mathbf{x}, \mathbf{f}_k\rangle|^2,$$

using that $1/\sqrt{n}$ is real. □

The vector $\hat{\mathbf{x}} = \Phi_n^*\mathbf{x} = (\langle \mathbf{x}, \mathbf{f}_0\rangle, \ldots, \langle \mathbf{x}, \mathbf{f}_{n-1}\rangle)$ is called the **discrete Fourier transform** of $\mathbf{x}$. (Conventions differ in where the factor $1/n$ goes and in the sign of the exponent.) Part (c) is the inversion formula $\mathbf{x} = (1/n)\Phi_n\hat{\mathbf{x}}$, and (d) is Parseval's identity for the transform. Both are instances of what we proved for general orthogonal bases; the only computation specific to Φ_n is Lemma 6.3.39.

Example 6.3.42 (A transform with $n = 4$). Let $\mathbf{x} = (1, 2, 3, 4)$. With $\omega = i$, $\langle \mathbf{x}, \mathbf{f}_k\rangle = \sum_j x_j i^{-jk}$:

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{f}_0\rangle &= 1 + 2 + 3 + 4 = 10,\\
\langle \mathbf{x}, \mathbf{f}_1\rangle &= 1 - 2i - 3 + 4i = -2 + 2i,\\
\langle \mathbf{x}, \mathbf{f}_2\rangle &= 1 - 2 + 3 - 4 = -2,\\
\langle \mathbf{x}, \mathbf{f}_3\rangle &= 1 + 2i - 3 - 4i = -2 - 2i.
\end{aligned}$$

By Proposition 6.3.41(c),

$$\mathbf{x} = \frac{1}{4}(10\mathbf{f}_0 + (-2+2i)\mathbf{f}_1 - 2\mathbf{f}_2 + (-2-2i)\mathbf{f}_3).$$

The first term is the constant vector $(5/2)\mathbf{f}_0$ whose entries are the mean of the entries of $\mathbf{x}$, and the other three terms add up to the deviations $(-3/2, -1/2, 1/2, 3/2)$ from the mean. Since $\mathbf{x}$ is real, $\langle \mathbf{x}, \mathbf{f}_3\rangle = \overline{\langle \mathbf{x}, \mathbf{f}_1\rangle} = \overline{\langle \mathbf{x}, \mathbf{f}_1\rangle}$. As a check of (d), $1 + 4 + 9 + 16 = 30$ and $(100 + 8 + 4 + 8)/4 = 30$.

The invertibility of Φ_n can also be proved without orthogonality, and comparing the two proofs shows what orthogonality buys.

Corollary 6.3.43 (Φ_n is invertible). The Fourier matrix Φ_n is invertible.

Proof (by orthogonality). ROUTINE

Uses. Proposition 6.3.41 (The Fourier basis)
Hint 1. Use Proposition 6.3.41(b).
Hint 2. It even gives the inverse.
Sketch. $(1/n)\overline{\Phi_n}$ is a two-sided inverse, by Proposition 6.3.41(b).

This is Proposition 6.3.41(b): $\Phi_n^{-1} = (1/n)\overline{\Phi_n}$. □

Proof (by the Vandermonde determinant). STANDARD

> *Uses.* Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.7 (Invertibility and the determinant), $\omega^m = 1$ if and only if $n \mid m$
> *Hint 1.* Row j of Φ_n is $\left(1, z_j, z_j^2, \ldots, z_j^{n-1}\right)$ with $z_j = \omega^j$. What kind of matrix is that?
> *Hint 2.* Φ_n is a Vandermonde matrix. Its determinant is nonzero when the nodes $z_0, \ldots, z_{n-1}$ are distinct; check that they are.
> *Sketch.* Φ_n is the Vandermonde matrix of the nodes $1, \omega, \ldots, \omega^{n-1}$ (it is symmetric, so the orientation does not matter). These are distinct, since $\omega^j = \omega^{j'}$ with $0 \le j < j' \le n-1$ would give $\omega^{j'-j} = 1$ with $0 < j' - j < n$. So $\det \Phi_n \neq 0$, and Φ_n is invertible.

Let $z_j = \omega^j$ for $j = 0, \ldots, n-1$. The entry of Φ_n in row j and column k is $\omega^{jk} = z_j^k$, so Φ_n is the Vandermonde matrix of the nodes $z_0, \ldots, z_{n-1}$; since Φ_n is symmetric, this holds whichever of rows or columns carries the powers in Theorem 5.5.21. The nodes are distinct: if $z_j = z_{j'}$ with $0 \le j < j' \le n-1$, then $\omega^{j'-j} = 1$ with $0 < j' - j < n$, which is impossible since $n \nmid (j' - j)$. By Theorem 5.5.21, $\det \Phi_n = \prod_{0 \le j < j' \le n-1} \left(z_{j'} - z_j\right) \neq 0$, and by Theorem 5.4.7, Φ_n is invertible. □

The Vandermonde proof uses only that the nodes ω^j are distinct, so it applies to any distinct nodes, but it says nothing about the inverse. The orthogonality proof uses the special position of the nodes on the unit circle, and in return it produces the inverse $(1/n)\overline{\Phi_n}$, the inversion formula, and Parseval's identity. It also shows that $|\det \Phi_n| = n^{n/2}$ (Exercise 6.3.60), which the product formula for the Vandermonde determinant hides.

Exercises

Exercise 6.3.44. ROUTINE In $\mathbb{C}^2$ with the dot product, let $\mathbf{u}_1 = \frac{1}{\sqrt{3}}(1, 1+i)$ and $\mathbf{u}_2 = \frac{1}{\sqrt{3}}(-1+i, 1)$.

(a) Show that $\mathcal{O} = \mathbf{u}_1, \mathbf{u}_2$ is an orthonormal basis of $\mathbb{C}^2$.

(b) Find $[\mathbf{v}]_{\mathcal{O}}$ for $\mathbf{v} = (2, 3i)$, and check your answer with Theorem 6.3.22(b).

(c) Compute $\langle \mathbf{u}_1, \mathbf{v} \rangle \mathbf{u}_1 + \langle \mathbf{u}_2, \mathbf{v} \rangle \mathbf{u}_2$, with the slots exchanged, and compare it with $\mathbf{v}$.

Solution on page 1485

Exercise 6.3.45. ROUTINE Find all unit vectors $\mathbf{x} \in \mathbb{R}^3$ that are orthogonal to both $(1, 2, 2)$ and $(2, 1, -2)$. Then find all 3×3 orthogonal matrices whose first two columns are $\frac{1}{3}(1, 2, 2)$ and $\frac{1}{3}(2, 1, -2)$.

Solution on page 1485

Exercise 6.3.46. STANDARD Let V be an inner product space. Prove or disprove each statement.

(a) If $\mathbf{u} \perp \mathbf{v}$ and $\mathbf{v} \perp \mathbf{w}$, then $\mathbf{u} \perp \mathbf{w}$.

(b) If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are nonzero with $\mathbf{v}_1 \perp \mathbf{v}_2$ and $\mathbf{v}_2 \perp \mathbf{v}_3$, then $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is linearly independent.

(c) If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is orthogonal and $c_1, \ldots, c_k \in \mathbb{F}$, then $c_1\mathbf{v}_1, \ldots, c_k\mathbf{v}_k$ is orthogonal.

(d) If $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ and $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u} \perp \mathbf{v}$.

(e) If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ spans V and $\mathbf{x} \perp \mathbf{v}_j$ for every j, then $\mathbf{x} = \mathbf{0}$.

Solution on page 1486

Exercise 6.3.47. STANDARD On $\mathbb{F}_2^4$, let $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4$. This formula is linear in each slot and symmetric, but it is not an inner product, since the scalars are not real or complex. Call $\mathbf{x}$ and $\mathbf{y}$ orthogonal if $\mathbf{x} \cdot \mathbf{y} = 0$.

(a) Find a nonzero $\mathbf{x} \in \mathbb{F}_2^4$ with $\mathbf{x} \cdot \mathbf{x} = 0$.

(b) Find three nonzero vectors in $\mathbb{F}_2^4$ that are pairwise orthogonal but linearly dependent. Which step of the first proof of Proposition 6.3.13 fails?

(c) Let $W = \operatorname{span}((1,1,0,0),(0,0,1,1))$. Show that $\left\{\mathbf{x} \in \mathbb{F}_2^4 \;\middle|\; \mathbf{x} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\right\} = W$. Which part of Proposition 6.3.9 fails?

Solution on page 1486

Exercise 6.3.48. STANDARD Give $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$ the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(B^* A)$ (Proposition 6.1.16).

(a) Show that every symmetric $S \in \mathbb{R}^{n\times n}$ is orthogonal to every skew-symmetric $K \in \mathbb{R}^{n\times n}$, and deduce that $\|A\|^2 = \left\|\left(A + A^{\mathsf{T}}\right)/2\right\|^2 + \left\|\left(A - A^{\mathsf{T}}\right)/2\right\|^2$ for every $A \in \mathbb{R}^{n\times n}$.

(b) Prove or disprove: every Hermitian $H \in \mathbb{C}^{n\times n}$ is orthogonal to every $K \in \mathbb{C}^{n\times n}$ with $K^* = -K$. Does the identity $\|A\|^2 = \|(A + A^*)/2\|^2 + \|(A - A^*)/2\|^2$ hold for every $A \in \mathbb{C}^{n\times n}$?

Solution on page 1487

Exercise 6.3.49. STANDARD Give $\mathbb{R}[x]_{\leq 3}$ the inner product $\langle p, q\rangle = \int_{-1}^{1} p(t)q(t)\,\mathrm{d}t$, as in Example 6.3.17.

(a) Find the polynomial $q = x^3 + ax^2 + bx + c$ that is orthogonal to 1, x, and x^2, and compute $\|q\|^2$.

(b) Show that $1, x, x^2 - 1/3, q$ is an orthogonal basis of $\mathbb{R}[x]_{\leq 3}$.

(c) Use Corollary 6.3.21 to write $p = 1 + x + x^2 + x^3$ in this basis, and check the result.

Solution on page 1487

Exercise 6.3.50. STANDARD The *Pauli matrices* are

$$\sigma_0 = I_2, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Give $\mathbb{C}^{2\times 2}$ the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(B^* A)$.

(a) Show that $\sigma_0/\sqrt{2}, \ldots, \sigma_3/\sqrt{2}$ is an orthonormal basis of $\mathbb{C}^{2\times 2}$.

(b) Deduce that every $A \in \mathbb{C}^{2\times 2}$ satisfies $A = \frac{1}{2}\sum_{a=0}^{3} \operatorname{tr}(\sigma_a A)\sigma_a$.

(c) Write $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ in this form.

Solution on page 1488

Exercise 6.3.51. STANDARD Let V be a vector space over $\mathbb{F}$ with a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$.

(a) Show that $\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}$ is an inner product on V for which $\mathcal{B}$ is orthonormal.

(b) Show that it is the only inner product on V for which $\mathcal{B}$ is orthonormal.

(c) For $V = \mathbb{R}^2$ and $\mathcal{B} = (1,0),(1,1)$, find the matrix M with $\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{B}} = \mathbf{y}^{\mathsf{T}} M \mathbf{x}$.

Solution on page 1489

Exercise 6.3.52. STANDARD

(a) In $\mathbb{R}^2$, let $\mathbf{w}_j = \sqrt{2/3}(\cos(2\pi j/3), \sin(2\pi j/3))$ for $j = 0, 1, 2$. Show that $\sum_{j=0}^{2} \langle \mathbf{v}, \mathbf{w}_j \rangle^2 = \|\mathbf{v}\|^2$ for every $\mathbf{v} \in \mathbb{R}^2$, although $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2$ is not orthonormal.

(b) Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be *unit* vectors in an inner product space V such that $\sum_{j=1}^{k} |\langle \mathbf{v}, \mathbf{v}_j \rangle|^2 = \|\mathbf{v}\|^2$ for every $\mathbf{v} \in V$. Show that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal basis of V.

Solution on page 1489

Exercise 6.3.53. STANDARD Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal list in V, and let $\mathbf{x}, \mathbf{y} \in V$.

(a) Show that $\sum_{j=1}^{k} |\langle \mathbf{x}, \mathbf{u}_j \rangle| |\langle \mathbf{y}, \mathbf{u}_j \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$.

(b) Show that $\sum_{j=1}^{k} |\langle \mathbf{x}, \mathbf{u}_j \rangle| \leq \sqrt{k} \|\mathbf{x}\|$, and find when equality holds.

Solution on page 1490

Exercise 6.3.54. STANDARD The $n \times n$ *Hilbert matrix* H_n has entries $h_{ij} = 1/(i + j - 1)$.

(a) Show that H_n is the Gram matrix of $1, x, \ldots, x^{n-1}$ in $\mathbb{R}[x]_{\leq n-1}$ with $\langle p, q \rangle = \int_0^1 p(t)q(t)\,dt$.

(b) Deduce that H_n is invertible and that $\mathbf{c}^{\mathsf{T}} H_n \mathbf{c} > 0$ for every nonzero $\mathbf{c} \in \mathbb{R}^n$.

Solution on page 1490

Exercise 6.3.55. STANDARD Find the error in the following argument, and decide whether the claim is true.

Claim. If the columns of $A \in \mathbb{R}^{n \times n}$ are orthogonal and nonzero, then the rows of A are orthogonal.

Argument. By Theorem 6.3.29, $A^{\mathsf{T}}A = D$ is diagonal with positive diagonal entries, so D is invertible. The columns are independent (Proposition 6.3.13), so A is invertible, and $A^{\mathsf{T}} = DA^{-1}$. Hence $AA^{\mathsf{T}} = ADA^{-1}$. A matrix similar to a diagonal matrix is diagonal, so AA^{T} is diagonal, which says that the rows of A are orthogonal.

Solution on page 1491

Exercise 6.3.56. STANDARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a list in V with Gram matrix G, let $S = (s_{ij}) \in \mathbb{F}^{n \times n}$, and let $\mathbf{w}_j = \sum_{i=1}^{n} s_{ij}\mathbf{v}_i$ for $j = 1, \ldots, n$.

(a) Show that the Gram matrix of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is S^*GS.

(b) Suppose $\mathcal{O} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal basis of V. Show that $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is an orthonormal basis of V if and only if S is unitary. Conclude that the change-of-coordinates matrix between two orthonormal bases (Section 4.5) is unitary.

Solution on page 1491

Exercise 6.3.57. STANDARD

(a) Show that $\det(A^*) = \overline{\det A}$ for every $A \in \mathbb{F}^{n \times n}$. Deduce that $|\det U| = 1$ for every unitary U, and that $\det Q = \pm 1$ for every real orthogonal Q.

(b) Show by example that a real matrix with determinant 1 need not be orthogonal.

(c) Show that an upper triangular unitary matrix is diagonal.

Solution on page 1492

Exercise 6.3.58. STANDARD Show that $U \in \mathbb{C}^{2 \times 2}$ is unitary if and only if

$$U = \begin{pmatrix} a & -\lambda\bar{c} \\ c & \lambda\bar{a} \end{pmatrix}$$

for some $a, c, \lambda \in \mathbb{C}$ with $|a|^2 + |c|^2 = 1$ and $|\lambda| = 1$, and that then $\det U = \lambda$.

Solution on page 1492

Exercise 6.3.59. ROUTINE Let $n = 3$, so that $\omega = -1/2 + \left(\sqrt{3}/2\right)i$.

(a) Write down Φ_3 and Φ_3^{-1}.

(b) Write $\mathbf{x} = (1, 2, 0)$ as a linear combination of $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$, and check Proposition 6.3.41(d).

Solution on page 1493

Exercise 6.3.60. STANDARD Index rows and columns by $0, \ldots, n-1$, as for Φ_n.

(a) Show that $\Phi_n^2 = nP$, where P has entries $p_{jl} = 1$ if $n \mid (j + l)$ and $p_{jl} = 0$ otherwise. Show that P is a permutation matrix with $P^2 = I_n$, and deduce that $\Phi_n^4 = n^2 I_n$.

(b) Show that $|\det \Phi_n| = n^{n/2}$.

Solution on page 1493

Exercise 6.3.61. HARD For $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{C}^n$, the *circulant matrix* C has entries $C_{jl} = c_{(j-l) \bmod n}$, with indices $0, \ldots, n-1$. For $n = 3$,

$$C = \begin{pmatrix} c_0 & c_2 & c_1 \\ c_1 & c_0 & c_2 \\ c_2 & c_1 & c_0 \end{pmatrix}.$$

(a) Show that $C\mathbf{f}_k = \langle \mathbf{c}, \mathbf{f}_k \rangle \mathbf{f}_k$ for $k = 0, \ldots, n-1$.

(b) Deduce that $C = \Phi_n D \Phi_n^{-1}$, where D is the diagonal matrix with entries $\langle \mathbf{c}, \mathbf{f}_0 \rangle, \ldots, \langle \mathbf{c}, \mathbf{f}_{n-1} \rangle$. Conclude that any two $n \times n$ circulant matrices commute, and that C is invertible if and only if $\langle \mathbf{c}, \mathbf{f}_k \rangle \neq 0$ for every k.

(c) Using (b), decide whether the following matrix is invertible, and find its determinant:

$$\begin{pmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Solution on page 1494

Exercise 6.3.62. HARD Let $n \geq 1$, let ω be as in Section 6.3.4, and let $r \in \{0, \ldots, n-1\}$.

(a) Let $p(x) = a_0 + a_1 x + \cdots + a_d x^d \in \mathbb{C}[x]$. Show that

$$\sum_{0 \leq k \leq d,\, n \mid (k-r)} a_k = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{-jr} p\left(\omega^j\right).$$

(b) Compute $\binom{12}{0} + \binom{12}{3} + \binom{12}{6} + \binom{12}{9} + \binom{12}{12}$ without adding these numbers.

(c) Show that for every $m \geq 0$, $\sum_{k \equiv 0 \pmod 3} \binom{m}{k} = (2^m + 2\cos(m\pi/3))/3$.

Solution on page 1494

Exercise 6.3.63. HARD A *Hadamard matrix* of order n is a matrix $H \in \mathbb{R}^{n\times n}$ whose entries are all ± 1 and whose columns are orthogonal; equivalently, $H^{\mathsf{T}}H = nI_n$.

(a) Show that if a Hadamard matrix of order $n \geq 3$ exists, then $4 \mid n$.

(b) Show that if H is a Hadamard matrix of order n, then $\begin{pmatrix} H & H \\ H & -H \end{pmatrix}$ is a Hadamard matrix of order $2n$. Deduce that Hadamard matrices of order 2^k exist for every $k \geq 0$.

Solution on page 1495

Exercise 6.3.64. CHALLENGE Let n be odd, let ω be as in Section 6.3.4, and let $g = \sum_{j=0}^{n-1} \omega^{j^2}$.

(a) Show that $|g|^2 = n$.

(b) Compute $|g|^2$ for $n = 4$ and for $n = 6$ (with the same definition of g), and explain where the proof of (a) used that n is odd.

Solution on page 1496

6.4 Orthogonal projection and Gram–Schmidt

Let W be a subspace of an inner product space V, and let $\mathbf{v} \in V$. Which vector of W is closest to $\mathbf{v}$? When W is a plane through $\mathbf{0}$ in $\mathbb{R}^3$, geometry gives the answer: drop the perpendicular from $\mathbf{v}$ to W, and take its foot. This section turns that answer into a theorem about every inner product space, provided W is finite-dimensional. The foot of the perpendicular is the *orthogonal projection* $\mathrm{Pr}_W(\mathbf{v})$. If W has an orthonormal basis, it is given by the formula from the proof of Bessel's inequality (Theorem 6.3.24), and the Gram–Schmidt algorithm produces an orthonormal basis of every finite-dimensional subspace. The leftover vectors $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ make up the orthogonal complement $W^{\perp}$, which turns out to be a complement of W in the sense of Section 4.4. Since distances also make sense between functions, the same theorem answers questions such as: which quadratic polynomial is closest to e^t on $[-1, 1]$?

Throughout, V is an inner product space over $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ (Notation 6.1.1). We do not assume that V is finite-dimensional. The subspaces we project onto will be finite-dimensional, but the ambient space is often a space of functions.

6.4.1 Projecting onto the span of an orthonormal list

Intuition (From a line to a plane). In Section 6.2.1 we split a vector $\mathbf{x}$ against a nonzero vector $\mathbf{y}$ as $\mathbf{x} = c\mathbf{y} + \mathbf{w}$, where $c\mathbf{y}$ lies on the line $\mathrm{span}(\mathbf{y})$ and $\mathbf{w} \perp \mathbf{y}$ (Figure 6.2.1). The vector $c\mathbf{y}$ is the foot of the perpendicular from $\mathbf{x}$ to the line, and the requirement $\mathbf{w} \perp \mathbf{y}$ forced $c = \langle \mathbf{x}, \mathbf{y} \rangle / \|\mathbf{y}\|^2$.

Now replace the line by a plane W in $\mathbb{R}^3$ spanned by orthonormal vectors $\mathbf{u}_1, \mathbf{u}_2$, and look for the foot $\mathbf{p}$ of the perpendicular from $\mathbf{v}$ to W. That is, we want $\mathbf{p} \in W$ such that $\mathbf{v} - \mathbf{p}$ is perpendicular to W. Write $\mathbf{p} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$. By Proposition 6.3.9(c), $\mathbf{v} - \mathbf{p} \perp W$ means that $\mathbf{v} - \mathbf{p} \perp \mathbf{u}_j$ for $j = 1, 2$, and taking the inner product with $\mathbf{u}_j$ isolates the coefficient a_j (Technique 6.3.12):

$$0 = \langle \mathbf{v} - \mathbf{p}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - a_j.$$

So $\mathbf{p} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2$. Each term is the shadow of $\mathbf{v}$ on one of the lines $\mathrm{span}(\mathbf{u}_1)$ and $\mathrm{span}(\mathbf{u}_2)$. So the foot of the perpendicular to the plane is the sum of the feet of the perpendiculars to the two lines (Figure 6.4.1).

The shadows can be computed separately and then added because $\mathbf{u}_1 \perp \mathbf{u}_2$. For a basis of W that is not orthogonal, the sum of the shadows is not the foot; see the example after Corollary 6.3.21.

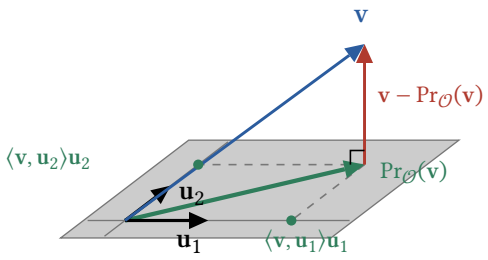


Figure 6.4.1. Projecting $\mathbf{v}$ onto the shaded plane $W = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$, for orthonormal $\mathbf{u}_1, \mathbf{u}_2$. The foot $\operatorname{Pr}_{\mathcal{O}}(\mathbf{v})$ of the perpendicular is the sum of the shadows $\langle \mathbf{v}, \mathbf{u}_1\rangle \mathbf{u}_1$ and $\langle \mathbf{v}, \mathbf{u}_2\rangle \mathbf{u}_2$ of $\mathbf{v}$ on the two axes, and $\mathbf{v} - \operatorname{Pr}_{\mathcal{O}}(\mathbf{v})$ is perpendicular to W.

Definition 6.4.1 (Projection onto the span of an orthonormal list). Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal list in V. The **projection onto** $\operatorname{span}(\mathcal{O})$ is the function $\operatorname{Pr}_{\mathcal{O}} : V \to V$ given by

$$\operatorname{Pr}_{\mathcal{O}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}_1\rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k\rangle \mathbf{u}_k.$$

If $k = 0$, the sum is empty and $\operatorname{Pr}_{\mathcal{O}}(\mathbf{v}) = \mathbf{0}$.

This is the vector $\mathbf{p}$ from the proof of Bessel's inequality (Theorem 6.3.24). As with coordinates (Remark 6.3.19), $\mathbf{v}$ goes in the first slot. For a single unit vector $\mathbf{u}$, $\operatorname{Pr}_{\mathbf{u}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}\rangle \mathbf{u}$ is the shadow of Technique 6.2.4, since there $c = \langle \mathbf{v}, \mathbf{u}\rangle / \|\mathbf{u}\|^2 = \langle \mathbf{v}, \mathbf{u}\rangle$.

Example 6.4.2 (A projection in $\mathbb{R}^3$). Take the orthonormal list $\mathcal{O} = \mathbf{u}_1, \mathbf{u}_2$ with $\mathbf{u}_1 = \frac{1}{3}(1, 2, 2)$ and $\mathbf{u}_2 = \frac{1}{3}(2, 1, -2)$ from Example 6.3.16(c), and $\mathbf{v} = (1, 2, 3)$. In Example 6.3.20 we computed $\langle \mathbf{v}, \mathbf{u}_1\rangle = 11/3$ and $\langle \mathbf{v}, \mathbf{u}_2\rangle = -2/3$. So

$$\operatorname{Pr}_{\mathcal{O}}(\mathbf{v}) = \frac{11}{9}(1, 2, 2) - \frac{2}{9}(2, 1, -2) = \frac{1}{9}(7, 20, 26),$$

and $\mathbf{v} - \operatorname{Pr}_{\mathcal{O}}(\mathbf{v}) = \frac{1}{9}(2, -2, 1)$. This vector is orthogonal to $(1, 2, 2)$ and to $(2, 1, -2)$, as the intuition above predicts. Its squared length is $9/81 = 1/9$, the deficit in Bessel's inequality found in Example 6.3.25.

Example 6.4.3 (Projecting a function). Give $C[-1, 1]$ the inner product $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt$ (Proposition 6.1.13). By Example 6.3.17, $\mathcal{O} = 1/\sqrt{2}, \sqrt{3/2}\,t$ is orthonormal. Since $\int_{-1}^{1} t^2\,dt = 2/3$ and $\int_{-1}^{1} t^3\,dt = 0$,

$$\begin{aligned}\operatorname{Pr}_{\mathcal{O}}(t^2) &= \left\langle t^2, \frac{1}{\sqrt{2}}\right\rangle \frac{1}{\sqrt{2}} + \left\langle t^2, \sqrt{\frac{3}{2}}\,t\right\rangle \sqrt{\frac{3}{2}}\,t \\ &= \frac{1}{2}\cdot\frac{2}{3} + 0 = \frac{1}{3}.\end{aligned}$$

So $t^2 - 1/3$ is orthogonal to 1 and to t. The polynomial $x^2 - 1/3$ of Example 6.3.17 arises in exactly this way, and Section 6.4.2 turns the computation into an algorithm.

The next proposition collects what we know about $\operatorname{Pr}_{\mathcal{O}}$. Most of it was proved along the way to Bessel's inequality.

Proposition 6.4.4 (Properties of $\operatorname{Pr}_{\mathcal{O}}$). Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal list in V, let $W = \operatorname{span}(\mathcal{O})$, and let $\mathbf{v} \in V$.

(a) $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) \in W$.

(b) $\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) \in W^{\perp}$.

(c) $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \mathbf{v}$ if and only if $\mathbf{v} \in W$.

(d) $\mathrm{Pr}_{\mathcal{O}}$ is linear.

(e) $\|\mathbf{v}\|^2 = \|\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\|^2 + \|\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\|^2$, and $\|\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\|^2 = \sum_{j=1}^{k} \left|\langle \mathbf{v}, \mathbf{u}_j \rangle\right|^2$. In particular $\|\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\| \le \|\mathbf{v}\|$, with equality if and only if $\mathbf{v} \in W$.

Proof. ROUTINE

Uses. Definition 6.4.1 (Projection onto the span of an orthonormal list), Definition 6.3.15 (Orthonormal list, orthonormal basis), Proposition 6.3.9 (Properties of $A^{\perp}$), Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 6.3.3 (Pythagorean theorem), Corollary 6.3.6 (Pythagoras for orthogonal lists), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. For (b), compute $\langle \mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}), \mathbf{u}_l \rangle$ for each l; then pass from the $\mathbf{u}_l$ to all of W.
Hint 2. For (c), one direction is (a), and the other is Proposition 6.3.18(a). For (d), each coefficient $\langle \mathbf{v}, \mathbf{u}_j \rangle$ depends linearly on $\mathbf{v}$.
Hint 3. For (e), apply the Pythagorean theorem to $\mathbf{v} = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) + (\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}))$, using (a) and (b).
Sketch. (a) $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ is a combination of the $\mathbf{u}_j$. (b) $\langle \mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}), \mathbf{u}_l \rangle = \langle \mathbf{v}, \mathbf{u}_l \rangle - \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \delta_{jl} = 0$, and Proposition 6.3.9(c) upgrades orthogonality to each $\mathbf{u}_l$ to orthogonality to W. (c) follows from (a) and Proposition 6.3.18(a). (d) follows from linearity of the inner product in the first slot. (e) is Pythagoras for the orthogonal pieces in (a) and (b), plus Corollary 6.3.6; equality means $\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \mathbf{0}$, which is (c).

(a) $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

(b) For each l, linearity in the first slot and orthonormality give

$$\begin{aligned} \langle \mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}), \mathbf{u}_l \rangle &= \langle \mathbf{v}, \mathbf{u}_l \rangle - \sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{u}_j \rangle \langle \mathbf{u}_j, \mathbf{u}_l \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_l \rangle - \langle \mathbf{v}, \mathbf{u}_l \rangle = 0. \end{aligned}$$

So $\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) \in \mathcal{O}^{\perp}$, and $\mathcal{O}^{\perp} = \mathrm{span}(\mathcal{O})^{\perp} = W^{\perp}$ by Proposition 6.3.9(c).

(c) If $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \mathbf{v}$, then $\mathbf{v} \in W$ by (a). Conversely, if $\mathbf{v} \in W$, then $\mathbf{v} = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ by Proposition 6.3.18(a).

(d) For $\mathbf{v}, \mathbf{v}' \in V$ and $a, b \in \mathbb{F}$, linearity in the first slot gives $\langle a\mathbf{v} + b\mathbf{v}', \mathbf{u}_j \rangle = a\langle \mathbf{v}, \mathbf{u}_j \rangle + b\langle \mathbf{v}', \mathbf{u}_j \rangle$ for each j. Multiplying by $\mathbf{u}_j$ and summing over j gives $\mathrm{Pr}_{\mathcal{O}}(a\mathbf{v} + b\mathbf{v}') = a\,\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) + b\,\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}')$.

(e) By (a) and (b), $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) \perp \mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$, and these two vectors add up to $\mathbf{v}$. So the first identity is Theorem 6.3.3(a). The second is Corollary 6.3.6 applied to $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$. Since $\|\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\|^2 \ge 0$, the first identity gives $\|\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\|^2 \le \|\mathbf{v}\|^2$, with equality if and only if $\|\mathbf{v} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})\| = 0$. By Proposition 6.1.27(a), that happens if and only if $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \mathbf{v}$, that is, by (c), if and only if $\mathbf{v} \in W$. □

Part (e) is (6.3.2), and its inequality is Bessel's inequality. The move behind it deserves a name, since it is the main source of upper bounds in this chapter.

Technique 6.4.5 (Projections shrink norms). Suppose $\mathbf{x} = \mathbf{p} + \mathbf{q}$, where $\mathbf{p}$ lies in a subspace W and $\mathbf{q} \perp W$, as for $\mathbf{p} = \mathrm{Pr}_{\mathcal{O}}(\mathbf{x})$ and $\mathbf{q} = \mathbf{x} - \mathrm{Pr}_{\mathcal{O}}(\mathbf{x})$. By Pythagoras,

$$\|\mathbf{x}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{q}\|^2.$$

So each piece is at most as long as $\mathbf{x}$. Moreover $\|\mathbf{p}\| = \|\mathbf{x}\|$ exactly when $\mathbf{q} = \mathbf{0}$, that is, when $\mathbf{x} \in W$, and $\|\mathbf{q}\| = \|\mathbf{x}\|$ exactly when $\mathbf{p} = \mathbf{0}$.

Reach for it when you need an upper bound for a sum $\sum_j |\langle \mathbf{x}, \mathbf{u}_j \rangle|^2$ over an orthonormal list, which is $\|\mathrm{Pr}_{\mathcal{O}}(\mathbf{x})\|^2$, or more generally when the quantity to be bounded is the length of one orthogonal piece of a vector whose length you know. The equality case turns the bound into a test for membership in W. The canonical example is Bessel's inequality (Theorem 6.3.24). In Theorem 6.4.33 the error $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ turns out to be the piece of $\mathbf{v} - \mathbf{w}$ orthogonal to W, for every $\mathbf{w} \in W$; and in infinite dimensions (Section 15.3) Bessel's inequality survives where Parseval's identity needs more work.
Used in. Theorem 6.4.33, Problem 6.2, Exercise 6.4.60.

The projection is determined by orthogonality

Part (b) of Proposition 6.4.4 is the property that made $\mathbf{p}$ the foot of the perpendicular in the intuition above. It determines $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ completely. This matters because the condition "$\mathbf{p} \in W$ and $\mathbf{v} - \mathbf{p} \perp W$" mentions only the subspace W, not the orthonormal list $\mathcal{O}$ that the formula uses.

Theorem 6.4.6 (The projection is characterized by orthogonality). Let $\mathcal{O}$ be an orthonormal list in V, let $W = \mathrm{span}(\mathcal{O})$, and let $\mathbf{v} \in V$. Then $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ is the unique vector $\mathbf{p} \in V$ such that

$$\mathbf{p} \in W \quad \text{and} \quad \mathbf{v} - \mathbf{p} \in W^{\perp}.$$

Proof (by two candidates). ROUTINE

Uses. Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$), Proposition 6.3.9 (Properties of $A^{\perp}$)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Existence is (a) and (b) of Proposition 6.4.4. For uniqueness, suppose that $\mathbf{p}$ and $\mathbf{p}'$ both have the two properties, and look at $\mathbf{p} - \mathbf{p}'$.
Hint 2. $\mathbf{p} - \mathbf{p}'$ lies in W. It is also the difference of $\mathbf{v} - \mathbf{p}'$ and $\mathbf{v} - \mathbf{p}$, which lie in $W^{\perp}$. What is $W \cap W^{\perp}$?
Sketch. $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ has both properties by Proposition 6.4.4. If $\mathbf{p}$ and $\mathbf{p}'$ both have them, then $\mathbf{p} - \mathbf{p}' \in W$, and $\mathbf{p} - \mathbf{p}' = (\mathbf{v} - \mathbf{p}') - (\mathbf{v} - \mathbf{p}) \in W^{\perp}$ because $W^{\perp}$ is a subspace. So $\mathbf{p} - \mathbf{p}' \in W \cap W^{\perp} = \{\mathbf{0}\}$.

By (a) and (b) of Proposition 6.4.4, $\mathbf{p} = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ has both properties. Suppose $\mathbf{p}$ and $\mathbf{p}'$ both have them. Since W is a subspace, $\mathbf{p} - \mathbf{p}' \in W$. Since $W^{\perp}$ is a subspace (Proposition 6.3.9(a)),

$$\mathbf{p} - \mathbf{p}' = (\mathbf{v} - \mathbf{p}') - (\mathbf{v} - \mathbf{p}) \in W^{\perp}.$$

So $\mathbf{p} - \mathbf{p}' \in W \cap W^{\perp}$, which is $\{\mathbf{0}\}$ by Proposition 6.3.9(d). Hence $\mathbf{p} = \mathbf{p}'$. □

Proof (by computing coordinates). ROUTINE

Uses. Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$), Definition 6.4.1 (Projection onto the span of an orthonormal list)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Let $\mathbf{p} \in W$ with $\mathbf{v} - \mathbf{p} \perp W$. Since $\mathbf{p} \in \mathrm{span}(\mathcal{O})$, Proposition 6.3.18(a) expresses $\mathbf{p}$ through the numbers $\langle \mathbf{p}, \mathbf{u}_j \rangle$.
Hint 2. Write $\langle \mathbf{p}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v} - \mathbf{p}, \mathbf{u}_j \rangle$.
Sketch. Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$. If $\mathbf{p} \in W$ and $\mathbf{v} - \mathbf{p} \perp W$, then $\mathbf{p} = \sum_j \langle \mathbf{p}, \mathbf{u}_j \rangle \mathbf{u}_j$ and $\langle \mathbf{p}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - 0$. So $\mathbf{p} = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$. Existence is Proposition 6.4.4(a) and (b).

Write $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$. Existence is (a) and (b) of Proposition 6.4.4. Conversely, let $\mathbf{p} \in W$ with $\mathbf{v} - \mathbf{p} \in W^{\perp}$. Since $\mathbf{u}_j \in W$, we have $\langle \mathbf{v} - \mathbf{p}, \mathbf{u}_j \rangle = 0$, so

$$\langle \mathbf{p}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v} - \mathbf{p}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle.$$

Since $\mathbf{p} \in \mathrm{span}(\mathcal{O})$, Proposition 6.3.18(a) gives

$$\mathbf{p} = \sum_{j=1}^{k} \langle \mathbf{p}, \mathbf{u}_j \rangle \mathbf{u}_j = \sum_{j=1}^{k} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v}),$$

as desired. □

The first proof never uses the formula for $\mathrm{Pr}_{\mathcal{O}}$. It shows that any two vectors with the two properties agree, whatever produced them, and all it needs is $W \cap W^{\perp} = \{\mathbf{0}\}$. The second proof computes the coordinates of

such a vector and recognizes them as those of $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$; this is how the lecture argues. The first proof is the one that generalizes, and it gives the following technique.

Technique 6.4.7 (Recognize a projection by its two conditions). To show that a vector $\mathbf{p}$ is the projection of $\mathbf{v}$ onto W, do not compute the projection. Check the two conditions of Theorem 6.4.6: $\mathbf{p} \in W$, and $\mathbf{v} - \mathbf{p} \perp W$. By Proposition 6.3.9(c), the second condition only has to be checked against a spanning list of W. Uniqueness then does the rest, much as a two-sided inverse is recognized by checking two products (Technique 2.5.16).

Reach for it when a formula or a vector is proposed as a projection, when projections onto related subspaces are compared, and when you need to know that a projection does not depend on the choices made to compute it. The canonical example is Corollary 6.4.8. Later examples are the projection onto $W^{\perp}$ (Proposition 6.4.29), the recognition of orthogonal projections among all linear maps (Proposition 6.4.30), and the projection formula of Section 6.5.

Used in. Corollary 6.4.8, Proposition 6.4.29, Proposition 6.4.30, Exercise 6.4.49, Theorem 6.5.5, Proposition 6.5.13, Proposition 6.5.17, Proposition 6.5.23, Corollary 6.6.29, Exercise 6.7.5, Exercise 6.7.13, Problem 6.2, Problem 6.4, Exercise 6.4.28, Exercise 6.4.42, Exercise 6.4.57, Exercise 6.4.58, Exercise 6.5.42, Exercise 8.1.32.

Corollary 6.4.8 (The projection depends only on the span). If $\mathcal{O}$ and $\mathcal{O}'$ are orthonormal lists in V with $\mathrm{span}(\mathcal{O}) = \mathrm{span}(\mathcal{O}')$, then $\mathrm{Pr}_{\mathcal{O}} = \mathrm{Pr}_{\mathcal{O}'}$.

Proof. ROUTINE

Uses. Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Let W be the common span and fix $\mathbf{v} \in V$. Which two properties does $\mathrm{Pr}_{\mathcal{O}'}(\mathbf{v})$ have?
Hint 2. Both properties mention only W. Apply the uniqueness in Theorem 6.4.6 for the list $\mathcal{O}$.
Sketch. With $W = \mathrm{span}(\mathcal{O}) = \mathrm{span}(\mathcal{O}')$, the vector $\mathrm{Pr}_{\mathcal{O}'}(\mathbf{v})$ lies in W and $\mathbf{v} - \mathrm{Pr}_{\mathcal{O}'}(\mathbf{v}) \in W^{\perp}$, by Proposition 6.4.4 for $\mathcal{O}'$. These are the conditions that characterize $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ in Theorem 6.4.6.

Let $W = \mathrm{span}(\mathcal{O}) = \mathrm{span}(\mathcal{O}')$ and $\mathbf{v} \in V$. By (a) and (b) of Proposition 6.4.4, applied to the list $\mathcal{O}'$, the vector $\mathbf{p} = \mathrm{Pr}_{\mathcal{O}'}(\mathbf{v})$ lies in W and satisfies $\mathbf{v} - \mathbf{p} \in W^{\perp}$. By Theorem 6.4.6, applied to the list $\mathcal{O}$, the only vector with these two properties is $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$. Hence $\mathrm{Pr}_{\mathcal{O}'}(\mathbf{v}) = \mathrm{Pr}_{\mathcal{O}}(\mathbf{v})$ for every $\mathbf{v}$. □

Example 6.4.9 (Two orthonormal bases, one projection). Let $W = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$. Both

$$\mathcal{O}: \quad \mathbf{a} = \frac{1}{\sqrt{2}}(1,-1,0), \quad \mathbf{b} = \frac{1}{\sqrt{6}}(1,1,-2),$$
$$\mathcal{O}': \quad \mathbf{a}' = \frac{1}{\sqrt{2}}(1,0,-1), \quad \mathbf{b}' = \frac{1}{\sqrt{6}}(1,-2,1)$$

are orthonormal lists of two vectors in the 2-dimensional plane W, so both are orthonormal bases of W (Corollary 6.3.14). For $\mathbf{v} = (3,0,0)$,

$$\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}) = \frac{3}{2}(1,-1,0) + \frac{3}{6}(1,1,-2) = (2,-1,-1),$$
$$\mathrm{Pr}_{\mathcal{O}'}(\mathbf{v}) = \frac{3}{2}(1,0,-1) + \frac{3}{6}(1,-2,1) = (2,-1,-1).$$

The individual shadows differ, because they depend on the axes, but their sums agree. As a check, $\mathbf{v} - (2,-1,-1) = (1,1,1)$ is orthogonal to W.

By hand it is often easier to work with an orthogonal list and postpone normalizing.

Corollary 6.4.10 (Projecting with an orthogonal list). Let $\mathbf{f}_1, \ldots, \mathbf{f}_k$ be an orthogonal list of nonzero vectors in V, and let $\mathcal{O}$ be the orthonormal list $\mathbf{f}_1/\|\mathbf{f}_1\|, \ldots, \mathbf{f}_k/\|\mathbf{f}_k\|$. Then

$$\Pr_{\mathcal{O}}(\mathbf{v}) = \sum_{j=1}^{k} \frac{\langle \mathbf{v}, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j \quad \text{for all } \mathbf{v} \in V.$$

Each term is unchanged if $\mathbf{f}_j$ is replaced by $c\mathbf{f}_j$ for a scalar $c \neq 0$.

Proof. ROUTINE

Uses. Definition 6.4.1 (Projection onto the span of an orthonormal list), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. Substitute $\mathbf{u}_j = \mathbf{f}_j/\|\mathbf{f}_j\|$ into Definition 6.4.1. Is $1/\|\mathbf{f}_j\|$ real?
Hint 2. For the last claim, take c out of both slots of $\langle \mathbf{v}, c\mathbf{f}_j \rangle$ and out of $\|c\mathbf{f}_j\|^2$.
Sketch. Since $\|\mathbf{f}_j\|$ is real, $\langle \mathbf{v}, \mathbf{f}_j/\|\mathbf{f}_j\| \rangle \mathbf{f}_j/\|\mathbf{f}_j\| = \langle \mathbf{v}, \mathbf{f}_j \rangle / \|\mathbf{f}_j\|^2 \, \mathbf{f}_j$. Replacing $\mathbf{f}_j$ by $c\mathbf{f}_j$ multiplies the term by $\bar{c}c/|c|^2 = 1$.

Since $\|\mathbf{f}_j\|$ is a positive real number, it equals its conjugate, and Proposition 6.1.4 gives

$$\left\langle \mathbf{v}, \frac{\mathbf{f}_j}{\|\mathbf{f}_j\|} \right\rangle \frac{\mathbf{f}_j}{\|\mathbf{f}_j\|} = \frac{\langle \mathbf{v}, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j.$$

Summing over j gives the formula. If $\mathbf{f}_j$ is replaced by $c\mathbf{f}_j$, the jth term becomes

$$\frac{\langle \mathbf{v}, c\mathbf{f}_j \rangle}{\|c\mathbf{f}_j\|^2} c\mathbf{f}_j = \frac{\bar{c}c}{|c|^2} \cdot \frac{\langle \mathbf{v}, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j = \frac{\langle \mathbf{v}, \mathbf{f}_j \rangle}{\|\mathbf{f}_j\|^2} \mathbf{f}_j,$$

by Proposition 6.1.4 and Proposition 6.1.27(b). □

For $k = 1$ this is the shadow $\langle \mathbf{v}, \mathbf{f}_1 \rangle / \|\mathbf{f}_1\|^2 \, \mathbf{f}_1$ of Technique 6.2.4, and for an orthogonal basis of V it is Corollary 6.3.21.

Orthogonal projection onto a subspace

By Corollary 6.4.8, $\Pr_{\mathcal{O}}$ depends only on the subspace spanned by $\mathcal{O}$. So we may name it after the subspace.

Definition 6.4.11 (Orthogonal projection). Let W be a subspace of V that has an orthonormal basis $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$. The **orthogonal projection onto** W is the map $\Pr_W = \Pr_{\mathcal{O}} : V \to V$, that is,

$$\Pr_W(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k.$$

By Corollary 6.4.8 this does not depend on the choice of the orthonormal basis $\mathcal{O}$. By Theorem 6.4.6, $\Pr_W(\mathbf{v})$ is the unique vector $\mathbf{p} \in W$ with $\mathbf{v} - \mathbf{p} \perp W$.

A basis is a finite list, so only a finite-dimensional subspace can have an orthonormal basis. Conversely, Theorem 6.4.20 below shows that every finite-dimensional subspace has one. So $\Pr_W$ is defined exactly for the finite-dimensional subspaces W of V, while V itself may be infinite-dimensional. In Example 6.4.3, for instance, $\Pr_W(t^2) = 1/3$ for $W = \operatorname{span}(1, t) \subseteq C[-1, 1]$.

Remark 6.4.12 (Why "orthogonal"). The map $P : \mathbb{R}^2 \to \mathbb{R}^2$, $P(x, y) = (x - y, 0)$, also sends every vector into the x-axis W and fixes every vector of W. It slides each point onto W along the direction $(1, 1)$, not along the perpendicular. For instance $P(0, 1) = (-1, 0)$, and $(0, 1) - P(0, 1) = (1, 1)$ is not orthogonal to W, whereas $\Pr_W(0, 1) = \mathbf{0}$. Maps like P, which project onto one subspace along another, exist in every vector space, with or without an inner product, and they belong with the direct sums of Section 4.4. Among them, the orthogonal projection is the one that moves every vector perpendicularly to W. Proposition 6.4.30 shows how to recognize it.

Exercise 6.4.13. ROUTINE In $\mathbb{C}^3$ with the dot product, let $\mathbf{u} = \frac{1}{\sqrt{2}}(1, i, 0)$ and $\mathbf{v} = (0, 1, 1)$.

(a) Compute $\mathrm{Pr}_{\mathbf{u}}(\mathbf{v})$, and check that $\mathbf{v} - \mathrm{Pr}_{\mathbf{u}}(\mathbf{v}) \perp \mathbf{u}$.

(b) A student uses $\langle \mathbf{u}, \mathbf{v} \rangle \mathbf{u}$ instead. Show that this vector does not have the property in (a).

Solution on page 1496

6.4.2 The Gram–Schmidt algorithm

To use Definition 6.4.11 we need orthonormal bases. A subspace usually arrives as the span of some list $\mathbf{v}_1, \ldots, \mathbf{v}_m$, such as the columns of a matrix or the monomials $1, x, \ldots, x^n$, and that list is rarely orthonormal. The Gram–Schmidt algorithm converts it into an orthonormal basis of the same span.

Intuition (Keep only what is new). Process the vectors one at a time, and keep an orthonormal list $\mathbf{u}_1, \ldots, \mathbf{u}_k$ with the same span as the vectors processed so far. When $\mathbf{v}_j$ arrives, split it as in Figure 6.4.1: $\mathbf{v}_j$ is its projection onto $\mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ plus the remainder

$$\mathbf{f} = \mathbf{v}_j - \mathrm{Pr}_{\mathbf{u}_1, \ldots, \mathbf{u}_k}(\mathbf{v}_j).$$

The projection lies in the span we already have, so it contributes nothing new. The remainder is orthogonal to $\mathbf{u}_1, \ldots, \mathbf{u}_k$ by Proposition 6.4.4(b), so after normalizing it can be appended to the orthonormal list, and it carries exactly the direction that $\mathbf{v}_j$ adds (Figure 6.4.2). If the remainder is $\mathbf{0}$, then $\mathbf{v}_j$ lies in the span of the earlier vectors and adds no new direction, so we skip it.

Algorithm 6.4.14 (Gram–Schmidt)

Input A list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in V.
Output An orthonormal list $\mathbf{u}_1, \ldots, \mathbf{u}_k$, where $k \le m$, with $\mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

Start with no vectors kept ($k = 0$) and $j = 1$.

1. **Project.** Subtract from $\mathbf{v}_j$ its projection onto the span of the vectors kept so far:
$$\mathbf{f} = \mathbf{v}_j - \sum_{i=1}^{k} \langle \mathbf{v}_j, \mathbf{u}_i \rangle \mathbf{u}_i.$$
2. **Keep or skip.** If $\mathbf{f} \neq \mathbf{0}$, increase k by 1, and set $\mathbf{f}_k = \mathbf{f}$ and $\mathbf{u}_k = \mathbf{f}_k / \|\mathbf{f}_k\|$. If $\mathbf{f} = \mathbf{0}$, keep nothing.
3. **Repeat.** If $j < m$, increase j by 1 and return to step 1. Otherwise output $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

The sum in step 1 is $\mathrm{Pr}_{\mathcal{O}}(\mathbf{v}_j)$ for the current list $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_k$; when $k = 0$ it is empty and $\mathbf{f} = \mathbf{v}_j$. We call $\mathbf{f}_1, \mathbf{f}_2, \ldots$ the **residuals**. If no vector is skipped, then $k = j$ after the jth pass, and the algorithm is the recursion

$$\begin{aligned} \mathbf{f}_j &= \mathbf{v}_j - \sum_{i=1}^{j-1} \langle \mathbf{v}_j, \mathbf{u}_i \rangle \mathbf{u}_i, \\ \mathbf{u}_j &= \frac{\mathbf{f}_j}{\|\mathbf{f}_j\|} \quad (j = 1, \ldots, m). \end{aligned} \tag{6.4.1}$$

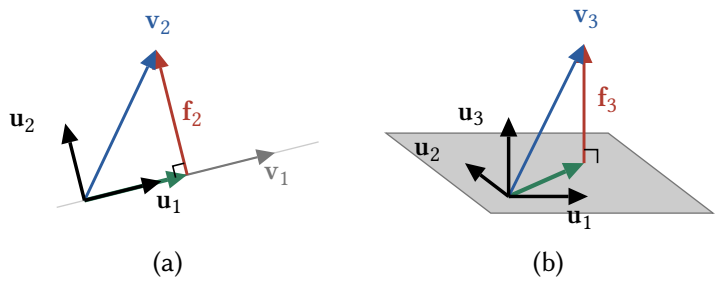


Figure 6.4.2. Two passes of Gram–Schmidt. (a) The residual $\mathbf{f}_2$ is what remains of $\mathbf{v}_2$ after subtracting its shadow on the line through $\mathbf{v}_1$; normalizing $\mathbf{v}_1$ and $\mathbf{f}_2$ gives $\mathbf{u}_1$ and $\mathbf{u}_2$. (b) In the next pass, $\mathbf{v}_3$ is projected onto the shaded plane $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2) = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$, and the residual $\mathbf{f}_3$, normalized, gives $\mathbf{u}_3$.

That the output is orthonormal and has the right span is a statement about what the algorithm keeps true from one pass to the next, so we prove it as in Technique 3.3.5.

Theorem 6.4.15 (Gram–Schmidt is correct). Run Algorithm 6.4.14 on a list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in V. For $0 \le j \le m$, let $\mathcal{O}_j$ be the list of vectors kept after the first j passes, so that $\mathcal{O}_0$ is empty.

(a) For every j, $\mathcal{O}_j$ is orthonormal and $\operatorname{span}(\mathcal{O}_j) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$.

(b) In pass j, the vector $\mathbf{v}_j$ is skipped if and only if $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$.

(c) The output is an orthonormal basis of $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

(d) If $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is linearly independent, then no vector is skipped, and the output $\mathbf{u}_1, \ldots, \mathbf{u}_m$ satisfies $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$ and $\langle \mathbf{v}_j, \mathbf{u}_j \rangle = \|\mathbf{f}_j\| > 0$ for every j.

Proof. STANDARD

Uses. Proposition 6.4.4 (Properties of $\operatorname{Pr}_{\mathcal{O}}$), Proposition 1.3.18 (Comparing spans), Proposition 6.3.13 (Orthogonal lists are independent), Theorem 1.5.5 (Independence is decided at zero), Definition 6.3.15 (Orthonormal list, orthonormal basis)

Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5), Induction on the length of a list (Technique 1.6.9), Double containment (Technique 1.3.14)

Hint 1. Prove (a) by induction on j. In pass j, let $W = \operatorname{span}(\mathcal{O}_{j-1})$. What do you know about $\mathbf{f} = \mathbf{v}_j - \operatorname{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j)$?

Hint 2. By Proposition 6.4.4, $\mathbf{f} \perp W$, and $\mathbf{f} = \mathbf{0}$ exactly when $\mathbf{v}_j \in W$. In the case $\mathbf{f} \neq \mathbf{0}$, the new unit vector is orthogonal to the old ones.

Hint 3. For the spans, compare with Proposition 1.3.18 in both directions, using $\mathbf{v}_j = \operatorname{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) + \|\mathbf{f}\|\mathbf{u}$.

Sketch. Induct on j; $\mathcal{O}_0$ is empty and spans $\{\mathbf{0}\}$. In pass j, let $W = \operatorname{span}(\mathcal{O}_{j-1}) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$. By Proposition 6.4.4, $\mathbf{f} \perp W$, and $\mathbf{f} = \mathbf{0}$ if and only if $\mathbf{v}_j \in W$; this gives (b). If $\mathbf{f} = \mathbf{0}$, nothing changes, and $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j) = W$. Otherwise $\mathbf{u} = \mathbf{f}/\|\mathbf{f}\|$ is a unit vector orthogonal to $\mathcal{O}_{j-1}$, and $\mathbf{u} \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$ and $\mathbf{v}_j = \operatorname{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) + \|\mathbf{f}\|\mathbf{u}$ give equal spans. (c) is (a) for $j = m$, with Proposition 6.3.13. For (d), independence rules out $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$, and $\langle \mathbf{v}_j, \mathbf{u}_j \rangle = \langle \Pr(\mathbf{v}_j) + \|\mathbf{f}_j\|\mathbf{u}_j, \mathbf{u}_j \rangle = \|\mathbf{f}_j\|$.

(a) We induct on j. For $j = 0$, the empty list is orthonormal, and both sides equal $\{\mathbf{0}\}$.

Let $1 \le j \le m$, and suppose that $\mathcal{O}_{j-1} = \mathbf{u}_1, \ldots, \mathbf{u}_k$ is orthonormal with $\operatorname{span}(\mathcal{O}_{j-1}) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$. Call this subspace W. Pass j computes $\mathbf{f} = \mathbf{v}_j - \operatorname{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j)$. By (a) and (c) of Proposition 6.4.4, $\operatorname{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) \in W$, and $\mathbf{f} = \mathbf{0}$ if and only if $\mathbf{v}_j \in W$.

Suppose $\mathbf{f} = \mathbf{0}$. Then $\mathcal{O}_j = \mathcal{O}_{j-1}$ is orthonormal. Since $\mathbf{v}_j \in W = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$, Proposition 1.3.18 gives $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}) = W = \operatorname{span}(\mathcal{O}_j)$.

Suppose $\mathbf{f} \neq \mathbf{0}$. Then $\mathcal{O}_j$ is $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}$ with $\mathbf{u} = \mathbf{f}/\|\mathbf{f}\|$, a unit vector. By Proposition 6.4.4(b), $\mathbf{f} \in W^\perp$, so $\langle \mathbf{u}, \mathbf{u}_i \rangle = \langle \mathbf{f}, \mathbf{u}_i \rangle / \|\mathbf{f}\| = 0$ for $i \le k$. Hence $\mathcal{O}_j$ is orthonormal. For the spans we prove both containments with Proposition 1.3.18. Each $\mathbf{u}_i$ with $i \le k$ lies in $W \subseteq \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$, and so does

$$\mathbf{u} = \frac{1}{\|\mathbf{f}\|}\Big(\mathbf{v}_j - \mathrm{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j)\Big),$$

because $\mathrm{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) \in W$. So $\mathrm{span}\big(\mathcal{O}_j\big) \subseteq \mathrm{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_j\big)$. Conversely, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ lie in $W = \mathrm{span}\big(\mathcal{O}_{j-1}\big) \subseteq \mathrm{span}\big(\mathcal{O}_j\big)$, and so does

$$\mathbf{v}_j = \mathrm{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) + \|\mathbf{f}\|\mathbf{u}.$$

So $\mathrm{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_j\big) \subseteq \mathrm{span}\big(\mathcal{O}_j\big)$, and the induction is complete.

(b) We saw in the proof of (a) that pass j keeps nothing exactly when $\mathbf{v}_j \in W = \mathrm{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\big)$.

(c) The output is $\mathcal{O}_m$. By (a) it is orthonormal, hence linearly independent (Proposition 6.3.13), and it spans $\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

(d) If $\mathbf{v}_j = c_1\mathbf{v}_1 + \cdots + c_{j-1}\mathbf{v}_{j-1}$ for some j, then $c_1\mathbf{v}_1 + \cdots + c_{j-1}\mathbf{v}_{j-1} - \mathbf{v}_j = \mathbf{0}$ is a linear relation with a nonzero coefficient, so the list is dependent (Theorem 1.5.5). So if the list is independent, (b) shows that nothing is skipped. Then $\mathcal{O}_j = \mathbf{u}_1, \ldots, \mathbf{u}_j$ for every j, and (a) gives the equality of spans. Finally,

$$\mathbf{v}_j = \mathrm{Pr}_{\mathcal{O}_{j-1}}(\mathbf{v}_j) + \|\mathbf{f}_j\|\mathbf{u}_j,$$

where the first term lies in $\mathrm{span}\big(\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}\big)$ and is therefore orthogonal to $\mathbf{u}_j$. Hence $\langle \mathbf{v}_j, \mathbf{u}_j \rangle = \|\mathbf{f}_j\|\langle \mathbf{u}_j, \mathbf{u}_j \rangle = \|\mathbf{f}_j\|$, which is positive because $\mathbf{f}_j \neq \mathbf{0}$. □

The invariant (a) is exactly what makes step 1 the right move. Because $\mathrm{span}\big(\mathcal{O}_{j-1}\big)$ is the span of all the earlier vectors, the residual is orthogonal to each earlier output, not only to the most recent one. Subtracting only the shadow on $\mathbf{u}_{j-1}$ would leave components along $\mathbf{u}_1, \ldots, \mathbf{u}_{j-2}$.

Remark 6.4.16 (Postponing the normalization). By Corollary 6.4.10, the projection in step 1 can be computed from the residuals instead of the normalized vectors:

$$\mathbf{f} = \mathbf{v}_j - \sum_{i=1}^{k} \frac{\langle \mathbf{v}_j, \mathbf{f}_i \rangle}{\|\mathbf{f}_i\|^2}\mathbf{f}_i,$$

and each $\mathbf{f}_i$ may be replaced by any nonzero multiple of itself without changing this sum. So by hand one runs the algorithm with residuals only, clears denominators in each residual as it appears, and normalizes once at the end. Square roots then appear only in the last step.

Example 6.4.17 (Gram–Schmidt in $\mathbb{R}^3$). Let $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (1, 0, 1)$, $\mathbf{v}_3 = (0, 1, 1)$. Following Remark 6.4.16:

$$\begin{aligned} \mathbf{f}_1 &= (1, 1, 0), \\ \mathbf{f}_2 &= (1, 0, 1) - \frac{1}{2}(1, 1, 0) = \frac{1}{2}(1, -1, 2), \end{aligned}$$

which we replace by $(1, -1, 2)$. Since $\langle \mathbf{v}_3, \mathbf{f}_1 \rangle = 1$ with $\|\mathbf{f}_1\|^2 = 2$, and $\langle \mathbf{v}_3, (1, -1, 2) \rangle = 1$ with $\|(1, -1, 2)\|^2 = 6$,

$$\begin{aligned} \mathbf{f}_3 &= (0, 1, 1) - \frac{1}{2}(1, 1, 0) - \frac{1}{6}(1, -1, 2) \\ &= \frac{2}{3}(-1, 1, 1), \end{aligned}$$

which we replace by $(-1, 1, 1)$. The rescaled residuals $(1, 1, 0)$, $(1, -1, 2)$, $(-1, 1, 1)$ are pairwise orthogonal, and normalizing them gives the orthonormal basis

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, 1, 0), \quad \mathbf{u}_2 = \frac{1}{\sqrt{6}}(1, -1, 2),$$

$$\mathbf{u}_3 = \frac{1}{\sqrt{3}}(-1,1,1)$$

of $\mathbb{R}^3$. As Theorem 6.4.15(d) predicts, $\langle \mathbf{v}_2, \mathbf{u}_2 \rangle = 3/\sqrt{6} = \sqrt{3/2}$ is the length of the unrescaled residual $\frac{1}{2}(1,-1,2)$.

$$\mathbf{f}_1 = (1,1,0), \quad \mathbf{f}_2 = \frac{1}{2}(1,-1,2),$$
$$\mathbf{f}_3 = \frac{2}{3}(-1,1,1)$$

Gram–Schmidt on the vectors of Example 6.4.17, one step at a time, keeping the residuals unnormalized as in Remark 6.4.16; normalizing them gives $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. In the interactive version, drag or type the inputs; a vector typed into the span of the earlier ones is skipped. (Interactive in the web version.)

Example 6.4.18 (A dependent list). Run the algorithm on $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (1,2,3)$, $\mathbf{v}_3 = (3,4,5)$, $\mathbf{v}_4 = (0,0,1)$ in $\mathbb{R}^3$, again with residuals only. First $\mathbf{f}_1 = (1,1,1)$, and

$$\mathbf{f}_2 = (1,2,3) - \frac{6}{3}(1,1,1) = (-1,0,1).$$

In the third pass, $\langle \mathbf{v}_3, \mathbf{f}_1 \rangle = 12$ and $\langle \mathbf{v}_3, \mathbf{f}_2 \rangle = 2$, so

$$\begin{aligned} &\mathbf{v}_3 - \frac{12}{3}(1,1,1) - \frac{2}{2}(-1,0,1) \\ &= (3-4+1, 4-4-0, 5-4-1) = \mathbf{0}. \end{aligned}$$

So $\mathbf{v}_3$ is skipped: by Theorem 6.4.15(b) it lies in $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$, and the computation even exhibits the relation $\mathbf{v}_3 = 4\mathbf{f}_1 + \mathbf{f}_2 = 4\mathbf{v}_1 + (\mathbf{v}_2 - 2\mathbf{v}_1) = 2\mathbf{v}_1 + \mathbf{v}_2$. In the fourth pass,

$$(0,0,1) - \frac{1}{3}(1,1,1) - \frac{1}{2}(-1,0,1) = \frac{1}{6}(1,-2,1),$$

which we replace by $(1,-2,1)$. The output, after normalizing, is $(1,1,1)/\sqrt{3}$, $(-1,0,1)/\sqrt{2}$, $(1,-2,1)/\sqrt{6}$: an orthonormal basis of $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_4) = \mathbb{R}^3$.

So Gram–Schmidt also detects linear dependence, much as row reduction does, and the number of vectors it keeps is the dimension of the span.

Example 6.4.19 (Legendre polynomials). Give $\mathbb{R}[x]_{\leq 3}$ the inner product $\langle p, q \rangle = \int_{-1}^{1} p(t)q(t)\,dt$ of Example 6.3.17, and run Gram–Schmidt on $1, x, x^2, x^3$. Since $\langle x^i, x^j \rangle = \int_{-1}^{1} t^{i+j}\,dt$ is $2/(i+j+1)$ when $i+j$ is even and 0 when $i+j$ is odd,

$$f_1 = 1, \quad f_2 = x - \frac{\langle x, 1 \rangle}{\|1\|^2} 1 = x,$$
$$f_3 = x^2 - \frac{\langle x^2, 1 \rangle}{\|1\|^2} 1 - \frac{\langle x^2, x \rangle}{\|x\|^2} x = x^2 - \frac{2/3}{2} = x^2 - \frac{1}{3},$$

which recovers the orthogonal basis of Example 6.3.17. The next pass gives $f_4 = x^3 - (3/5)x$, the polynomial of Exercise 6.3.49. By Theorem 6.4.15(d), $\operatorname{span}(f_1, \ldots, f_j) = \operatorname{span}(1, x, \ldots, x^{j-1})$. So $f_{j+1} = x^j - \Pr(x^j)$, where the projection is onto the polynomials of degree less than j: it has degree j, leading coefficient 1, and is orthogonal to every polynomial of degree less than j. Rescaled so that their value at 1 is 1, these are the **Legendre polynomials**

$$P_0 = 1, \quad P_1 = x,$$

$$P_2 = \frac{1}{2}(3x^2 - 1), \quad P_3 = \frac{1}{2}(5x^3 - 3x).$$

The output depends on the inner product. The lecture's example uses $\langle p, q\rangle = \int_0^1 p(t)q(t)\,dt$ instead, and Gram–Schmidt applied to $1, x$ then gives $1, \sqrt{3}(2x-1)$ (compare Exercise 6.4.46).

Now we can show that the hypothesis of Definition 6.4.11 holds for every finite-dimensional subspace.

Theorem 6.4.20 (Orthonormal bases exist).

(a) Every finite-dimensional inner product space has an orthonormal basis.

(b) In a finite-dimensional inner product space, every orthonormal list extends to an orthonormal basis.

(c) Every finite-dimensional subspace W of an inner product space V has an orthonormal basis. In particular, Pr_W is defined.

Proof. ROUTINE

Uses. Theorem 6.4.15 (Gram–Schmidt is correct), Corollary 1.6.24 (Bases exist in finite dimensions), Definition 6.4.11 (Orthogonal projection), Definition 6.1.2 (Inner product)
Hint 1. Run Gram–Schmidt on a basis, or on any spanning list.
Hint 2. For (b), put the orthonormal list in front of a basis of V. What does Gram–Schmidt do to vectors that are already orthonormal?
Sketch. (a) Gram–Schmidt turns a basis of V into an orthonormal basis of its span, which is V (Theorem 6.4.15(c)). (b) Run it on $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{b}_1, \ldots, \mathbf{b}_n$ for a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of V: the first k residuals are $\mathbf{u}_1, \ldots, \mathbf{u}_k$ themselves, and the output is an orthonormal basis of V that starts with them. (c) The inner product restricted to W is an inner product on W; apply (a).

(a) Let V be finite-dimensional. It has a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ (Corollary 1.6.24). By Theorem 6.4.15(c), Gram–Schmidt turns it into an orthonormal basis of $\mathrm{span}(\mathbf{b}_1, \ldots, \mathbf{b}_n) = V$.

(b) Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be orthonormal, and let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V. Run Gram–Schmidt on the list $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{b}_1, \ldots, \mathbf{b}_n$. In pass $j \le k$, the vectors kept so far are $\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}$ (by induction on j), so the residual is

$$\mathbf{u}_j - \sum_{i=1}^{j-1} \langle \mathbf{u}_j, \mathbf{u}_i\rangle \mathbf{u}_i = \mathbf{u}_j,$$

a unit vector, which is kept unchanged. So the output begins with $\mathbf{u}_1, \ldots, \mathbf{u}_k$, and by Theorem 6.4.15(c) it is an orthonormal basis of the span of the whole list, which is V because the list contains a basis.

(c) The inner product of V, restricted to pairs of vectors of W, satisfies the axioms of Definition 6.1.2 on W. So W is a finite-dimensional inner product space, and it has an orthonormal basis by (a). Then Pr_W is defined by Definition 6.4.11. □

Part (b) is the orthonormal version of extending an independent list to a basis (Theorem 4.1.16). Together with Parseval's identity, part (a) completes the picture sketched after Theorem 6.3.22: every n-dimensional inner product space has an orthonormal basis, and its coordinates carry the inner product to the dot product of $\mathbb{F}^n$.

Exercise 6.4.21. ROUTINE In $\mathbb{C}^3$ with the dot product, apply Gram–Schmidt to $\mathbf{v}_1 = (1, i, 0)$, $\mathbf{v}_2 = (1, 1, 1)$. What goes wrong if $\langle \mathbf{v}_2, \mathbf{u}_1\rangle$ is replaced by $\langle \mathbf{u}_1, \mathbf{v}_2\rangle$ in step 1?

Solution on page 1497

6.4.3 Orthogonal complements

In Orthogonality to a set we defined $A^\perp$ for every subset A of V. When A is a subspace, $A^\perp$ deserves a name of its own, which we now justify.

Definition 6.4.22 (Orthogonal complement). Let W be a subspace of V. The **orthogonal complement** of W is the subspace

$$W^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W\}$$

of Definition 6.3.8.

By Proposition 6.3.9, $W^\perp$ is a subspace of V, $W \cap W^\perp = \{\mathbf{0}\}$, and $U \subseteq W$ implies $W^\perp \subseteq U^\perp$. By Exercise 6.3.11, $W \subseteq (W^\perp)^\perp$. Note that $W^\perp$ depends on the ambient space: the perp of the x-axis is a line in $\mathbb{R}^2$ and a plane in $\mathbb{R}^3$.

Example 6.4.23 (Orthogonal complements).

(a) In $\mathbb{R}^3$, $\operatorname{span}((1,1,1))^\perp$ is the plane $W = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 0\}$ of Example 6.4.9 (Example 6.3.10). Conversely $W^\perp = \operatorname{span}((1,1,1))$. Indeed $(1,1,1)$ is orthogonal to the spanning vectors $(1,-1,0)$ and $(1,1,-2)$ of W, so $\operatorname{span}((1,1,1)) \subseteq W^\perp$, and Theorem 6.4.26 below gives equality.

(b) For $A \in \mathbb{R}^{m\times n}$, let $R \subseteq \mathbb{R}^n$ be the span of the rows of A. By Example 6.3.10 and Proposition 6.3.9(c), $\ker A = R^\perp$: solving $A\mathbf{x} = \mathbf{0}$ means finding the vectors orthogonal to every row.

(c) In $C[-1,1]$ with $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt$, every odd function is orthogonal to every even function (Example 6.3.2(c)). Exercise 6.4.54 shows that the odd functions form exactly the orthogonal complement of the even functions.

The main theorem says that $W^\perp$ really is a complement of W, as soon as W is finite-dimensional. The proof is already contained in Proposition 6.4.4: every vector is its projection onto W plus something orthogonal to W.

Theorem 6.4.24 (Orthogonal decomposition). Let W be a finite-dimensional subspace of V. Every $\mathbf{v} \in V$ can be written in exactly one way as

$$\mathbf{v} = \mathbf{w} + \mathbf{w}' \quad \text{with } \mathbf{w} \in W \text{ and } \mathbf{w}' \in W^\perp,$$

namely with $\mathbf{w} = \operatorname{Pr}_W(\mathbf{v})$ and $\mathbf{w}' = \mathbf{v} - \operatorname{Pr}_W(\mathbf{v})$. That is, $V = W \oplus W^\perp$.

Proof. ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.4.4 (Properties of $\operatorname{Pr}_{\mathcal{O}}$), Proposition 6.3.9 (Properties of $A^\perp$), Definition 4.4.4 (Direct sum), Theorem 4.4.7 (Criteria for direct sums)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For existence, Pr_W is defined because W has an orthonormal basis. Which properties of $\operatorname{Pr}_W(\mathbf{v})$ do you need?
Hint 2. For uniqueness, suppose $\mathbf{w} + \mathbf{w}' = \mathbf{x} + \mathbf{x}'$ with $\mathbf{w}, \mathbf{x} \in W$ and $\mathbf{w}', \mathbf{x}' \in W^\perp$, and rearrange.
Sketch. By Theorem 6.4.20, Pr_W is defined, and $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) + (\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}))$ has its first piece in W and its second in $W^\perp$ by Proposition 6.4.4. If $\mathbf{w} + \mathbf{w}' = \mathbf{x} + \mathbf{x}'$ are two such decompositions, then $\mathbf{w} - \mathbf{x} = \mathbf{x}' - \mathbf{w}'$ lies in $W \cap W^\perp = \{\mathbf{0}\}$.

By Theorem 6.4.20, W has an orthonormal basis, so Pr_W is defined. For $\mathbf{v} \in V$, (a) and (b) of Proposition 6.4.4 give $\operatorname{Pr}_W(\mathbf{v}) \in W$ and $\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) \in W^\perp$, and these two vectors add up to $\mathbf{v}$. So a decomposition exists.

Suppose $\mathbf{v} = \mathbf{w} + \mathbf{w}' = \mathbf{x} + \mathbf{x}'$ with $\mathbf{w}, \mathbf{x} \in W$ and $\mathbf{w}', \mathbf{x}' \in W^\perp$. Then

$$\mathbf{w} - \mathbf{x} = \mathbf{x}' - \mathbf{w}'.$$

The left side lies in W and the right side in $W^\perp$, since both are subspaces. So both sides lie in $W \cap W^\perp = \{\mathbf{0}\}$ (Proposition 6.3.9(d)), and $\mathbf{w} = \mathbf{x}$, $\mathbf{w}' = \mathbf{x}'$. In the language of Section 4.4, $V = W + W^\perp$ and $W \cap W^\perp = \{\mathbf{0}\}$, so the sum is direct (Theorem 4.4.7, Definition 4.4.4). □

Uniqueness can also be read off from Theorem 6.4.6: if $\mathbf{v} = \mathbf{w} + \mathbf{w}'$ with $\mathbf{w} \in W$ and $\mathbf{w}' \in W^\perp$, then $\mathbf{w} \in W$ and $\mathbf{v} - \mathbf{w} = \mathbf{w}' \perp W$, so $\mathbf{w} = \operatorname{Pr}_W(\mathbf{v})$. In $\mathbb{R}^3$ the theorem says that every vector is the sum of a vector in a given plane through $\mathbf{0}$ and a vector along the normal line, which is how one splits a force into the part along a surface and the part pushing against it.

Note what was assumed: W is finite-dimensional, but V need not be. When W is infinite-dimensional the theorem can fail. Exercise 6.4.63 gives a proper subspace W of an inner product space V with $W^\perp = \{\mathbf{0}\}$, so that $W + W^\perp = W \neq V$; Section 15.3 explains what survives in infinite dimensions.

Corollary 6.4.25 (Dimension of the orthogonal complement). Let V be finite-dimensional and W a subspace of V. Then $\dim W^\perp = \dim V - \dim W$. More precisely, if $\mathbf{u}_1, \dots, \mathbf{u}_n$ is an orthonormal basis of V whose first k vectors form an orthonormal basis of W, then $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ is an orthonormal basis of $W^\perp$.

Proof (by the dimension formula). ROUTINE

Uses. Theorem 6.4.24 (Orthogonal decomposition), Theorem 4.4.19 (Dimension of a sum), Proposition 6.3.9 (Properties of $A^\perp$), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. W and $W^\perp$ are finite-dimensional because V is. What does Theorem 6.4.24 say about $W + W^\perp$ and $W \cap W^\perp$?
Hint 2. Apply $\dim\big(W + W^\perp\big) = \dim W + \dim W^\perp - \dim\big(W \cap W^\perp\big)$.
Sketch. $V = W + W^\perp$ and $W \cap W^\perp = \{\mathbf{0}\}$, so the dimension formula gives $\dim V = \dim W + \dim W^\perp - 0$.

Subspaces of the finite-dimensional space V are finite-dimensional (Proposition 4.1.21). By Theorem 6.4.24, $W + W^\perp = V$, and $W \cap W^\perp = \{\mathbf{0}\}$ by Proposition 6.3.9(d). By Theorem 4.4.19,

$$\begin{aligned} \dim V &= \dim\big(W + W^\perp\big) \\ &= \dim W + \dim W^\perp - \dim\big(W \cap W^\perp\big) \\ &= \dim W + \dim W^\perp, \end{aligned}$$

which gives the dimension formula. The statement about bases is proved in the second approach. □

Proof (by extending an orthonormal basis). ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 6.3.9 (Properties of $A^\perp$), Definition 6.3.15 (Orthonormal list, orthonormal basis)
Techniques. Double containment (Technique 1.3.14), Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Take an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of W and extend it to an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of V with Theorem 6.4.20(b). Guess $W^\perp$.
Hint 2. Show $\operatorname{span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n) = W^\perp$ by double containment. For $\subseteq$, test against the spanning list $\mathbf{u}_1, \dots, \mathbf{u}_k$ of W. For $\supseteq$, expand $\mathbf{v} \in W^\perp$ in the basis.
Sketch. Extend an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of W to one of V. Each $\mathbf{u}_j$ with $j > k$ is orthogonal to $\mathbf{u}_1, \dots, \mathbf{u}_k$, hence to W. Conversely, $\mathbf{v} \in W^\perp$ has $\langle \mathbf{v}, \mathbf{u}_j \rangle = 0$ for $j \le k$, so its expansion $\sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$ involves only $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$. So $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ is an orthonormal basis of $W^\perp$, and $\dim W^\perp = n - k$.

By Theorem 6.4.20, W has an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$, and it extends to an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of V, where $n = \dim V$. (Conversely, every orthonormal basis of V whose first k vectors span W arises this way.) Let $U = \operatorname{span}(\mathbf{u}_{k+1}, \dots, \mathbf{u}_n)$. We show $U = W^\perp$ by double containment.

For $j > k$, $\mathbf{u}_j$ is orthogonal to $\mathbf{u}_1, \dots, \mathbf{u}_k$, so $\mathbf{u}_j \in W^\perp$ by Proposition 6.3.9(c). As $W^\perp$ is a subspace, $U \subseteq W^\perp$. Conversely, let $\mathbf{v} \in W^\perp$. Then $\langle \mathbf{v}, \mathbf{u}_j \rangle = 0$ for $j \le k$, so by Proposition 6.3.18(b),

$$\mathbf{v} = \sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j = \sum_{j=k+1}^{n} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j \in U.$$

So $W^\perp = U$. The list $\mathbf{u}_{k+1}, \dots, \mathbf{u}_n$ is orthonormal and spans $W^\perp$, so it is an orthonormal basis of $W^\perp$, and $\dim W^\perp = n - k = \dim V - \dim W$. □

The first proof counts. The second builds the basis of $W^\perp$, and it shows that, in coordinates with respect to an orthonormal basis adapted to W, the decomposition $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) + (\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}))$ simply splits the coordinate vector into its first k entries and the rest.

For the matrix A of Example 6.4.23(b), the corollary gives $\dim R = n - \dim\ker A$, and rank–nullity (Theorem 4.2.11) gives $\dim \operatorname{im} A = n - \dim\ker A$. So the rows of a real matrix span a space of the same dimension as its columns. This is Theorem 4.3.8 for real matrices, obtained here by a dimension count.

Since $W \subseteq \left(W^\perp\right)^\perp$ always holds, the next question is whether taking the complement twice returns W. In $\mathbb{R}^3$ the complement of a plane is its normal line, and the complement of that line is the plane again.

Theorem 6.4.26 (The double complement). Let W be a finite-dimensional subspace of V. Then $\left(W^\perp\right)^\perp = W$.

Proof (by the orthogonal decomposition). STANDARD

Uses. Theorem 6.4.24 (Orthogonal decomposition), Proposition 6.3.9 (Properties of $A^\perp$), Exercise 6.3.11
Techniques. Double containment (Technique 1.3.14)
Hint 1. One containment, $W \subseteq \left(W^\perp\right)^\perp$, holds for every subset (Exercise 6.3.11). For the other, take $\mathbf{v} \in \left(W^\perp\right)^\perp$ and decompose it with Theorem 6.4.24.
Hint 2. Write $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) + \mathbf{w}'$ with $\mathbf{w}' \in W^\perp$. Show that $\mathbf{w}'$ also lies in $\left(W^\perp\right)^\perp$.
Sketch. $W \subseteq \left(W^\perp\right)^\perp$ always. Conversely, if $\mathbf{v} \in \left(W^\perp\right)^\perp$, write $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) + \mathbf{w}'$ with $\mathbf{w}' \in W^\perp$. Both $\mathbf{v}$ and $\operatorname{Pr}_W(\mathbf{v}) \in W$ lie in the subspace $\left(W^\perp\right)^\perp$, so $\mathbf{w}' \in W^\perp \cap \left(W^\perp\right)^\perp = \{\mathbf{0}\}$, and $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) \in W$.

By Exercise 6.3.11, $W \subseteq \left(W^\perp\right)^\perp$. Conversely, let $\mathbf{v} \in \left(W^\perp\right)^\perp$. By Theorem 6.4.24, $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) + \mathbf{w}'$ with $\mathbf{w}' \in W^\perp$. Here $\mathbf{v} \in \left(W^\perp\right)^\perp$ by assumption and $\operatorname{Pr}_W(\mathbf{v}) \in W \subseteq \left(W^\perp\right)^\perp$, so $\mathbf{w}' = \mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) \in \left(W^\perp\right)^\perp$, since $\left(W^\perp\right)^\perp$ is a subspace. So $\mathbf{w}'$ lies in $W^\perp$ and in $\left(W^\perp\right)^\perp$, and Proposition 6.3.9(d), applied to the subspace $W^\perp$, gives $\mathbf{w}' = \mathbf{0}$. Hence $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) \in W$. □

Proof (by counting dimensions, when V is finite-dimensional). ROUTINE

Uses. Corollary 6.4.25 (Dimension of the orthogonal complement), Proposition 4.1.21 (Dimensions of subspaces), Exercise 6.3.11
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. You know $W \subseteq \left(W^\perp\right)^\perp$. Compare dimensions.
Hint 2. Apply Corollary 6.4.25 twice: once to W and once to $W^\perp$.
Sketch. $W \subseteq \left(W^\perp\right)^\perp$, and $\dim\left(W^\perp\right)^\perp = n - (n - \dim W) = \dim W$ by Corollary 6.4.25 applied twice. A subspace of the same finite dimension as a space containing it is that space.

Suppose V is finite-dimensional, with $n = \dim V$. By Corollary 6.4.25, applied to W and then to $W^\perp$,

$$\dim\left(W^\perp\right)^\perp = n - \dim W^\perp = n - (n - \dim W) = \dim W.$$

Since $W \subseteq \left(W^\perp\right)^\perp$ (Exercise 6.3.11) and the dimensions agree, $W = \left(W^\perp\right)^\perp$ by Proposition 4.1.21. □

The second proof is shorter but needs V finite-dimensional, so that $W^\perp$ has a dimension at all. The first needs only that W is finite-dimensional, which is the setting of the approximation problems below, where V is a space of functions.

Corollary 6.4.27 (Testing whether a subspace is everything). Let W be a finite-dimensional subspace of V. Then $W = V$ if and only if $W^\perp = \{\mathbf{0}\}$.

Proof. ROUTINE

Uses. Theorem 6.4.24 (Orthogonal decomposition), Proposition 6.3.9 (Properties of $A^\perp$)
Hint 1. One direction is $V^\perp = \{\mathbf{0}\}$ (Proposition 6.3.9(d)).
Hint 2. For the other, decompose an arbitrary $\mathbf{v} \in V$ with Theorem 6.4.24.
Sketch. If $W = V$, then $W^\perp = V^\perp = \{\mathbf{0}\}$. If $W^\perp = \{\mathbf{0}\}$, then every $\mathbf{v}$ equals $\operatorname{Pr}_W(\mathbf{v}) + \mathbf{0} \in W$.

If $W = V$, then $W^\perp = V^\perp = \{\mathbf{0}\}$ by Proposition 6.3.9(d). Conversely, suppose $W^\perp = \{\mathbf{0}\}$, and let $\mathbf{v} \in V$. By Theorem 6.4.24, $\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) \in W^\perp = \{\mathbf{0}\}$, so $\mathbf{v} = \operatorname{Pr}_W(\mathbf{v}) \in W$. Hence $W = V$. □

So to show that a finite-dimensional subspace W is all of V, it suffices to show that the only vector orthogonal to W is $\mathbf{0}$.

Exercise 6.4.28. ROUTINE Let $W = \{\mathbf{x} \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\}$. Find $W^\perp$, a formula for $\mathrm{Pr}_{W^\perp}$, and the vector of W closest to $\mathbf{e}_1$.

Solution on page 1497

The projection as a linear map

We now collect the properties of Pr_W as a map $V \to V$.

Proposition 6.4.29 (The orthogonal projection as a linear map). Let W be a finite-dimensional subspace of V, and let $P = \mathrm{Pr}_W$.

(a) P is linear, $\operatorname{im} P = W$, and $\ker P = W^\perp$.

(b) $P^2 = P$. In fact $P(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$.

(c) $\langle P\mathbf{v}, \mathbf{x}\rangle = \langle P\mathbf{v}, P\mathbf{x}\rangle = \langle \mathbf{v}, P\mathbf{x}\rangle$ for all $\mathbf{v}, \mathbf{x} \in V$.

(d) If V is finite-dimensional, then $\mathrm{id}_V - P = \mathrm{Pr}_{W^\perp}$.

Proof. ROUTINE

Uses. Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$), Theorem 6.4.6 (The projection is characterized by orthogonality), Definition 6.4.11 (Orthogonal projection), Theorem 6.4.20 (Orthonormal bases exist), Proposition 4.1.21 (Dimensions of subspaces), Exercise 6.3.11
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. For $\ker P$, apply Theorem 6.4.6: when is $\mathbf{0}$ the vector $\mathbf{p}$ with $\mathbf{p} \in W$ and $\mathbf{v} - \mathbf{p} \perp W$?
Hint 2. For (c), split the vector in the second slot as $\mathbf{x} = P\mathbf{x} + (\mathbf{x} - P\mathbf{x})$, and the vector in the first slot similarly.
Hint 3. For (d), check that $\mathbf{v} - P\mathbf{v}$ satisfies the two conditions that characterize $\mathrm{Pr}_{W^\perp}(\mathbf{v})$.
Sketch. (a) Linearity is Proposition 6.4.4(d). The image lies in W and contains every $\mathbf{w} = P\mathbf{w}$. $P\mathbf{v} = \mathbf{0}$ if and only if $\mathbf{0} \in W$ and $\mathbf{v} - \mathbf{0} \perp W$, that is, $\mathbf{v} \in W^\perp$. (b) $P\mathbf{v} \in W$ is fixed. (c) $\langle P\mathbf{v}, \mathbf{x} - P\mathbf{x}\rangle = 0$ and $\langle \mathbf{v} - P\mathbf{v}, P\mathbf{x}\rangle = 0$. (d) $\mathbf{v} - P\mathbf{v} \in W^\perp$, and $\mathbf{v} - (\mathbf{v} - P\mathbf{v}) = P\mathbf{v} \in W$ is orthogonal to $W^\perp$.

(a) P is linear by Proposition 6.4.4(d), and $\operatorname{im} P \subseteq W$ by part (a) of the same proposition. If $\mathbf{w} \in W$, then $P\mathbf{w} = \mathbf{w}$ by part (c), so $W \subseteq \operatorname{im} P$. For the kernel: by Theorem 6.4.6, $P\mathbf{v} = \mathbf{0}$ if and only if $\mathbf{0}$ satisfies $\mathbf{0} \in W$ and $\mathbf{v} - \mathbf{0} \in W^\perp$, that is, if and only if $\mathbf{v} \in W^\perp$.

(b) For $\mathbf{v} \in V$, $P\mathbf{v} \in W$, so $P(P\mathbf{v}) = P\mathbf{v}$ by Proposition 6.4.4(c).

(c) Let $\mathbf{v}, \mathbf{x} \in V$. Since $P\mathbf{v} \in W$ and $\mathbf{x} - P\mathbf{x} \in W^\perp$,

$$\langle P\mathbf{v}, \mathbf{x}\rangle = \langle P\mathbf{v}, P\mathbf{x}\rangle + \langle P\mathbf{v}, \mathbf{x} - P\mathbf{x}\rangle = \langle P\mathbf{v}, P\mathbf{x}\rangle.$$

Similarly, since $\mathbf{v} - P\mathbf{v} \in W^\perp$ and $P\mathbf{x} \in W$,

$$\langle \mathbf{v}, P\mathbf{x}\rangle = \langle P\mathbf{v}, P\mathbf{x}\rangle + \langle \mathbf{v} - P\mathbf{v}, P\mathbf{x}\rangle = \langle P\mathbf{v}, P\mathbf{x}\rangle.$$

(d) Since V is finite-dimensional, so is $W^\perp$ (Proposition 4.1.21), so $\mathrm{Pr}_{W^\perp}$ is defined (Theorem 6.4.20). Let $\mathbf{v} \in V$ and $\mathbf{p} = \mathbf{v} - P\mathbf{v}$. Then $\mathbf{p} \in W^\perp$ by Proposition 6.4.4(b), and $\mathbf{v} - \mathbf{p} = P\mathbf{v} \in W \subseteq \left(W^\perp\right)^\perp$ (Exercise 6.3.11). These are the two conditions of Theorem 6.4.6 for the subspace $W^\perp$, so $\mathbf{p} = \mathrm{Pr}_{W^\perp}(\mathbf{v})$. □

In (d), finite-dimensionality of V was used only to know that $\mathrm{Pr}_{W^\perp}$ is defined. For any V, the map $\mathbf{v} \mapsto \mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ is linear, has image $W^\perp$, and picks out the $W^\perp$-component in Theorem 6.4.24.

Parts (b) and (c) characterize orthogonal projections among all linear maps. Part (b) alone does not: the map P of Remark 6.4.12 satisfies $P^2 = P$, but $\langle P\mathbf{e}_2, \mathbf{e}_1\rangle = \langle (-1, 0), (1, 0)\rangle = -1$ while $\langle \mathbf{e}_2, P\mathbf{e}_1\rangle = \langle \mathbf{e}_2, \mathbf{e}_1\rangle = 0$.

Proposition 6.4.30 (Recognizing orthogonal projections). Let V be finite-dimensional and $P \in \mathrm{L}(V)$. Then $P = \mathrm{Pr}_W$ for some subspace W of V if and only if

$$P^2 = P \quad \text{and} \quad \langle P\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, P\mathbf{x}\rangle \text{ for all } \mathbf{v}, \mathbf{x} \in V.$$

In that case $W = \operatorname{im} P$.

Proof. STANDARD

Uses. Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.4.6 (The projection is characterized by orthogonality), Theorem 6.4.20 (Orthonormal bases exist), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. One direction is Proposition 6.4.29(a)–(c). For the other, the only candidate is $W = \operatorname{im} P$. Check the two conditions that characterize $\operatorname{Pr}_W(\mathbf{v})$ for the vector $P\mathbf{v}$.
Hint 2. Every vector of W has the form $P\mathbf{x}$. Move P across the inner product: $\langle \mathbf{v} - P\mathbf{v}, P\mathbf{x}\rangle = \langle P(\mathbf{v} - P\mathbf{v}), \mathbf{x}\rangle$.
Sketch. If $P = \operatorname{Pr}_W$, then $P^2 = P$, the symmetry holds, and $\operatorname{im} P = W$, by Proposition 6.4.29. Conversely, let $W = \operatorname{im} P$. Then $P\mathbf{v} \in W$, and for every $P\mathbf{x} \in W$, $\langle \mathbf{v} - P\mathbf{v}, P\mathbf{x}\rangle = \langle P\mathbf{v} - P^2\mathbf{v}, \mathbf{x}\rangle = 0$. So $P\mathbf{v} = \operatorname{Pr}_W(\mathbf{v})$ by Theorem 6.4.6.

If $P = \operatorname{Pr}_W$, then $P^2 = P$ and $\langle P\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, P\mathbf{x}\rangle$ by (b) and (c) of Proposition 6.4.29, and $\operatorname{im} P = W$ by part (a).

Conversely, suppose P has the two properties, and let $W = \operatorname{im} P$. It is a subspace of the finite-dimensional space V, so it is finite-dimensional (Proposition 4.1.21) and Pr_W is defined (Theorem 6.4.20). Let $\mathbf{v} \in V$. Then $P\mathbf{v} \in W$. Every vector of W has the form $P\mathbf{x}$ for some $\mathbf{x} \in V$, and by the second property, applied to the vectors $\mathbf{v} - P\mathbf{v}$ and $\mathbf{x}$,

$$\begin{aligned}\langle \mathbf{v} - P\mathbf{v}, P\mathbf{x}\rangle &= \langle P(\mathbf{v} - P\mathbf{v}), \mathbf{x}\rangle \\ &= \langle P\mathbf{v} - P^2\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{0}, \mathbf{x}\rangle = 0.\end{aligned}$$

So $\mathbf{v} - P\mathbf{v} \in W^\perp$. By Theorem 6.4.6, $P\mathbf{v} = \operatorname{Pr}_W(\mathbf{v})$. As $\mathbf{v}$ was arbitrary, $P = \operatorname{Pr}_W$. □

The second condition says that P can be moved from one slot of the inner product to the other. In the language of Section 6.6, it says that P equals its adjoint, $P^* = P$, that is, P is self-adjoint (Definition 6.6.26). So the orthogonal projections are exactly the operators with $P^2 = P = P^*$. For $V = \mathbb{F}^n$ with the dot product, Proposition 6.3.27(d) shows that the condition says that the matrix of P is Hermitian.

Projection matrices

In $\mathbb{F}^n$ with the dot product, Pr_W is a linear map $\mathbb{F}^n \to \mathbb{F}^n$, so it is multiplication by a matrix. The matrix is built from an orthonormal basis of W and its conjugate transpose.

Theorem 6.4.31 (The matrix of an orthogonal projection). Let W be a subspace of $\mathbb{F}^n$ with the dot product, let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be an orthonormal basis of W, and let $Q \in \mathbb{F}^{n\times k}$ be the matrix with columns $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Then

$$\operatorname{Pr}_W(\mathbf{x}) = QQ^*\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{F}^n,$$

and $QQ^* = \mathbf{u}_1\mathbf{u}_1^* + \cdots + \mathbf{u}_k\mathbf{u}_k^*$. Moreover $Q^*Q = I_k$, while the $n \times n$ matrix $P = QQ^*$ satisfies $P^2 = P$ and $P^* = P$.

Proof. ROUTINE

Uses. Definition 6.3.26 (Conjugate transpose), Proposition 2.4.8 (Four views of the product), Definition 6.1.8 (Dot product), Theorem 6.3.29 (Orthonormal columns), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Compute $Q^*\mathbf{x}$ entry by entry. Row i of Q^* is $\mathbf{u}_i^*$.
Hint 2. The entries of $Q^*\mathbf{x}$ are the coefficients $\langle \mathbf{x}, \mathbf{u}_i\rangle$, and multiplying by Q combines the columns of Q with these coefficients.
Sketch. The ith entry of $Q^*\mathbf{x}$ is $\mathbf{u}_i^*\mathbf{x} = \mathbf{x} \cdot \mathbf{u}_i = \langle \mathbf{x}, \mathbf{u}_i\rangle$, so $Q(Q^*\mathbf{x}) = \sum_i \langle \mathbf{x}, \mathbf{u}_i\rangle \mathbf{u}_i = \operatorname{Pr}_W(\mathbf{x})$. Columns times rows gives $QQ^* = \sum_i \mathbf{u}_i\mathbf{u}_i^*$. Then $Q^*Q = I_k$ is Theorem 6.3.29, so $(QQ^*)^2 = Q(Q^*Q)Q^* = QQ^*$, and $(QQ^*)^* = QQ^*$.

Row i of Q^* is $\mathbf{u}_i^*$ (Definition 6.3.26), so the ith entry of $Q^*\mathbf{x}$ is $\mathbf{u}_i^*\mathbf{x} = \mathbf{x} \cdot \mathbf{u}_i = \langle \mathbf{x}, \mathbf{u}_i\rangle$ (Definition 6.1.8). By the column view of Proposition 2.4.8, $Q(Q^*\mathbf{x})$ is the combination of the columns of Q with these coefficients:

$$QQ^*\mathbf{x} = \sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{u}_i \rangle \mathbf{u}_i = \mathrm{Pr}_W(\mathbf{x}).$$

The columns-times-rows view of Proposition 2.4.8 gives $QQ^* = \sum_i \mathbf{u}_i \mathbf{u}_i^*$, since column i of Q is $\mathbf{u}_i$ and row i of Q^* is $\mathbf{u}_i^*$. Next, $Q^*Q = I_k$ by Theorem 6.3.29. Hence, by associativity and Proposition 6.3.27,

$$(QQ^*)^2 = Q(Q^*Q)Q^* = QQ^*,$$
$$(QQ^*)^* = (Q^*)^*Q^* = QQ^*.$$

□

The two products of Q and Q^* play different roles. $Q^*Q = I_k$ records that the columns are orthonormal. QQ^* is the projection onto their span, and it equals I_n only when $k = n$, that is, when Q is unitary. This is the identification promised after Theorem 6.3.29. Each term $\mathbf{u}_i \mathbf{u}_i^*$ is the projection onto the line $\mathrm{span}(\mathbf{u}_i)$, so Pr_W is the sum of the projections onto the lines of an orthonormal basis, as in Figure 6.4.1. By Corollary 6.4.10, for an orthogonal basis $\mathbf{f}_1, \ldots, \mathbf{f}_k$ of W the matrix is

$$\sum_{i=1}^{k} \frac{\mathbf{f}_i \mathbf{f}_i^*}{\mathbf{f}_i^* \mathbf{f}_i}.$$

In particular, the projection onto the line spanned by a nonzero $\mathbf{a} \in \mathbb{F}^n$ has matrix $\mathbf{a}\mathbf{a}^*/(\mathbf{a}^*\mathbf{a})$. For a general basis of W the formula needs an inverse, $A(A^*A)^{-1}A^*$, which is the subject of Section 6.5.

Example 6.4.32 (Projecting onto a plane in $\mathbb{R}^3$). Let W be the plane $x_1 + x_2 + x_3 = 0$ of Example 6.4.9, and $\mathbf{j} = (1, 1, 1)$, so that $W^\perp = \mathrm{span}(\mathbf{j})$ (Example 6.4.23(a)). By Proposition 6.4.29(d) and the formula for lines,

$$\mathrm{Pr}_W = I_3 - \frac{\mathbf{j}\mathbf{j}^\mathsf{T}}{3} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

For instance $\mathrm{Pr}_W(3, 0, 0) = (2, -1, -1)$, as in Example 6.4.9. The same matrix comes from Theorem 6.4.31 with Q built from either orthonormal basis of Example 6.4.9. Projecting onto $W^\perp$ and subtracting needed one rank-one matrix instead of two, a trick worth remembering when $W^\perp$ is smaller than W.

6.4.4 Best approximation

We return to the question that opened the section: which vector of W is closest to $\mathbf{v}$?

Intuition (The right triangle). Let W be a plane through $\mathbf{0}$ in $\mathbb{R}^3$, and $\mathbf{w}$ any point of W. The three points $\mathbf{v}$, $\mathrm{Pr}_W(\mathbf{v})$, and $\mathbf{w}$ form a triangle with a right angle at $\mathrm{Pr}_W(\mathbf{v})$: one leg, $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$, is perpendicular to W, and the other, $\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w}$, lies in W (Figure 6.4.3). The distance from $\mathbf{v}$ to $\mathbf{w}$ is the hypotenuse, which is longer than the leg $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ unless the other leg is $\mathbf{0}$. So no point of W is closer to $\mathbf{v}$ than $\mathrm{Pr}_W(\mathbf{v})$, and every other point of W is farther away.

This is the shadow argument of Figure 6.2.1, with the line replaced by a subspace. It uses only Pythagoras, so it works in every inner product space.

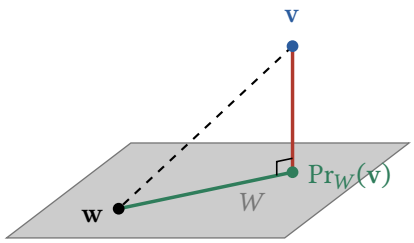


Figure 6.4.3. For every $\mathbf{w} \in W$, the points $\mathbf{v}$, $\mathrm{Pr}_W(\mathbf{v})$, $\mathbf{w}$ form a right triangle. Its hypotenuse $\mathbf{v} - \mathbf{w}$ is longer than the leg $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ unless $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$.

Theorem 6.4.33 (Best approximation). Let W be a finite-dimensional subspace of V, and let $\mathbf{v} \in V$. Then

$$\|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{w}\| \quad \text{for every } \mathbf{w} \in W,$$

with equality if and only if $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$. That is, $\mathrm{Pr}_W(\mathbf{v})$ is the unique vector of W closest to $\mathbf{v}$. If $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is an orthonormal basis of W, the least distance is given by

$$\|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2 = \|\mathbf{v}\|^2 - \sum_{j=1}^{k} |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2.$$

Proof (by Pythagoras). ROUTINE

Uses. Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.1.27 (Basic properties of the norm), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Orthogonal decomposition (Technique 6.2.4), Projections shrink norms (Technique 6.4.5)
Hint 1. Write $\mathbf{v} - \mathbf{w} = (\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})) + (\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w})$. Where do the two pieces lie?
Hint 2. The first piece is in $W^\perp$ and the second is in W, so they are orthogonal. Apply the Pythagorean theorem.
Hint 3. The formula for the least distance is Proposition 6.4.4(e).
Sketch. For $\mathbf{w} \in W$, $\mathbf{v} - \mathbf{w} = (\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})) + (\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w})$, with the first piece in $W^\perp$ and the second in W. By Pythagoras, $\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2 + \|\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w}\|^2$, which is at least $\|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2$, with equality exactly when $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$. The distance formula is Proposition 6.4.4(e).

Pr_W is defined by Theorem 6.4.20. Let $\mathbf{w} \in W$. Then

$$\mathbf{v} - \mathbf{w} = (\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})) + (\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w}),$$

where $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v}) \in W^\perp$ by Proposition 6.4.4(b), and $\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w} \in W$ because both terms lie in W. So the two pieces are orthogonal, and by Theorem 6.3.3(a),

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2 + \|\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w}\|^2 \\ &\geq \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2. \end{aligned} \tag{6.4.2}$$

Taking nonnegative square roots gives the inequality. By (6.4.2), equality holds if and only if $\|\mathrm{Pr}_W(\mathbf{v}) - \mathbf{w}\| = 0$, that is (Proposition 6.1.27(a)), if and only if $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$. Finally, by Proposition 6.4.4(e),

$$\begin{aligned} \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|^2 &= \|\mathbf{v}\|^2 - \|\mathrm{Pr}_W(\mathbf{v})\|^2 \\ &= \|\mathbf{v}\|^2 - \sum_{j=1}^{k} |\langle \mathbf{v}, \mathbf{u}_j \rangle|^2. \end{aligned}$$

□

Proof (by completing the square). STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.1.4 (Sesquilinearity), Corollary 6.3.6 (Pythagoras for orthogonal lists), Theorem 6.4.20 (Orthonormal bases exist), Definition 6.4.11 (Orthogonal projection)
Techniques. A sum of nonnegative terms (Technique 6.1.9)

Hint 1. Fix an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of W, write $\mathbf{w} = \sum_j c_j \mathbf{u}_j$ and $a_j = \langle \mathbf{v}, \mathbf{u}_j \rangle$, and expand $\|\mathbf{v} - \mathbf{w}\|^2$ as a function of $c_1, \dots, c_k$.
Hint 2. You should get $\|\mathbf{v}\|^2 + \sum_j \big(|c_j|^2 - 2\,\mathrm{Re}(\overline{c_j} a_j) \big)$. Complete the square in each c_j, using $|c - a|^2 = |c|^2 - 2\,\mathrm{Re}(\overline{c}a) + |a|^2$.
Sketch. With $\mathbf{w} = \sum_j c_j \mathbf{u}_j$ and $a_j = \langle \mathbf{v}, \mathbf{u}_j \rangle$: $\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 - 2\,\mathrm{Re}\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \sum_j \big(|c_j|^2 - 2\,\mathrm{Re}(\overline{c_j} a_j) \big) = \|\mathbf{v}\|^2 - \sum_j |a_j|^2 + \sum_j |c_j - a_j|^2$. The last sum is nonnegative and vanishes exactly when every $c_j = a_j$, that is, when $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$.

By Theorem 6.4.20, W has an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$. Let $a_j = \langle \mathbf{v}, \mathbf{u}_j \rangle$, so that $\mathrm{Pr}_W(\mathbf{v}) = \sum_j a_j \mathbf{u}_j$. Every $\mathbf{w} \in W$ is $\mathbf{w} = \sum_j c_j \mathbf{u}_j$ for unique scalars c_j. By Proposition 6.1.29 with $a = -1$,

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 - 2\,\mathrm{Re}\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2.$$

By Proposition 6.1.4, $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_j \overline{c_j} \langle \mathbf{v}, \mathbf{u}_j \rangle = \sum_j \overline{c_j} a_j$, and by Corollary 6.3.6, $\|\mathbf{w}\|^2 = \sum_j |c_j|^2$. For complex numbers c and a, $|c - a|^2 = (c - a)\overline{(c - a)} = |c|^2 - 2\,\mathrm{Re}(\overline{c}a) + |a|^2$. Hence

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= \|\mathbf{v}\|^2 + \sum_{j=1}^{k} \big(|c_j|^2 - 2\,\mathrm{Re}(\overline{c_j} a_j) \big) \\ &= \|\mathbf{v}\|^2 - \sum_{j=1}^{k} |a_j|^2 + \sum_{j=1}^{k} |c_j - a_j|^2. \end{aligned}$$

The first two terms do not depend on $\mathbf{w}$, and the last sum is a sum of nonnegative terms, which is 0 if and only if $c_j = a_j$ for every j. So $\|\mathbf{v} - \mathbf{w}\|^2$ is smallest exactly when $\mathbf{w} = \sum_j a_j \mathbf{u}_j = \mathrm{Pr}_W(\mathbf{v})$, and the smallest value is $\|\mathbf{v}\|^2 - \sum_j |a_j|^2$. □

The first proof starts from the candidate $\mathrm{Pr}_W(\mathbf{v})$ and shows that it wins. The second does not need to know the answer in advance: completing the square produces the minimizing coefficients $c_j = \langle \mathbf{v}, \mathbf{u}_j \rangle$ and the minimum at the same time, and it shows that the squared error grows like $\sum_j |c_j - a_j|^2$ as $\mathbf{w}$ moves away from the optimum. In the language of Technique 6.4.5, the first proof says that for every $\mathbf{w} \in W$, the error $\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$ is the piece of $\mathbf{v} - \mathbf{w}$ orthogonal to W: moving $\mathbf{w}$ around W changes only the piece in W, and dropping that piece can only shorten a vector.

The theorem gives a third reason why Pr_W does not depend on the orthonormal basis used to compute it: "the closest vector of W" is defined by distances alone. Putting Theorem 6.4.6 and Theorem 6.4.33 together, for $\mathbf{p} \in W$ the following are equivalent:

(i) $\mathbf{p} = \mathrm{Pr}_W(\mathbf{v})$, the formula computed from an orthonormal basis of W;
(ii) $\mathbf{v} - \mathbf{p} \perp W$;
(iii) $\|\mathbf{v} - \mathbf{p}\| \le \|\mathbf{v} - \mathbf{w}\|$ for every $\mathbf{w} \in W$.

In proofs one usually checks (ii), in computations one uses (i), and in applications the reason to care is (iii).

Definition 6.4.34 (Distance to a subspace). Let W be a finite-dimensional subspace of V and $\mathbf{v} \in V$. The **distance** from $\mathbf{v}$ to W is

$$\mathrm{dist}(\mathbf{v}, W) = \min_{\mathbf{w} \in W} \|\mathbf{v} - \mathbf{w}\| = \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|.$$

The minimum exists by Theorem 6.4.33. So $\mathrm{dist}(\mathbf{v}, W) = 0$ if and only if $\mathbf{v} \in W$, and by Theorem 6.4.24, $\mathrm{dist}(\mathbf{v}, W)$ is the length of the component of $\mathbf{v}$ in $W^\perp$. When V is finite-dimensional, that component is $\mathrm{Pr}_{W^\perp}(\mathbf{v})$ (Proposition 6.4.29(d)).

Example 6.4.35 (Distance to a hyperplane). Let $\mathbf{n} \in \mathbb{F}^n$ be nonzero and $H = \{\mathbf{n}\}^\perp = \{\mathbf{x} \in \mathbb{F}^n \mid \mathbf{x} \cdot \mathbf{n} = 0\}$. Then $H = \mathrm{span}(\mathbf{n})^\perp$ (Proposition 6.3.9(c)), so $H^\perp = \mathrm{span}(\mathbf{n})$ by Theorem 6.4.26. By Proposition 6.4.29(d), $\mathbf{v} - \mathrm{Pr}_H(\mathbf{v}) = \mathrm{Pr}_{\mathrm{span}(\mathbf{n})}(\mathbf{v}) = \big(\langle \mathbf{v}, \mathbf{n} \rangle / \|\mathbf{n}\|^2 \big) \mathbf{n}$, and therefore

$$\mathrm{dist}(\mathbf{v}, H) = \frac{|\langle \mathbf{v}, \mathbf{n} \rangle|}{\|\mathbf{n}\|}.$$

For the plane $2x_1 - x_2 + 2x_3 = 0$ in $\mathbb{R}^3$ and $\mathbf{v} = (1,1,1)$, with $\mathbf{n} = (2,-1,2)$: $\langle \mathbf{v}, \mathbf{n} \rangle = 3$ and $\|\mathbf{n}\| = 3$, so $\operatorname{dist}(\mathbf{v}, H) = 1$, and the closest point of the plane is $\mathbf{v} - (3/9)\mathbf{n} = (1/3, 4/3, 1/3)$. This is the point-to-plane formula of multivariable calculus, now proved in every dimension and over $\mathbb{C}$.

Exercise 6.4.36. ROUTINE Find the point of the plane $W = \operatorname{span}((1,1,0),(1,0,1))$ in $\mathbb{R}^3$ closest to $\mathbf{v} = (0,0,3)$, and the distance from $\mathbf{v}$ to W. Check your answer with Example 6.4.35.

Solution on page 1497

Approximating functions

In a space of functions, Theorem 6.4.33 solves approximation problems. The inner product decides what "close" means. For $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\,dt$, the distance $\|f - p\| = \left(\int_{-1}^{1} (f(t) - p(t))^2 \,dt \right)^{1/2}$ measures the error on average over the interval: it is small when $f - p$ is small on most of $[-1,1]$, even if it is larger at a few points.

Example 6.4.37 (The quadratic closest to e^t). In $C[-1,1]$ with $\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\,dt$, let W be the subspace of polynomial functions of degree at most 2. Which $p \in W$ makes $\int_{-1}^{1} \left(e^t - p(t)\right)^2 dt$ as small as possible? By Theorem 6.4.33 the answer is $p = \operatorname{Pr}_W(e^t)$. The orthogonal basis $1, t, t^2 - 1/3$ of Example 6.4.19, whose squared norms are 2, $2/3$, and $8/45$, and Corollary 6.4.10 give

$$p = \frac{\langle e^t, 1 \rangle}{2} + \frac{\langle e^t, t \rangle}{2/3} t + \frac{\langle e^t, t^2 - 1/3 \rangle}{8/45} \left(t^2 - \frac{1}{3} \right).$$

Integrating by parts,

$$\int_{-1}^{1} e^t \,dt = e - e^{-1},$$
$$\int_{-1}^{1} t e^t \,dt = 2e^{-1},$$
$$\int_{-1}^{1} t^2 e^t \,dt = e - 5e^{-1},$$

so $\langle e^t, t^2 - 1/3 \rangle = e - 5e^{-1} - (e - e^{-1})/3 = (2/3)(e - 7e^{-1})$. Therefore

$$p(t) = \frac{e - e^{-1}}{2} + 3e^{-1}t + \frac{15}{4}(e - 7e^{-1})\left(t^2 - \frac{1}{3}\right)$$
$$\approx 0.9963 + 1.1036t + 0.5367t^2.$$

Compare the Taylor polynomial $T(t) = 1 + t + t^2/2$, which is also quadratic. Figure 6.4.4 plots both errors. Near 0, T is better: it agrees with e^t to second order there. Over the whole interval, p is better. Its squared distance from e^t is $\int_{-1}^{1} (e^t - p)^2 \approx 0.0014$, against $\int_{-1}^{1} (e^t - T)^2 \approx 0.0090$ for T, and even its largest error on $[-1,1]$ is smaller: about 0.082 at $t = 1$, against 0.218 for T.

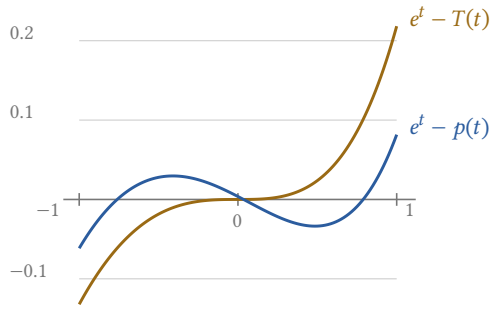


Figure 6.4.4. The errors $e^t - p(t)$ of the best quadratic approximation in the sense of Example 6.4.37, and $e^t - T(t)$ of the Taylor polynomial $T(t) = 1 + t + t^2/2$, on $[-1, 1]$. The Taylor error is tiny near 0 and large near the ends; the projection spreads the error over the interval and changes sign three times.

The error $g(t) = e^t - p(t)$ is orthogonal to 1, t, and t^2. It has at most three zeros, since $g''' = e^t > 0$ and, by Rolle's theorem, four zeros of g would give a zero of g'''. With this, the argument of Exercise 6.4.61 shows that g changes sign at least three times in $(-1, 1)$, as Figure 6.4.4 shows. The same recipe works for any continuous f and any degree n: run Gram–Schmidt on $1, t, \ldots, t^n$ once, and then each approximation costs $n + 1$ integrals. A function need not be differentiable to have a best approximation; Exercise 6.4.44 finds the quadratic closest to $|t|$, which has no Taylor polynomial at 0.

ASIDE: *Fourier truncation* BEYOND 21-242

Take the inner product $\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$ on $C([0, 2\pi], \mathbb{C})$ and the orthonormal functions $e_k(t) = \cos kt + i \sin kt$ of Aside (Fourier coefficients), and let $W_N = \operatorname{span}(e_{-N}, \ldots, e_N)$, the *trigonometric polynomials* of degree at most N. By Theorem 6.4.33, the trigonometric polynomial of degree at most N closest to f is the *partial Fourier sum*

$$S_N f = \operatorname{Pr}_{W_N}(f) = \sum_{k=-N}^{N} c_k e_k, \quad c_k = \langle f, e_k\rangle,$$

and $\|f - S_N f\|^2 = \|f\|^2 - \sum_{k=-N}^{N} |c_k|^2$. Since $W_N \subseteq W_{N+1}$, the error can only decrease as N grows. For $f(t) = t$, integration by parts gives $c_0 = \pi$ and $c_k = i/k$ for $k \neq 0$, so

$$S_N f(t) = \pi - 2\sum_{k=1}^{N} \frac{\sin kt}{k},$$

$$\|f - S_N f\|^2 = \frac{4\pi^2}{3} - \pi^2 - 2\sum_{k=1}^{N} \frac{1}{k^2}.$$

The left side is nonnegative, so $\sum_{k=1}^{N} 1/k^2 \leq \pi^2/6$ for every N. In fact the error tends to 0 as $N \to \infty$ (Section 15.3), which gives Euler's formula $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$.

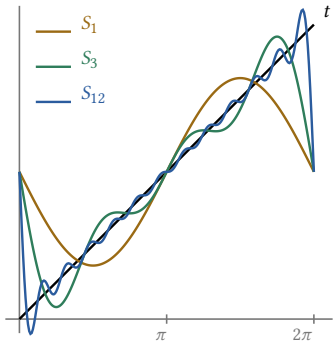


Figure 6.4.5. The function t on $[0, 2\pi]$ and its partial Fourier sums S_1, S_3, S_{12}. Each S_N is the closest trigonometric polynomial of degree at most N in the mean-square sense, and the fit improves with N except near the ends of the interval.

Best approximation in $\mathbb{F}^m$ is the subject of Section 6.5. There W is the image of a matrix A, and $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b})$ is the closest one can come to solving an inconsistent system $A\mathbf{x} = \mathbf{b}$.

6.4.5 QR factorization BEYOND 21-242

QR factorization is Gram–Schmidt in matrix form. It is not covered in lecture; see LADR 7D.

Gram–Schmidt applied to the columns of a matrix factors the matrix. If A has linearly independent columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and Gram–Schmidt turns them into $\mathbf{u}_1, \ldots, \mathbf{u}_n$, then $\mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j) = \mathrm{span}(\mathbf{a}_1, \ldots, \mathbf{a}_j)$ for each j. So each $\mathbf{a}_j$ is a combination of $\mathbf{u}_1, \ldots, \mathbf{u}_j$ only, and recording the coefficients produces an upper triangular matrix.

Theorem 6.4.38 (QR factorization). Let $A \in \mathbb{F}^{m\times n}$ have linearly independent columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Then there are unique matrices $Q \in \mathbb{F}^{m\times n}$ with orthonormal columns and $R \in \mathbb{F}^{n\times n}$ upper triangular with positive real diagonal entries such that $A = QR$. The columns of Q are the output $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of Gram–Schmidt applied to $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and $r_{ij} = \langle \mathbf{a}_j, \mathbf{u}_i \rangle$ for $i \le j$.

Proof. HARD

Uses. Theorem 6.4.15 (Gram–Schmidt is correct), Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 2.4.8 (Four views of the product), Lemma 3.3.32 (Products and inverses of triangular matrices), Theorem 6.3.29 (Orthonormal columns), Exercise 6.3.57, Proposition 6.3.27 (Properties of the conjugate transpose)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For existence, by Theorem 6.4.15(d), $\mathbf{a}_j \in \mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$. Expand it with Proposition 6.3.18(a), and read the coefficients as column j of R.
Hint 2. For uniqueness, suppose $QR = Q'R'$. Then $Q = Q'U$ with $U = R'R^{-1}$. What kind of matrix is U, and what does $Q^*Q = I_n$ say about it?
Hint 3. An upper triangular unitary matrix is diagonal (Exercise 6.3.57(c)).
Sketch. Existence: $\mathbf{a}_j = \sum_{i\le j}\langle \mathbf{a}_j, \mathbf{u}_i\rangle \mathbf{u}_i$, so column j of A is Q times the column $(r_{1j}, \ldots, r_{jj}, 0, \ldots, 0)$, and $r_{jj} = \|\mathbf{f}_j\| > 0$ by Theorem 6.4.15(d). Uniqueness: if $QR = Q'R'$, then $U = R'R^{-1}$ is upper triangular with positive diagonal and $Q = Q'U$; then $I_n = Q^*Q = U^*U$, so U is unitary and upper triangular, hence diagonal with entries of modulus 1, hence $U = I_n$.

Existence. By Theorem 6.4.15(d), Gram–Schmidt applied to $\mathbf{a}_1, \ldots, \mathbf{a}_n$ skips nothing, and its output $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is orthonormal with $\mathbf{a}_j \in \mathrm{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$. By Proposition 6.3.18(a),

$$\mathbf{a}_j = \sum_{i=1}^{j} \langle \mathbf{a}_j, \mathbf{u}_i \rangle \mathbf{u}_i.$$

Let Q have columns $\mathbf{u}_1, \dots, \mathbf{u}_n$, and let $r_{ij} = \langle \mathbf{a}_j, \mathbf{u}_i \rangle$ for $i \le j$ and $r_{ij} = 0$ for $i > j$. Then R is upper triangular, $r_{jj} = \|\mathbf{f}_j\| > 0$ by Theorem 6.4.15(d), and by the column view of Proposition 2.4.8, column j of QR is $\sum_i r_{ij}\mathbf{u}_i = \mathbf{a}_j$. So $A = QR$.

Uniqueness. Suppose $A = QR = Q'R'$, where both pairs have the stated properties. An upper triangular matrix R with nonzero diagonal entries is invertible with upper triangular inverse, whose diagonal entries are $1/r_{jj}$: write $R = DN$ with D the diagonal matrix of the r_{jj} and $N = D^{-1}R$ unit upper triangular, and apply Lemma 3.3.32 to get $R^{-1} = N^{-1}D^{-1}$. So $U = R'R^{-1}$ is upper triangular with diagonal entries $r'_{jj}/r_{jj} > 0$, again by Lemma 3.3.32. Multiplying $QR = Q'R'$ on the right by R^{-1} gives $Q = Q'U$. Since Q and Q' have orthonormal columns, Theorem 6.3.29 and Proposition 6.3.27 give

$$I_n = Q^*Q = U^*Q'^*Q'U = U^*U.$$

So the square matrix U is unitary and upper triangular. By Exercise 6.3.57(c) it is diagonal with diagonal entries of modulus 1, and these entries are positive, so $U = I_n$. Hence $R' = UR = R$ and $Q' = QU^{-1} = Q$.□

The factorization repackages several earlier facts. The columns of Q are an orthonormal basis of $\operatorname{im} A$, so $\operatorname{Pr}_{\operatorname{im} A} = QQ^*$ by Theorem 6.4.31. If A is square, then Q is unitary, so $|\det A| = |\det Q| \det R = r_{11}\cdots r_{nn}$ (Theorem 5.4.10, Proposition 5.3.15, Exercise 6.3.57), and the system $A\mathbf{x} = \mathbf{b}$ becomes $R\mathbf{x} = Q^*\mathbf{b}$, which is solved by back substitution. Section 6.5 uses the same reduction for inconsistent systems.

Example 6.4.39 (A QR factorization). Let A have the columns $\mathbf{a}_1 = (1,1,0)$, $\mathbf{a}_2 = (1,0,1)$, $\mathbf{a}_3 = (0,1,1)$ of Example 6.4.17. The output of Gram–Schmidt was $\mathbf{u}_1 = (1,1,0)/\sqrt{2}$, $\mathbf{u}_2 = (1,-1,2)/\sqrt{6}$, $\mathbf{u}_3 = (-1,1,1)/\sqrt{3}$, and the entries of R are the inner products $r_{ij} = \langle \mathbf{a}_j, \mathbf{u}_i \rangle$:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = QR,$$

where

$$Q = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix},$$

$$R = \begin{pmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 3/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 2/\sqrt{3} \end{pmatrix}.$$

The diagonal entries $\sqrt{2}, 3/\sqrt{6}, 2/\sqrt{3}$ are the lengths of the residuals, and their product is $2 = |\det A|$.

In floating-point arithmetic, Algorithm 6.4.14 as written can lose orthogonality badly when the columns are nearly dependent. Practical codes reorder the same computation (*modified Gram–Schmidt*) or build Q from reflections instead (Section 9.3).

Exercises

Exercise 6.4.40. ROUTINE In $\mathbb{R}^4$, let $W = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ with $\mathbf{v}_1 = (1,1,1,1)$, $\mathbf{v}_2 = (2,0,0,2)$, $\mathbf{v}_3 = (0,2,0,0)$.

(a) Find an orthonormal basis of W with Gram–Schmidt.

(b) Find $\operatorname{Pr}_W(\mathbf{b})$ for $\mathbf{b} = (1,2,3,4)$, and $\operatorname{dist}(\mathbf{b}, W)$.

(c) Find $W^{\perp}$.

Solution on page 1498

Exercise 6.4.41. ROUTINE Let $W = \operatorname{span}((1,0,1),(1,1,0))$ in $\mathbb{R}^3$. Find the matrix P of Pr_W in two ways: (a) by Gram–Schmidt and Theorem 6.4.31, and (b) from a vector $\mathbf{n}$ spanning $W^\perp$. Check that $P^2 = P$, $P^{\mathsf{T}} = P$, and $\operatorname{tr} P = 2$.

Solution on page 1498

Exercise 6.4.42. ROUTINE To project $\mathbf{v} = (1,2,3)$ onto $W = \operatorname{span}(\mathbf{a},\mathbf{b})$, where $\mathbf{a} = (1,0,0)$ and $\mathbf{b} = (1,1,0)$, a student adds the shadows of $\mathbf{v}$ on the two spanning vectors:

$$\frac{\langle \mathbf{v},\mathbf{a}\rangle}{\|\mathbf{a}\|^2}\mathbf{a} + \frac{\langle \mathbf{v},\mathbf{b}\rangle}{\|\mathbf{b}\|^2}\mathbf{b} = (1,0,0) + \frac{3}{2}(1,1,0) = \left(\frac{5}{2},\frac{3}{2},0\right).$$

Find the error, show that the answer fails the test of Theorem 6.4.6, and compute $\mathrm{Pr}_W(\mathbf{v})$ correctly.

Solution on page 1499

Exercise 6.4.43. ROUTINE Run Gram–Schmidt on $\mathbf{v}_1 = (1,0,1,0)$, $\mathbf{v}_2 = (1,1,0,0)$, $\mathbf{v}_3 = (0,1,-1,0)$, $\mathbf{v}_4 = (0,0,0,1)$ in $\mathbb{R}^4$. Which vector is skipped? Use the computation to write the skipped vector as a combination of the earlier ones, and give $\dim \operatorname{span}(\mathbf{v}_1,\dots,\mathbf{v}_4)$.

Solution on page 1499

Exercise 6.4.44. STANDARD In $C[-1,1]$ with $\langle f,g\rangle = \int_{-1}^{1} f(t)g(t)\,dt$, find the polynomial functions of degree at most 0, at most 1, and at most 2 closest to $f(t) = |t|$, and the squared distance from f to each.

Solution on page 1500

Exercise 6.4.45. STANDARD Find the minimum of $\int_{-1}^{1} \left(t^3 - a - bt - ct^2\right)^2 dt$ over all $a,b,c \in \mathbb{R}$, and the values of a,b,c that attain it.

Solution on page 1501

Exercise 6.4.46. STANDARD Give $\mathbb{R}[x]_{\le 2}$ the inner product $\langle p,q\rangle = \int_0^1 p(t)q(t)\,dt$ used in lecture.

(a) Apply Gram–Schmidt to $1, x, x^2$, and show that the output is 1, $\sqrt{3}(2x-1)$, $\sqrt{5}\left(6x^2-6x+1\right)$.

(b) Find $\min_{a,b\in\mathbb{R}} \int_0^1 \left(t^2 - a - bt\right)^2 dt$.

Solution on page 1501

Exercise 6.4.47. STANDARD Let U and W be subspaces of V.

(a) Show that $(U+W)^\perp = U^\perp \cap W^\perp$.

(b) Suppose V is finite-dimensional. Show that $(U \cap W)^\perp = U^\perp + W^\perp$.

Solution on page 1502

Exercise 6.4.48. STANDARD Let V be a finite-dimensional inner product space and U, W subspaces of V. Prove or disprove each statement.

(a) If $V = U \oplus W$, then $W = U^\perp$.

(b) If every vector of U is orthogonal to every vector of W and $U + W = V$, then $W = U^\perp$.

(c) $\mathrm{Pr}_U \mathrm{Pr}_W = \mathrm{Pr}_W \mathrm{Pr}_U$.

(d) If $\|\mathrm{Pr}_W(\mathbf{v})\| = \|\mathbf{v}\|$, then $\mathbf{v} \in W$.

(e) Every $P \in \mathrm{L}(V)$ with $P^2 = P$ satisfies $\|P\mathbf{v}\| \leq \|\mathbf{v}\|$ for all $\mathbf{v} \in V$.

Solution on page 1502

Exercise 6.4.49. STANDARD Let U and W be finite-dimensional subspaces of V.

(a) Suppose every vector of U is orthogonal to every vector of W. Show that $\mathrm{Pr}_{U+W} = \mathrm{Pr}_U + \mathrm{Pr}_W$ and $\mathrm{Pr}_U \mathrm{Pr}_W = 0$.

(b) Conversely, show that if $\mathrm{Pr}_U \mathrm{Pr}_W = 0$, then every vector of U is orthogonal to every vector of W.

Solution on page 1503

Exercise 6.4.50. STANDARD Find the error in the following argument, and decide whether the claim is true.

Claim. If U and W are finite-dimensional subspaces of V, then $\mathrm{Pr}_{U+W} = \mathrm{Pr}_U + \mathrm{Pr}_W$.

Argument. Let $\mathbf{v} \in V$ and $\mathbf{p} = \mathrm{Pr}_U(\mathbf{v}) + \mathrm{Pr}_W(\mathbf{v})$, which lies in $U + W$. We have $\mathbf{v} - \mathrm{Pr}_U(\mathbf{v}) \perp U$ and $\mathrm{Pr}_W(\mathbf{v}) \perp U$, so $\mathbf{v} - \mathbf{p} = (\mathbf{v} - \mathrm{Pr}_U(\mathbf{v})) - \mathrm{Pr}_W(\mathbf{v}) \perp U$. In the same way $\mathbf{v} - \mathbf{p} \perp W$, so $\mathbf{v} - \mathbf{p} \perp U + W$. By Theorem 6.4.6, $\mathbf{p} = \mathrm{Pr}_{U+W}(\mathbf{v})$.

Solution on page 1503

Exercise 6.4.51. STANDARD Let W be a k-dimensional subspace of $\mathbb{F}^n$, and let P be the matrix of Pr_W.

(a) Show that $\operatorname{rank} P = \operatorname{tr} P = k$.

(b) Show that $p_{ii} = \|P\mathbf{e}_i\|^2$ for each i, and deduce that $\sum_{i=1}^n \|\mathrm{Pr}_W(\mathbf{e}_i)\|^2 = k$.

(c) Deduce that some standard basis vector $\mathbf{e}_i$ satisfies $\operatorname{dist}(\mathbf{e}_i, W)^2 \leq 1 - k/n$.

Solution on page 1504

Exercise 6.4.52. STANDARD Let W be a finite-dimensional subspace of V and $\mathbf{v} \in V$ with $\mathbf{v} \notin W$.

(a) Show that

$$\operatorname{dist}(\mathbf{v}, W) = \max\bigl\{|\langle \mathbf{v}, \mathbf{x}\rangle| \;\bigm|\; \mathbf{x} \in W^\perp, \|\mathbf{x}\| = 1\bigr\},$$

and find the $\mathbf{x}$ attaining the maximum.

(b) In $C[-1,1]$ with $\langle f, g\rangle = \int_{-1}^1 fg$, find the largest value of $\int_{-1}^1 t^3 g(t)\,\mathrm{d}t$ over all g with $\int_{-1}^1 g^2 = 1$ that are orthogonal to $1, t, t^2$.

Solution on page 1504

Exercise 6.4.53. STANDARD

(a) Let U be a finite-dimensional subspace of V, and let $\mathbf{x}_0, \mathbf{y} \in V$. Show that among the vectors of $\mathbf{x}_0 + U = \{\mathbf{x}_0 + \mathbf{u} \mid \mathbf{u} \in U\}$, the one closest to $\mathbf{y}$ is $\mathbf{x}_0 + \mathrm{Pr}_U(\mathbf{y} - \mathbf{x}_0)$, and that it is unique.

(b) Find the solution of smallest norm of the system $x_1 + x_2 = 2$, $x_2 + x_3 = 2$ in $\mathbb{R}^3$, and check that it is orthogonal to every solution of the homogeneous system.

Solution on page 1505

Exercise 6.4.54. STANDARD In $C[-1,1]$ with $\langle f,g\rangle = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$, let E be the subspace of even functions and O the subspace of odd functions.

(a) Show that $E^\perp = O$ and $O^\perp = E$.

(b) Let $f \in C[-1,1]$ and $f_e(t) = (f(t)+f(-t))/2$. Show that f_e is the unique even function closest to f, although E is infinite-dimensional.

Solution on page 1505

Exercise 6.4.55. STANDARD

(a) Let $\mathbf{u}_1,\dots,\mathbf{u}_m$ and $\mathbf{u}'_1,\dots,\mathbf{u}'_m$ be orthonormal lists in V with $\operatorname{span}(\mathbf{u}_1,\dots,\mathbf{u}_j) = \operatorname{span}(\mathbf{u}'_1,\dots,\mathbf{u}'_j)$ for every j. Show that $\mathbf{u}'_j = \lambda_j \mathbf{u}_j$ for scalars λ_j with $|\lambda_j| = 1$.

(b) Let $\mathbf{v}_1,\dots,\mathbf{v}_m$ be linearly independent. Show that the output of Gram–Schmidt is the only orthonormal list $\mathbf{u}_1,\dots,\mathbf{u}_m$ such that, for every j, $\operatorname{span}(\mathbf{u}_1,\dots,\mathbf{u}_j) = \operatorname{span}(\mathbf{v}_1,\dots,\mathbf{v}_j)$ and $\langle \mathbf{v}_j, \mathbf{u}_j\rangle$ is a positive real number.

Solution on page 1506

Exercise 6.4.56. STANDARD Give $\mathbb{R}[x]$ the inner product $\langle p,q\rangle = \int_{-1}^{1} p(t)q(t)\,\mathrm{d}t$, and let $n \geq 1$. Among all polynomials of degree n with leading coefficient 1, show that exactly one has the smallest norm, namely $x^n - \operatorname{Pr}_W(x^n)$ with $W = \mathbb{R}[x]_{\leq n-1}$, the $(n+1)$st Gram–Schmidt residual of $1, x, x^2, \dots$. Find the smallest value of $\int_{-1}^{1} p(t)^2\,\mathrm{d}t$ for $n = 2$ and $n = 3$.

Solution on page 1506

Exercise 6.4.57. HARD Let V be finite-dimensional and let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$. Show that the following are equivalent:

(i) P is an orthogonal projection, that is, $P = \operatorname{Pr}_{\operatorname{im} P}$;

(ii) every vector of $\ker P$ is orthogonal to every vector of $\operatorname{im} P$;

(iii) $\|P\mathbf{v}\| \leq \|\mathbf{v}\|$ for every $\mathbf{v} \in V$.

Solution on page 1506

Exercise 6.4.58. HARD Let V be finite-dimensional, let U and W be subspaces, and let $P = \operatorname{Pr}_U$ and $Q = \operatorname{Pr}_W$. Show that PQ is an orthogonal projection if and only if $PQ = QP$, and that in this case $PQ = \operatorname{Pr}_{U\cap W}$.

Solution on page 1507

Exercise 6.4.59. HARD Let $\mathbf{v}_1,\dots,\mathbf{v}_m$ be linearly independent in V, with Gram matrix G (Definition 6.3.28). Run Gram–Schmidt, with residuals $\mathbf{f}_j$ and output $\mathbf{u}_j$, and let $R \in \mathbb{F}^{m\times m}$ have entries $r_{ij} = \langle \mathbf{v}_j, \mathbf{u}_i\rangle$.

(a) Show that R is upper triangular with diagonal entries $\|\mathbf{f}_1\|,\dots,\|\mathbf{f}_m\|$, and that $G = R^*R$.

(b) Deduce that $\det G = \|\mathbf{f}_1\|^2 \cdots \|\mathbf{f}_m\|^2$, and that $\operatorname{dist}(\mathbf{v}_m, \operatorname{span}(\mathbf{v}_1,\dots,\mathbf{v}_{m-1}))^2 = \det G_m / \det G_{m-1}$, where G_j is the Gram matrix of $\mathbf{v}_1,\dots,\mathbf{v}_j$ and $\det G_0 = 1$.

(c) (*Hadamard's inequality.*) Show that $\det G \leq \|\mathbf{v}_1\|^2 \cdots \|\mathbf{v}_m\|^2$, with equality if and only if the list is orthogonal. Deduce that $|\det A| \leq \|\mathbf{a}_1\| \cdots \|\mathbf{a}_n\|$ for every $A \in \mathbb{F}^{n\times n}$ with columns $\mathbf{a}_1,\dots,\mathbf{a}_n$.

Solution on page 1507

Exercise 6.4.60. HARD Let V be finite-dimensional, and let $W_1,\dots,W_m$ be subspaces of V with

$$\mathrm{Pr}_{W_1} + \cdots + \mathrm{Pr}_{W_m} = \mathrm{id}_V.$$

Show that every vector of W_i is orthogonal to every vector of W_j for $i \neq j$, and that every $\mathbf{v} \in V$ can be written in exactly one way as $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_m$ with $\mathbf{w}_i \in W_i$.

Solution on page 1508

Exercise 6.4.61. HARD Let $n \geq 1$, and let $p \in \mathbb{R}[x]$ be a polynomial of degree n with $\int_{-1}^{1} p(t)q(t)\,dt = 0$ for every polynomial q of degree less than n. (For example, the Legendre polynomial P_n of Example 6.4.19.) Show that p has n distinct roots, all in the open interval $(-1, 1)$.

Solution on page 1509

Exercise 6.4.62. CHALLENGE Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ satisfy $\langle \mathbf{v}_i, \mathbf{v}_j \rangle < 0$ for all $i \neq j$. Show that $m \leq n + 1$, and find $n + 1$ such vectors in $\mathbb{R}^n$.

Solution on page 1509

Exercise 6.4.63. HARD BEYOND 21-242 Give $\mathbb{R}[x]$ the inner product $\langle p, q \rangle = \sum_{k \geq 0} p_k q_k$, where $p = \sum_k p_k x^k$ and $q = \sum_k q_k x^k$, and let $W = \{p \in \mathbb{R}[x] \mid p(1) = 0\}$.

(a) Show that W is a subspace with $W \neq \mathbb{R}[x]$, but $W^\perp = \{\mathbf{0}\}$.

(b) Deduce that $\mathbb{R}[x] \neq W \oplus W^\perp$ and $\left(W^\perp\right)^\perp \neq W$.

(c) Show that no vector of W is closest to the constant polynomial 1, although $\inf_{w \in W} \|1 - w\| = 0$.

Solution on page 1510

Exercise 6.4.64. ROUTINE BEYOND 21-242 Let $A = \begin{pmatrix} 1 & 3 \\ 2 & 3 \\ 2 & 0 \end{pmatrix}$.

(a) Find the QR factorization of A.

(b) Use it to find $\mathrm{Pr}_{\mathrm{im}\, A}(\mathbf{e}_1)$.

(c) Check that $A^\mathsf{T} A = R^\mathsf{T} R$.

Solution on page 1511

6.5 Least squares

A linear system with more equations than unknowns usually has no solution, and that is the normal situation with measured data. Suppose the length of a spring is measured under five different loads. Hooke's law says that the five data points lie on a line $y = c_0 + c_1 x$, which gives five linear equations in the two unknowns c_0, c_1, and small measurement errors make them inconsistent. When no $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{b}$, the natural substitute is an $\mathbf{x}$ for which $A\mathbf{x}$ is as close to $\mathbf{b}$ as possible. As $\mathbf{x}$ varies, $A\mathbf{x}$ runs over the column space $\mathrm{im}\, A$, so this asks for the point of $\mathrm{im}\, A$ closest to $\mathbf{b}$, and Theorem 6.4.33 says that this point is $\mathrm{Pr}_{\mathrm{im}\, A}(\mathbf{b})$. This section turns the observation into a computation, the *normal equations* $A^* A\mathbf{x} = A^*\mathbf{b}$, which find the best $\mathbf{x}$ without an orthonormal basis of $\mathrm{im}\, A$, and applies it to fitting lines and curves to data.

Throughout, $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{F}^{m \times n}$ has columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and $\mathbf{b} \in \mathbb{F}^m$. The spaces $\mathbb{F}^m$ and $\mathbb{F}^n$ carry the dot product (Definition 6.1.8), so $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^* \mathbf{u}$ and $\|\mathbf{u}\|^2 = \sum_i |u_i|^2$. Only Section 6.5.5 and the exercises built on it use other inner products.

6.5.1 The least-squares problem

Intuition (Solving as well as possible). As $\mathbf{x}$ ranges over $\mathbb{F}^n$, the vector $A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ ranges over the column space $\operatorname{im} A = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ (Proposition 2.6.18). So $A\mathbf{x} = \mathbf{b}$ has a solution exactly when $\mathbf{b} \in \operatorname{im} A$. If $\mathbf{b} \notin \operatorname{im} A$, every $A\mathbf{x}$ misses $\mathbf{b}$, and the best we can ask for is that the miss $\|\mathbf{b} - A\mathbf{x}\|$ be as small as possible. That is, we want the point of the subspace $\operatorname{im} A$ nearest to $\mathbf{b}$.

Picture $m = 3$, with $\operatorname{im} A$ a plane through $\mathbf{0}$ and $\mathbf{b}$ off the plane (Figure 6.5.1). The nearest point of the plane is the foot of the perpendicular from $\mathbf{b}$, which is $\operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. It lies in $\operatorname{im} A$, so it equals $A\hat{\mathbf{x}}$ for some $\hat{\mathbf{x}}$, and that $\hat{\mathbf{x}}$ is the best we can do. So the problem splits into a geometric step, projecting $\mathbf{b}$ onto $\operatorname{im} A$, and an algebraic one, solving $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$, a system that is consistent by construction. The normal equations of Section 6.5.2 merge the two steps.

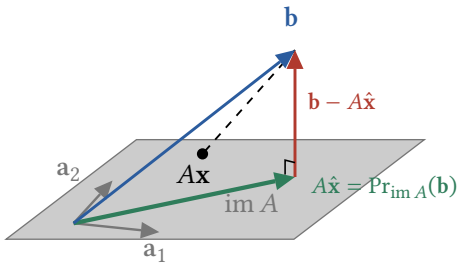


Figure 6.5.1. The column space $\operatorname{im} A = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_2)$ (shaded) and a vector $\mathbf{b} \notin \operatorname{im} A$. The point of $\operatorname{im} A$ closest to $\mathbf{b}$ is the foot $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ of the perpendicular, and the residual $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to $\operatorname{im} A$. Every other point $A\mathbf{x}$ of $\operatorname{im} A$ is farther from $\mathbf{b}$ (dashed).

Definition 6.5.1 (Least-squares solution). Let $A \in \mathbb{F}^{m\times n}$ and $\mathbf{b} \in \mathbb{F}^m$. A **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}} \in \mathbb{F}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\| \quad \text{for every } \mathbf{x} \in \mathbb{F}^n.$$

For such an $\hat{\mathbf{x}}$, the vector $\mathbf{b} - A\hat{\mathbf{x}}$ is the **residual**, and $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ is the **least-squares error**.

The name comes from the quantity being minimized. In coordinates,

$$\|\mathbf{b} - A\mathbf{x}\|^2 = \sum_{i=1}^{m} |b_i - (A\mathbf{x})_i|^2,$$

the sum of the squares of the errors in the m equations. If $A\mathbf{x} = \mathbf{b}$ has a solution, its least-squares solutions are exactly its solutions: the solutions have error 0, and by definiteness only they do. The course writes $A\mathbf{x} \approx \mathbf{b}$ for the problem of finding a least-squares solution.

The intuition above becomes the following proposition, which is Theorem 6.4.33 read for the subspace $\operatorname{im} A$.

Proposition 6.5.2 (Least squares is projection). Let $A \in \mathbb{F}^{m\times n}$ and $\mathbf{b} \in \mathbb{F}^m$.

(a) $\hat{\mathbf{x}} \in \mathbb{F}^n$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$.

(b) Least-squares solutions exist. If $\hat{\mathbf{x}}$ is one, the set of all of them is $\hat{\mathbf{x}} + \ker A = \{\hat{\mathbf{x}} + \mathbf{z} \mid \mathbf{z} \in \ker A\}$.

(c) The least-squares solution is unique if and only if the columns of A are linearly independent.

(d) The least-squares error is $\operatorname{dist}(\mathbf{b}, \operatorname{im} A)$.

Proof. ROUTINE

Uses. Proposition 2.6.18 (The column space is spanned by the columns), Theorem 6.4.20 (Orthonormal bases exist), Theorem 6.4.33 (Best approximation), Corollary 2.6.31 (Fibers are translates of the kernel), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Corollary 2.6.24 (The columns decide), Definition 6.4.34 (Distance to a subspace)
Hint 1. Every vector of $\operatorname{im} A$ has the form $A\mathbf{x}$. So minimizing $\|\mathbf{b} - A\mathbf{x}\|$ over $\mathbf{x} \in \mathbb{F}^n$ is minimizing $\|\mathbf{b} - \mathbf{w}\|$ over $\mathbf{w} \in \operatorname{im} A$.
Hint 2. For (a), use both halves of Theorem 6.4.33: the inequality and its equality case.
Hint 3. For (b), $\operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ lies in $\operatorname{im} A$. And $A\mathbf{x} = A\hat{\mathbf{x}}$ exactly when $\mathbf{x} - \hat{\mathbf{x}} \in \ker A$.
Sketch. Let $\mathbf{p} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. If $A\hat{\mathbf{x}} = \mathbf{p}$, then $\|\mathbf{b} - \mathbf{p}\| \le \|\mathbf{b} - A\mathbf{x}\|$ for all $\mathbf{x}$ by Theorem 6.4.33, since $A\mathbf{x} \in \operatorname{im} A$. Conversely, a least-squares solution does at least as well as any $\mathbf{x}_0$ with $A\mathbf{x}_0 = \mathbf{p}$, so $A\hat{\mathbf{x}}$ is a point of $\operatorname{im} A$ at least as close to $\mathbf{b}$ as $\mathbf{p}$, and the equality case forces $A\hat{\mathbf{x}} = \mathbf{p}$. The least-squares solutions are then the solutions of the consistent system $A\mathbf{x} = \mathbf{p}$, which form $\hat{\mathbf{x}} + \ker A$; this is a single point exactly when $\ker A = \{\mathbf{0}\}$.

The column space $W = \operatorname{im} A$ is a subspace of the finite-dimensional space $\mathbb{F}^m$, so it has an orthonormal basis and Pr_W is defined (Theorem 6.4.20). Let $\mathbf{p} = \operatorname{Pr}_W(\mathbf{b})$. By Proposition 2.6.18, every vector of W has the form $A\mathbf{x}$.

(a) Suppose $A\hat{\mathbf{x}} = \mathbf{p}$, and let $\mathbf{x} \in \mathbb{F}^n$. Since $A\mathbf{x} \in W$, Theorem 6.4.33 gives

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \|\mathbf{b} - \mathbf{p}\| \le \|\mathbf{b} - A\mathbf{x}\|.$$

Conversely, suppose $\hat{\mathbf{x}}$ is a least-squares solution. Since $\mathbf{p} \in W$, there is $\mathbf{x}_0$ with $A\mathbf{x}_0 = \mathbf{p}$, and so $\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}_0\| = \|\mathbf{b} - \mathbf{p}\|$. On the other hand, $A\hat{\mathbf{x}} \in W$, so $\|\mathbf{b} - \mathbf{p}\| \le \|\mathbf{b} - A\hat{\mathbf{x}}\|$ by Theorem 6.4.33. So equality holds, and by the equality case of Theorem 6.4.33, $A\hat{\mathbf{x}} = \mathbf{p}$.

(b) Since $\mathbf{p} \in W$, there is $\hat{\mathbf{x}}$ with $A\hat{\mathbf{x}} = \mathbf{p}$, and it is a least-squares solution by (a). By (a), the least-squares solutions are exactly the solutions of the consistent system $A\mathbf{x} = \mathbf{p}$, and by Corollary 2.6.31 (applied to T_A) these form the set $\hat{\mathbf{x}} + \ker A$.

(c) By (b), the least-squares solution is unique if and only if $\ker A = \{\mathbf{0}\}$, that is, if and only if T_A is injective (Theorem 2.6.11). By Corollary 2.6.24, this happens if and only if the columns of A are linearly independent.

(d) By (a) and Definition 6.4.34, $\|\mathbf{b} - A\hat{\mathbf{x}}\| = \|\mathbf{b} - \operatorname{Pr}_W(\mathbf{b})\| = \operatorname{dist}(\mathbf{b}, W)$. □

By (b), the set of least-squares solutions has the same shape as the solution set of a consistent system (Proposition 3.1.19), and for a good reason: it is the solution set of the consistent system $A\mathbf{x} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. Note also that the vector $A\hat{\mathbf{x}}$ is always unique, even when $\hat{\mathbf{x}}$ is not.

Example 6.5.3 (Three points and a line). No line passes through the three points $(0, 1), (1, 1), (2, 4)$. Let us find the line $y = c_0 + c_1 x$ that comes closest, in the sense that the sum of the squared vertical gaps between the points and the line is as small as possible. The line passes through all three points exactly when

$$c_0 + 0c_1 = 1, \quad c_0 + 1c_1 = 1, \quad c_0 + 2c_1 = 4,$$

that is, when $A\mathbf{c} = \mathbf{y}$ for the unknown $\mathbf{c} = (c_0, c_1)$, where

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}.$$

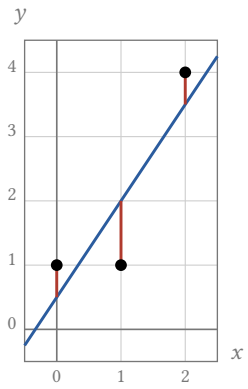
The first two equations force $c_1 = 0$ and $c_0 = 1$, and then the third says $1 = 4$, so the system is inconsistent. The ith entry of $\mathbf{y} - A\mathbf{c}$ is the vertical gap $y_i - (c_0 + c_1 x_i)$ at the ith point, so the best line comes from the least-squares solution of $A\mathbf{c} = \mathbf{y}$.

The column space is the plane in $\mathbb{R}^3$ spanned by $\mathbf{1} = (1, 1, 1)$ and $\mathbf{x} = (0, 1, 2)$. These are not orthogonal, so we first run Gram–Schmidt (Algorithm 6.4.14): $\mathbf{x} - (\langle \mathbf{x}, \mathbf{1} \rangle / 3)\mathbf{1} = (0, 1, 2) - (1, 1, 1) = (-1, 0, 1)$. With the orthogonal basis $\mathbf{1}, (-1, 0, 1)$ of $\operatorname{im} A$, Corollary 6.4.10 gives

$$\operatorname{Pr}_{\operatorname{im} A}(\mathbf{y}) = \frac{6}{3}(1, 1, 1) + \frac{3}{2}(-1, 0, 1) = \left(\frac{1}{2}, 2, \frac{7}{2}\right).$$

Solving $A\mathbf{c} = (1/2, 2, 7/2)$ gives $c_0 = 1/2$ and $c_1 = 3/2$, and the third equation checks: $1/2 + 2 \cdot 3/2 = 7/2$. The columns of A are independent, so by Proposition 6.5.2 the least-squares line is unique: it is $y = \frac{1}{2} + \frac{3}{2}x$.

The residual is $\mathbf{y} - A\mathbf{c} = (1/2, -1, 1/2)$. Its entries are the three vertical gaps, the red segments in the plot below, and the sum of their squares is $3/2$. In $\mathbb{R}^3$ it is a single vector, orthogonal to the plane $\operatorname{im} A$: indeed $(1/2, -1, 1/2) \cdot (1, 1, 1) = 0$ and $(1/2, -1, 1/2) \cdot (0, 1, 2) = 0$. Minimizing the sum of the squared gaps in the plot is the same as minimizing the length of one vector in $\mathbb{R}^3$.



The line of Example 6.5.3 and its residuals (red). In the interactive version, "Column space" shows the same fit in $\mathbb{R}^3$: the data vector $\mathbf{y}$, its projection $A\mathbf{c}$ onto the plane spanned by $\mathbf{1}$ and $\mathbf{x}$, and the residual, whose three coordinates are the three red gaps in the plot. Drag the points and watch the residual vector stay perpendicular to the plane. (Interactive in the web version.)

The computation in Example 6.5.3 took a detour. We wanted the coefficients $\mathbf{c}$, but we first computed an orthogonal basis of $\operatorname{im} A$ and the vector $A\mathbf{c}$, and only then solved for $\mathbf{c}$. The normal equations go to $\mathbf{c}$ directly.

6.5.2 The normal equations

Intuition (Why the residual is perpendicular). Here is a direct reason, without Theorem 6.4.33, why the best $A\mathbf{x}$ must be the foot of the perpendicular. Take any candidate $A\mathbf{x}$ and its residual $\mathbf{r} = \mathbf{b} - A\mathbf{x}$. Suppose $\mathbf{r}$ is not perpendicular to $\operatorname{im} A$, so that $\langle \mathbf{r}, \mathbf{w} \rangle \neq 0$ for some $\mathbf{w} \in \operatorname{im} A$; replacing $\mathbf{w}$ by a scalar multiple, we may take $\langle \mathbf{r}, \mathbf{w} \rangle$ real and positive. Slide the candidate a little in the direction $\mathbf{w}$, to $A\mathbf{x} + t\mathbf{w}$ with $t > 0$, which is still in $\operatorname{im} A$. Its distance to $\mathbf{b}$ satisfies

$$\|\mathbf{r} - t\mathbf{w}\|^2 = \|\mathbf{r}\|^2 - 2t\langle \mathbf{r}, \mathbf{w} \rangle + t^2\|\mathbf{w}\|^2$$

by Proposition 6.1.29. For small $t > 0$ the term $-2t\langle \mathbf{r}, \mathbf{w} \rangle$ beats $t^2\|\mathbf{w}\|^2$, so the new candidate is strictly closer. A candidate whose residual has any component along $\operatorname{im} A$ can be improved by sliding toward that component. So the best candidate, if there is one, has a residual orthogonal to all of $\operatorname{im} A$, which is the condition of Theorem 6.4.6.

The condition $\mathbf{r} \perp \operatorname{im} A$ is checked one column at a time, since $\operatorname{im} A$ is spanned by the columns. And "$\mathbf{r} \perp \mathbf{a}_j$ for every j" is a single matrix equation: $\mathbf{a}_j^* \mathbf{r}$ is the jth entry of $A^* \mathbf{r}$. With $\mathbf{r} = \mathbf{b} - A\mathbf{x}$, the condition $A^*(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$ is a linear system for $\mathbf{x}$. Here "normal" means perpendicular: the equations say that the residual is normal to the column space.

The column-by-column test deserves to be recorded for any matrix.

Lemma 6.5.4 (The orthogonal complement of a column space). For every $M \in \mathbb{F}^{m\times n}$, $(\operatorname{im} M)^\perp = \ker M^*$. That is, a vector $\mathbf{y} \in \mathbb{F}^m$ is orthogonal to every column of M if and only if $M^*\mathbf{y} = \mathbf{0}$.

Proof. ROUTINE

Uses. Definition 6.3.26 (Conjugate transpose), Proposition 2.4.8 (Four views of the product), Definition 6.1.8 (Dot product), Proposition 6.3.9 (Properties of $A^\perp$), Proposition 2.6.18 (The column space is spanned by the columns)
Hint 1. Compute the jth entry of $M^*\mathbf{y}$. What is row j of M^*?
Hint 2. The jth entry is $\mathbf{m}_j^*\mathbf{y} = \langle \mathbf{y}, \mathbf{m}_j\rangle$, where $\mathbf{m}_j$ is column j of M. Then pass from the columns to their span with Proposition 6.3.9(c).
Sketch. Row j of M^* is $\mathbf{m}_j^*$, so the jth entry of $M^*\mathbf{y}$ is $\mathbf{m}_j^*\mathbf{y} = \langle \mathbf{y}, \mathbf{m}_j\rangle$. So $M^*\mathbf{y} = \mathbf{0}$ if and only if $\mathbf{y}$ is orthogonal to each column, and by Proposition 6.3.9(c) that means $\mathbf{y} \perp \operatorname{span}(\mathbf{m}_1, \ldots, \mathbf{m}_n) = \operatorname{im} M$.

Let $\mathbf{m}_1, \ldots, \mathbf{m}_n$ be the columns of M. By Definition 6.3.26, row j of M^* is $\mathbf{m}_j^*$, so by Proposition 2.4.8 and Definition 6.1.8 the jth entry of $M^*\mathbf{y}$ is

$$\mathbf{m}_j^*\mathbf{y} = \mathbf{y} \cdot \mathbf{m}_j = \langle \mathbf{y}, \mathbf{m}_j\rangle.$$

Hence $M^*\mathbf{y} = \mathbf{0}$ if and only if $\mathbf{y} \perp \mathbf{m}_j$ for every j. By Proposition 6.3.9(c) and Proposition 2.6.18, that holds if and only if $\mathbf{y} \perp \operatorname{span}(\mathbf{m}_1, \ldots, \mathbf{m}_n) = \operatorname{im} M$. □

For real M this says $(\operatorname{im} M)^\perp = \ker M^\mathsf{T}$; applied to M^T, it is the statement $\ker A = R^\perp$ of Example 6.4.23(b), where R is the span of the rows. Theorem 6.6.21 extends the lemma to linear maps between arbitrary inner product spaces.

Theorem 6.5.5 (The normal equations). Let $A \in \mathbb{F}^{m\times n}$, $\mathbf{b} \in \mathbb{F}^m$, and $\hat{\mathbf{x}} \in \mathbb{F}^n$. The following are equivalent:

(i) $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$;
(ii) $\mathbf{b} - A\hat{\mathbf{x}} \perp \operatorname{im} A$, that is, the residual is orthogonal to every column of A;
(iii) $A^*A\hat{\mathbf{x}} = A^*\mathbf{b}$.

In particular, the **normal equations** $A^*A\mathbf{x} = A^*\mathbf{b}$ always have a solution, and their solutions are exactly the least-squares solutions of $A\mathbf{x} = \mathbf{b}$.

Proof (by projection). ROUTINE

Uses. Proposition 6.5.2 (Least squares is projection), Theorem 6.4.6 (The projection is characterized by orthogonality), Theorem 6.4.20 (Orthonormal bases exist), Lemma 6.5.4 (The orthogonal complement of a column space)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. By Proposition 6.5.2(a), (i) says that $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. Which two conditions characterize this projection?
Hint 2. $A\hat{\mathbf{x}}$ lies in $\operatorname{im} A$ no matter what, so only the second condition of Theorem 6.4.6 is in question.
Hint 3. For (ii) ⇔ (iii), apply Lemma 6.5.4 to the vector $\mathbf{b} - A\hat{\mathbf{x}}$.
Sketch. (i) holds iff $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ (Proposition 6.5.2(a)). Since $A\hat{\mathbf{x}} \in \operatorname{im} A$ automatically, Theorem 6.4.6 says this holds iff $\mathbf{b} - A\hat{\mathbf{x}} \perp \operatorname{im} A$, which is (ii). By Lemma 6.5.4, (ii) holds iff $A^*(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, which is (iii). Existence is Proposition 6.5.2(b).

(i) ⇔ (ii). By Proposition 6.5.2(a), (i) holds if and only if $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. Fix an orthonormal basis of $\operatorname{im} A$ (Theorem 6.4.20). By Theorem 6.4.6, $\operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ is the unique vector $\mathbf{p}$ with $\mathbf{p} \in \operatorname{im} A$ and $\mathbf{b} - \mathbf{p} \perp \operatorname{im} A$. The vector $\mathbf{p} = A\hat{\mathbf{x}}$ lies in $\operatorname{im} A$ in any case, so $A\hat{\mathbf{x}} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ if and only if $\mathbf{b} - A\hat{\mathbf{x}} \perp \operatorname{im} A$.

(ii) ⇔ (iii). By Lemma 6.5.4, $\mathbf{b} - A\hat{\mathbf{x}} \perp \operatorname{im} A$ if and only if $A^*(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, that is, $A^*A\hat{\mathbf{x}} = A^*\mathbf{b}$.

Finally, least-squares solutions exist by Proposition 6.5.2(b), and by (i) ⇔ (iii) they are exactly the solutions of the normal equations. □

The second proof does not use projections at all. It makes the sliding argument of the intuition precise, and the move it uses applies to many minimization problems.

Technique 6.5.6 (Perturb a minimizer). Suppose $\hat{\mathbf{x}}$ minimizes a real-valued quantity $f(\mathbf{x})$ over a vector space. Then for every direction $\mathbf{h}$, the function $t \mapsto f(\hat{\mathbf{x}} + t\mathbf{h})$ of a real variable t has a minimum at $t = 0$. When f is the squared norm of a vector that depends linearly (or affinely) on $\mathbf{x}$, this function is a

quadratic polynomial $\alpha + \beta t + \gamma t^2$ with $\gamma \geq 0$, and a minimum at $t = 0$ forces the linear coefficient β to vanish: if $\beta \neq 0$, a small t of the opposite sign makes $\beta t + \gamma t^2 < 0$. Doing this for every $\mathbf{h}$ gives an equation that $\hat{\mathbf{x}}$ must satisfy. Choosing $\mathbf{h}$ cleverly, usually equal to the vector you want to show is zero, turns "$\beta = 0$" into a statement $\|\mathbf{s}\|^2 = 0$ (Technique 6.1.36). In the other direction, if the linear coefficient vanishes for every $\mathbf{h}$, the expansion shows that $f(\hat{\mathbf{x}} + \mathbf{h}) \geq f(\hat{\mathbf{x}})$, so $\hat{\mathbf{x}}$ is a minimizer.

This is "the derivative vanishes at a minimum" without derivatives, and with the converse built in. Reach for it when a vector is defined as the minimizer (or maximizer) of something and you need an equation it satisfies. The canonical example is the second proof of Theorem 6.5.5. The same move gives the minimal-norm solution in Section 6.5.7, and in Section 12.3 it shows that a maximizer of $\langle T\mathbf{v}, \mathbf{v}\rangle$ over unit vectors is an eigenvector.

Used in. Theorem 6.5.5, Exercise 6.5.39, Exercise 6.7.6, Exercise 6.7.13.

Proof of Theorem 6.5.5 (by perturbation). STANDARD

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.3.27 (Properties of the conjugate transpose), Definition 6.1.2 (Inner product), Lemma 6.5.4 (The orthogonal complement of a column space), Proposition 6.5.2 (Least squares is projection)
Techniques. Perturb a minimizer (Technique 6.5.6), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Let $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ and $\mathbf{s} = A^*\mathbf{r}$. Expand $\|\mathbf{b} - A(\hat{\mathbf{x}} + \mathbf{h})\|^2 = \|\mathbf{r} - A\mathbf{h}\|^2$ for an arbitrary $\mathbf{h} \in \mathbb{F}^n$.
Hint 2. Move A across the inner product: $\langle \mathbf{r}, A\mathbf{h}\rangle = \langle A^*\mathbf{r}, \mathbf{h}\rangle = \langle \mathbf{s}, \mathbf{h}\rangle$. You should get $\|\mathbf{r}\|^2 - 2\,\mathrm{Re}\langle \mathbf{s}, \mathbf{h}\rangle + \|A\mathbf{h}\|^2$.
Hint 3. If $\mathbf{s} = \mathbf{0}$, the middle term is gone. If $\hat{\mathbf{x}}$ is a minimizer, take $\mathbf{h} = t\mathbf{s}$ with $t > 0$ small.
Sketch. With $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ and $\mathbf{s} = A^*\mathbf{r}$, for every $\mathbf{h}$: $\|\mathbf{b} - A(\hat{\mathbf{x}} + \mathbf{h})\|^2 = \|\mathbf{r}\|^2 - 2\,\mathrm{Re}\langle \mathbf{s}, \mathbf{h}\rangle + \|A\mathbf{h}\|^2$. If $\mathbf{s} = \mathbf{0}$, this is at least $\|\mathbf{r}\|^2$, so $\hat{\mathbf{x}}$ is a minimizer. If $\hat{\mathbf{x}}$ is a minimizer, put $\mathbf{h} = t\mathbf{s}$: then $2t\|\mathbf{s}\|^2 \leq t^2\|A\mathbf{s}\|^2$ for all $t > 0$, so $\|\mathbf{s}\|^2 \leq (t/2)\|A\mathbf{s}\|^2$ for all $t > 0$, which forces $\mathbf{s} = \mathbf{0}$. So (i) $\Leftrightarrow$ (iii); and (ii) $\Leftrightarrow$ (iii) is Lemma 6.5.4.

Let $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$ and $\mathbf{s} = A^*\mathbf{r} = A^*\mathbf{b} - A^*A\hat{\mathbf{x}}$, so that (iii) says $\mathbf{s} = \mathbf{0}$. Every $\mathbf{x} \in \mathbb{F}^n$ has the form $\hat{\mathbf{x}} + \mathbf{h}$, and $\mathbf{b} - A(\hat{\mathbf{x}} + \mathbf{h}) = \mathbf{r} - A\mathbf{h}$. By Proposition 6.3.27(d) and conjugate symmetry,

$$\langle \mathbf{r}, A\mathbf{h}\rangle = \overline{\langle A\mathbf{h}, \mathbf{r}\rangle} = \overline{\langle \mathbf{h}, A^*\mathbf{r}\rangle} = \langle \mathbf{s}, \mathbf{h}\rangle,$$

so Proposition 6.1.29 gives, for every $\mathbf{h} \in \mathbb{F}^n$,

$$\|\mathbf{b} - A(\hat{\mathbf{x}} + \mathbf{h})\|^2 = \|\mathbf{r}\|^2 - 2\,\mathrm{Re}\langle \mathbf{s}, \mathbf{h}\rangle + \|A\mathbf{h}\|^2. \tag{6.5.1}$$

(iii) $\Rightarrow$ (i). If $\mathbf{s} = \mathbf{0}$, then (6.5.1) reads $\|\mathbf{b} - A(\hat{\mathbf{x}} + \mathbf{h})\|^2 = \|\mathbf{r}\|^2 + \|A\mathbf{h}\|^2 \geq \|\mathbf{r}\|^2$ for every $\mathbf{h}$. So $\hat{\mathbf{x}}$ is a least-squares solution.

(i) $\Rightarrow$ (iii). Suppose $\hat{\mathbf{x}}$ is a least-squares solution, and take $\mathbf{h} = t\mathbf{s}$ with t real and positive. Then $\langle \mathbf{s}, \mathbf{h}\rangle = t\|\mathbf{s}\|^2$, and (6.5.1) together with minimality gives

$$\|\mathbf{r}\|^2 \leq \|\mathbf{r}\|^2 - 2t\|\mathbf{s}\|^2 + t^2\|A\mathbf{s}\|^2, \quad \text{that is,} \quad 2\|\mathbf{s}\|^2 \leq t\|A\mathbf{s}\|^2$$

after dividing by $t > 0$. If $\mathbf{s} \neq \mathbf{0}$, choose $t > 0$ so small that $t\|A\mathbf{s}\|^2 < 2\|\mathbf{s}\|^2$ (any $t > 0$ works if $A\mathbf{s} = \mathbf{0}$); this contradicts the inequality. So $\mathbf{s} = \mathbf{0}$.

(ii) $\Leftrightarrow$ (iii) is Lemma 6.5.4 applied to $\mathbf{r}$, as in the first proof. The perturbation argument characterizes the least-squares solutions but does not produce one, so existence still comes from Proposition 6.5.2(b); Exercise 6.5.45 gives an algebraic existence proof instead. □

The first proof is the picture of Figure 6.5.1: least squares is projection, and projections are recognized by orthogonality. The second is the calculus approach of "set the gradient to zero", done without derivatives: $\mathbf{s} = A^*(\mathbf{b} - A\hat{\mathbf{x}})$ is, up to a factor -2, the gradient of $\|\mathbf{b} - A\mathbf{x}\|^2$. It does not use Theorem 6.4.33, and it gives a sharper statement. If $\hat{\mathbf{x}}$ solves the normal equations, then by (6.5.1) with $\mathbf{h} = \mathbf{x} - \hat{\mathbf{x}}$,

$$\begin{aligned}\|\mathbf{b} - A\mathbf{x}\|^2 &= \|\mathbf{b} - A\hat{\mathbf{x}}\|^2 \\ &\quad + \|A(\mathbf{x} - \hat{\mathbf{x}})\|^2\end{aligned} \tag{6.5.2}$$

for every $\mathbf{x} \in \mathbb{F}^n$. This is Pythagoras for the right triangle of Figure 6.5.1 with vertices $\mathbf{b}$, $A\hat{\mathbf{x}}$, $A\mathbf{x}$, and it says exactly how fast the error grows as $\mathbf{x}$ moves away from $\hat{\mathbf{x}}$.

Example 6.5.7 (Three points, by the normal equations). Return to Example 6.5.3, with $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$ and $\mathbf{y} = (1, 1, 4)$. The entries of $A^{\mathsf{T}}A$ are the dot products of the columns $\mathbf{1} = (1, 1, 1)$ and $\mathbf{x} = (0, 1, 2)$ (Theorem 6.3.29), and $A^{\mathsf{T}}\mathbf{y}$ lists their dot products with $\mathbf{y}$:

$$A^{\mathsf{T}}A = \begin{pmatrix} \mathbf{1}\cdot\mathbf{1} & \mathbf{x}\cdot\mathbf{1} \\ \mathbf{1}\cdot\mathbf{x} & \mathbf{x}\cdot\mathbf{x} \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}, \qquad A^{\mathsf{T}}\mathbf{y} = \begin{pmatrix} \mathbf{y}\cdot\mathbf{1} \\ \mathbf{y}\cdot\mathbf{x} \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}.$$

The normal equations $3c_0 + 3c_1 = 6$, $3c_0 + 5c_1 = 9$ give $2c_1 = 3$, so $c_1 = 3/2$ and $c_0 = 1/2$, as before. This time no orthogonal basis was needed: only the inner products of the columns with each other and with $\mathbf{y}$.

Exercise 6.5.8. ROUTINE In $\mathbb{C}^3$, let $\mathbf{a} = (1, i, 1)$ and $\mathbf{b} = (1, 1, 0)$, and let A be the 3×1 matrix with column $\mathbf{a}$.

(a) Find the least-squares solution $\hat{z} \in \mathbb{C}$ of $Az = \mathbf{b}$, that is, of $z\mathbf{a} = \mathbf{b}$, and check that the residual is orthogonal to $\mathbf{a}$.

(b) A student solves $A^{\mathsf{T}}Az = A^{\mathsf{T}}\mathbf{b}$ instead. Show that the resulting z is not a least-squares solution.

Solution on page 1511

When is the solution unique?

Intuition (Same kernel). The coefficient matrix A^*A of the normal equations is the Gram matrix of the columns of A (Theorem 6.3.29). By Proposition 6.5.2(c) and Theorem 6.5.5, the normal equations have exactly one solution precisely when the columns of A are independent. Since A^*A is square, we expect it to be invertible exactly when the columns of A are independent, and the underlying statement is that A^*A and A have the same kernel.

One containment is free: if $A\mathbf{x} = \mathbf{0}$, then $A^*A\mathbf{x} = \mathbf{0}$. For the other we must get from $A^*(A\mathbf{x}) = \mathbf{0}$ back to $A\mathbf{x} = \mathbf{0}$, and A^* can have a large kernel, so we cannot cancel it. Two things rescue us. Multiplying by $\mathbf{x}^*$ turns $\mathbf{x}^*A^*A\mathbf{x}$ into $\|A\mathbf{x}\|^2$, so the norm vanishes (Technique 6.1.36). Geometrically, $A^*(A\mathbf{x}) = \mathbf{0}$ says that $A\mathbf{x}$ is orthogonal to $\operatorname{im} A$ (Lemma 6.5.4), while $A\mathbf{x}$ lies in $\operatorname{im} A$.

Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$). Let $A \in \mathbb{F}^{m\times n}$.

(a) $\ker(A^*A) = \ker A$.

(b) $\operatorname{rank}(A^*A) = \operatorname{rank} A$.

(c) A^*A is invertible if and only if the columns of A are linearly independent, that is, if and only if $\operatorname{rank} A = n$.

Proof (by the norm). STANDARD

Uses. Proposition 6.3.27 (Properties of the conjugate transpose), Definition 6.1.8 (Dot product), Definition 6.1.2 (Inner product), Theorem 4.2.11 (Rank–nullity), Corollary 3.5.25 (The invertible matrix theorem), Corollary 2.6.24 (The columns decide), Theorem 3.5.12 (First Fundamental Theorem, completed)
Techniques. Show a vector is zero via its norm (Technique 6.1.36), Double containment (Technique 1.3.14)
Hint 1. For (a), the containment $\ker A \subseteq \ker(A^*A)$ is immediate. For the other, suppose $A^*A\mathbf{x} = \mathbf{0}$ and multiply on the left by $\mathbf{x}^*$.
Hint 2. $\mathbf{x}^*A^*A\mathbf{x} = (A\mathbf{x})^*(A\mathbf{x}) = \|A\mathbf{x}\|^2$.
Hint 3. For (b), both matrices have n columns; apply rank–nullity to each. For (c), a square matrix is invertible exactly when its kernel is $\{\mathbf{0}\}$.

Sketch. If $A^*A\mathbf{x} = \mathbf{0}$, then $\|A\mathbf{x}\|^2 = \mathbf{x}^*A^*A\mathbf{x} = 0$, so $A\mathbf{x} = \mathbf{0}$; the reverse containment is clear. Both matrices have n columns, so rank–nullity turns equal kernels into equal ranks. The square matrix A^*A is invertible iff its kernel, which is $\ker A$, is $\{\mathbf{0}\}$, iff the columns of A are independent.

(a) If $A\mathbf{x} = \mathbf{0}$, then $A^*A\mathbf{x} = A^*\mathbf{0} = \mathbf{0}$. Conversely, suppose $A^*A\mathbf{x} = \mathbf{0}$. By Proposition 6.3.27(b) and (a), $\mathbf{x}^*A^* = (A\mathbf{x})^*$, so, identifying 1×1 matrices with scalars (Definition 6.1.8),

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*(A^*A\mathbf{x}) = \mathbf{x}^*\mathbf{0} = 0.$$

By definiteness, $A\mathbf{x} = \mathbf{0}$. So $\ker(A^*A) = \ker A$.

(b) Both $A^*A \in \mathbb{F}^{n\times n}$ and A have n columns, so by Theorem 4.2.11 and (a),

$$\begin{aligned}\operatorname{rank}(A^*A) &= n - \dim\ker(A^*A)\\ &= n - \dim\ker A = \operatorname{rank} A.\end{aligned}$$

(c) The matrix A^*A is square, so it is invertible if and only if $\ker(A^*A) = \{\mathbf{0}\}$ (Corollary 3.5.25). By (a), that is $\ker A = \{\mathbf{0}\}$, which holds if and only if the columns of A are linearly independent (Corollary 2.6.24), and if and only if $\operatorname{rank} A = n$ (Theorem 3.5.12). □

Proof (by orthogonality). ROUTINE

Uses. Lemma 6.5.4 (The orthogonal complement of a column space), Proposition 6.3.9 (Properties of $A^\perp$), Theorem 4.2.11 (Rank–nullity), Corollary 3.5.25 (The invertible matrix theorem), Corollary 2.6.24 (The columns decide)
Techniques. Double containment (Technique 1.3.14)
Hint 1. Suppose $A^*(A\mathbf{x}) = \mathbf{0}$. What does Lemma 6.5.4 say about the vector $A\mathbf{x}$?
Hint 2. $A\mathbf{x}$ lies in $\operatorname{im} A$ and in $(\operatorname{im} A)^\perp$.
Sketch. If $A^*(A\mathbf{x}) = \mathbf{0}$, then $A\mathbf{x} \in \ker A^* = (\operatorname{im} A)^\perp$ by Lemma 6.5.4, and also $A\mathbf{x} \in \operatorname{im} A$. So $A\mathbf{x} \in \operatorname{im} A \cap (\operatorname{im} A)^\perp = \{\mathbf{0}\}$. Parts (b) and (c) follow from (a) as in the first proof.

As before, $\ker A \subseteq \ker(A^*A)$. Conversely, let $A^*A\mathbf{x} = \mathbf{0}$, and put $\mathbf{y} = A\mathbf{x}$. Then $A^*\mathbf{y} = \mathbf{0}$, so $\mathbf{y} \in (\operatorname{im} A)^\perp$ by Lemma 6.5.4. But also $\mathbf{y} \in \operatorname{im} A$. Hence $\mathbf{y} \in \operatorname{im} A \cap (\operatorname{im} A)^\perp = \{\mathbf{0}\}$ (Proposition 6.3.9(d)), that is, $A\mathbf{x} = \mathbf{0}$. This proves (a), and (b) and (c) follow from (a) exactly as in the first proof. □

The first proof is the one that generalizes to operators: the same computation shows $\ker(T^*T) = \ker T$ for a linear map T between inner product spaces (Section 6.6). It is also Proposition 6.3.30(b) for the columns of A, whose Gram matrix is A^*A. The second proof is Theorem 6.5.5 with $\mathbf{b} = \mathbf{0}$: the least-squares solutions of $A\mathbf{x} = \mathbf{0}$ are the solutions of $A\mathbf{x} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{0}) = \mathbf{0}$.

Remark 6.5.10 (The conjugate is essential). Both proofs used that $\mathbf{v}^*\mathbf{v} = \|\mathbf{v}\|^2$ vanishes only for $\mathbf{v} = \mathbf{0}$. Without the conjugate the proposition fails. For $A = \begin{pmatrix}1\\ i\end{pmatrix}$ over $\mathbb{C}$, $\operatorname{rank} A = 1$ but $A^{\mathsf{T}}A = 1 + i^2 = 0$. Over $\mathbb{F}_2$, where there is no inner product at all (Section 6.1), $A = \begin{pmatrix}1\\ 1\end{pmatrix}$ has $A^{\mathsf{T}}A = 1 + 1 = 0$. Over $\mathbb{R}$, $A^* = A^{\mathsf{T}}$, and the proposition says $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$.

Corollary 6.5.11 (The least-squares solution for independent columns). If the columns of $A \in \mathbb{F}^{m\times n}$ are linearly independent, then for every $\mathbf{b} \in \mathbb{F}^m$ the system $A\mathbf{x} = \mathbf{b}$ has exactly one least-squares solution,

$$\hat{\mathbf{x}} = (A^*A)^{-1}A^*\mathbf{b}.$$

If moreover $m = n$, then A is invertible and $\hat{\mathbf{x}} = A^{-1}\mathbf{b}$.

Proof. ROUTINE

Uses. Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$), Theorem 6.5.5 (The normal equations), Corollary 3.5.25 (The invertible matrix theorem), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 2.5.17 (Algebra of inverses)
Hint 1. By Proposition 6.5.9(c), A^*A is invertible. What does that say about the normal equations?
Hint 2. For $m = n$, write $(A^*A)^{-1} = A^{-1}(A^*)^{-1}$.

Sketch. A^*A is invertible (Proposition 6.5.9(c)), so the normal equations have the single solution $(A^*A)^{-1}A^*\mathbf{b}$, and by Theorem 6.5.5 it is the unique least-squares solution. If A is square, it is invertible, and $(A^*A)^{-1}A^* = A^{-1}(A^*)^{-1}A^* = A^{-1}$.

By Proposition 6.5.9(c), A^*A is invertible, so the normal equations $A^*A\mathbf{x} = A^*\mathbf{b}$ have exactly one solution, $(A^*A)^{-1}A^*\mathbf{b}$. By Theorem 6.5.5, the least-squares solutions are exactly the solutions of the normal equations.

If $m = n$, then A is a square matrix with independent columns, so it is invertible (Corollary 2.6.24, Corollary 3.5.25), and so is A^* (Proposition 6.3.27(c)). By Proposition 2.5.17, $(A^*A)^{-1} = A^{-1}(A^*)^{-1}$, so $\hat{\mathbf{x}} = A^{-1}(A^*)^{-1}A^*\mathbf{b} = A^{-1}\mathbf{b}$, which is no surprise: $A\mathbf{x} = \mathbf{b}$ then has an exact solution. □

The formula is good for theory, but to compute $\hat{\mathbf{x}}$ by hand one solves the normal equations by row reduction rather than inverting A^*A, as in Example 6.5.7.

Example 6.5.12 (Dependent columns). Let

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

The second column is twice the first, so rank $A = 1$ and the least-squares solution is not unique. The normal equations are

$$A^{\mathsf{T}}A = \begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix},$$

so they reduce to $6x_1 + 12x_2 = 3$ (the second equation is twice the first). Their solutions are $\hat{\mathbf{x}} = (1/2, 0) + t(-2, 1)$ for $t \in \mathbb{R}$, a line in $\mathbb{R}^2$ parallel to $\ker A = \operatorname{span}((-2, 1))$, as Proposition 6.5.2(b) predicts. Each of them gives the same vector $A\hat{\mathbf{x}} = (1/2)(1, 1, 2)$, which is the projection of $\mathbf{b}$ onto the line $\operatorname{im} A = \operatorname{span}((1, 1, 2))$, namely $(3/6)(1, 1, 2)$. Section 6.5.7 singles out one of these solutions.

6.5.3 Projection onto a column space

Theorem 6.4.31 computes the matrix of Pr_W from an orthonormal basis of W. The normal equations give a formula from any basis, with no Gram–Schmidt step.

Proposition 6.5.13 (Projection onto a column space). Let $A \in \mathbb{F}^{m\times n}$ have linearly independent columns. Then

$$\mathrm{Pr}_{\operatorname{im} A}(\mathbf{b}) = A(A^*A)^{-1}A^*\mathbf{b} \quad \text{for every } \mathbf{b} \in \mathbb{F}^m.$$

That is, the matrix of $\mathrm{Pr}_{\operatorname{im} A}$ is $P = A(A^*A)^{-1}A^*$.

Proof (by the two conditions). ROUTINE

Uses. Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$), Theorem 6.4.6 (The projection is characterized by orthogonality), Lemma 6.5.4 (The orthogonal complement of a column space), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. A^*A is invertible by Proposition 6.5.9(c). Let $\mathbf{p} = A(A^*A)^{-1}A^*\mathbf{b}$. Check the two conditions of Theorem 6.4.6.
Hint 2. $\mathbf{p} = A\mathbf{c}$ for $\mathbf{c} = (A^*A)^{-1}A^*\mathbf{b}$, so $\mathbf{p} \in \operatorname{im} A$. For the other condition, compute $A^*(\mathbf{b} - \mathbf{p})$ and use Lemma 6.5.4.
Sketch. With $\mathbf{c} = (A^*A)^{-1}A^*\mathbf{b}$, the vector $\mathbf{p} = A\mathbf{c}$ lies in $\operatorname{im} A$, and $A^*(\mathbf{b} - \mathbf{p}) = A^*\mathbf{b} - (A^*A)(A^*A)^{-1}A^*\mathbf{b} = \mathbf{0}$, so $\mathbf{b} - \mathbf{p} \perp \operatorname{im} A$ (Lemma 6.5.4). By Theorem 6.4.6, $\mathbf{p} = \mathrm{Pr}_{\operatorname{im} A}(\mathbf{b})$.

By Proposition 6.5.9(c), A^*A is invertible, so P is defined. Let $\mathbf{b} \in \mathbb{F}^m$ and $\mathbf{p} = P\mathbf{b} = A\mathbf{c}$, where $\mathbf{c} = (A^*A)^{-1}A^*\mathbf{b}$. Then $\mathbf{p} \in \operatorname{im} A$, and

$$A^*(\mathbf{b} - \mathbf{p}) = A^*\mathbf{b} - (A^*A)(A^*A)^{-1}A^*\mathbf{b}$$

$$= A^*\mathbf{b} - A^*\mathbf{b} = \mathbf{0},$$

so $\mathbf{b} - \mathbf{p} \in (\operatorname{im} A)^\perp$ by Lemma 6.5.4. These are the two conditions of Theorem 6.4.6 (for an orthonormal basis of $\operatorname{im} A$, which exists by Theorem 6.4.20), so $\mathbf{p} = \operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$. □

Proof (by reduction to an orthonormal basis). STANDARD

> *Uses.* Theorem 6.4.31 (The matrix of an orthogonal projection), Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.3.29 (Orthonormal columns), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 2.5.17 (Algebra of inverses), Corollary 3.5.25 (The invertible matrix theorem), Corollary 2.6.24 (The columns decide)
> *Techniques.* Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
> *Hint 1.* Let Q have orthonormal columns forming a basis of $\operatorname{im} A$, so that $\operatorname{Pr}_{\operatorname{im} A} = QQ^*$. How many columns does Q have?
> *Hint 2.* Each column of A is fixed by QQ^*, so $A = QS$ with $S = Q^*A \in \mathbb{F}^{n\times n}$. Show that S is invertible, and substitute $A = QS$ into $A(A^*A)^{-1}A^*$.
> *Sketch.* The columns of A are a basis of $\operatorname{im} A$, so an orthonormal basis of $\operatorname{im} A$ has n vectors; let $Q \in \mathbb{F}^{m\times n}$ have them as columns, so $\operatorname{Pr}_{\operatorname{im} A} = QQ^*$ and $Q^*Q = I_n$. Since QQ^* fixes $\operatorname{im} A$, $A = QQ^*A = QS$ with $S = Q^*A$, which is invertible because $S\mathbf{x} = \mathbf{0}$ gives $A\mathbf{x} = \mathbf{0}$. Then $A^*A = S^*S$ and $A(A^*A)^{-1}A^* = QSS^{-1}(S^*)^{-1}S^*Q^* = QQ^*$.

The columns of A are a basis of $\operatorname{im} A$, so $\dim \operatorname{im} A = n$. By Theorem 6.4.20, $\operatorname{im} A$ has an orthonormal basis, which therefore has n vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Let $Q \in \mathbb{F}^{m\times n}$ have these columns. By Theorem 6.4.31, $\operatorname{Pr}_{\operatorname{im} A} = QQ^*$, and $Q^*Q = I_n$ by Theorem 6.3.29. Since $\operatorname{Pr}_{\operatorname{im} A}$ fixes every vector of $\operatorname{im} A$ (Proposition 6.4.29(b)), it fixes each column of A, so

$$A = QQ^*A = QS \quad \text{with } S = Q^*A \in \mathbb{F}^{n\times n}.$$

The square matrix S is invertible: if $S\mathbf{x} = \mathbf{0}$, then $A\mathbf{x} = QS\mathbf{x} = \mathbf{0}$, so $\mathbf{x} = \mathbf{0}$ because the columns of A are independent (Corollary 2.6.24), and Corollary 3.5.25 applies. So S^* is invertible too (Proposition 6.3.27(c)). Now $A^*A = S^*Q^*QS = S^*S$, so $(A^*A)^{-1} = S^{-1}(S^*)^{-1}$ by Proposition 2.5.17, and

$$A(A^*A)^{-1}A^* = QSS^{-1}(S^*)^{-1}S^*Q^* = QQ^* = \operatorname{Pr}_{\operatorname{im} A},$$

as desired. □

The first proof is Technique 6.4.7, the method Section 6.4 promised for this formula. The second shows how the formula relates to QQ^*: an arbitrary basis of $\operatorname{im} A$ is $A = QS$ for an invertible change-of-basis matrix S, and the factor $(A^*A)^{-1} = S^{-1}(S^*)^{-1}$ undoes it. When the columns of A are already orthonormal, $A^*A = I_n$ and the formula is AA^*, which is Theorem 6.4.31. For a single nonzero column $\mathbf{a}$ it is $\mathbf{a}(\mathbf{a}^*\mathbf{a})^{-1}\mathbf{a}^* = \mathbf{a}\mathbf{a}^*/(\mathbf{a}^*\mathbf{a})$, the projection onto a line from Section 6.4.

Some consequences can be read off directly. The matrix P satisfies $P^2 = A(A^*A)^{-1}(A^*A)(A^*A)^{-1}A^* = P$ and $P^* = P$, since A^*A is Hermitian, as Proposition 6.4.30 requires. The matrix $I_m - P$ is the projection onto $(\operatorname{im} A)^\perp = \ker A^*$ (Proposition 6.4.29(d), Lemma 6.5.4); it sends $\mathbf{b}$ to the residual. The formula needs independent columns, since otherwise A^*A is not invertible. For a matrix with dependent columns, apply it to the matrix of pivot columns of A, which are a basis of $\operatorname{im} A$ (Corollary 3.3.19). In statistics, P is called the *hat matrix*, because it turns the data $\mathbf{y}$ into the fitted values $\hat{\mathbf{y}} = P\mathbf{y}$.

Example 6.5.14 (A projection matrix from a non-orthogonal basis). The plane $W = \{\mathbf{x} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$ of Example 6.4.32 has the basis $(1, -1, 0)$, $(0, 1, -1)$, which is not orthogonal. With these as the columns of A,

$$A^{\mathsf{T}}A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \left(A^{\mathsf{T}}A\right)^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

and

$$A\left(A^{\mathsf{T}}A\right)^{-1}A^{\mathsf{T}} = \frac{1}{3}\begin{pmatrix}1&0\\-1&1\\0&-1\end{pmatrix}\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}1&-1&0\\0&1&-1\end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix}2&1\\-1&1\\-1&-2\end{pmatrix}\begin{pmatrix}1&-1&0\\0&1&-1\end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix},$$

the matrix found in Example 6.4.32 from the normal vector $(1,1,1)$.

Exercise 6.5.15. ROUTINE Let $\mathbf{a} = (1,2,2)$ and $\mathbf{b} = (3,0,3)$. Find the matrix P of the projection onto $\operatorname{span}(\mathbf{a})$ with Proposition 6.5.13, the least-squares solution of $x\mathbf{a} = \mathbf{b}$, and the least-squares error.

Solution on page 1512

6.5.4 Linear regression

The most common source of inconsistent systems is data. Suppose we have measured two real quantities m times, with results $(x_1, y_1), \ldots, (x_m, y_m)$, and have reason to expect y to be a linear function of x.

Definition 6.5.16 (Least-squares line). Let $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^2$ be data points. A **least-squares line**, or **regression line**, for the data is a line $y = c_0 + c_1 x$ for which

$$\sum_{i=1}^{m} (y_i - c_0 - c_1 x_i)^2$$

is as small as possible. The numbers $y_i - (c_0 + c_1 x_i)$ are the **residuals** of the line.

To turn this into a least-squares problem, let $\mathbf{1} = (1, \ldots, 1)$, $\mathbf{x} = (x_1, \ldots, x_m)$, and $\mathbf{y} = (y_1, \ldots, y_m)$ in $\mathbb{R}^m$, and let $A \in \mathbb{R}^{m\times 2}$ be the matrix with columns $\mathbf{1}$ and $\mathbf{x}$. For $\mathbf{c} = (c_0, c_1)$, the ith entry of $A\mathbf{c}$ is $c_0 + c_1 x_i$, so the sum in Definition 6.5.16 is $\|\mathbf{y} - A\mathbf{c}\|^2$. Hence $y = c_0 + c_1 x$ is a least-squares line if and only if $\mathbf{c}$ is a least-squares solution of $A\mathbf{c} = \mathbf{y}$, and the residuals of the line are the entries of the residual vector $\mathbf{y} - A\mathbf{c}$. (The lecture writes the line as $y = a + bx$, so $a = c_0$ and $b = c_1$. We keep the letter $\mathbf{b}$ for right-hand sides, and the indices match the interactive figures.)

As in Example 6.2.19, the answer is best expressed through the means $m_x = (x_1 + \cdots + x_m)/m$ and $m_y = (y_1 + \cdots + y_m)/m$ and the centered data

$$\tilde{\mathbf{x}} = \mathbf{x} - m_x\mathbf{1} = (x_1 - m_x, \ldots, x_m - m_x),$$
$$\tilde{\mathbf{y}} = \mathbf{y} - m_y\mathbf{1} = \left(y_1 - m_y, \ldots, y_m - m_y\right).$$

Proposition 6.5.17 (The least-squares line). Let $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^2$ be data points whose x_i are not all equal. Then there is exactly one least-squares line $y = c_0 + c_1 x$, given by

$$c_1 = \frac{\sum_{i=1}^{m}(x_i - m_x)\left(y_i - m_y\right)}{\sum_{i=1}^{m}(x_i - m_x)^2} = \frac{\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}}}{\tilde{\mathbf{x}} \cdot \tilde{\mathbf{x}}},$$
$$c_0 = m_y - c_1 m_x.$$

In particular, the line passes through the point $\left(m_x, m_y\right)$, and its residuals add up to 0. If all x_i are equal, there are infinitely many least-squares lines.

Proof (by solving the normal equations). ROUTINE

> *Uses.* Theorem 6.5.5 (The normal equations), Theorem 6.3.29 (Orthonormal columns), Technique 6.1.9 (A sum of nonnegative terms), Proposition 6.5.2 (Least squares is projection), Corollary 2.6.24 (The columns decide)
> *Hint 1.* Write out $A^{\mathsf{T}}A$ and $A^{\mathsf{T}}\mathbf{y}$ with sums: the entries are $\mathbf{1}\cdot\mathbf{1} = m$, $\mathbf{1}\cdot\mathbf{x} = \sum_i x_i$, and so on.
> *Hint 2.* Divide the first normal equation by m to get $c_0 = m_y - c_1 m_x$, and substitute it into the second.
> *Hint 3.* You will need $\sum_i (x_i - m_x)^2 = \sum_i x_i^2 - m m_x^2$ and $\sum_i (x_i - m_x)\big(y_i - m_y\big) = \sum_i x_i y_i - m m_x m_y$.
> *Sketch.* The normal equations are $m c_0 + (\sum x_i) c_1 = \sum y_i$ and $(\sum x_i) c_0 + \big(\sum x_i^2\big) c_1 = \sum x_i y_i$. The first is $c_0 + c_1 m_x = m_y$; substituting into the second gives $c_1\big(\sum x_i^2 - m m_x^2\big) = \sum x_i y_i - m m_x m_y$, that is, $c_1 \|\tilde{\mathbf{x}}\|^2 = \tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}$, and $\|\tilde{\mathbf{x}}\|^2 > 0$. The first equation also says that $\big(m_x, m_y\big)$ lies on the line, and that $\mathbf{1}\cdot(\mathbf{y} - A\mathbf{c}) = 0$.

By Theorem 6.3.29, the entries of $A^{\mathsf{T}}A$ and $A^{\mathsf{T}}\mathbf{y}$ are dot products:

$$A^{\mathsf{T}}A = \begin{pmatrix} m & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{y} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{pmatrix}.$$

So the normal equations are

$$m c_0 + \left(\sum_i x_i\right) c_1 = \sum_i y_i,$$
$$\left(\sum_i x_i\right) c_0 + \left(\sum_i x_i^2\right) c_1 = \sum_i x_i y_i.$$

Dividing the first by m gives $c_0 + c_1 m_x = m_y$, that is, $c_0 = m_y - c_1 m_x$. Substituting this into the second, and using $\sum_i x_i = m m_x$, gives

$$c_1\left(\sum_i x_i^2 - m m_x^2\right) = \sum_i x_i y_i - m m_x m_y.$$

Expanding the squares and products,

$$\begin{aligned} &\sum_i (x_i - m_x)^2 \\ &= \sum_i x_i^2 - 2 m_x \sum_i x_i + m m_x^2 \\ &= \sum_i x_i^2 - m m_x^2 \end{aligned}$$

and

$$\begin{aligned} &\sum_i (x_i - m_x)\big(y_i - m_y\big) \\ &= \sum_i x_i y_i - m_x \sum_i y_i - m_y \sum_i x_i + m m_x m_y \\ &= \sum_i x_i y_i - m m_x m_y, \end{aligned}$$

so the second equation reads $c_1 \|\tilde{\mathbf{x}}\|^2 = \tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}$. Since the x_i are not all equal, $\tilde{\mathbf{x}} \neq \mathbf{0}$, and $\|\tilde{\mathbf{x}}\|^2 > 0$ (Technique 6.1.9). So the normal equations have exactly one solution, the one in the statement, and by Theorem 6.5.5 it is the unique least-squares solution. The first normal equation, $m_y = c_0 + c_1 m_x$, says that $\big(m_x, m_y\big)$ lies on the line. It is also the equation $\mathbf{1}\cdot(\mathbf{y} - A\mathbf{c}) = 0$, so the residuals add up to 0.

If all x_i are equal, then $\mathbf{x}$ is a multiple of $\mathbf{1}$, the columns of A are dependent, and by Proposition 6.5.2 the least-squares solutions form a translate of the line $\ker A$. Different vectors $\mathbf{c}$ give different lines, so there are infinitely many least-squares lines. □

Proof (by centering, which orthogonalizes the columns). STANDARD

Uses. Corollary 6.4.10 (Projecting with an orthogonal list), Proposition 6.3.13 (Orthogonal lists are independent), Proposition 6.5.2 (Least squares is projection), Theorem 6.5.5 (The normal equations)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Compute $\mathbf{1}\cdot\tilde{\mathbf{x}}$. What does Gram–Schmidt do to the list $\mathbf{1},\mathbf{x}$?
Hint 2. $\mathbf{1}$ and $\tilde{\mathbf{x}}$ are an orthogonal basis of $\operatorname{im} A$. Project $\mathbf{y}$ onto $\operatorname{im} A$ with Corollary 6.4.10, and rewrite the result in terms of $\mathbf{1}$ and $\mathbf{x}$.
Sketch. $\mathbf{1}\cdot\tilde{\mathbf{x}}=\sum_i(x_i-m_x)=0$, and $\tilde{\mathbf{x}}=\mathbf{x}-m_x\mathbf{1}$ is the second vector Gram–Schmidt produces from $\mathbf{1},\mathbf{x}$. So $\mathbf{1},\tilde{\mathbf{x}}$ is an orthogonal basis of $\operatorname{im} A$, and $\operatorname{Pr}_{\operatorname{im} A}(\mathbf{y})=m_y\mathbf{1}+c_1\tilde{\mathbf{x}}=\big(m_y-c_1m_x\big)\mathbf{1}+c_1\mathbf{x}$ with $c_1=(\mathbf{y}\cdot\tilde{\mathbf{x}})/\|\tilde{\mathbf{x}}\|^2=(\tilde{\mathbf{y}}\cdot\tilde{\mathbf{x}})/\|\tilde{\mathbf{x}}\|^2$. This is $A\mathbf{c}$, so $\mathbf{c}$ is the least-squares solution, unique since the columns are independent.

First, $\mathbf{1}\cdot\tilde{\mathbf{x}}=\sum_i(x_i-m_x)=mm_x-mm_x=0$, so $\tilde{\mathbf{x}}\perp\mathbf{1}$. In fact $\tilde{\mathbf{x}}=\mathbf{x}-\big(\langle\mathbf{x},\mathbf{1}\rangle/\|\mathbf{1}\|^2\big)\mathbf{1}$ is what Gram–Schmidt produces from $\mathbf{1},\mathbf{x}$. Since the x_i are not all equal, $\tilde{\mathbf{x}}\neq\mathbf{0}$; so $\mathbf{1},\tilde{\mathbf{x}}$ is an orthogonal list of nonzero vectors, hence independent (Proposition 6.3.13). It spans $\operatorname{im} A=\operatorname{span}(\mathbf{1},\mathbf{x})$, because $\tilde{\mathbf{x}}\in\operatorname{span}(\mathbf{1},\mathbf{x})$ and $\mathbf{x}=\tilde{\mathbf{x}}+m_x\mathbf{1}$. So $\mathbf{1},\mathbf{x}$ is independent too.

By Corollary 6.4.10, with $\mathbf{y}\cdot\mathbf{1}=mm_y$ and $\mathbf{y}\cdot\tilde{\mathbf{x}}=\big(\tilde{\mathbf{y}}+m_y\mathbf{1}\big)\cdot\tilde{\mathbf{x}}=\tilde{\mathbf{y}}\cdot\tilde{\mathbf{x}}$,

$$\begin{aligned}\operatorname{Pr}_{\operatorname{im} A}(\mathbf{y})&=\frac{\mathbf{y}\cdot\mathbf{1}}{m}\mathbf{1}+\frac{\mathbf{y}\cdot\tilde{\mathbf{x}}}{\|\tilde{\mathbf{x}}\|^2}\tilde{\mathbf{x}}=m_y\mathbf{1}+c_1(\mathbf{x}-m_x\mathbf{1})\\&=\big(m_y-c_1m_x\big)\mathbf{1}+c_1\mathbf{x}=A\mathbf{c},\end{aligned}$$

with $c_1=(\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{x}}\|^2$ and $c_0=m_y-c_1m_x$. By Proposition 6.5.2, $\mathbf{c}$ is a least-squares solution, and the only one, since the columns of A are independent. The residual $\mathbf{y}-\operatorname{Pr}_{\operatorname{im} A}(\mathbf{y})$ is orthogonal to $\mathbf{1}$ (Theorem 6.5.5), so its entries add up to 0, and $c_0+c_1m_x=m_y$ says that $\big(m_x,m_y\big)$ is on the line. The case of equal x_i is handled as in the first proof. □

The second proof explains the shape of the formula. Centering the x-data is Gram–Schmidt applied to the columns $\mathbf{1},\mathbf{x}$, and once the columns are orthogonal, the projection splits into two one-dimensional problems: the best constant fit to the data is the mean m_y, and the slope is the coefficient of the shadow of $\tilde{\mathbf{y}}$ on $\tilde{\mathbf{x}}$. In the plane of the data, centering moves the origin to $\big(m_x,m_y\big)$, and c_1 is then the slope of the best line through the new origin (Exercise 6.5.20). When the y_i are not all equal, Example 6.2.19 gives $\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}=r\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|$, where r is the correlation coefficient, so $c_1=r\|\tilde{\mathbf{y}}\|/\|\tilde{\mathbf{x}}\|$: the slope is the correlation scaled by the ratio of the spreads. Exercise 6.5.40 shows that the squared least-squares error is $\big(1-r^2\big)\|\tilde{\mathbf{y}}\|^2$.

Example 6.5.18 (A spring). A spring hangs from a hook, and its length y (in cm) is measured under loads $x=0,1,2,3,4$ (in newtons):

$$(0,10.1),\quad(1,10.9),\quad(2,12.2),\quad(3,12.8),\quad(4,14.0).$$

Hooke's law predicts $y=c_0+c_1x$, where c_0 is the natural length and $1/c_1$ is the stiffness. Here $m_x=2$ and $m_y=60/5=12$, so

$$\tilde{\mathbf{x}}=(-2,-1,0,1,2),\quad\tilde{\mathbf{y}}=(-1.9,-1.1,0.2,0.8,2.0).$$

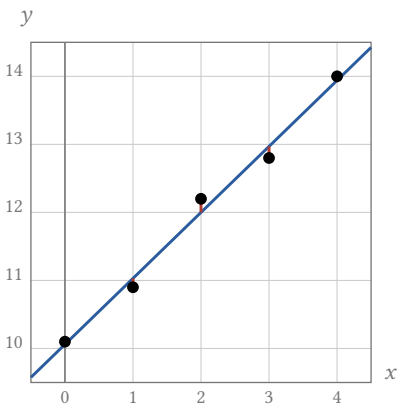
Then $\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}=3.8+1.1+0+0.8+4.0=9.7$ and $\tilde{\mathbf{x}}\cdot\tilde{\mathbf{x}}=10$, so by Proposition 6.5.17

$$c_1=0.97,\quad c_0=12-0.97\cdot2=10.06,$$

and the least-squares line is $y=10.06+0.97x$: a natural length of about 10.06 cm and a stiffness of about $1/0.97\approx1.03$ N/cm. In matrix form, the normal equations were

$$A^{\mathsf{T}}A\mathbf{c}=\begin{pmatrix}5&10\\10&30\end{pmatrix}\begin{pmatrix}c_0\\c_1\end{pmatrix}=\begin{pmatrix}60\\129.7\end{pmatrix}=A^{\mathsf{T}}\mathbf{y}.$$

The residuals are 0.04, −0.13, 0.20, −0.17, 0.06. They add up to 0, as Proposition 6.5.17 says, and the least-squares error is $\sqrt{0.091}\approx0.30$ cm.



The spring data of Example 6.5.18 and the least-squares line $y = 10.06 + 0.97x$; the residuals (red) are short. In the interactive version, drag a point to see the normal equations, the coefficients $\mathbf{c}$, and the sum of squared residuals change; add or remove points; and switch to a parabola or a cubic (Other models). (Interactive in the web version.)

Remark 6.5.19 (The calculus derivation). Statistics courses usually find the least-squares line with calculus. The function $S(c_0, c_1) = \sum_i (y_i - c_0 - c_1 x_i)^2$ has partial derivatives

$$\frac{\partial S}{\partial c_0} = -2 \sum_i (y_i - c_0 - c_1 x_i),$$
$$\frac{\partial S}{\partial c_1} = -2 \sum_i x_i (y_i - c_0 - c_1 x_i),$$

and setting both to 0 gives $\mathbf{1} \cdot (\mathbf{y} - A\mathbf{c}) = 0$ and $\mathbf{x} \cdot (\mathbf{y} - A\mathbf{c}) = 0$. These are the normal equations, in form (ii) of Theorem 6.5.5. The second proof of Theorem 6.5.5 is this computation without derivatives, and it shows at the same time that the critical point is a minimum, which the calculus argument has to check separately.

Exercise 6.5.20. ROUTINE

(a) Show that the line through the origin $y = cx$ that best fits data $(x_1, y_1), \ldots, (x_m, y_m)$, with the x_i not all 0, in the least-squares sense has $c = \left(\sum_i x_i y_i\right) / \left(\sum_i x_i^2\right)$.

(b) Currents $I = 0.5, 1.0, 1.5, 2.0$ (in amperes) through a resistor produce voltages $V = 1.0, 2.1, 2.9, 4.1$ (in volts). Ohm's law says $V = RI$. Estimate R.

Solution on page 1512

Other models

Nothing in the line-fitting problem was special to lines. What made it a least-squares problem is that the unknown coefficients enter linearly. A model

$$y \approx c_0 f_0(x) + c_1 f_1(x) + \cdots + c_n f_n(x),$$

with fixed functions $f_0, \ldots, f_n$ and unknown coefficients c_j, leads to the system $A\mathbf{c} \approx \mathbf{y}$ with entries $a_{ij} = f_j(x_i)$, however nonlinear the f_j are. The most important case is fitting a polynomial of degree at most d, where $f_j(x) = x^j$ and row i of A is $\left(1, x_i, x_i^2, \ldots, x_i^d\right)$. For $m = d + 1$ data points this is the square Vandermonde matrix of Definition 3.4.22.

Proposition 6.5.21 (Polynomial fits). Let $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^2$, let $d \geq 0$, and let $A \in \mathbb{R}^{m \times (d+1)}$ have rows $\left(1, x_i, x_i^2, \dots, x_i^d\right)$.

(a) A polynomial $p = c_0 + c_1 x + \cdots + c_d x^d$ of degree at most d minimizes $\sum_i (y_i - p(x_i))^2$ among all such polynomials if and only if $\mathbf{c} = (c_0, \dots, c_d)$ is a least-squares solution of $A\mathbf{c} = \mathbf{y}$.

(b) The minimizing polynomial is unique if and only if at least $d + 1$ of the numbers $x_1, \dots, x_m$ are distinct.

(c) If $x_1, \dots, x_m$ are distinct and $m = d + 1$, the minimizer is the interpolating polynomial of Theorem 3.4.24, and every residual is 0.

Proof. STANDARD

Uses. Definition 1.2.17 (Evaluation), Definition 1.2.13 (Polynomials), Proposition 6.5.2 (Least squares is projection), Lemma 3.4.23 (Too many roots), Corollary 3.4.19 (More unknowns than equations), Theorem 3.4.24 (Polynomial interpolation)
Techniques. Test injectivity at **0** (Technique 2.6.12)
Hint 1. The ith entry of $A\mathbf{c}$ is $p(x_i)$. By Proposition 6.5.2(c), uniqueness means $\ker A = \{\mathbf{0}\}$; what does $A\mathbf{c} = \mathbf{0}$ say about p?
Hint 2. If $d + 1$ of the x_i are distinct, use Lemma 3.4.23. If at most d are distinct, the system $A\mathbf{c} = \mathbf{0}$ has repeated equations; how many different ones?
Sketch. (a) $(A\mathbf{c})_i = p(x_i)$, so the sum is $\|\mathbf{y} - A\mathbf{c}\|^2$. (b) Uniqueness means $\ker A = \{\mathbf{0}\}$, and $A\mathbf{c} = \mathbf{0}$ says p vanishes at every x_i. With $d + 1$ distinct x_i, Lemma 3.4.23 forces $p = \mathbf{0}$. With $k \leq d$ distinct values, $A\mathbf{c} = \mathbf{0}$ is equivalent to a system of k equations in $d + 1 > k$ unknowns, which has a nonzero solution. (c) The interpolating polynomial has error 0, so it is a minimizer, the only one by (b).

(a) By Definition 1.2.17, the ith entry of $A\mathbf{c}$ is $c_0 + c_1 x_i + \cdots + c_d x_i^d = p(x_i)$, so $\sum_i (y_i - p(x_i))^2 = \|\mathbf{y} - A\mathbf{c}\|^2$. Polynomials of degree at most d correspond bijectively to their coefficient vectors $\mathbf{c} \in \mathbb{R}^{d+1}$ (Definition 1.2.13), so minimizing over p is minimizing over $\mathbf{c}$.

(b) By (a) and Proposition 6.5.2(c), the minimizer is unique if and only if $\ker A = \{\mathbf{0}\}$. A vector $\mathbf{c}$ lies in $\ker A$ exactly when the corresponding p satisfies $p(x_i) = 0$ for every i.

Suppose $d + 1$ of the x_i are distinct. If $\mathbf{c} \in \ker A$, then p has degree at most d and vanishes at $d + 1$ distinct points, so $p = \mathbf{0}$ by Lemma 3.4.23, and $\mathbf{c} = \mathbf{0}$.

Suppose instead that the x_i take only $k \leq d$ distinct values $z_1, \dots, z_k$. Rows of A with the same x_i are equal, so $A\mathbf{c} = \mathbf{0}$ has the same solutions as the system of k equations $c_0 + c_1 z_l + \cdots + c_d z_l^d = 0$, $l = 1, \dots, k$, in the $d + 1 > k$ unknowns $c_0, \dots, c_d$. By Corollary 3.4.19 it has a nonzero solution, so $\ker A \neq \{\mathbf{0}\}$.

(c) By Theorem 3.4.24 there is a polynomial p of degree at most $d = m - 1$ with $p(x_i) = y_i$ for every i. Its residuals are all 0, so it is a minimizer, and by (b) it is the only one. □

So the polynomial of degree at most d closest to the data is unique as soon as there are $d + 1$ distinct x-values, and it interpolates the data when there are exactly that many points. With more points it does not interpolate in general, and it is usually the better model: the high-degree interpolating polynomial passes through every measurement error as well.

Example 6.5.22 (A thrown ball). A ball is thrown upward, and its height h (in meters) is read off a video every half second:

$$t = 0, 0.5, 1, 1.5, 2, 2.5, \quad h = 1.0, 5.9, 8.1, 7.9, 5.4, 0.4.$$

Without air resistance, $h = h_0 + v_0 t - (g/2)t^2$ with initial height h_0, initial speed v_0, and gravitational acceleration g, so we fit $h \approx c_0 + c_1 t + c_2 t^2$. Row i of A is $\left(1, t_i, t_i^2\right)$, and the entries of $A^{\mathsf{T}}A$ and $A^{\mathsf{T}}\mathbf{h}$ are the sums $\sum_i t_i^k$ and $\sum_i t_i^k h_i$:

$$A^{\mathsf{T}}A = \begin{pmatrix} 6 & 7.5 & 13.75 \\ 7.5 & 13.75 & 28.125 \\ 13.75 & 28.125 & 61.1875 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{h} = \begin{pmatrix} 28.7 \\ 34.7 \\ 51.45 \end{pmatrix}.$$

There are six distinct t_i, so by Proposition 6.5.21 the solution is unique; row reduction gives

$$c_0 = \frac{59}{56} \approx 1.054, \quad c_1 = \frac{16699}{1400} \approx 11.93,$$
$$c_2 = -\frac{683}{140} \approx -4.879.$$

So the ball left the hand at about 1.05 m with speed about 11.9 m/s, and $g \approx -2c_2 \approx 9.76\text{m/s}^2$, close to the true 9.81. The residuals are all smaller than 0.11 m in absolute value (Figure 6.5.2).

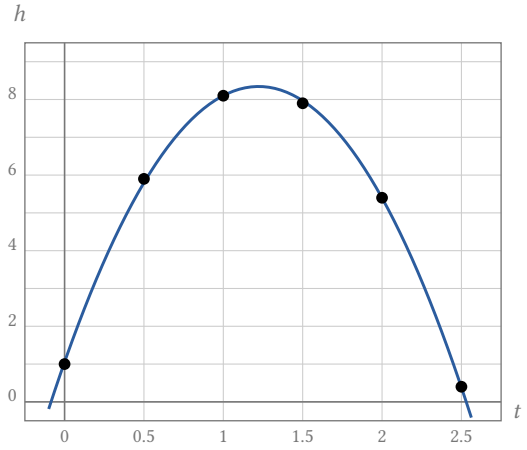


Figure 6.5.2. The heights of Example 6.5.22 and the least-squares parabola $h \approx 1.054 + 11.93t - 4.879t^2$. The residuals (red) are barely visible.

Two further variations use the same machinery. With several explanatory variables, say $y \approx c_0 + c_1u + c_2v$ with data (u_i, v_i, y_i), the matrix A has the columns $\mathbf{1}$, $\mathbf{u}$, $\mathbf{v}$; this is *multiple regression*. And some nonlinear models become linear after a change of variables. Exponential growth $y = Ce^{kx}$ is not linear in C and k, but $\ln y = \ln C + kx$ is linear in $\ln C$ and k, so one fits a line to the points $(x_i, \ln y_i)$. This minimizes the squared errors of $\ln y$, not of y, which roughly means relative rather than absolute errors; it is a different problem with a slightly different answer. How to choose the functions f_j, and how to keep a fit with many of them from following the noise, is taken up in Section 16.1.

6.5.5 Normal equations in an inner product space

The only place where the proof of Theorem 6.5.5 used $\mathbb{F}^m$ was in writing the inner products of the columns as the entries of A^*A. In a general inner product space, the Gram matrix (Definition 6.3.28) takes the place of A^*A.

Proposition 6.5.23 (Normal equations for a spanning list). Let V be an inner product space, let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ have Gram matrix G, let $W = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$, and let $\mathbf{v} \in V$. For $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{F}^n$,

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \operatorname{Pr}_W(\mathbf{v}) \quad \text{if and only if} \quad G\mathbf{c} = \mathbf{d},$$

where $\mathbf{d} = (\langle \mathbf{v}, \mathbf{v}_1 \rangle, \dots, \langle \mathbf{v}, \mathbf{v}_n \rangle)$. The system $G\mathbf{c} = \mathbf{d}$ always has a solution, and the solution is unique if and only if $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent.

Proof. ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.3.9 (Properties of $A^\perp$), Definition 6.3.28 (Gram matrix), Proposition 6.3.30 (Gram matrices detect independence), Corollary 3.5.25 (The invertible matrix theorem)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7), Gram matrices (Technique 6.3.31)

Hint 1. Let $\mathbf{w} = \sum_j c_j \mathbf{v}_j$. By Theorem 6.4.6, $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$ if and only if $\mathbf{v} - \mathbf{w} \perp W$. It suffices to test against $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
Hint 2. Expand $\langle \mathbf{w}, \mathbf{v}_i \rangle = \sum_j c_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle$ and recognize the ith entry of $G\mathbf{c}$.
Sketch. With $\mathbf{w} = \sum_j c_j \mathbf{v}_j$: $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$ iff $\langle \mathbf{v} - \mathbf{w}, \mathbf{v}_i \rangle = 0$ for all i, iff $\langle \mathbf{w}, \mathbf{v}_i \rangle = \langle \mathbf{v}, \mathbf{v}_i \rangle$ for all i, and $\langle \mathbf{w}, \mathbf{v}_i \rangle = \sum_j g_{ij} c_j = (G\mathbf{c})_i$. A solution exists because $\mathrm{Pr}_W(\mathbf{v}) \in W$ is some combination of the $\mathbf{v}_j$; it is unique iff G is invertible, iff the list is independent (Proposition 6.3.30).

The subspace W is finite-dimensional, so Pr_W is defined (Theorem 6.4.20). Let $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in W$. By Theorem 6.4.6, $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$ if and only if $\mathbf{v} - \mathbf{w} \perp W$, and by Proposition 6.3.9(c) it suffices to test this against the spanning list: $\langle \mathbf{v} - \mathbf{w}, \mathbf{v}_i \rangle = 0$, that is, $\langle \mathbf{w}, \mathbf{v}_i \rangle = \langle \mathbf{v}, \mathbf{v}_i \rangle$, for $i = 1, \ldots, n$. By linearity in the first slot and Definition 6.3.28,

$$\langle \mathbf{w}, \mathbf{v}_i \rangle = \sum_{j=1}^{n} c_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \sum_{j=1}^{n} g_{ij} c_j = (G\mathbf{c})_i.$$

So $\mathbf{w} = \mathrm{Pr}_W(\mathbf{v})$ if and only if $G\mathbf{c} = \mathbf{d}$.

Since $\mathrm{Pr}_W(\mathbf{v}) \in W$, it equals $\sum_j c_j \mathbf{v}_j$ for some $\mathbf{c}$, which then solves $G\mathbf{c} = \mathbf{d}$. The solution is unique if and only if the square matrix G is invertible (Corollary 3.5.25), that is, by Proposition 6.3.30, if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly independent. □

For $V = \mathbb{F}^m$ with the dot product and $\mathbf{v}_j = \mathbf{a}_j$, the columns of A, the Gram matrix is $G = A^*A$ (Theorem 6.3.29) and $d_i = \langle \mathbf{b}, \mathbf{a}_i \rangle = \mathbf{a}_i^* \mathbf{b}$ is the ith entry of $A^*\mathbf{b}$. So $G\mathbf{c} = \mathbf{d}$ is the system $A^*A\mathbf{c} = A^*\mathbf{b}$, and by Proposition 6.5.2(a) the proposition is Theorem 6.5.5 again. The general version applies when V is a space of functions, where it finds best approximations without orthogonalizing first.

Example 6.5.24 (The line closest to t^2). In $C[0,1]$ with $\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$ (Proposition 6.1.13), which line $c_0 + c_1 t$ minimizes $\int_0^1 \left(t^2 - c_0 - c_1 t\right)^2 dt$? Here $W = \mathrm{span}(1, t)$, and

$$G = \begin{pmatrix} \langle 1,1 \rangle & \langle t,1 \rangle \\ \langle 1,t \rangle & \langle t,t \rangle \end{pmatrix} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix},$$

$$\mathbf{d} = \begin{pmatrix} \langle t^2, 1 \rangle \\ \langle t^2, t \rangle \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/4 \end{pmatrix}.$$

The system $c_0 + c_1/2 = 1/3$, $c_0/2 + c_1/3 = 1/4$ gives $c_1 = 1$ and $c_0 = -1/6$. So the best line is $t - 1/6$, and its squared error is

$$\int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt = \frac{1}{180}.$$

The chord t through the endpoints of the graph has squared error $\int_0^1 \left(t^2 - t\right)^2 dt = 1/30$, six times as large: the best line is the chord lowered by $1/6$.

For polynomials of higher degree the Gram matrix of $1, t, \ldots, t^n$ is the Hilbert matrix of Exercise 6.3.54. It is invertible, but so close to singular that solving $G\mathbf{c} = \mathbf{d}$ numerically loses most of the accuracy for moderate n. That is why Approximating functions projects onto orthogonal polynomials instead.

ASIDE: *Weighted least squares* BEYOND 21-242

Sometimes the equations of an inconsistent system deserve different amounts of trust, for instance measurements made with instruments of different precision. Then one minimizes a weighted sum $\sum_i w_i |b_i - (A\mathbf{x})_i|^2$ with weights $w_i > 0$. With $D = \mathrm{diag}(w_1, \ldots, w_m)$, this sum is $\|\mathbf{b} - A\mathbf{x}\|_D^2$ for the weighted dot product $\langle \mathbf{u}, \mathbf{v} \rangle_D = \mathbf{v}^* D \mathbf{u}$ of Example 6.1.21. So the argument of Proposition 6.5.2, run in the inner product space $\left(\mathbb{F}^m, \langle \cdot, \cdot \rangle_D\right)$, shows that the minimizers are the $\mathbf{x}$ with $A\mathbf{x} = \mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b})$, where the projection is now taken with respect to $\langle \cdot, \cdot \rangle_D$. Since $\langle \mathbf{a}_j, \mathbf{a}_i \rangle_D = \mathbf{a}_i^* D \mathbf{a}_j$ and $\langle \mathbf{b}, \mathbf{a}_i \rangle_D = \mathbf{a}_i^* D \mathbf{b}$, Proposition 6.5.23 turns this into the **weighted normal equations**

$$A^*DA\mathbf{x} = A^*D\mathbf{b}.$$

The same works with D replaced by any positive definite matrix M, since $\mathbf{v}^*M\mathbf{u}$ is then an inner product (Proposition 6.1.22). In statistics one takes $w_i = 1/\sigma_i^2$, where σ_i is the standard deviation of the error in the i th measurement, so that precise measurements count for more. Exercise 6.5.48 gives a second derivation and a computation.

6.5.6 Least squares via QR

BEYOND 21-242

QR factorization is not covered in lecture (Section 6.4.5).

If $A = QR$ is the QR factorization of a matrix with independent columns (Theorem 6.4.38), the normal equations simplify, because $Q^*Q = I_n$.

Proposition 6.5.25 (Least squares via QR). Let $A \in \mathbb{F}^{m\times n}$ have linearly independent columns, and let $A = QR$ be its QR factorization. Then the least-squares solution of $A\mathbf{x} = \mathbf{b}$ is the unique solution of the upper triangular system

$$R\hat{\mathbf{x}} = Q^*\mathbf{b}.$$

Proof. ROUTINE

Uses. Theorem 6.4.38 (QR factorization), Theorem 6.3.29 (Orthonormal columns), Proposition 6.3.27 (Properties of the conjugate transpose), Corollary 3.5.25 (The invertible matrix theorem), Theorem 6.5.5 (The normal equations), Corollary 2.6.24 (The columns decide)
Hint 1. Substitute $A = QR$ into A^*A and $A^*\mathbf{b}$, using $Q^*Q = I_n$.
Hint 2. You get $R^*R\mathbf{x} = R^*Q^*\mathbf{b}$. Why may you cancel R^*?
Sketch. $A^*A = R^*Q^*QR = R^*R$ and $A^*\mathbf{b} = R^*Q^*\mathbf{b}$. R is invertible, since $R\mathbf{x} = \mathbf{0}$ gives $A\mathbf{x} = \mathbf{0}$, and then so is R^*. Multiplying by $(R^*)^{-1}$, the normal equations become $R\mathbf{x} = Q^*\mathbf{b}$.

By Theorem 6.4.38, Q has orthonormal columns, so $Q^*Q = I_n$ (Theorem 6.3.29), and by Proposition 6.3.27(b),

$$A^*A = R^*Q^*QR = R^*R, \quad A^*\mathbf{b} = R^*Q^*\mathbf{b}.$$

The square matrix R is invertible: if $R\mathbf{x} = \mathbf{0}$, then $A\mathbf{x} = QR\mathbf{x} = \mathbf{0}$, so $\mathbf{x} = \mathbf{0}$ by independence of the columns of A (Corollary 2.6.24, Corollary 3.5.25). Hence R^* is invertible (Proposition 6.3.27(c)), and $R^*R\mathbf{x} = R^*Q^*\mathbf{b}$ holds if and only if $R\mathbf{x} = Q^*\mathbf{b}$. By Theorem 6.5.5, the least-squares solution is the unique solution of $R\mathbf{x} = Q^*\mathbf{b}$. □

The two sides have a meaning. $Q^*\mathbf{b}$ is the coordinate vector of $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b}) = QQ^*\mathbf{b}$ with respect to the orthonormal basis formed by the columns of Q, and R converts coordinates with respect to the columns of A into coordinates with respect to the columns of Q, since $A\mathbf{x} = Q(R\mathbf{x})$. So $R\hat{\mathbf{x}} = Q^*\mathbf{b}$ says that $A\hat{\mathbf{x}}$ and $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b})$ have the same coordinates. The system is upper triangular and is solved by back substitution.

Example 6.5.26 (Least squares by QR). For $A = \begin{pmatrix} 1 & 3 \\ 2 & 3 \\ 2 & 0 \end{pmatrix}$, Exercise 6.4.64 found $Q = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$ and $R = \begin{pmatrix} 3 & 3 \\ 0 & 3 \end{pmatrix}$. Let $\mathbf{b} = (3, 3, 3)$. Then $Q^{\mathsf{T}}\mathbf{b} = (1/3)(3 + 6 + 6, 6 + 3 - 6) = (5, 1)$, and $R\hat{\mathbf{x}} = (5, 1)$ gives $3\hat{x}_2 = 1$ and $3\hat{x}_1 + 3\hat{x}_2 = 5$, so $\hat{\mathbf{x}} = (4/3, 1/3)$. The residual $\mathbf{b} - A\hat{\mathbf{x}} = (2/3, -2/3, 1/3)$ is orthogonal to both columns of A, and the least-squares error is 1.

In exact arithmetic there is no reason to prefer QR to the normal equations. In floating-point arithmetic there is: forming A^*A can destroy information that A still contains. Let $A = \begin{pmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$ with $\varepsilon = 10^{-8}$. Its columns are independent, and $A^{\mathsf{T}}A = \begin{pmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{pmatrix}$. But $\varepsilon^2 = 10^{-16}$ is smaller than half the spacing between double-precision numbers near 1 (about $2.2 \cdot 10^{-16}$), so the computer stores $1 + \varepsilon^2$ as 1 and $A^{\mathsf{T}}A$

as the singular matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. A QR factorization computed from A itself, with reflections as in Section 9.3, does not square the small entries. Numerical libraries solve least-squares problems this way, or with the singular value decomposition (Section 9.2).

6.5.7 The pseudoinverse BEYOND 21-242

The pseudoinverse is not covered in lecture; see LADR 6C. Section 9.2 completes the picture with the SVD.

When the columns of A are dependent, the least-squares solutions form a translate $\hat{\mathbf{x}} + \ker A$ of the kernel (Example 6.5.12). A natural way to pick one is to take the shortest.

Proposition 6.5.27 (The least-squares solution of minimal norm). Let $A \in \mathbb{F}^{m\times n}$ and $\mathbf{b} \in \mathbb{F}^m$, and let $K = \ker A$.

(a) Among the least-squares solutions of $A\mathbf{x} = \mathbf{b}$ there is exactly one of smallest norm, $\mathbf{x}^+$. It is the only least-squares solution in $K^\perp$, and $\mathbf{x}^+ = \hat{\mathbf{x}} - \mathrm{Pr}_K(\hat{\mathbf{x}})$ for every least-squares solution $\hat{\mathbf{x}}$.

(b) The map $\mathbb{F}^m \to \mathbb{F}^n$, $\mathbf{b} \mapsto \mathbf{x}^+$, is linear.

Proof. STANDARD

Uses. Proposition 6.5.2 (Least squares is projection), Theorem 6.4.24 (Orthogonal decomposition), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.3.9 (Properties of $A^\perp$), Theorem 6.5.5 (The normal equations)
Techniques. Orthogonal decomposition (Technique 6.2.4), Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Fix one least-squares solution $\hat{\mathbf{x}}$, so that all of them form $\hat{\mathbf{x}} + K$. Split $\hat{\mathbf{x}} = \mathrm{Pr}_K(\hat{\mathbf{x}}) + \mathbf{q}$ with $\mathbf{q} \in K^\perp$ (Theorem 6.4.24).
Hint 2. Every least-squares solution is $\mathbf{q} + \mathbf{z}$ with $\mathbf{z} \in K$. Compute $\|\mathbf{q} + \mathbf{z}\|^2$ with Pythagoras.
Hint 3. For (b), $\mathbf{x}^+$ is the only vector $\mathbf{x}$ with $\mathbf{x} \in K^\perp$ and $A^*A\mathbf{x} = A^*\mathbf{b}$. Check that $a\mathbf{x}^+ + \mathbf{x}'^+$ has both properties for $a\mathbf{b} + \mathbf{b}'$.
Sketch. The least-squares solutions are $\hat{\mathbf{x}} + K = \mathbf{q} + K$, where $\mathbf{q} = \hat{\mathbf{x}} - \mathrm{Pr}_K(\hat{\mathbf{x}}) \in K^\perp$. For $\mathbf{z} \in K$, $\|\mathbf{q} + \mathbf{z}\|^2 = \|\mathbf{q}\|^2 + \|\mathbf{z}\|^2$, smallest exactly at $\mathbf{z} = \mathbf{0}$. A least-squares solution in $K^\perp$ differs from $\mathbf{q}$ by a vector of $K \cap K^\perp = \{\mathbf{0}\}$. Linearity: $\mathbf{x}^+$ is characterized by $\mathbf{x}^+ \in K^\perp$ and the normal equations, and both conditions are preserved by linear combinations.

(a) Let $\hat{\mathbf{x}}$ be a least-squares solution (Proposition 6.5.2(b)). The kernel K is a subspace of the finite-dimensional space $\mathbb{F}^n$, so by Theorem 6.4.24 $\hat{\mathbf{x}} = \mathrm{Pr}_K(\hat{\mathbf{x}}) + \mathbf{q}$ with $\mathbf{q} = \hat{\mathbf{x}} - \mathrm{Pr}_K(\hat{\mathbf{x}}) \in K^\perp$. By Proposition 6.5.2(b), the least-squares solutions are the vectors $\hat{\mathbf{x}} + \mathbf{z} = \mathbf{q} + (\mathrm{Pr}_K(\hat{\mathbf{x}}) + \mathbf{z})$ with $\mathbf{z} \in K$; as $\mathbf{z}$ runs over K, so does $\mathrm{Pr}_K(\hat{\mathbf{x}}) + \mathbf{z}$. So the least-squares solutions are the vectors $\mathbf{q} + \mathbf{z}'$ with $\mathbf{z}' \in K$. Since $\mathbf{q} \perp \mathbf{z}'$, Theorem 6.3.3 gives

$$\|\mathbf{q} + \mathbf{z}'\|^2 = \|\mathbf{q}\|^2 + \|\mathbf{z}'\|^2 \geq \|\mathbf{q}\|^2,$$

with equality if and only if $\mathbf{z}' = \mathbf{0}$. So $\mathbf{x}^+ = \mathbf{q}$ is the unique least-squares solution of smallest norm, and $\mathbf{q} \in K^\perp$. If $\mathbf{x}$ is any least-squares solution in $K^\perp$, then $\mathbf{x} - \mathbf{q}$ lies in K (both are in $\mathbf{q} + K$) and in $K^\perp$ (a subspace), so $\mathbf{x} = \mathbf{q}$ by Proposition 6.3.9(d). Since $\hat{\mathbf{x}}$ was arbitrary, $\mathbf{x}^+ = \hat{\mathbf{x}} - \mathrm{Pr}_K(\hat{\mathbf{x}})$ for every least-squares solution $\hat{\mathbf{x}}$.

(b) By (a) and Theorem 6.5.5, $\mathbf{x}^+$ is the only vector $\mathbf{x}$ with $\mathbf{x} \in K^\perp$ and $A^*A\mathbf{x} = A^*\mathbf{b}$. Let $\mathbf{b}, \mathbf{b}' \in \mathbb{F}^m$ have minimal solutions $\mathbf{x}^+, \mathbf{x}'^+$, and let $a \in \mathbb{F}$. Then $a\mathbf{x}^+ + \mathbf{x}'^+ \in K^\perp$, since $K^\perp$ is a subspace, and $A^*A\big(a\mathbf{x}^+ + \mathbf{x}'^+\big) = aA^*\mathbf{b} + A^*\mathbf{b}' = A^*(a\mathbf{b} + \mathbf{b}')$. So $a\mathbf{x}^+ + \mathbf{x}'^+$ is the minimal least-squares solution for $a\mathbf{b} + \mathbf{b}'$, which is linearity. □

Definition 6.5.28 (Pseudoinverse). The **pseudoinverse** of $A \in \mathbb{F}^{m\times n}$ is the matrix $A^+ \in \mathbb{F}^{n\times m}$ of the linear map $\mathbf{b} \mapsto \mathbf{x}^+$ of Proposition 6.5.27. That is, $A^+\mathbf{b}$ is the least-squares solution of $A\mathbf{x} = \mathbf{b}$ of smallest norm.

It is also called the *Moore–Penrose inverse*. By Proposition 6.5.2(a), $AA^+ = \mathrm{Pr}_{\mathrm{im}\, A}$. If the columns of A are independent, the least-squares solution is unique, so $A^+ = (A^*A)^{-1}A^*$ by Corollary 6.5.11; in particular

$A^+ = A^{-1}$ when A is invertible. Exercise 6.5.49 shows $A^+A = \mathrm{Pr}_{(\ker A)^\perp}$. LADR defines A^+ differently, as the inverse of the restriction of T_A to $(\ker A)^\perp$, composed with $\mathrm{Pr}_{\mathrm{im}\,A}$; by Proposition 6.5.27(a) it is the same matrix.

Example 6.5.29 (A pseudoinverse). Let $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{pmatrix}$ as in Example 6.5.12, and $\mathbf{a} = (1, 1, 2)$, so that $A\mathbf{x} = (x_1 + 2x_2)\mathbf{a}$ and $\mathrm{im}\,A = \mathrm{span}(\mathbf{a})$. For $\mathbf{b} \in \mathbb{R}^3$, $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b}) = ((\mathbf{b}\cdot\mathbf{a})/6)\mathbf{a}$, so the least-squares solutions are the $\mathbf{x}$ with $x_1 + 2x_2 = (\mathbf{b}\cdot\mathbf{a})/6$. Here $\ker A = \mathrm{span}((-2, 1))$, and $(\ker A)^\perp = \mathrm{span}((1, 2))$. The least-squares solution in $(\ker A)^\perp$ is $s(1, 2)$ with $5s = (\mathbf{b}\cdot\mathbf{a})/6$. So

$$\mathbf{x}^+ = \frac{\mathbf{b}\cdot\mathbf{a}}{30}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A^+ = \frac{1}{30}\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

For $\mathbf{b} = (1, 2, 0)$, $\mathbf{x}^+ = (1/10, 1/5)$, the point of the line of solutions $(1/2, 0) + t(-2, 1)$ of Example 6.5.12 closest to $\mathbf{0}$.

Computing A^+ from the definition means solving a least-squares problem for every $\mathbf{b}$. The singular value decomposition gives A^+ by a formula, and Section 9.2 finishes the story there.

Exercises

Exercise 6.5.30. ROUTINE Find the least-squares solution and the least-squares error of the system

$$x_1 + x_2 = 2, \quad x_1 - x_2 = 0, \quad 2x_1 + x_2 = 5,$$

and check that the residual is orthogonal to both columns of the coefficient matrix.

Solution on page 1512

Exercise 6.5.31. ROUTINE Find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ in $\mathbb{C}^3$, where

$$A = \begin{pmatrix} 1 & i \\ 1 & -i \\ 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

and the least-squares error.

Solution on page 1513

Exercise 6.5.32. ROUTINE A used car of a certain model sells for the following prices y (in thousands of dollars) at ages x (in years):

$$(1, 21), \quad (2, 19), \quad (3, 16), \quad (4, 15), \quad (5, 12).$$

Find the least-squares line and its residuals, and use it to predict the price at age 6. What does it predict at age 11, and what does that say about extrapolating?

Solution on page 1513

Exercise 6.5.33. STANDARD For the data $(-1, 2)$, $(0, 1)$, $(1, 1)$, $(2, 4)$:

(a) find the least-squares line and its squared error;

(b) find the least-squares parabola $y = c_0 + c_1x + c_2x^2$ and its squared error;

(c) explain without computing why the error in (b) can be no larger than the one in (a).

Solution on page 1514

Exercise 6.5.34. STANDARD Let $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$, the matrix of Example 6.5.3. Find all $\mathbf{b} \in \mathbb{R}^3$ for which $(1, 1)$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$, and describe them geometrically. In terms of data, which data sets $(0, y_1), (1, y_2), (2, y_3)$ have the least-squares line $y = 1 + x$?

Solution on page 1514

Exercise 6.5.35. STANDARD Let $A \in \mathbb{F}^{m\times n}$ and $\mathbf{b}, \mathbf{b}' \in \mathbb{F}^m$. Prove or disprove:

(a) If $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ and $\hat{\mathbf{x}}'$ is one of $A\mathbf{x} = \mathbf{b}'$, then $\hat{\mathbf{x}} + \hat{\mathbf{x}}'$ is a least-squares solution of $A\mathbf{x} = \mathbf{b} + \mathbf{b}'$.

(b) If $U \in \mathbb{F}^{m\times m}$ is unitary, then $UA\mathbf{x} = U\mathbf{b}$ and $A\mathbf{x} = \mathbf{b}$ have the same least-squares solutions.

(c) If $G \in \mathbb{F}^{m\times m}$ is invertible, then $GA\mathbf{x} = G\mathbf{b}$ and $A\mathbf{x} = \mathbf{b}$ have the same least-squares solutions.

(d) If $S \in \mathbb{F}^{n\times n}$ is invertible, then $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $S^{-1}\hat{\mathbf{x}}$ is a least-squares solution of $(AS)\mathbf{y} = \mathbf{b}$.

Solution on page 1515

Exercise 6.5.36. ROUTINE Let $A \in \mathbb{F}^{m\times n}$ have independent columns. A student computes

$$A(A^*A)^{-1}A^* = AA^{-1}(A^*)^{-1}A^* = I_m,$$

and concludes from Proposition 6.5.13 that $\mathrm{Pr}_{\mathrm{im}\, A} = I_m$, so that every vector of $\mathbb{F}^m$ lies in $\mathrm{im}\, A$ and every system $A\mathbf{x} = \mathbf{b}$ is consistent. Find the error. For which A is the conclusion correct?

Solution on page 1515

Exercise 6.5.37. ROUTINE

(a) Let $A \in \mathbb{F}^{m\times n}$, and suppose $\mathbf{1} = (1, \ldots, 1) \in \mathrm{im}\, A$. Show that the residual of a least-squares solution of $A\mathbf{x} = \mathbf{b}$ has entries adding up to 0, so that the entries of $A\hat{\mathbf{x}}$ have the same mean as those of $\mathbf{b}$.

(b) Part (a) applies to every polynomial fit. Show by an example that the residuals of the least-squares line through the origin, $y = cx$, need not add up to 0.

Solution on page 1515

Exercise 6.5.38. STANDARD Let $A \in \mathbb{F}^{m\times n}$ have independent columns, and let $P = A(A^*A)^{-1}A^*$. Without using Proposition 6.5.13, show:

(a) $P^2 = P$ and $P^* = P$;

(b) $PA = A$, and $\mathrm{im}\, P = \mathrm{im}\, A$;

(c) $\mathrm{tr}\, P = n$.

(d) Compute P for the regression matrix A with columns $\mathbf{1}$ and $\mathbf{x} = (1, 2, 3, 4)$, and check that $P\mathbf{1} = \mathbf{1}$, $P\mathbf{x} = \mathbf{x}$, and $\mathrm{tr}\, P = 2$.

Solution on page 1516

Exercise 6.5.39. STANDARD Let $A \in \mathbb{F}^{m\times n}$, $\mathbf{a}, \mathbf{b} \in \mathbb{F}^m$, and let $A' \in \mathbb{F}^{m\times(n+1)}$ be A with $\mathbf{a}$ appended as a last column. Let E and E' be the least-squares errors of $A\mathbf{x} = \mathbf{b}$ and $A'\mathbf{x}' = \mathbf{b}$, and let $\mathbf{r}$ be the residual of a least-squares solution of $A\mathbf{x} = \mathbf{b}$. Show that $E' \le E$, with $E' < E$ if and only if $\langle \mathbf{r}, \mathbf{a} \rangle \neq 0$.

Solution on page 1516

Exercise 6.5.40. STANDARD Let $(x_1, y_1), \ldots, (x_m, y_m)$ be data with the x_i not all equal and the y_i not all equal, let r be their correlation coefficient (Example 6.2.19), and let $y = c_0 + c_1 x$ be their least-squares line.

(a) Show that the residual vector is $\tilde{\mathbf{y}} - c_1 \tilde{\mathbf{x}}$.

(b) Show that the squared least-squares error is $\left(1 - r^2\right)\|\tilde{\mathbf{y}}\|^2$. So the line accounts for the fraction r^2 of the spread $\|\tilde{\mathbf{y}}\|^2$ of the y-values about their mean.

(c) Compute r^2 for the spring data of Example 6.5.18.

Solution on page 1517

Exercise 6.5.41. STANDARD Let $(x_1, y_1), \ldots, (x_m, y_m)$ be data with the x_i not all equal and the y_i not all equal, with correlation coefficient r. Let $y = c_0 + c_1 x$ be the least-squares line of y on x, and $x = d_0 + d_1 y$ the least-squares line of x on y (the roles of the coordinates exchanged).

(a) Show that $c_1 d_1 = r^2$.

(b) Show that both lines pass through $\left(m_x, m_y\right)$, and that they are the same line if and only if all data points lie on one line.

(c) Find both lines for the data $(0, 0), (1, 2), (2, 1), (3, 3)$.

Solution on page 1517

Exercise 6.5.42. STANDARD Let $y = c_0 + c_1 x$ be the least-squares line for data $(x_1, y_1), \ldots, (x_m, y_m)$ with the x_i not all equal. Let $\alpha, \gamma \neq 0$ and $\beta, \delta \in \mathbb{R}$, and change units and origins: $x_i' = \alpha x_i + \beta$ and $y_i' = \gamma y_i + \delta$. Show that the least-squares line for the new data is the image of the old line under $(x, y) \mapsto (\alpha x + \beta, \gamma y + \delta)$, namely

$$y' = \gamma\left(c_0 + c_1 \frac{x' - \beta}{\alpha}\right) + \delta.$$

Solution on page 1518

Exercise 6.5.43. STANDARD In $C[0, 1]$ with $\langle f, g \rangle = \int_0^1 f(t) g(t)\, \mathrm{d}t$, find the polynomial p of degree at most 2 closest to $f(t) = \sqrt{t}$, using Proposition 6.5.23, and compute $\|f - p\|^2$. Compare with the best line.

Solution on page 1518

Exercise 6.5.44. STANDARD Over any field F, the equations $A^{\mathsf{T}} A\mathbf{x} = A^{\mathsf{T}}\mathbf{b}$ make sense for $A \in F^{m \times n}$ and $\mathbf{b} \in F^m$, even though there is no inner product. Prove or disprove:

(a) For every $A \in \mathbb{F}_2^{m \times n}$ and $\mathbf{b} \in \mathbb{F}_2^m$, the system $A^{\mathsf{T}} A\mathbf{x} = A^{\mathsf{T}}\mathbf{b}$ has a solution.

(b) If $A \in \mathbb{F}_2^{m \times n}$ has independent columns, then $A^{\mathsf{T}} A$ is invertible.

(c) Both statements hold over $\mathbb{R}$.

Solution on page 1519

Exercise 6.5.45. HARD Give a second proof that the normal equations $A^* A\mathbf{x} = A^*\mathbf{b}$ always have a solution, by showing that $\operatorname{im}(A^* A) = \operatorname{im} A^*$.

Solution on page 1519

Solution on page 1520

Exercise 6.5.46. HARD Let $A \in \mathbb{F}^{m \times n}$ have linearly independent rows, and let $\mathbf{b} \in \mathbb{F}^m$.

(a) Show that AA^* is invertible.

(b) Show that $\mathbf{x}_0 = A^*(AA^*)^{-1}\mathbf{b}$ solves $A\mathbf{x} = \mathbf{b}$ and lies in $(\ker A)^\perp$.

(c) Show that $\mathbf{x}_0$ is the unique solution of $A\mathbf{x} = \mathbf{b}$ of smallest norm.

(d) Find the solution of smallest norm of $x_1 + x_2 + x_3 = 3$, $x_1 - x_3 = 1$.

Solution on page 1520

Exercise 6.5.47. HARD Let $(x_1, y_1), \dots, (x_m, y_m)$ be data with the x_i not all equal. For $i < j$ with $x_i \neq x_j$, let $s_{ij} = (y_j - y_i)/(x_j - x_i)$ be the slope of the line through the ith and jth points. Show that the slope of the least-squares line is a weighted average of these slopes:

$$c_1 = \frac{\sum_{i<j} (x_j - x_i)^2 s_{ij}}{\sum_{i<j} (x_j - x_i)^2},$$

where pairs with $x_i = x_j$ are omitted (they would have weight 0). Deduce that c_1 lies between the smallest and the largest of the s_{ij}.

Solution on page 1520

Exercise 6.5.48. STANDARD BEYOND 21-242

(a) Let $w_1, \dots, w_m > 0$, let $W = \operatorname{diag}(\sqrt{w_1}, \dots, \sqrt{w_m})$, and let $D = W^2$. Show that $\sum_i w_i |b_i - (A\mathbf{x})_i|^2 = \|W\mathbf{b} - WA\mathbf{x}\|^2$, and deduce from Theorem 6.5.5 that the minimizers of this sum are the solutions of $A^*DA\mathbf{x} = A^*D\mathbf{b}$.

(b) Four measurements give $(0,1), (1,2), (2,3), (3,8)$, but the last one was made with an instrument whose errors have twice the standard deviation of the others, so it gets weight $1/4$. Find the ordinary and the weighted least-squares lines.

Solution on page 1521

Exercise 6.5.49. STANDARD BEYOND 21-242 Let $A \in \mathbb{F}^{m \times n}$ with pseudoinverse A^+ (Definition 6.5.28).

(a) Show that $A^+A = \Pr_{(\ker A)^\perp}$.

(b) Deduce the *Penrose equations* $AA^+A = A$, $A^+AA^+ = A^+$, $(AA^+)^* = AA^+$, and $(A^+A)^* = A^+A$.

(c) Find A^+ for $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and check the four equations.

Solution on page 1522

Exercise 6.5.50. CHALLENGE BEYOND 21-242 Least squares measures vertical distances to the line. *Orthogonal regression* measures perpendicular distances instead. Let $\mathbf{p}_1, \dots, \mathbf{p}_m \in \mathbb{R}^2$ be points, not all equal, with centroid $\mathbf{g}$, and write $\tilde{\mathbf{p}}_i = \mathbf{p}_i - \mathbf{g} = (\tilde{x}_i, \tilde{y}_i)$, $S_{xx} = \sum_i \tilde{x}_i^2$, $S_{yy} = \sum_i \tilde{y}_i^2$, $S_{xy} = \sum_i \tilde{x}_i \tilde{y}_i$. For a line ℓ and a point $\mathbf{p}$, let $\operatorname{dist}(\mathbf{p}, \ell) = \min_{\mathbf{q} \in \ell} \|\mathbf{p} - \mathbf{q}\|$, and let $D(\ell) = \sum_i \operatorname{dist}(\mathbf{p}_i, \ell)^2$.

(a) Every line is $\ell = \{\mathbf{p} \mid \langle \mathbf{p}, \mathbf{n}\rangle = c\}$ for a unit vector $\mathbf{n}$ and some $c \in \mathbb{R}$. Show that $\operatorname{dist}(\mathbf{p}, \ell) = |\langle \mathbf{p}, \mathbf{n}\rangle - c|$, and that for fixed $\mathbf{n}$, $D(\ell)$ is smallest when ℓ passes through $\mathbf{g}$.

(b) For a line through $\mathbf{g}$ with unit direction $\mathbf{u}$, show that $D(\ell) = \sum_i \|\tilde{\mathbf{p}}_i\|^2 - \sum_i \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2$.

(c) With $\mathbf{u} = (\cos\theta, \sin\theta)$, show that $\sum_i \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2 = \big(S_{xx} + S_{yy}\big)/2 + \Big(\big(S_{xx} - S_{yy}\big)/2\Big)\cos 2\theta + S_{xy}\sin 2\theta$. Deduce that, if $\mathbf{w} = \Big(\big(S_{xx} - S_{yy}\big)/2, S_{xy}\Big) \neq \mathbf{0}$, the lines minimizing D are exactly the line through $\mathbf{g}$ with $(\cos 2\theta, \sin 2\theta) = \mathbf{w}/\|\mathbf{w}\|$.

(d) Find the best line in this sense for $(0,0), (1,2), (2,1), (3,3)$, and compare it with the two lines of Exercise 6.5.41(c).

Solution on page 1522

6.6 Adjoints

The conjugate transpose moves a matrix from one side of the dot product to the other: $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y})$ (Proposition 6.3.27(d)). That identity is what made the normal equations and the projection matrices of the last two sections work. This section asks for the same thing for a linear map $T : V \to W$ between inner product spaces: a map $T^* : W \to V$ with $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle$ for all $\mathbf{v}$ and $\mathbf{w}$. When V and W are finite-dimensional there is exactly one such map, the *adjoint* of T. It is defined without coordinates, and in orthonormal coordinates it is the conjugate transpose, so it is what the transpose of a matrix really is. Its kernel and image are the orthogonal complements of the image and kernel of T, which explains the orthogonality among the four fundamental subspaces of a matrix. The operators that equal their own adjoint, the self-adjoint operators, are the subject of Chapter 8. To construct T^* we first answer a simpler question: which linear functions $V \to \mathbb{F}$ are inner products with a fixed vector?

Throughout, V and W are inner product spaces over $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ (Notation 6.1.1), and we write $T\mathbf{v}$ for $T(\mathbf{v})$.

6.6.1 Linear functionals and the Riesz representation theorem

Definition 6.6.1 (Linear functional). A **linear functional** on a vector space V over a field F is a linear map $\varphi : V \to F$.

The linear functionals on V form the vector space $\mathrm{L}(V, F)$, the *dual space* of V, which Section 13.1 studies for its own sake. Here we need only a few examples.

Example 6.6.2 (Linear functionals).

(a) For scalars $a_1, \ldots, a_n \in F$, the map $\varphi(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n$ is a linear functional on F^n. Every linear functional on F^n has this form, with $a_j = \varphi(\mathbf{e}_j)$, since $\varphi(\mathbf{x}) = \varphi\big(\sum_j x_j\mathbf{e}_j\big) = \sum_j x_j\varphi(\mathbf{e}_j)$ (Corollary 2.2.17).

(b) The trace $\operatorname{tr} : F^{n\times n} \to F$ (Proposition 2.4.27) and evaluation $p \mapsto p(c)$ on $F[x]$ (Example 2.1.9) are linear functionals, and so is $f \mapsto \int_0^1 f(t)\,dt$ on $C[0,1]$.

(c) Let V be an inner product space and $\mathbf{u} \in V$. Then $\mathbf{v} \mapsto \langle \mathbf{v}, \mathbf{u}\rangle$ is a linear functional on V, by linearity in the first slot. With $\mathbf{u}$ in the first slot instead, $\mathbf{v} \mapsto \langle \mathbf{u}, \mathbf{v}\rangle$ is not linear when $\mathbb{F} = \mathbb{C}$, since $\langle \mathbf{u}, i\mathbf{v}\rangle = -i\langle \mathbf{u}, \mathbf{v}\rangle$ (Proposition 6.1.4).

The Riesz representation theorem says that in a finite-dimensional inner product space, (c) is the only example: every linear functional is an inner product with a fixed vector.

Intuition (Level sets and the normal vector). Start with $V = \mathbb{R}^2$ and a nonzero functional $\varphi(\mathbf{x}) = a_1x_1 + a_2x_2$. Here $\varphi(\mathbf{x}) = \mathbf{x} \cdot \mathbf{u}$ with $\mathbf{u} = (a_1, a_2)$, so a representing vector exists, and the picture shows why it had

to. The level sets $\{\mathbf{x} \mid \varphi(\mathbf{x}) = c\}$ are parallel lines, the translates of the line $\ker\varphi$ (Corollary 2.6.31). A vector $\mathbf{u}$ with $\varphi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{u}\rangle$ for all $\mathbf{x}$ must be orthogonal to every vector of $\ker\varphi$, so it lies on the line $(\ker\varphi)^\perp$, normal to the level sets. That fixes $\mathbf{u}$ up to a scalar, and one value of φ fixes the scalar: $\varphi(\mathbf{u}) = \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{u}\|^2$. Geometrically, $\langle \mathbf{x}, \mathbf{u}\rangle$ is $\|\mathbf{u}\|$ times the signed distance from $\mathbf{x}$ to the line $\ker\varphi$ (Example 6.4.35), so $\mathbf{u}$ points in the direction in which φ grows fastest, and $\|\mathbf{u}\|$ is the rate of growth (Figure 6.6.1). In the language of calculus, $\mathbf{u}$ is the gradient of φ.

The same reasoning works in every finite-dimensional inner product space: the representing vector is normal to the kernel. Over $\mathbb{C}$ there is one twist. In $\mathbb{C}^n$, the functional $\varphi(\mathbf{x}) = \sum_j a_j x_j$ equals $\mathbf{x} \cdot \mathbf{u} = \sum_j x_j \overline{u_j}$ only for $u_j = \overline{a_j}$. The coefficients of φ enter $\mathbf{u}$ conjugated, because $\mathbf{u}$ sits in the conjugate-linear slot.

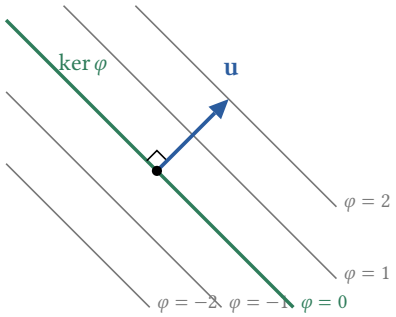


Figure 6.6.1. The level sets $\varphi = c$ of $\varphi(\mathbf{x}) = x_1 + x_2$ on $\mathbb{R}^2$ are parallel lines. The vector $\mathbf{u} = (1, 1)$ with $\varphi(\mathbf{x}) = \langle \mathbf{x}, \mathbf{u}\rangle$ is perpendicular to all of them, and its tip lies on the level set $\varphi = \|\mathbf{u}\|^2 = 2$.

The lecture does not prove the next theorem; it constructs adjoints of matrices directly, as in the second proof of Theorem 6.6.15 below. We include it because it explains what an adjoint is, and because its two proofs illustrate two ways of producing a vector with prescribed inner products.

Theorem 6.6.3 (Riesz representation theorem). BEYOND 21-242 Let V be a finite-dimensional inner product space, and let φ be a linear functional on V. There is exactly one vector $\mathbf{u} \in V$ such that

$$\varphi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}\rangle \quad \text{for all } \mathbf{v} \in V.$$

If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is an orthonormal basis of V, then $\mathbf{u} = \overline{\varphi(\mathbf{e}_1)}\mathbf{e}_1 + \cdots + \overline{\varphi(\mathbf{e}_n)}\mathbf{e}_n$.

Proof (with an orthonormal basis). STANDARD

Uses. Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.1.4 (Sesquilinearity), Definition 6.3.15 (Orthonormal list, orthonormal basis), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 6.1.37 (Inner products distinguish vectors)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Take an inner product to isolate a coefficient (Technique 6.3.12), Check it on a basis (Technique 2.2.7)
Hint 1. For uniqueness, suppose $\mathbf{u}$ and $\mathbf{u}'$ both work, and compare $\langle \mathbf{v}, \mathbf{u}\rangle$ with $\langle \mathbf{v}, \mathbf{u}'\rangle$. For existence, first find the only possible candidate.
Hint 2. If $\mathbf{u}$ works, its coefficient on $\mathbf{e}_j$ is $\langle \mathbf{u}, \mathbf{e}_j\rangle = \overline{\langle \mathbf{e}_j, \mathbf{u}\rangle} = \overline{\varphi(\mathbf{e}_j)}$.
Hint 3. To check the candidate, note that φ and $\mathbf{v} \mapsto \langle \mathbf{v}, \mathbf{u}\rangle$ are both linear in $\mathbf{v}$, so it suffices to compare them on the basis.
Sketch. Uniqueness: if $\langle \mathbf{v}, \mathbf{u}\rangle = \langle \mathbf{v}, \mathbf{u}'\rangle$ for all $\mathbf{v}$, then $\mathbf{u} = \mathbf{u}'$ by Proposition 6.1.37. Existence: for $\mathbf{u} = \sum_j \overline{\varphi(\mathbf{e}_j)}\mathbf{e}_j$, pulling the coefficients out of the second slot conjugates them back, so $\langle \mathbf{e}_k, \mathbf{u}\rangle = \sum_j \varphi(\mathbf{e}_j)\langle \mathbf{e}_k, \mathbf{e}_j\rangle = \varphi(\mathbf{e}_k)$. Two linear functionals that agree on a basis are equal.

Uniqueness. Suppose $\langle \mathbf{v}, \mathbf{u}\rangle = \langle \mathbf{v}, \mathbf{u}'\rangle$ for all $\mathbf{v} \in V$. Taking conjugates, $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}', \mathbf{v}\rangle$ for all $\mathbf{v}$, so $\mathbf{u} = \mathbf{u}'$ by Proposition 6.1.37.

Existence. By Theorem 6.4.20, V has an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Let $\mathbf{u} = \sum_{j=1}^n \overline{\varphi(\mathbf{e}_j)}\mathbf{e}_j$, and let $\psi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}\rangle$, a linear functional by linearity in the first slot. For each k, Proposition 6.1.4 and orthonormality give

$$\psi(\mathbf{e}_k) = \left\langle \mathbf{e}_k, \sum_{j=1}^{n} \overline{\varphi(\mathbf{e}_j)}\mathbf{e}_j \right\rangle = \sum_{j=1}^{n} \varphi(\mathbf{e}_j)\langle \mathbf{e}_k, \mathbf{e}_j\rangle = \varphi(\mathbf{e}_k),$$

since the conjugate of $\overline{\varphi(\mathbf{e}_j)}$ is $\varphi(\mathbf{e}_j)$ and $\langle \mathbf{e}_k, \mathbf{e}_j\rangle = \delta_{kj}$. So the linear maps φ and ψ agree on a basis of V, and $\varphi = \psi$ by Corollary 2.2.6. That is, $\varphi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u}\rangle$ for every $\mathbf{v}$, as desired. □

Proof (via the orthogonal complement of the kernel). STANDARD

> *Uses.* Proposition 2.6.8 (Kernels and images are subspaces), Proposition 4.1.21 (Dimensions of subspaces), Corollary 6.4.27 (Testing whether a subspace is everything), Proposition 6.3.9 (Properties of $A^{\perp}$), Proposition 6.1.4 (Sesquilinearity)
> *Techniques.* Orthogonal decomposition (Technique 6.2.4)
> *Hint 1.* If $\mathbf{u}$ works, it is orthogonal to every vector of $\ker\varphi$. Where does that put it?
> *Hint 2.* For $\varphi \neq 0$, pick $\mathbf{z} \neq \mathbf{0}$ in $(\ker\varphi)^{\perp}$ (why does one exist?) and look for $\mathbf{u}$ as a multiple of $\mathbf{z}$. Testing at $\mathbf{v} = \mathbf{z}$ determines the multiple.
> *Hint 3.* To check the candidate at a general $\mathbf{v}$, subtract the multiple of $\mathbf{z}$ that lies on the same level set of φ: $\mathbf{v} - (\varphi(\mathbf{v})/\varphi(\mathbf{z}))\mathbf{z} \in \ker\varphi$.
> *Sketch.* Uniqueness is as in the first proof. If $\varphi \neq 0$, then $\ker\varphi \neq V$, so some $\mathbf{z} \neq \mathbf{0}$ is orthogonal to $\ker\varphi$ (Corollary 6.4.27), and $\varphi(\mathbf{z}) \neq 0$. Every $\mathbf{v}$ is $(\varphi(\mathbf{v})/\varphi(\mathbf{z}))\mathbf{z}$ plus a vector of $\ker\varphi$, which is orthogonal to $\mathbf{z}$. So $\langle \mathbf{v}, \mathbf{z}\rangle = (\varphi(\mathbf{v})/\varphi(\mathbf{z}))\|\mathbf{z}\|^2$, and $\mathbf{u} = \left(\overline{\varphi(\mathbf{z})}/\|\mathbf{z}\|^2\right)\mathbf{z}$ works.

Uniqueness is proved as in the first proof. For existence, if $\varphi = 0$, take $\mathbf{u} = \mathbf{0}$. Suppose $\varphi \neq 0$, and let $K = \ker\varphi$. It is a subspace of V (Proposition 2.6.8), finite-dimensional (Proposition 4.1.21), and $K \neq V$. By Corollary 6.4.27, $K^{\perp} \neq \{\mathbf{0}\}$, so we may choose $\mathbf{z} \in K^{\perp}$ with $\mathbf{z} \neq \mathbf{0}$. Since $K \cap K^{\perp} = \{\mathbf{0}\}$ (Proposition 6.3.9(d)), $\mathbf{z} \notin K$, that is, $\varphi(\mathbf{z}) \neq 0$. Let

$$\mathbf{u} = \frac{\overline{\varphi(\mathbf{z})}}{\|\mathbf{z}\|^2}\mathbf{z}.$$

Let $\mathbf{v} \in V$ and $c = \varphi(\mathbf{v})/\varphi(\mathbf{z})$. Then $\varphi(\mathbf{v} - c\mathbf{z}) = \varphi(\mathbf{v}) - c\varphi(\mathbf{z}) = 0$, so $\mathbf{v} - c\mathbf{z} \in K$. As $\mathbf{z} \in K^{\perp}$, conjugate symmetry gives $\langle \mathbf{v} - c\mathbf{z}, \mathbf{z}\rangle = 0$, and therefore

$$\langle \mathbf{v}, \mathbf{z}\rangle = c\langle \mathbf{z}, \mathbf{z}\rangle = \frac{\varphi(\mathbf{v})}{\varphi(\mathbf{z})}\|\mathbf{z}\|^2.$$

Pulling the real number $\|\mathbf{z}\|^2$ and the scalar $\overline{\varphi(\mathbf{z})}$ out of the second slot (Proposition 6.1.4),

$$\langle \mathbf{v}, \mathbf{u}\rangle = \frac{\varphi(\mathbf{z})}{\|\mathbf{z}\|^2}\langle \mathbf{v}, \mathbf{z}\rangle = \varphi(\mathbf{v}),$$

as desired. □

The first proof computes $\mathbf{u}$, and uniqueness shows that the formula, although written with an orthonormal basis, does not depend on which one. The second proof chooses no basis. It says where $\mathbf{u}$ lives, on the line $(\ker\varphi)^{\perp}$ normal to the level sets, and it uses finite-dimensionality only to know that $\ker\varphi$ has a nonzero orthogonal complement. In infinite-dimensional spaces, where no finite orthonormal basis exists, that is the step that can sometimes be rescued (Section 15.3).

Remark 6.6.4 (Where the conjugate comes from). The representing vector must sit in the second slot, since $\mathbf{v} \mapsto \langle \mathbf{u}, \mathbf{v}\rangle$ is not linear over $\mathbb{C}$ (Example 6.6.2(c)). Moving the coefficients $\varphi(\mathbf{e}_j)$ into the second slot conjugates them, which is why $\mathbf{u}$ has coefficients $\overline{\varphi(\mathbf{e}_j)}$. For the same reason the assignment $\mathbf{u} \mapsto \langle \cdot, \mathbf{u}\rangle$, from V to the linear functionals on V, is additive but conjugates scalars: $\langle \cdot, c\mathbf{u}\rangle = \overline{c}\langle \cdot, \mathbf{u}\rangle$. So Theorem 6.6.3 matches the vectors of V one to one with the linear functionals on V, by a map that is linear when $\mathbb{F} = \mathbb{R}$ and conjugate-linear when $\mathbb{F} = \mathbb{C}$. Section 13.2 returns to this identification.

Example 6.6.5 (Representing evaluation at 0). Give $\mathbb{R}[x]_{\le 1}$ the inner product $\langle p, q\rangle = \int_0^1 p(t)q(t)\,dt$, and let $\varphi(p) = p(0)$. Gram–Schmidt on $1, x$ gives the orthonormal basis $e_1 = 1, e_2 = \sqrt{3}(2x-1)$ (Example 6.4.19). Here $\varphi(e_1) = 1$ and $\varphi(e_2) = -\sqrt{3}$, so Theorem 6.6.3 gives the representing polynomial

$$q = 1 - \sqrt{3}\cdot\sqrt{3}(2x-1) = 4 - 6x.$$

Check on the basis $1, x$: $\int_0^1 (4-6t)\,dt = 4 - 3 = 1 = \varphi(1)$, and $\int_0^1 t(4-6t)\,dt = 2 - 2 = 0 = \varphi(x)$. So $\int_0^1 p(t)(4-6t)\,dt = p(0)$ for every p of degree at most 1.

The answer depends on the space, not only on the formula for φ. On $\mathbb{R}[x]_{\le 2}$ with the same inner product, evaluation at 0 is represented by $9 - 36x + 30x^2$, as one checks on $1, x, x^2$. The representing polynomials also grow: $\|4-6x\|^2 = \langle q, q\rangle = \varphi(q) = q(0) = 4$, and $\|9 - 36x + 30x^2\|^2 = 9$. (On $\mathbb{R}[x]_{\le n}$ the representing polynomial has norm $n+1$; Exercise 6.6.49 shows that its norm is at least $\sqrt{2n+1}$.) A polynomial that integrates against every p to give $p(0)$ has to concentrate its weight near $t = 0$, and higher degrees allow more concentration.

Exercise 6.6.6. ROUTINE

(a) Find $\mathbf{u} \in \mathbb{C}^3$ with $2z_1 - iz_2 + (1+i)z_3 = \mathbf{z}\cdot\mathbf{u}$ for all $\mathbf{z} \in \mathbb{C}^3$.

(b) With the inner product of Example 6.6.5 on $\mathbb{R}[x]_{\le 1}$, find the polynomial q with $\langle p, q\rangle = p(1)$ for all p.

Solution on page 1523

Finite-dimensionality cannot be dropped from Theorem 6.6.3.

Non-example 6.6.7 (Riesz fails in infinite dimensions). BEYOND 21-242 Let P be the space of real polynomial functions on $[0,1]$, a subspace of $C[0,1]$, with $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$ (Proposition 6.1.13). We show that the linear functional $\varphi(p) = p(0)$ is not of the form $\langle \cdot, q\rangle$ for any $q \in P$. Suppose that $\int_0^1 p(t)q(t)\,dt = p(0)$ for every $p \in P$. Apply this to $p(t) = tq(t)$, which vanishes at 0:

$$\int_0^1 tq(t)^2\,dt = 0.$$

The integrand is continuous and nonnegative on $[0,1]$, so it is identically 0 (Lemma 6.1.14). Hence $q(t) = 0$ for $0 < t \le 1$, and $q(0) = 0$ by continuity. But then $p(0) = \langle p, q\rangle = 0$ for every p, which fails for $p = 1$.

Restricted to the polynomials of degree at most n, the same functional is represented by a polynomial of norm $n+1$ (Example 6.6.5). As n grows, these representatives pile up near $t = 0$ without converging to any polynomial. Section 15.3 explains which functionals do have representatives in infinite dimensions.

6.6.2 The adjoint

Intuition (Moving a map across the inner product). For a matrix A, read the identity $(A\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot(A^*\mathbf{y})$ as follows. Fix $\mathbf{y}$. The left side is a measurement of the output $A\mathbf{x}$: for a unit vector $\mathbf{y}$, it is the component of $A\mathbf{x}$ along $\mathbf{y}$. The right side computes the same number from the input $\mathbf{x}$ alone, as its component along the vector $A^*\mathbf{y}$. So A^* turns a measurement on the output side into a measurement on the input side.

For a linear map $T : V \to W$ we ask for the same conversion. Fix $\mathbf{w} \in W$. The function $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{w}\rangle$ is a linear functional on V, and when V is finite-dimensional, Theorem 6.6.3 says that it is $\mathbf{v} \mapsto \langle \mathbf{v}, \mathbf{u}\rangle$ for exactly one $\mathbf{u} \in V$. Calling this vector $T^*\mathbf{w}$ defines a function $T^* : W \to V$. The definition records the property that T^* is meant to have, without presupposing that such a function exists.

Definition 6.6.8 (Adjoint). Let V and W be inner product spaces and $T \in \mathrm{L}(V, W)$. A function $S : W \to V$ is an **adjoint** of T if

$$\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S\mathbf{w}\rangle \quad \text{for all } \mathbf{v} \in V \text{ and } \mathbf{w} \in W.$$

By Proposition 6.6.9 below, T has at most one adjoint. When it exists, it is denoted T^* ("T star").

The inner product on the left is that of W, and the one on the right is that of V; we use the same notation for both. Taking conjugates of both sides gives the equivalent form

$$\langle \mathbf{w}, T\mathbf{v}\rangle = \langle T^*\mathbf{w}, \mathbf{v}\rangle,$$

which is how the lecture writes the definition. So T can be moved out of either slot into the other, at the price of a star. We did not ask S to be linear; linearity comes for free.

Proposition 6.6.9 (Adjoints are unique and linear). Let V and W be inner product spaces and $T \in \mathrm{L}(V, W)$.

(a) T has at most one adjoint.

(b) Every adjoint of T is linear. So $T^* \in \mathrm{L}(W, V)$ whenever T^* exists.

Proof. ROUTINE

Uses. Definition 6.6.8 (Adjoint), Proposition 6.1.37 (Inner products distinguish vectors), Proposition 6.1.4 (Sesquilinearity)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), let S and S' be adjoints and fix $\mathbf{w}$. What do you know about $\langle \mathbf{v}, S\mathbf{w}\rangle$ and $\langle \mathbf{v}, S'\mathbf{w}\rangle$ as $\mathbf{v}$ varies?
Hint 2. Proposition 6.1.37 says that a vector is determined by its inner products with all vectors. Here the unknown vectors sit in the second slot, so conjugate first.
Hint 3. For (b), compute $\langle \mathbf{v}, S(a\mathbf{w} + b\mathbf{w}')\rangle$ with the defining identity and Proposition 6.1.4, and compare it with $\langle \mathbf{v}, aS\mathbf{w} + bS\mathbf{w}'\rangle$.
Sketch. (a) For each $\mathbf{w}$, $\langle \mathbf{v}, S\mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S'\mathbf{w}\rangle$ for every $\mathbf{v}$, so $S\mathbf{w} = S'\mathbf{w}$ by Proposition 6.1.37. (b) Both $\langle \mathbf{v}, S(a\mathbf{w} + b\mathbf{w}')\rangle$ and $\langle \mathbf{v}, aS\mathbf{w} + bS\mathbf{w}'\rangle$ equal $\overline{a}\langle T\mathbf{v}, \mathbf{w}\rangle + \overline{b}\langle T\mathbf{v}, \mathbf{w}'\rangle$ for every $\mathbf{v}$; apply Proposition 6.1.37 again.

(a) Let $S, S' : W \to V$ be adjoints of T, and fix $\mathbf{w} \in W$. For every $\mathbf{v} \in V$,

$$\langle \mathbf{v}, S\mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S'\mathbf{w}\rangle.$$

Taking conjugates, $\langle S\mathbf{w}, \mathbf{v}\rangle = \langle S'\mathbf{w}, \mathbf{v}\rangle$ for every $\mathbf{v} \in V$, so $S\mathbf{w} = S'\mathbf{w}$ by Proposition 6.1.37. As $\mathbf{w}$ was arbitrary, $S = S'$.

(b) Let S be an adjoint of T, and let $\mathbf{w}, \mathbf{w}' \in W$ and $a, b \in \mathbb{F}$. For every $\mathbf{v} \in V$, the defining identity and Proposition 6.1.4 give

$$\begin{aligned}\langle \mathbf{v}, S(a\mathbf{w} + b\mathbf{w}')\rangle &= \langle T\mathbf{v}, a\mathbf{w} + b\mathbf{w}'\rangle \\ &= \overline{a}\langle T\mathbf{v}, \mathbf{w}\rangle + \overline{b}\langle T\mathbf{v}, \mathbf{w}'\rangle \\ &= \overline{a}\langle \mathbf{v}, S\mathbf{w}\rangle + \overline{b}\langle \mathbf{v}, S\mathbf{w}'\rangle \\ &= \langle \mathbf{v}, aS\mathbf{w} + bS\mathbf{w}'\rangle.\end{aligned}$$

As in (a), taking conjugates and applying Proposition 6.1.37 gives $S(a\mathbf{w} + b\mathbf{w}') = aS\mathbf{w} + bS\mathbf{w}'$. □

Neither part used a dimension: uniqueness and linearity hold in every inner product space, and this is the order in which the lecture proceeds. In practice it means that adjoints are found by guessing a formula and checking the defining identity; uniqueness does the rest.

Example 6.6.10 (The adjoint of a matrix). Let $A \in \mathbb{F}^{m\times n}$, and give $\mathbb{F}^n$ and $\mathbb{F}^m$ the dot product. By Proposition 6.3.27(d), $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y})$ for all $\mathbf{x}$ and $\mathbf{y}$; that is, T_{A^*} is an adjoint of T_A. By Proposition 6.6.9,

$$(T_A)^* = T_{A^*}.$$

Over $\mathbb{R}$ this is $T_{A^{\mathsf{T}}}$. Over $\mathbb{C}$ the transpose would be wrong. For $A = (i) \in \mathbb{C}^{1\times 1}$, T_A is multiplication by i on $\mathbb{C}$, and $(ix)\cdot y = ix\overline{y}$, while $x\cdot(iy) = x\overline{iy} = -ix\overline{y}$. The adjoint of multiplication by i is multiplication by $\overline{i} = -i$, in line with $A^* = \left(\overline{i}\right)$.

So the star on a matrix and the star on a map agree, as long as $\mathbb{F}^n$ and $\mathbb{F}^m$ carry the dot product. The adjoint depends on the inner products as well as on the map (Exercise 6.6.31).

Technique 6.6.11 (Find the adjoint by isolating the input). To compute $T^*\mathbf{w}$, start from $\langle T\mathbf{v}, \mathbf{w}\rangle$, substitute the formula for T, and rearrange until the expression reads $\langle \mathbf{v}, \mathbf{u}\rangle$, with $\mathbf{v}$ alone in the first slot and $\mathbf{u}$ built from $\mathbf{w}$ only. Then $T^*\mathbf{w} = \mathbf{u}$ by Proposition 6.6.9. Scalars that move into the second slot pick up a bar. The same move proves identities: to show that a given map S is T^*, check $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S\mathbf{w}\rangle$ for all $\mathbf{v}$ and $\mathbf{w}$, without ever writing a formula for T^*. This is the analogue for adjoints of Technique 2.5.16 and Technique 6.4.7.

Reach for it when computing the adjoint of an explicit map, and when proving identities such as $(RT)^* = T^*R^*$. The canonical example is Example 6.6.13.

Used in. Exercise 6.6.17, Proposition 6.6.18, Exercise 6.6.31, Exercise 6.6.33, Proposition 8.4.10.

Example 6.6.12 (An adjoint from $\mathbb{C}^3$ to $\mathbb{C}^2$). Let $T : \mathbb{C}^3 \to \mathbb{C}^2$, $T(z_1, z_2, z_3) = (z_1 + iz_3, 2z_2 - z_3)$, with dot products. For $\mathbf{z} \in \mathbb{C}^3$ and $\mathbf{w} \in \mathbb{C}^2$,

$$\begin{aligned}\langle T\mathbf{z}, \mathbf{w}\rangle &= (z_1 + iz_3)\overline{w_1} + (2z_2 - z_3)\overline{w_2}\\ &= z_1\overline{w_1} + z_2\cdot 2\overline{w_2} + z_3(\overline{iw_1} - \overline{w_2})\\ &= z_1\overline{w_1} + z_2\overline{2w_2} + z_3\overline{-iw_1 - w_2},\end{aligned}$$

which is $\langle \mathbf{z}, (w_1, 2w_2, -iw_1 - w_2)\rangle$. So $T^*(w_1, w_2) = (w_1, 2w_2, -iw_1 - w_2)$. As Example 6.6.10 predicts, this is multiplication by the conjugate transpose of the matrix of T:

$$\begin{pmatrix}1 & 0 & i\\ 0 & 2 & -1\end{pmatrix}^* = \begin{pmatrix}1 & 0\\ 0 & 2\\ -i & -1\end{pmatrix}.$$

Example 6.6.13 (Rank-one maps). Fix $\mathbf{u} \in V$ and $\mathbf{x} \in W$, and let $T\mathbf{v} = \langle \mathbf{v}, \mathbf{u}\rangle\mathbf{x}$. For $\mathbf{v} \in V$ and $\mathbf{w} \in W$,

$$\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle\langle \mathbf{x}, \mathbf{w}\rangle = \left\langle \mathbf{v}, \overline{\langle \mathbf{x}, \mathbf{w}\rangle}\mathbf{u}\right\rangle = \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{x}\rangle\mathbf{u}\rangle,$$

using Proposition 6.1.4 to move the scalar $\langle \mathbf{x}, \mathbf{w}\rangle$ into the second slot. So

$$T^*\mathbf{w} = \langle \mathbf{w}, \mathbf{x}\rangle\mathbf{u}.$$

The map T reads off the component of its input along $\mathbf{u}$ and writes the result along $\mathbf{x}$; its adjoint reads along $\mathbf{x}$ and writes along $\mathbf{u}$. For $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ with dot products, $T\mathbf{v} = (\mathbf{u}^*\mathbf{v})\mathbf{x} = \mathbf{x}\mathbf{u}^*\mathbf{v}$ is multiplication by the rank-one matrix $\mathbf{x}\mathbf{u}^*$, and T^* is multiplication by $\mathbf{u}\mathbf{x}^* = (\mathbf{x}\mathbf{u}^*)^*$. No dimension assumption was needed.

Example 6.6.14 (Multiplication operators). Let $V = C([0,1], \mathbb{C})$ with $\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt$ (Proposition 6.1.13). For a continuous function $h : [0,1] \to \mathbb{C}$, let $M_h f = hf$, the pointwise product. Then

$$\langle hf, g\rangle = \int_0^1 h(t)f(t)\overline{g(t)}\,dt = \int_0^1 f(t)\overline{\overline{h(t)}g(t)}\,dt = \left\langle f, \overline{h}g\right\rangle,$$

so M_h has an adjoint, $M_h^* = M_{\overline{h}}$, although V is infinite-dimensional. This is the continuous version of $\operatorname{diag}(d_1, \ldots, d_n)^* = \operatorname{diag}\left(\overline{d_1}, \ldots, \overline{d_n}\right)$: M_h scales the value at each point t by $h(t)$, as a diagonal matrix scales each coordinate.

Existence, and the matrix of the adjoint

In finite dimensions adjoints always exist, and orthonormal bases compute them.

Theorem 6.6.15 (Existence of adjoints). Let V and W be finite-dimensional inner product spaces and $T \in \mathrm{L}(V, W)$.

(a) T has exactly one adjoint T^*, and $T^* \in \mathrm{L}(W, V)$.

(b) If $\mathcal{B} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathcal{C} = \mathbf{f}_1, \ldots, \mathbf{f}_m$ are orthonormal bases of V and W, then

$$[T^*]_{\mathcal{B}\mathcal{C}} = \big([T]_{\mathcal{C}\mathcal{B}}\big)^*.$$

Proof (via the Riesz representation theorem). STANDARD BEYOND 21-242

Uses. Theorem 6.6.3 (Riesz representation theorem), Proposition 6.6.9 (Adjoints are unique and linear), Proposition 6.3.18 (Coordinates in an orthonormal basis), Definition 2.3.1 (Matrix of a linear map), Definition 6.3.26 (Conjugate transpose)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. For (a), fix $\mathbf{w}$ and look at $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{w}\rangle$. What kind of function of $\mathbf{v}$ is it?
Hint 2. Apply Theorem 6.6.3 to that functional and call the resulting vector $T^*\mathbf{w}$. Linearity and uniqueness are Proposition 6.6.9.
Hint 3. For (b), the (i, j) entry of $[T]_{\mathcal{C}\mathcal{B}}$ is the ith coordinate of $T\mathbf{e}_j$ in the orthonormal basis $\mathcal{C}$, which is an inner product (Proposition 6.3.18).
Sketch. For fixed $\mathbf{w}$, $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{w}\rangle$ is a linear functional on V; Theorem 6.6.3 represents it as $\langle \cdot, \mathbf{u}\rangle$, and $T^*\mathbf{w} = \mathbf{u}$ satisfies the defining identity. In orthonormal bases, $[T]_{\mathcal{C}\mathcal{B}}$ has entries $\langle T\mathbf{e}_j, \mathbf{f}_i\rangle$, and $[T^*]_{\mathcal{B}\mathcal{C}}$ has entries $\langle T^*\mathbf{f}_i, \mathbf{e}_j\rangle = \overline{\langle T\mathbf{e}_j, \mathbf{f}_i\rangle}$.

(a) Fix $\mathbf{w} \in W$. The function $\varphi_{\mathbf{w}} : V \to \mathbb{F}$, $\varphi_{\mathbf{w}}(\mathbf{v}) = \langle T\mathbf{v}, \mathbf{w}\rangle$, is a linear functional, since T is linear and the inner product is linear in its first slot. By Theorem 6.6.3 there is exactly one vector of V, which we call $T^*\mathbf{w}$, with

$$\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle \quad \text{for all } \mathbf{v} \in V.$$

Doing this for every $\mathbf{w}$ defines a function $T^* : W \to V$ that satisfies Definition 6.6.8. By Proposition 6.6.9 it is the only adjoint of T, and it is linear.

(b) Let $A = [T]_{\mathcal{C}\mathcal{B}}$. Its jth column is $\big[T\mathbf{e}_j\big]_{\mathcal{C}}$, so by Proposition 6.3.18(b) its (i, j) entry is $a_{ij} = \langle T\mathbf{e}_j, \mathbf{f}_i\rangle$. In the same way, the (j, i) entry of $[T^*]_{\mathcal{B}\mathcal{C}}$ is the jth coordinate of $T^*\mathbf{f}_i$ in $\mathcal{B}$, namely

$$\langle T^*\mathbf{f}_i, \mathbf{e}_j\rangle = \overline{\langle \mathbf{e}_j, T^*\mathbf{f}_i\rangle} = \overline{\langle T\mathbf{e}_j, \mathbf{f}_i\rangle} = \overline{a_{ij}}.$$

So $[T^*]_{\mathcal{B}\mathcal{C}} = A^*$ (Definition 6.3.26). □

Proof (via the conjugate transpose). ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Theorem 2.3.20 (Linear maps correspond to matrices), Theorem 2.3.9 (The matrix of a map computes the map), Theorem 6.3.22 (Parseval's identity), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 6.6.9 (Adjoints are unique and linear)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Fix orthonormal bases. In coordinates, inner products become dot products (Theorem 6.3.22), and T becomes multiplication by $A = [T]_{\mathcal{C}\mathcal{B}}$. For matrices you already know an adjoint.
Hint 2. Let $S \in \mathrm{L}(W, V)$ be the map with $[S]_{\mathcal{B}\mathcal{C}} = A^*$, and compute $\langle T\mathbf{v}, \mathbf{w}\rangle$ in coordinates.
Sketch. Let S be the linear map with $[S]_{\mathcal{B}\mathcal{C}} = A^*$. With $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ and $\mathbf{y} = [\mathbf{w}]_{\mathcal{C}}$, Parseval's identity gives $\langle T\mathbf{v}, \mathbf{w}\rangle = (A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y}) = \langle \mathbf{v}, S\mathbf{w}\rangle$. So S is an adjoint, it is the only one by Proposition 6.6.9, and (b) holds by construction.

Orthonormal bases of V and W exist by Theorem 6.4.20. We prove (a) and (b) together, for arbitrary orthonormal bases $\mathcal{B}$ and $\mathcal{C}$. Let $A = [T]_{\mathcal{C}\mathcal{B}} \in \mathbb{F}^{m \times n}$. By Theorem 2.3.20 there is a linear map $S : W \to V$ with $[S]_{\mathcal{B}\mathcal{C}} = A^*$. Let $\mathbf{v} \in V$ and $\mathbf{w} \in W$, and write $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ and $\mathbf{y} = [\mathbf{w}]_{\mathcal{C}}$. By Theorem 2.3.9, $[T\mathbf{v}]_{\mathcal{C}} = A\mathbf{x}$ and $[S\mathbf{w}]_{\mathcal{B}} = A^*\mathbf{y}$. Parseval's identity (Theorem 6.3.22(a)) computes the inner products of V and of W as dot products of coordinate vectors, so

$$\langle T\mathbf{v}, \mathbf{w}\rangle = (A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y}) = \langle \mathbf{v}, S\mathbf{w}\rangle,$$

where the middle step is Proposition 6.3.27(d). So S is an adjoint of T. By Proposition 6.6.9 it is the only one, so $T^* = S \in \mathrm{L}(W, V)$ and $[T^*]_{\mathcal{BC}} = A^*$. □

The second proof is the lecture's route: orthonormal coordinates turn V and W into $\mathbb{F}^n$ and $\mathbb{F}^m$ with their dot products, and there the adjoint is the conjugate transpose (Example 6.6.10). The first proof never chooses a basis of W, so it shows that (a) holds whenever V alone is finite-dimensional. Without any finite-dimensionality, adjoints can fail to exist: on the space P of Non-example 6.6.7, the operator $Tp = p(0) \cdot 1$ has none, since an adjoint would give $\langle p, T^*1\rangle = \langle Tp, 1\rangle = p(0)$ for every p, so that T^*1 would represent evaluation at 0.

Intuition (Reading and writing). Here is a formula for T^* that needs only an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V. Every $\mathbf{v} \in V$ is $\sum_j \langle \mathbf{v}, \mathbf{e}_j\rangle \mathbf{e}_j$ (Proposition 6.3.18), so

$$T\mathbf{v} = \sum_{j=1}^{n} \langle \mathbf{v}, \mathbf{e}_j\rangle T\mathbf{e}_j.$$

That is, T is a sum of n rank-one maps, the jth of which reads the $\mathbf{e}_j$-coordinate of its input and writes it along $T\mathbf{e}_j$. By Example 6.6.13, the adjoint of the jth map reads along $T\mathbf{e}_j$ and writes along $\mathbf{e}_j$, and the adjoint of a sum is the sum of the adjoints (Proposition 6.6.18(a) below). So

$$T^*\mathbf{w} = \sum_{j=1}^{n} \langle \mathbf{w}, T\mathbf{e}_j\rangle \mathbf{e}_j.$$

One can also read this off directly: the $\mathbf{e}_j$-coordinate of $T^*\mathbf{w}$ is $\langle T^*\mathbf{w}, \mathbf{e}_j\rangle = \overline{\langle \mathbf{e}_j, T^*\mathbf{w}\rangle} = \overline{\langle T\mathbf{e}_j, \mathbf{w}\rangle} = \langle \mathbf{w}, T\mathbf{e}_j\rangle$. For $T = T_A$ on $\mathbb{F}^n$ with the standard basis, the vectors $T\mathbf{e}_j$ are the columns $\mathbf{a}_j$ of A, and the formula says that the jth entry of $A^*\mathbf{w}$ is $\mathbf{w} \cdot \mathbf{a}_j$: the rows of A^* are the conjugated columns of A.

The hypothesis that the bases are orthonormal in Theorem 6.6.15(b) cannot be dropped. In other bases the matrix of T^* is usually not the conjugate transpose of the matrix of T.

Example 6.6.16 (The adjoint of differentiation). Give $V = \mathbb{R}[x]_{\leq 1}$ the inner product $\langle p, q\rangle = \int_0^1 p(t)q(t)\,\mathrm{d}t$, and let $D(a + bx) = b$ be differentiation (Example 2.1.9). In the basis $\mathcal{B} = 1, x$,

$$[D]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad ([D]_{\mathcal{B}})^{\mathsf{T}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is the matrix of the map $a + bx \mapsto ax$. That map is not D^*: for $p = x$ and $q = 1$, $\langle Dp, q\rangle = \langle 1, 1\rangle = 1$, while $\langle p, 1 \cdot x\rangle = \langle x, x\rangle = 1/3$. The basis $\mathcal{B}$ is not orthonormal, since $\langle 1, x\rangle = 1/2$. To find D^*, use the orthonormal basis $\mathcal{O} = e_1, e_2$ with $e_1 = 1$ and $e_2 = \sqrt{3}(2x - 1)$ of Example 6.6.5. Since $De_1 = 0$ and $De_2 = 2\sqrt{3} = 2\sqrt{3}e_1$,

$$[D]_{\mathcal{O}} = \begin{pmatrix} 0 & 2\sqrt{3} \\ 0 & 0 \end{pmatrix}, \quad [D^*]_{\mathcal{O}} = \begin{pmatrix} 0 & 0 \\ 2\sqrt{3} & 0 \end{pmatrix}$$

by Theorem 6.6.15(b). So $D^*e_1 = 2\sqrt{3}e_2 = 6(2x - 1)$ and $D^*e_2 = 0$. Writing $a + bx = (a + b/2)e_1 + \big(b/\big(2\sqrt{3}\big)\big)e_2$ gives

$$D^*(a + bx) = (2a + b)(6x - 3), \quad [D^*]_{\mathcal{B}} = \begin{pmatrix} -6 & -3 \\ 12 & 6 \end{pmatrix}.$$

As a check, for $p = a + bx$ and $q = c + dx$, $\langle Dp, q\rangle = \int_0^1 b(c + dt)\,\mathrm{d}t = b(c + d/2)$, and since $\int_0^1 (6t - 3)\,\mathrm{d}t = 0$ and $\int_0^1 t(6t - 3)\,\mathrm{d}t = 1/2$, $\langle p, D^*q\rangle = (2c + d)b/2$, the same number.

So $[D^*]_{\mathcal{B}}$ has nothing to do with $([D]_{\mathcal{B}})^{\mathsf{T}}$. Nor is D^* a differentiation-like map, such as $-D$: integrating by parts, $\langle Dp, q\rangle = p(1)q(1) - p(0)q(0) - \langle p, Dq\rangle$, and the boundary terms are what D^* has to absorb (compare

Exercise 6.6.39, where there are no boundary terms). Exercise 6.6.46 gives the general rule for a basis that is not orthonormal: $[T^*]_{\mathcal{B}} = G^{-1}([T]_{\mathcal{B}})^* G$, where G is the Gram matrix of $\mathcal{B}$.

Intuition (What the adjoint is). In coordinates. In orthonormal bases, T^* is the conjugate transpose (Theorem 6.6.15(b)). The transpose of a matrix is defined by moving entries around, so it depends on the basis, and in other bases it describes some unrelated map (Example 6.6.16). The adjoint is defined by the inner product alone, and the conjugate transpose is what it looks like in orthonormal coordinates. This is the sense in which the adjoint is the right generalization of the transpose. What makes the transpose useful, in the normal equations, in A^*A, in projection matrices, is the identity $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y})$, and that identity is the definition of T^*. The index shuffle is a consequence that holds only in orthonormal bases. Over $\mathbb{C}$ the conjugate is forced, as Example 6.6.10 shows.

As a measurement. For a unit vector $\mathbf{w}$, $\langle T\mathbf{v}, \mathbf{w}\rangle$ is the component of the output $T\mathbf{v}$ along $\mathbf{w}$, and the definition says that it is also the component of the input $\mathbf{v}$ along $T^*\mathbf{w}$. By Cauchy–Schwarz, among unit vectors $\mathbf{v}$ the real part of $\langle T\mathbf{v}, \mathbf{w}\rangle$ is largest when $\mathbf{v}$ points along $T^*\mathbf{w}$ (Exercise 6.6.42). So $T^*\mathbf{w}$ points toward the inputs that T pushes furthest in the direction $\mathbf{w}$, and its length says how far.

Reading and writing. By Example 6.6.13, T^* exchanges the vectors along which T reads and writes. Every T is a sum of such rank-one maps, and T^* exchanges the roles in each.

Without an inner product. A matrix $A \in F^{m\times n}$ over any field still has a transpose, and Section 13.1 gives it a meaning: it is the matrix of the dual map, which acts on linear functionals rather than on vectors. In an inner product space, Theorem 6.6.3 identifies functionals with vectors, and the dual map becomes the adjoint (Section 13.2).

ASIDE: *Gradients flow backward through the adjoint* BEYOND 21-242

Over $\mathbb{R}$, the derivative of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{v}$ is a linear functional, $\mathbf{h} \mapsto \nabla f(\mathbf{v}) \cdot \mathbf{h}$, and its Riesz vector is the gradient: $f(\mathbf{v} + \mathbf{h}) \approx f(\mathbf{v}) + \nabla f(\mathbf{v}) \cdot \mathbf{h}$ for small $\mathbf{h}$. Now let $f(\mathbf{v}) = g(A\mathbf{v})$ for a matrix A and a differentiable $g : \mathbb{R}^m \to \mathbb{R}$. Then

$$f(\mathbf{v} + \mathbf{h}) = g(A\mathbf{v} + A\mathbf{h}) \approx g(A\mathbf{v}) + \nabla g(A\mathbf{v}) \cdot (A\mathbf{h}) = f(\mathbf{v}) + \left(A^{\mathsf{T}} \nabla g(A\mathbf{v})\right) \cdot \mathbf{h},$$

so $\nabla f(\mathbf{v}) = A^{\mathsf{T}} \nabla g(A\mathbf{v})$. The map A carries inputs forward, and its adjoint carries gradients backward. A neural network composes many linear maps with simple nonlinear ones, and training it requires the gradient of a loss with respect to the inputs of every layer. These are computed from the last layer to the first, multiplying by transposed weight matrices along the way; this is *backpropagation* (Section 16.6).

Exercise 6.6.17. ROUTINE Give $\mathbb{F}^{n\times n}$ the Frobenius inner product $\langle X, Y\rangle = \operatorname{tr}(Y^*X)$ (Proposition 6.1.16), and fix $A \in \mathbb{F}^{n\times n}$.

(a) Find the adjoint of $R_A(X) = XA$.

(b) Find the adjoint of $X \mapsto X^{\mathsf{T}}$.

(c) For $\mathbb{F} = \mathbb{C}$, explain why $X \mapsto X^*$ has no adjoint in the sense of Definition 6.6.8.

Solution on page 1523

6.6.3 The algebra of adjoints

The rules for computing with adjoints are those of Proposition 6.3.27 for matrices. We state them with as few hypotheses as possible: each part asks only that the adjoints on its right-hand side exist, which is automatic in finite dimensions (Theorem 6.6.15).

Proposition 6.6.18 (Properties of the adjoint). Let U, V, W be inner product spaces, let $S, T \in \mathrm{L}(V, W)$ and $R \in \mathrm{L}(W, U)$, and let $a \in \mathbb{F}$. Suppose that S, T, and R have adjoints. Then:

(a) $S + T$ has an adjoint, and $(S+T)^* = S^* + T^*$;

(b) aT has an adjoint, and $(aT)^* = \overline{a}T^*$;

(c) T^* has an adjoint, and $(T^*)^* = T$;

(d) RT has an adjoint, and $(RT)^* = T^*R^*$;

(e) id_V has an adjoint, and $\mathrm{id}_V^* = \mathrm{id}_V$;

(f) if T is invertible and T^{-1} has an adjoint, then T^* is invertible and $(T^*)^{-1} = \left(T^{-1}\right)^*$.

Proof (from the defining identity). ROUTINE

Uses. Definition 6.6.8 (Adjoint), Proposition 6.6.9 (Adjoints are unique and linear), Proposition 6.1.4 (Sesquilinearity), Definition 2.5.12 (Invertible)
Techniques. Find the adjoint by isolating the input (Technique 6.6.11), Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Each part names a candidate for the adjoint. Check the defining identity $\langle Xv, w\rangle = \langle v, Yw\rangle$ for the candidate Y and let uniqueness do the rest.
Hint 2. For (b), pulling $\overline{a}$ out of the second slot produces $\overline{\overline{a}} = a$. For (c), the defining identity of an adjoint of $T^* : W \to V$ reads $\langle T^*w, v\rangle = \langle w, Sv\rangle$; conjugate the identity for T.
Hint 3. For (f), apply (d) and (e) to $T^{-1}T = \mathrm{id}_V$ and $TT^{-1} = \mathrm{id}_W$.
Sketch. (a) $\langle (S+T)v, w\rangle = \langle v, S^*w\rangle + \langle v, T^*w\rangle$. (b) $\langle aTv, w\rangle = a\langle v, T^*w\rangle = \langle v, \overline{a}T^*w\rangle$. (c) $\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle$. (d) $\langle RTv, u\rangle = \langle Tv, R^*u\rangle = \langle v, T^*R^*u\rangle$. (e) is immediate. (f) By (d) and (e), $T^*\left(T^{-1}\right)^* = \left(T^{-1}T\right)^* = \mathrm{id}_V$ and $\left(T^{-1}\right)^*T^* = \left(TT^{-1}\right)^* = \mathrm{id}_W$.

In each part we check the defining identity for the proposed adjoint; by Proposition 6.6.9 it is then the adjoint. Let $v \in V$, $w \in W$, and $u \in U$.

(a) $\langle (S+T)v, w\rangle = \langle Sv, w\rangle + \langle Tv, w\rangle = \langle v, S^*w\rangle + \langle v, T^*w\rangle = \langle v, (S^*+T^*)w\rangle$.

(b) By Proposition 6.1.4, $\langle v, \overline{a}T^*w\rangle = a\langle v, T^*w\rangle$, so $\langle (aT)v, w\rangle = a\langle Tv, w\rangle = a\langle v, T^*w\rangle = \langle v, \overline{a}T^*w\rangle$.

(c) An adjoint of $T^* : W \to V$ is a function $S : V \to W$ with $\langle T^*w, v\rangle = \langle w, Sv\rangle$ for all w, v. By conjugate symmetry and the identity for T,

$$\langle T^*w, v\rangle = \overline{\langle v, T^*w\rangle} = \overline{\langle Tv, w\rangle} = \langle w, Tv\rangle,$$

so $S = T$ works.

(d) $\langle RTv, u\rangle = \langle Tv, R^*u\rangle = \langle v, T^*R^*u\rangle$, using the identity for R and then for T.

(e) $\langle \mathrm{id}_V v, v'\rangle = \langle v, \mathrm{id}_V v'\rangle$ for all $v, v' \in V$.

(f) By (d) and (e), applied to $T^{-1}T = \mathrm{id}_V$ and to $TT^{-1} = \mathrm{id}_W$,

$$T^*\left(T^{-1}\right)^* = \left(T^{-1}T\right)^* = \mathrm{id}_V, \quad \left(T^{-1}\right)^*T^* = \left(TT^{-1}\right)^* = \mathrm{id}_W.$$

So $\left(T^{-1}\right)^*$ is a two-sided inverse of T^* (Definition 2.5.12). □

Proof (via matrices, in finite dimensions). ROUTINE

Uses. Theorem 6.6.15 (Existence of adjoints), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.4.12 (Matrices of sums and scalar multiples), Theorem 2.4.2 (The matrix of a composite), Proposition 2.5.20 (Inverses and matrices), Proposition 6.3.27 (Properties of the conjugate transpose), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Assume U, V, W are finite-dimensional and fix orthonormal bases. By Theorem 6.6.15(b), every adjoint has the conjugate transpose of the original matrix.
Hint 2. Two maps are equal if their matrices are (Theorem 2.3.20). Translate each identity into one about matrices and quote Proposition 6.3.27.

Sketch. In orthonormal bases, $[X^*] = [X]^*$ for every map X. So, for instance, $\big[(RT)^*\big] = [RT]^* = ([R][T])^* = [T]^*[R]^* = [T^*][R^*] = [T^*R^*]$, and the maps agree because their matrices do. The other parts translate in the same way into (a)–(c) of Proposition 6.3.27.

Suppose U, V, W are finite-dimensional, so that all the adjoints in question exist (Theorem 6.6.15), and fix orthonormal bases of the three spaces (Theorem 6.4.20). We write $[X]$ for the matrix of a map with respect to these bases. By Theorem 6.6.15(b), $[X^*] = [X]^*$ for every linear map X between two of the spaces, and by Theorem 2.3.20, two maps with the same matrix are equal. Now, using Proposition 2.4.12, Theorem 2.4.2, and Proposition 6.3.27,

$$\begin{aligned}\big[(S+T)^*\big] &= ([S]+[T])^* = [S]^* + [T]^* = [S^* + T^*],\\ \big[(aT)^*\big] &= (a[T])^* = \overline{a}[T]^* = [\overline{a}T^*],\\ \big[(T^*)^*\big] &= \big([T]^*\big)^* = [T],\\ \big[(RT)^*\big] &= ([R][T])^* = [T]^*[R]^* = [T^*R^*],\end{aligned}$$

and $[\mathrm{id}_V^*] = I^* = I = [\mathrm{id}_V]$. For (f), $\big[T^{-1}\big] = [T]^{-1}$ (Proposition 2.5.20), and $\big([T]^{-1}\big)^* = \big([T]^*\big)^{-1}$ by Proposition 6.3.27(c), so $\big[\big(T^{-1}\big)^*\big] = \big([T]^*\big)^{-1} = [T^*]^{-1}$, which is the matrix of $(T^*)^{-1}$. □

The two proofs run in opposite directions. The second derives the rules for maps from those for matrices, and needs finite dimensions and orthonormal bases. The first needs neither, and applied to T_A, T_B on $\mathbb{F}^n$ it reproves Proposition 6.3.27 without a single index: for instance $(AB)^* = B^*A^*$ because $(T_AT_B)^* = T_B^*T_A^*$. Part (c) is the lecture's observation that T is an adjoint of T^*.

By (a)–(c), the map $T \mapsto T^*$ from $\mathrm{L}(V,W)$ to $\mathrm{L}(W,V)$ is additive, conjugates scalars, and is undone by the same operation in the other direction. In finite dimensions it is therefore a bijection; it is linear when $\mathbb{F} = \mathbb{R}$ and conjugate-linear when $\mathbb{F} = \mathbb{C}$, just like $z \mapsto \overline{z}$.

Exercise 6.6.19. ROUTINE Let V be finite-dimensional, $T \in \mathrm{L}(V)$, and $c_0, \dots, c_k \in \mathbb{F}$. Show that

$$\big(c_0\,\mathrm{id} + c_1T + \cdots + c_kT^k\big)^* = \overline{c_0}\,\mathrm{id} + \overline{c_1}T^* + \cdots + \overline{c_k}(T^*)^k.$$

Solution on page 1524

6.6.4 Kernels and images

For a matrix M we already know two orthogonality relations. One is $(\mathrm{im}\,M)^\perp = \ker M^*$ (Lemma 6.5.4). The other is $\ker M = (\mathrm{im}\,M^*)^\perp$: solving $M\mathbf{x} = \mathbf{0}$ means finding the vectors orthogonal to the conjugated rows of M, which are the columns of M^* (Example 6.3.10). Both were proved by looking at entries. For linear maps between inner product spaces they follow from a single move, which the rest of the book uses constantly.

Technique 6.6.20 (Move an operator across the inner product). When a map acts on one argument of an inner product, move it to the other argument as its adjoint: $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle$ and $\langle \mathbf{w}, T\mathbf{v}\rangle = \langle T^*\mathbf{w}, \mathbf{v}\rangle$. Do this so that the map lands on the vector you know more about, or so that two vectors you want to compare end up in the same inner product. Typical outcomes:

- "$\mathbf{w}$ is orthogonal to $T\mathbf{v}$ for every $\mathbf{v}$" becomes "$T^*\mathbf{w}$ is orthogonal to every vector", so $T^*\mathbf{w} = \mathbf{0}$;
- $\langle T^*T\mathbf{v}, \mathbf{v}\rangle$ becomes $\langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$, which is where Technique 6.1.36 takes over;
- a hypothesis on T becomes a statement about T^*, or, for a self-adjoint T, about T itself acting on the other vector.

Reach for it whenever an inner product contains $T\mathbf{v}$ and the hypotheses or the goal concern T^*, or concern the vector in the other slot. The canonical example is part (a) of Theorem 6.6.21. In Chapter 8

the same move shows that the eigenvalues of a self-adjoint operator are real and that its eigenvectors for different eigenvalues are orthogonal.

Used in. Theorem 6.6.21, Proposition 6.6.23, Proposition 6.6.30, Exercise 6.6.38, Exercise 6.6.40, Exercise 6.6.42, Exercise 6.6.44, Exercise 6.7.8, Exercise 6.7.9, Exercise 6.7.10, Exercise 6.7.14, Problem 6.3, Problem 6.7, Problem 6.10, Problem 6.14, Problem 6.18, Proposition 8.1.1, Exercise 8.1.6, Proposition 8.1.13, Lemma 8.1.17, Exercise 8.1.28, Theorem 8.2.5, Lemma 8.2.10, Exercise 8.2.44, Lemma 8.3.8, Proposition 8.3.22, Exercise 8.3.28, Theorem 8.4.4, Lemma 8.4.14, Exercise 6.6.51, Exercise 8.1.29.

Theorem 6.6.21 (Kernels and images of adjoints). Let V and W be inner product spaces, and let $T \in \mathrm{L}(V, W)$ have an adjoint. Then

(a) $\ker T^* = (\operatorname{im} T)^\perp$;

(b) $\ker T = (\operatorname{im} T^*)^\perp$.

If moreover V is finite-dimensional, then also

(c) $\operatorname{im} T^* = (\ker T)^\perp$;

(d) $\operatorname{im} T = (\ker T^*)^\perp$;

and there are orthogonal decompositions $V = \ker T \oplus \operatorname{im} T^*$ and $W = \operatorname{im} T \oplus \ker T^*$.

Intuition. Part (a) is a chain of rewritings. A vector $\mathbf{w}$ is in $(\operatorname{im} T)^\perp$ when $\langle T\mathbf{v}, \mathbf{w}\rangle = 0$ for every $\mathbf{v}$; moving T across, that says $\langle \mathbf{v}, T^*\mathbf{w}\rangle = 0$ for every $\mathbf{v}$, and the only vector orthogonal to everything is $\mathbf{0}$. Part (b) is (a) for T^*, whose adjoint is T. Parts (c) and (d) take orthogonal complements of (b) and (a), which is allowed once the subspaces are finite-dimensional (Theorem 6.4.26).

Proof. STANDARD

Uses. Definition 6.6.8 (Adjoint), Definition 6.3.8 (Vectors orthogonal to a set), Proposition 6.1.4 (Sesquilinearity), Definition 6.1.2 (Inner product), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.4.26 (The double complement), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.2.11 (Rank–nullity), Theorem 6.4.24 (Orthogonal decomposition)

Techniques. Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36)

Hint 1. For (a), write out what $\mathbf{w} \in (\operatorname{im} T)^\perp$ means, move T to the other slot, and decide when a vector is orthogonal to every vector of V.

Hint 2. For (b), apply (a) to T^*; what is the adjoint of T^*?

Hint 3. For (c) and (d), take orthogonal complements in (b) and (a). Why are $\operatorname{im} T^*$ and $\operatorname{im} T$ finite-dimensional when V is?

Sketch. (a) $\mathbf{w} \in \ker T^*$ iff $\langle \mathbf{v}, T^*\mathbf{w}\rangle = 0$ for all $\mathbf{v}$ (take $\mathbf{v} = T^*\mathbf{w}$ for one direction) iff $\langle T\mathbf{v}, \mathbf{w}\rangle = 0$ for all $\mathbf{v}$ iff $\mathbf{w} \perp \operatorname{im} T$. (b) is (a) for T^*, since $(T^*)^* = T$. If V is finite-dimensional, so are $\operatorname{im} T^* \subseteq V$ and $\operatorname{im} T$, and Theorem 6.4.26 turns (b) and (a) into (c) and (d). The decompositions are Theorem 6.4.24 for $\ker T$ and $\operatorname{im} T$.

(a) Let $\mathbf{w} \in W$. We claim that

$$\begin{aligned} T^*\mathbf{w} = \mathbf{0} &\iff \langle \mathbf{v}, T^*\mathbf{w}\rangle = 0 \text{ for all } \mathbf{v} \in V \\ &\iff \langle T\mathbf{v}, \mathbf{w}\rangle = 0 \text{ for all } \mathbf{v} \in V \\ &\iff \mathbf{w} \in (\operatorname{im} T)^\perp. \end{aligned}$$

In the first line, if $T^*\mathbf{w} = \mathbf{0}$, then every $\langle \mathbf{v}, T^*\mathbf{w}\rangle$ is 0 by Proposition 6.1.4(b); conversely, taking $\mathbf{v} = T^*\mathbf{w}$ gives $\|T^*\mathbf{w}\|^2 = 0$, so $T^*\mathbf{w} = \mathbf{0}$ by definiteness. The second equivalence is the defining identity of T^*. For the third, the elements of $\operatorname{im} T$ are the vectors $T\mathbf{v}$, and $\langle \mathbf{w}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v}, \mathbf{w}\rangle}$ vanishes exactly when $\langle T\mathbf{v}, \mathbf{w}\rangle$ does (Definition 6.3.8). So $\ker T^* = (\operatorname{im} T)^\perp$.

(b) By Proposition 6.6.18(c), T^* has an adjoint, namely T. Applying (a) to $T^* \in \mathrm{L}(W, V)$ gives $\ker T = \ker (T^*)^* = (\operatorname{im} T^*)^\perp$.

(c) Suppose V is finite-dimensional. Then its subspace $\operatorname{im} T^*$ is finite-dimensional (Proposition 4.1.21), so by (b) and Theorem 6.4.26,

$$(\ker T)^\perp = \left((\operatorname{im} T^*)^\perp\right)^\perp = \operatorname{im} T^*.$$

(d) By Theorem 4.2.11, $\dim \operatorname{im} T \le \dim V$, so $\operatorname{im} T$ is a finite-dimensional subspace of W. By (a) and Theorem 6.4.26, $(\ker T^*)^\perp = \left((\operatorname{im} T)^\perp\right)^\perp = \operatorname{im} T$.

Finally, $\ker T$ and $\operatorname{im} T$ are finite-dimensional, so Theorem 6.4.24 gives $V = \ker T \oplus (\ker T)^\perp = \ker T \oplus \operatorname{im} T^*$ by (c), and $W = \operatorname{im} T \oplus (\operatorname{im} T)^\perp = \operatorname{im} T \oplus \ker T^*$ by (a). □

For $T = T_M$, part (a) is Lemma 6.5.4 and part (b) is the computation of Example 6.3.10. The lecture proves the weaker statement that $\operatorname{im} T^*$ is orthogonal to $\ker T$, which is half of (b) and (c): if $\mathbf{v} = T^*\mathbf{w}$ and $T\mathbf{z} = \mathbf{0}$, then $\langle \mathbf{v}, \mathbf{z}\rangle = \langle T^*\mathbf{w}, \mathbf{z}\rangle = \langle \mathbf{w}, T\mathbf{z}\rangle = 0$.

In infinite dimensions, (a) and (b) survive but (c) and (d) can fail, because a subspace can have a zero orthogonal complement without being everything (Exercise 6.4.63).

Corollary 6.6.22 (T and T^* have the same rank). Let V and W be finite-dimensional and $T \in \mathrm{L}(V, W)$.

(a) The restriction $T\restriction_{\operatorname{im} T^*}$ is injective with image $\operatorname{im} T$. So T maps $\operatorname{im} T^*$ bijectively onto $\operatorname{im} T$, and likewise T^* maps $\operatorname{im} T$ bijectively onto $\operatorname{im} T^*$.

(b) $\operatorname{rank} T^* = \operatorname{rank} T$.

(c) For every $A \in \mathbb{F}^{m\times n}$, $\operatorname{rank} A^* = \operatorname{rank} A$.

Proof (by restriction). STANDARD

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Definition 4.2.1 (Restriction), Proposition 6.3.9 (Properties of $A^\perp$), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 6.6.18 (Properties of the adjoint), Example 6.6.10 (The adjoint of a matrix), Theorem 4.2.7 (The image of T_A)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. The kernel of $T\restriction_{\operatorname{im} T^*}$ is $\ker T \cap \operatorname{im} T^*$. Use Theorem 6.6.21(c).
Hint 2. For the image, decompose $\mathbf{v} = \mathbf{z} + \mathbf{r}$ with $\mathbf{z} \in \ker T$ and $\mathbf{r} \in \operatorname{im} T^*$, and compute $T\mathbf{v}$.
Hint 3. For (c), apply (b) to T_A, using $(T_A)^* = T_{A^*}$.
Sketch. $\ker(T\restriction_{\operatorname{im} T^*}) = \ker T \cap \operatorname{im} T^* = \ker T \cap (\ker T)^\perp = \{\mathbf{0}\}$. Writing $\mathbf{v} = \mathbf{z} + \mathbf{r}$ with $\mathbf{z} \in \ker T$ and $\mathbf{r} \in \operatorname{im} T^*$ gives $T\mathbf{v} = T\mathbf{r}$, so the restriction already hits all of $\operatorname{im} T$. An injective map preserves dimension, so $\operatorname{rank} T = \dim \operatorname{im} T^* = \operatorname{rank} T^*$. The statement for T^* follows by symmetry, and (c) is (b) for T_A.

(a) Let $R = T\restriction_{\operatorname{im} T^*} : \operatorname{im} T^* \to W$. Its kernel is $\ker T \cap \operatorname{im} T^*$, which by Theorem 6.6.21(c) is $\ker T \cap (\ker T)^\perp = \{\mathbf{0}\}$ (Proposition 6.3.9(d)). So R is injective. Clearly $\operatorname{im} R \subseteq \operatorname{im} T$. Conversely, let $\mathbf{v} \in V$. By Theorem 6.6.21, $\mathbf{v} = \mathbf{z} + \mathbf{r}$ with $\mathbf{z} \in \ker T$ and $\mathbf{r} \in \operatorname{im} T^*$, so $T\mathbf{v} = T\mathbf{z} + T\mathbf{r} = R\mathbf{r} \in \operatorname{im} R$. Hence $\operatorname{im} R = \operatorname{im} T$, and R is a bijection from $\operatorname{im} T^*$ onto $\operatorname{im} T$. Since T^* has adjoint T (Proposition 6.6.18(c)), the same argument applied to T^* shows that T^* maps $\operatorname{im} T$ bijectively onto $\operatorname{im} T^*$.

(b) By (a) and Proposition 4.2.6(a), applied to the injective map R, $\dim \operatorname{im} T = \dim \operatorname{im} R = \dim \operatorname{im} T^*$.

(c) With dot products, $(T_A)^* = T_{A^*}$ (Example 6.6.10), and $\operatorname{rank} A = \dim \operatorname{im} T_A$ (Theorem 4.2.7). So (c) is (b) for $T = T_A$. □

Proof (by counting dimensions). ROUTINE

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 4.2.11 (Rank–nullity), Example 6.6.10 (The adjoint of a matrix)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Compute $\dim \operatorname{im} T^*$ from Theorem 6.6.21(c) and Corollary 6.4.25.
Hint 2. Compare with rank–nullity for T.
Sketch. $\operatorname{rank} T^* = \dim (\ker T)^\perp = \dim V - \dim \ker T = \operatorname{rank} T$. This proves (b), and (c) as before; it does not give the bijection in (a).

By Theorem 6.6.21(c), Corollary 6.4.25, and Theorem 4.2.11,

$$\operatorname{rank} T^* = \dim (\ker T)^\perp = \dim V - \dim \ker T = \operatorname{rank} T.$$

This is (b), and (c) follows as in the first proof. This argument does not give (a). □

The counting proof is shorter; the first proof shows where the equality comes from. A linear map T throws away $\ker T$ and is injective on the rest of V, and Theorem 6.6.21 identifies "the rest" as $\operatorname{im} T^*$. The two bijections in (a) are not inverse to each other in general: for $T = 2\mathrm{id}_V$, both are multiplication by 2.

The four fundamental subspaces

For a matrix $A \in \mathbb{F}^{m\times n}$, apply Theorem 6.6.21 to $T_A : \mathbb{F}^n \to \mathbb{F}^m$, whose adjoint is T_{A^*} (Example 6.6.10). It gives the orthogonal decompositions

$$\mathbb{F}^n = \ker A \oplus \operatorname{im} A^*, \quad \mathbb{F}^m = \operatorname{im} A \oplus \ker A^*.$$

Over $\mathbb{R}$, $A^* = A^\mathsf{T}$, and these are the four fundamental subspaces of Definition 4.3.12: the null space $\ker A$ and the row space $\operatorname{im} A^\mathsf{T}$ are orthogonal complements in $\mathbb{R}^n$, and the column space $\operatorname{im} A$ and the left null space $\ker A^\mathsf{T}$ are orthogonal complements in $\mathbb{R}^m$. So $\mathbb{R}^n = \operatorname{im} A^\mathsf{T} \oplus \ker A$. Every $\mathbf{x} \in \mathbb{R}^n$ splits as $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$ with $\mathbf{x}_r$ in the row space and $\mathbf{x}_n \in \ker A$, and $A\mathbf{x} = A\mathbf{x}_r$; by Corollary 6.6.22, A maps the row space bijectively onto the column space. In particular $\operatorname{rank} A^\mathsf{T} = \operatorname{rank} A$, which is Theorem 4.3.8 for real matrices. Figure 6.6.2 shows the picture.

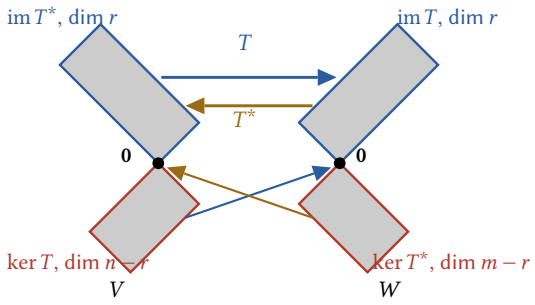


Figure 6.6.2. The four subspaces of $T \in \mathrm{L}(V, W)$, with $n = \dim V$, $m = \dim W$, and $r = \operatorname{rank} T$. In V, the subspaces $\operatorname{im} T^*$ and $\ker T$ are orthogonal complements, and in W so are $\operatorname{im} T$ and $\ker T^*$. The map T (blue) sends $\ker T$ to $\mathbf{0}$ and $\operatorname{im} T^*$ bijectively onto $\operatorname{im} T$; its adjoint T^* (ochre) sends $\ker T^*$ to $\mathbf{0}$ and $\operatorname{im} T$ bijectively onto $\operatorname{im} T^*$. For $T = T_A$ with A real, these are the row space, null space, column space, and left null space of A.

Over $\mathbb{C}$ the orthogonal complement of $\ker A$ is $\operatorname{im} A^*$, the span of the conjugated rows, and the conjugation matters. For $A = \begin{pmatrix} 1 & i \end{pmatrix}$, $\ker A = \operatorname{span}((-i, 1))$, and the conjugated row $(1, -i)$ is orthogonal to $(-i, 1)$: $(1, -i) \cdot (-i, 1) = \overline{-i} + (-i) = 0$. The row $(1, i)$ itself is not: $(1, i) \cdot (-i, 1) = i + i = 2i$. (The rank of A^T still equals the rank of A, since conjugating every entry does not change which columns are independent.)

The lecture's other result about adjoints generalizes Proposition 6.5.9 from matrices to maps, with the same two proofs.

Proposition 6.6.23 (T^*T and T). Let V and W be inner product spaces, and let $T \in \mathrm{L}(V, W)$ have an adjoint.

(a) $\ker(T^*T) = \ker T$.

(b) If V is finite-dimensional, then $\operatorname{rank}(T^*T) = \operatorname{rank} T$ and $\operatorname{im}(T^*T) = \operatorname{im} T^*$.

Proof (by the norm). ROUTINE

Uses. Definition 6.6.8 (Adjoint), Definition 6.1.2 (Inner product), Theorem 4.2.11 (Rank–nullity), Corollary 6.6.22 (T and T^* have the same rank), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36), Double containment (Technique 1.3.14)
Hint 1. One containment in (a) is immediate. For the other, suppose $T^*T\mathbf{v} = \mathbf{0}$ and take the inner product with $\mathbf{v}$.
Hint 2. Move T^* across: $\langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle$.
Hint 3. For (b), use rank–nullity for T^*T and T, which have the same domain, and then compare $\operatorname{im}(T^*T) \subseteq \operatorname{im} T^*$ by dimension.
Sketch. If $T^*T\mathbf{v} = \mathbf{0}$, then $\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = 0$, so $T\mathbf{v} = \mathbf{0}$; the other containment is clear. Equal kernels and rank–nullity give $\operatorname{rank}(T^*T) = \operatorname{rank} T$, which equals $\operatorname{rank} T^*$ by Corollary 6.6.22; since $\operatorname{im}(T^*T) \subseteq \operatorname{im} T^*$, the images agree.

(a) If $T\mathbf{v} = \mathbf{0}$, then $T^*T\mathbf{v} = T^*\mathbf{0} = \mathbf{0}$. Conversely, suppose $T^*T\mathbf{v} = \mathbf{0}$. By the defining identity of T^* with $\mathbf{w} = T\mathbf{v}$,

$$\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle \mathbf{v}, T^*T\mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{0}\rangle = 0,$$

so $T\mathbf{v} = \mathbf{0}$ by definiteness. Hence $\ker(T^*T) = \ker T$.

(b) Both T^*T and T have domain V, so by Theorem 4.2.11 and (a),

$$\operatorname{rank}(T^*T) = \dim V - \dim\ker(T^*T) = \dim V - \dim\ker T = \operatorname{rank} T.$$

By Corollary 6.6.22 this is $\operatorname{rank} T^* = \dim\operatorname{im} T^*$, and $\operatorname{im}(T^*T) \subseteq \operatorname{im} T^*$. A subspace of $\operatorname{im} T^*$ of the same dimension is all of $\operatorname{im} T^*$ (Proposition 4.1.21(b)), so $\operatorname{im}(T^*T) = \operatorname{im} T^*$. □

Proof (by orthogonality). ROUTINE

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Proposition 6.3.9 (Properties of $A^\perp$)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint. Suppose $T^*(T\mathbf{v}) = \mathbf{0}$. Which two subspaces of W contain $T\mathbf{v}$?
Sketch. If $T^*(T\mathbf{v}) = \mathbf{0}$, then $T\mathbf{v} \in \ker T^* = (\operatorname{im} T)^\perp$ by Theorem 6.6.21(a), and also $T\mathbf{v} \in \operatorname{im} T$, so $T\mathbf{v} \in \operatorname{im} T \cap (\operatorname{im} T)^\perp = \{\mathbf{0}\}$. Part (b) follows from (a) as in the first proof.

As before, $\ker T \subseteq \ker(T^*T)$. Conversely, if $T^*(T\mathbf{v}) = \mathbf{0}$, then $T\mathbf{v} \in \ker T^* = (\operatorname{im} T)^\perp$ by Theorem 6.6.21(a), and $T\mathbf{v} \in \operatorname{im} T$. So $T\mathbf{v} \in \operatorname{im} T \cap (\operatorname{im} T)^\perp = \{\mathbf{0}\}$ (Proposition 6.3.9(d)). This proves (a), and (b) follows from (a) exactly as in the first proof. □

The first proof is the lecture's. Both proofs move T^* somewhere useful: the first next to $T\mathbf{v}$, to produce a norm, and the second into the description of $\ker T^*$ as an orthogonal complement.

The next example shows the theorem at work on a map that is not self-adjoint.

Example 6.6.24 (An oblique projection and its adjoint). Let $P : \mathbb{R}^2 \to \mathbb{R}^2$, $P(x, y) = (x - y, 0)$, the projection onto the x-axis along the line $y = x$ from Remark 6.4.12. Its matrix is $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, so P^* has matrix $\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$, that is, $P^*(x, y) = (x, -x)$. Then

$$\operatorname{im} P = \operatorname{span}(\mathbf{e}_1), \qquad \ker P = \operatorname{span}((1, 1)),$$
$$\operatorname{im} P^* = \operatorname{span}((1, -1)) = (\ker P)^\perp, \quad \ker P^* = \operatorname{span}(\mathbf{e}_2) = (\operatorname{im} P)^\perp,$$

as Theorem 6.6.21 predicts. Moreover $(P^*)^2 = \left(P^2\right)^* = P^*$, so P^* is again a projection: it moves each point onto the line $y = -x$ along the vertical direction (Figure 6.6.3). In general, the adjoint of the projection onto U along X is the projection onto $X^\perp$ along $U^\perp$ (Exercise 6.6.47). The two coincide exactly when $X = U^\perp$, which is when the projection is orthogonal.

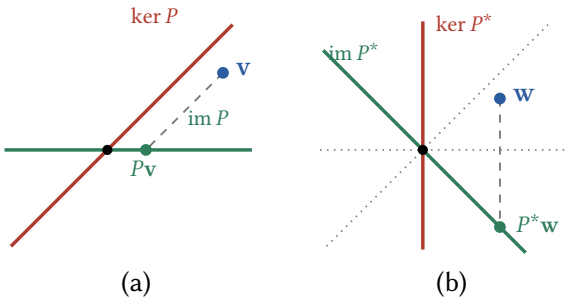


Figure 6.6.3. (a) The projection $P(x, y) = (x - y, 0)$ moves each point onto $\operatorname{im} P$ (the x-axis) along $\ker P$ (the line $y = x$). (b) Its adjoint $P^*(x, y) = (x, -x)$ moves each point onto $\operatorname{im} P^* = (\ker P)^\perp$ along $\ker P^* = (\operatorname{im} P)^\perp$. The dotted lines are $\operatorname{im} P$ and $\ker P$ from (a).

Exercise 6.6.25. ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 3 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Find bases of $\ker A$, $\operatorname{im} A^\mathsf{T}$, $\operatorname{im} A$, and $\ker A^\mathsf{T}$, and check that $\mathbb{R}^3 = \ker A \oplus \operatorname{im} A^\mathsf{T}$ and $\mathbb{R}^3 = \operatorname{im} A \oplus \ker A^\mathsf{T}$ are orthogonal decompositions.

Solution on page 1524

6.6.5 Self-adjoint operators

For an operator $T \in \mathrm{L}(V)$, the maps T and T^* act on the same space, so they can be compared. The operators that equal their adjoint are the subject of Chapter 8. Here we define them, collect examples, and characterize them over $\mathbb{C}$.

Definition 6.6.26 (Self-adjoint operator). Let V be an inner product space. An operator $T \in \mathrm{L}(V)$ is **self-adjoint** if

$$\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle \quad \text{for all } \mathbf{v}, \mathbf{w} \in V,$$

that is, if T has an adjoint and $T^* = T$.

By Example 6.6.10, T_A is self-adjoint for the dot product on $\mathbb{F}^n$ exactly when $A^* = A$, that is, when A is Hermitian (Definition 6.3.26), and over $\mathbb{R}$ when A is symmetric. One says "self-adjoint" for operators and "Hermitian" for matrices; in orthonormal coordinates they are the same condition (Corollary 6.6.28).

Example 6.6.27 (Self-adjoint operators).

(a) The matrices $\begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$ are Hermitian, so they give self-adjoint operators on $\mathbb{C}^2$ and $\mathbb{R}^2$.

(b) Let W be a finite-dimensional subspace of V. By Proposition 6.4.29(c), $\langle \mathrm{Pr}_W \mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, \mathrm{Pr}_W \mathbf{x}\rangle$ for all $\mathbf{v}, \mathbf{x}$, so Pr_W is self-adjoint.

(c) For a continuous $h : [0, 1] \to \mathbb{R}$, the multiplication operator M_h on $C([0, 1], \mathbb{C})$ is self-adjoint, since $M_h^* = M_{\overline{h}} = M_h$ (Example 6.6.14).

(d) If $T \in \mathrm{L}(V, W)$ has an adjoint, then $T^*T \in \mathrm{L}(V)$ and $TT^* \in \mathrm{L}(W)$ are self-adjoint: $(T^*T)^* = T^*(T^*)^* = T^*T$ by (c) and (d) of Proposition 6.6.18, and similarly for TT^*. In particular A^*A is Hermitian for every matrix A, as we knew from its description as a Gram matrix (Theorem 6.3.29). Similarly $T + T^*$ is self-adjoint for every $T \in \mathrm{L}(V)$ with an adjoint.

Some operators that are not self-adjoint: the rotation $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ of Example 6.1.39, with $J^\mathsf{T} = -J$; multiplication by i on $\mathbb{C}$, whose adjoint is multiplication by $-i$ (Example 6.6.10); the oblique projection of Example 6.6.24; and T_A for the complex symmetric matrix $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$, for which $A^\mathsf{T} = A$ but $A^* \neq A$ (Exercise 6.6.34).

Corollary 6.6.28 (Self-adjointness in an orthonormal basis). Let V be finite-dimensional, let $\mathcal{O}$ be an orthonormal basis of V, and let $T \in \mathrm{L}(V)$. Then T is self-adjoint if and only if $[T]_{\mathcal{O}}$ is Hermitian.

Proof. ROUTINE

Uses. Theorem 6.6.15 (Existence of adjoints), Theorem 2.3.20 (Linear maps correspond to matrices), Definition 6.6.26 (Self-adjoint operator)
Hint. Compare the matrices of T and T^* in the basis $\mathcal{O}$.
Sketch. $[T^*]_{\mathcal{O}} = \big([T]_{\mathcal{O}}\big)^*$ by Theorem 6.6.15(b), and maps with equal matrices are equal. So $T^* = T$ iff $\big([T]_{\mathcal{O}}\big)^* = [T]_{\mathcal{O}}$.

By Theorem 6.6.15, T^* exists and $[T^*]_{\mathcal{O}} = \big([T]_{\mathcal{O}}\big)^*$. By Theorem 2.3.20, $T^* = T$ if and only if $[T^*]_{\mathcal{O}} = [T]_{\mathcal{O}}$, that is, if and only if $\big([T]_{\mathcal{O}}\big)^* = [T]_{\mathcal{O}}$. □

Orthonormality is essential here. The projection Pr of $\mathbb{R}^2$ onto the x-axis is self-adjoint by (b) above, but in the basis $\mathbf{b}_1 = (1,0)$, $\mathbf{b}_2 = (1,1)$ we have $\mathrm{Pr}\,\mathbf{b}_1 = \mathbf{b}_1$ and $\mathrm{Pr}\,\mathbf{b}_2 = (1,0) = \mathbf{b}_1$, so its matrix is $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, which is not symmetric.

The next result completes Proposition 6.4.30, which characterized orthogonal projections without naming the adjoint.

Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents). Let V be finite-dimensional and $P \in \mathrm{L}(V)$. Then P is an orthogonal projection, that is, $P = \mathrm{Pr}_W$ for some subspace W of V, if and only if

$$P^2 = P = P^*.$$

In that case $W = \operatorname{im} P$.

Proof (from the earlier characterization). ROUTINE

Uses. Proposition 6.4.30 (Recognizing orthogonal projections), Definition 6.6.26 (Self-adjoint operator)
Hint. Read the second condition of Proposition 6.4.30 in the language of Definition 6.6.26.
Sketch. The condition $\langle P\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, P\mathbf{x}\rangle$ for all $\mathbf{v}, \mathbf{x}$ in Proposition 6.4.30 says exactly that P is self-adjoint.

The second condition in Proposition 6.4.30, $\langle P\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, P\mathbf{x}\rangle$ for all $\mathbf{v}, \mathbf{x} \in V$, says exactly that P is self-adjoint (Definition 6.6.26). So the corollary is Proposition 6.4.30 restated. □

Proof (via kernels and images). ROUTINE

Uses. Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.6.21 (Kernels and images of adjoints), Theorem 6.4.6 (The projection is characterized by orthogonality), Theorem 6.4.20 (Orthonormal bases exist), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. One direction is (b) and (c) of Proposition 6.4.29. For the other, let $W = \operatorname{im} P$ and split $\mathbf{v} = P\mathbf{v} + (\mathbf{v} - P\mathbf{v})$. Where does each piece lie?
Hint 2. $P(\mathbf{v} - P\mathbf{v}) = \mathbf{0}$, and for a self-adjoint P, Theorem 6.6.21(a) says $\ker P = (\operatorname{im} P)^{\perp}$.
Sketch. Orthogonal projections satisfy $P^2 = P$ and are self-adjoint (Proposition 6.4.29). Conversely, $P\mathbf{v} \in \operatorname{im} P$ and $P(\mathbf{v} - P\mathbf{v}) = P\mathbf{v} - P^2\mathbf{v} = \mathbf{0}$, so $\mathbf{v} - P\mathbf{v} \in \ker P = \ker P^* = (\operatorname{im} P)^{\perp}$. These are the two conditions that characterize $\mathrm{Pr}_{\operatorname{im} P}(\mathbf{v})$.

If $P = \mathrm{Pr}_W$, then $P^2 = P$ by Proposition 6.4.29(b), P is self-adjoint by (c) of the same proposition, and $\operatorname{im} P = W$ by part (a).

Conversely, suppose $P^2 = P = P^*$, and let $W = \operatorname{im} P$, a finite-dimensional subspace (Proposition 4.1.21), so that Pr_W is defined (Theorem 6.4.20). Let $\mathbf{v} \in V$. Then $P\mathbf{v} \in W$, and $P(\mathbf{v} - P\mathbf{v}) = P\mathbf{v} - P^2\mathbf{v} = \mathbf{0}$, so by Theorem 6.6.21(a),

$$\mathbf{v} - P\mathbf{v} \in \ker P = \ker P^* = (\operatorname{im} P)^{\perp} = W^{\perp}.$$

By Theorem 6.4.6, $P\mathbf{v} = \mathrm{Pr}_W(\mathbf{v})$. □

The second proof separates the roles of the two conditions. Any P with $P^2 = P$ projects onto $\operatorname{im} P$ along $\ker P$, since $\mathbf{v} = P\mathbf{v} + (\mathbf{v} - P\mathbf{v})$ with $\mathbf{v} - P\mathbf{v} \in \ker P$; this is how the oblique projections of Remark 6.4.12 and

Example 6.6.24 work. Self-adjointness is what makes $\ker P = (\operatorname{im} P)^{\perp}$, so that the direction of projection is perpendicular to the target.

Over $\mathbb{C}$, self-adjointness can be read off the "diagonal" values $\langle T\mathbf{v}, \mathbf{v}\rangle$ alone. This is the operator version of the fact that a complex number z is real exactly when $z = \overline{z}$.

Proposition 6.6.30 (Self-adjoint means real values). Let V be a finite-dimensional complex inner product space and $T \in \mathrm{L}(V)$. Then T is self-adjoint if and only if $\langle T\mathbf{v}, \mathbf{v}\rangle \in \mathbb{R}$ for every $\mathbf{v} \in V$.

Proof. STANDARD

Uses. Theorem 6.6.15 (Existence of adjoints), Definition 6.6.8 (Adjoint), Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v}, \mathbf{v}\rangle$ determine T), Notation 6.1.1 (Real and complex scalars), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. A complex number z is real iff $z - \overline{z} = 0$. Write $\overline{\langle T\mathbf{v}, \mathbf{v}\rangle}$ as an inner product with T^* in it.
Hint 2. $\overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle \mathbf{v}, T\mathbf{v}\rangle = \langle T^*\mathbf{v}, \mathbf{v}\rangle$. So $\langle T\mathbf{v}, \mathbf{v}\rangle - \overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle (T - T^*)\mathbf{v}, \mathbf{v}\rangle$.
Hint 3. Over $\mathbb{C}$, an operator S with $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$ is 0 (Proposition 6.1.41).
Sketch. For every $\mathbf{v}$, $\langle T\mathbf{v}, \mathbf{v}\rangle - \overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle T\mathbf{v}, \mathbf{v}\rangle - \langle T^*\mathbf{v}, \mathbf{v}\rangle = \langle (T - T^*)\mathbf{v}, \mathbf{v}\rangle$. So all values $\langle T\mathbf{v}, \mathbf{v}\rangle$ are real iff $\langle (T - T^*)\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$, iff $T - T^* = 0$ by Proposition 6.1.41.

By Theorem 6.6.15, T has an adjoint. For every $\mathbf{v} \in V$, conjugate symmetry and the conjugated form of the defining identity give

$$\overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle \mathbf{v}, T\mathbf{v}\rangle = \langle T^*\mathbf{v}, \mathbf{v}\rangle,$$

and therefore

$$\langle T\mathbf{v}, \mathbf{v}\rangle - \overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle (T - T^*)\mathbf{v}, \mathbf{v}\rangle.$$

A complex number is real exactly when it equals its conjugate (Notation 6.1.1). So $\langle T\mathbf{v}, \mathbf{v}\rangle \in \mathbb{R}$ for every $\mathbf{v}$ if and only if $\langle (T - T^*)\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$. By Proposition 6.1.41, applied to $T - T^* \in \mathrm{L}(V)$ (Proposition 2.1.23), this holds if and only if $T - T^* = 0$, that is, T is self-adjoint. □

Finite-dimensionality was used only to know that T^* exists. Exercise 6.1.53 proves the same equivalence on every complex inner product space, without adjoints, by polarization. Over $\mathbb{R}$ the proposition fails, since every $\langle T\mathbf{v}, \mathbf{v}\rangle$ is real; the rotation J is not self-adjoint. Over $\mathbb{R}$, the values $\langle T\mathbf{v}, \mathbf{v}\rangle$ see only the self-adjoint part $(T + T^*)/2$ of T (Exercise 6.1.52).

Intuition (Self-adjoint operators as real numbers). The analogy between $T \mapsto T^*$ and $z \mapsto \overline{z}$ runs through Chapter 8. Both are additive, conjugate scalars, and undo themselves, and adjoints also reverse products, which for numbers is invisible. Self-adjoint operators play the role of real numbers: their values $\langle T\mathbf{v}, \mathbf{v}\rangle$ are real (Proposition 6.6.30), and every operator on a complex space splits uniquely as $A + iB$ with A and B self-adjoint, like $z = x + iy$ (Exercise 6.6.37). Operators of the form T^*T play the role of $|z|^2 = \overline{z}z$: they are self-adjoint with $\langle T^*T\mathbf{v}, \mathbf{v}\rangle = \|T\mathbf{v}\|^2 \geq 0$ (Section 8.3). Operators with $T^*T = \mathrm{id}$ play the role of the unit circle; they are the maps that preserve lengths (Exercise 6.6.40, Section 8.4). The analogy is imperfect because operators need not commute, and it works best for the normal operators, those with $T^*T = TT^*$ (Section 8.1).

The main theorem about self-adjoint operators, proved in Section 8.2, is that on a finite-dimensional space they are exactly the operators that are diagonal, with real diagonal entries, in some orthonormal basis. Geometrically, a self-adjoint operator stretches V by real factors along mutually perpendicular axes. An orthogonal projection is the case in which every factor is 0 or 1. The oblique projection of Example 6.6.24 also fixes one line and kills another, but the two lines are not perpendicular, and that is exactly the failure of self-adjointness.

Exercises

Exercise 6.6.31. ROUTINE Let $T : \mathbb{C}^2 \to \mathbb{C}^2$, $T(z_1, z_2) = (z_1 + iz_2, (1-i)z_1)$.

(a) Find T^* when $\mathbb{C}^2$ has the dot product.

(b) Find T^* when $\mathbb{C}^2$ has the weighted inner product $\langle \mathbf{z}, \mathbf{w} \rangle = z_1\overline{w_1} + 2z_2\overline{w_2}$ (Example 6.1.21).

(c) In (b), is T^* multiplication by the conjugate transpose of the matrix of T? Why does this not contradict Theorem 6.6.15(b)?

Solution on page 1524

Exercise 6.6.32. ROUTINE Give $\mathbb{R}[x]_{\leq 2}$ the inner product $\langle p, q \rangle = \int_{-1}^{1} p(t)q(t)\,dt$. Find the polynomials $q_1, q_2 \in \mathbb{R}[x]_{\leq 2}$ such that, for every $p \in \mathbb{R}[x]_{\leq 2}$,

(a) $\langle p, q_1 \rangle = p'(0)$;

(b) $\langle p, q_2 \rangle = \int_0^1 p(t)\,dt$.

Solution on page 1525

Exercise 6.6.33. ROUTINE On $\mathbb{F}^n$ with the dot product ($n \geq 2$), let

$$R(x_1, ..., x_n) = (0, x_1, ..., x_{n-1}), \quad C(x_1, ..., x_n) = (x_n, x_1, ..., x_{n-1}).$$

(a) Find R^*, and compute R^*R and RR^*. Identify each as an orthogonal projection.

(b) Find C^*, and show that $C^* = C^{-1}$.

Solution on page 1525

Exercise 6.6.34. ROUTINE Let $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$, and give $\mathbb{C}^2$ the dot product.

(a) Find $\ker A$, $\operatorname{im} A$, $\ker A^*$, and $\operatorname{im} A^*$, and check the four statements of Theorem 6.6.21.

(b) Show that $\ker A$ is not orthogonal to $\operatorname{im} A^{\mathsf{T}}$, although $A^{\mathsf{T}} = A$.

(c) Is T_A self-adjoint?

Solution on page 1526

Exercise 6.6.35. STANDARD Let V be finite-dimensional and $T \in \mathrm{L}(V)$ with $T^*T = T$. Show that T is an orthogonal projection.

Solution on page 1526

Exercise 6.6.36. STANDARD Let V be a finite-dimensional inner product space, and let $S, T \in \mathrm{L}(V)$ be self-adjoint. Prove or disprove each statement.

(a) $S + T$ is self-adjoint.

(b) ST is self-adjoint.

(c) STS is self-adjoint.

(d) If $\mathbb{F} = \mathbb{C}$, then iT is self-adjoint.

(e) If T is invertible, then T^{-1} is self-adjoint.

Solution on page 1526

Exercise 6.6.37. STANDARD Let V be a finite-dimensional complex inner product space.

(a) Show that every $T \in \mathrm{L}(V)$ can be written in exactly one way as $T = A + iB$ with $A, B \in \mathrm{L}(V)$ self-adjoint.

(b) Find A and B for $T = T_M$ on $\mathbb{C}^2$ with the dot product, where $M = \begin{pmatrix} 1 & 2 \\ 0 & i \end{pmatrix}$.

Solution on page 1527

Exercise 6.6.38. STANDARD Let V be finite-dimensional, $T \in \mathrm{L}(V)$, and U a subspace of V. Show that T maps U into U if and only if T^* maps $U^\perp$ into $U^\perp$. Deduce that if T is self-adjoint and maps U into U, then T also maps $U^\perp$ into $U^\perp$.

Solution on page 1527

Exercise 6.6.39. STANDARD Let V be the real vector space spanned by $1, \cos t, \sin t, \cos 2t, \sin 2t$ (as functions on $[0, 2\pi]$), with $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)\,\mathrm{d}t$, an inner product by the proof of Proposition 6.1.13. Let $Df = f'$.

(a) Show that D maps V into V and that $D^* = -D$.

(b) Deduce that $\langle Df, f\rangle = 0$ for every $f \in V$, and that D is not self-adjoint.

(c) Find the matrix of D in the basis $1, \cos t, \sin t, \cos 2t, \sin 2t$. Using that this basis is orthogonal with $\|\cos kt\| = \|\sin kt\|$, explain how the matrix also shows $D^* = -D$.

Solution on page 1527

Exercise 6.6.40. STANDARD Let V and W be finite-dimensional and $T \in \mathrm{L}(V, W)$. Show that the following are equivalent:

(i) $\|T\mathbf{v}\| = \|\mathbf{v}\|$ for all $\mathbf{v} \in V$;
(ii) $\langle T\mathbf{v}, T\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$;
(iii) $T^*T = \mathrm{id}_V$.

What does (iii) say about A when $T = T_A$ with $A \in \mathbb{F}^{m\times n}$ and dot products?

Solution on page 1528

Exercise 6.6.41. STANDARD Let V and W be finite-dimensional and $T \in \mathrm{L}(V, W)$ with $T^*T = \mathrm{id}_V$. A student argues: "Since $T^*T = \mathrm{id}_V$, T^* is the inverse of T, so also $TT^* = \mathrm{id}_W$."

(a) Find the error, and give a counterexample to the conclusion.

(b) Show that $TT^* = \mathrm{Pr}_{\mathrm{im}\, T}$.

(c) Show that the conclusion is correct when $\dim V = \dim W$.

Solution on page 1528

Exercise 6.6.42. STANDARD Let V and W be inner product spaces, let $T \in \mathrm{L}(V, W)$ have an adjoint, and let $\mathbf{w} \in W$.

(a) Show that $\mathrm{Re}\langle T\mathbf{v}, \mathbf{w}\rangle \le \|T^*\mathbf{w}\|$ for every unit vector $\mathbf{v} \in V$, with equality for $\mathbf{v} = T^*\mathbf{w}/\|T^*\mathbf{w}\|$ when $T^*\mathbf{w} \neq \mathbf{0}$. So $T^*\mathbf{w}$ points along the input that T pushes furthest in the direction of $\mathbf{w}$.

(b) Let $c \ge 0$. Show that $\|T\mathbf{v}\| \le c\|\mathbf{v}\|$ for all $\mathbf{v} \in V$ if and only if $\|T^*\mathbf{w}\| \le c\|\mathbf{w}\|$ for all $\mathbf{w} \in W$.

Solution on page 1529

Exercise 6.6.43. STANDARD Let V be finite-dimensional and $T \in \mathrm{L}(V)$.

(a) Show that every vector of $\operatorname{im} T$ is orthogonal to every vector of $\operatorname{im} T^*$ if and only if $T^2 = 0$.

(b) Deduce that if $T^2 = 0$, then $\operatorname{rank} T \le (\dim V)/2$.

(c) Find a nonzero operator T on $\mathbb{R}^3$ with $T^2 = 0$, and check (a) for it.

Solution on page 1529

Exercise 6.6.44. STANDARD Let V be finite-dimensional and $T \in \mathrm{L}(V)$.

(a) Show that $\operatorname{rank}(TT^*) = \operatorname{rank}(T^*T) = \operatorname{rank} T$.

(b) Suppose $T^*T = TT^*$. Show that $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for all $\mathbf{v} \in V$, and deduce that $\ker T = \ker T^*$ and $\operatorname{im} T = \operatorname{im} T^*$.

(c) Find an operator T on $\mathbb{R}^2$ with $\ker T \neq \ker T^*$.

Solution on page 1529

Exercise 6.6.45. STANDARD On $V = \mathbb{R}[x]_{\le 2}$, let $(Sp)(x) = p(1-x)$.

(a) Show that S is self-adjoint for $\langle p, q\rangle = \int_0^1 p(t)q(t)\,dt$, but not for $\langle p, q\rangle = \int_0^2 p(t)q(t)\,dt$.

(b) For the first inner product, show that $(\mathrm{id} + S)/2$ is the orthogonal projection onto the subspace E of polynomials with $p(1-x) = p(x)$, and find a basis of E.

Solution on page 1530

Exercise 6.6.46. HARD Let V be finite-dimensional with a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$, not necessarily orthonormal, and let G be its Gram matrix, $g_{ij} = \langle \mathbf{b}_j, \mathbf{b}_i\rangle$ (Definition 6.3.28).

(a) Show that $\langle \mathbf{v}, \mathbf{w}\rangle = [\mathbf{w}]_{\mathcal{B}}^* G [\mathbf{v}]_{\mathcal{B}}$ for all $\mathbf{v}, \mathbf{w} \in V$.

(b) Show that $[T^*]_{\mathcal{B}} = G^{-1}([T]_{\mathcal{B}})^* G$ for every $T \in \mathrm{L}(V)$.

(c) Recompute $[D^*]_{\mathcal{B}}$ of Example 6.6.16 with (b).

(d) A student claims that $[T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^*$ for every basis, "since $\langle \mathbf{v}, \mathbf{w}\rangle = [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}}$, so the second proof of Theorem 6.6.15 applies." Which step fails, and for which bases is the claim true?

Solution on page 1530

Exercise 6.6.47. HARD Let V be finite-dimensional, and suppose $V = U \oplus X$ for subspaces U and X (Definition 4.4.4). Let $P \in \mathrm{L}(V)$ be the projection onto U along X: $P(\mathbf{u} + \mathbf{x}) = \mathbf{u}$ for $\mathbf{u} \in U$ and $\mathbf{x} \in X$.

(a) Show that $P^2 = P$, $\operatorname{im} P = U$, and $\ker P = X$.

(b) Show that $V = X^\perp \oplus U^\perp$ and that P^* is the projection onto $X^\perp$ along $U^\perp$.

(c) Deduce that P is an orthogonal projection if and only if $X = U^\perp$.

Solution on page 1531

Exercise 6.6.48. HARD

(a) Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Find an inner product on $\mathbb{R}^2$ for which T_A is self-adjoint. (Look for $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}} M \mathbf{x}$ as in Proposition 6.1.22.)

(b) Show that the rotation T_J, $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, is not self-adjoint for any inner product on $\mathbb{R}^2$.

Solution on page 1531

Exercise 6.6.49. HARD BEYOND 21-242 Let $P_n = \mathbb{R}[x]_{\le n}$ with $\langle p, q\rangle = \int_0^1 p(t)q(t)\,dt$, and let $q_n \in P_n$ represent evaluation at 0: $\langle p, q_n\rangle = p(0)$ for all $p \in P_n$ (Theorem 6.6.3).

(a) Show that $\|q_n\| \ge |p(0)|/\|p\|$ for every nonzero $p \in P_n$.

(b) Using $p(t) = (1-t)^n$, deduce that $\|q_n\| \ge \sqrt{2n+1}$.

(c) Deduce again (Non-example 6.6.7) that no polynomial q satisfies $\int_0^1 p(t)q(t)\,dt = p(0)$ for all polynomials p.

Solution on page 1532

Exercise 6.6.50. HARD Let $A \in \mathbb{C}^{n\times n}$ satisfy $A^*A = A^2$. Show that $A^* = A$.

Solution on page 1532

Exercise 6.6.51. CHALLENGE BEYOND 21-242 Let V be a finite-dimensional inner product space and $T \in \mathrm{L}(V)$ with $\|T\mathbf{v}\| \le \|\mathbf{v}\|$ for all $\mathbf{v} \in V$.

(a) Show that $T\mathbf{v} = \mathbf{v}$ if and only if $T^*\mathbf{v} = \mathbf{v}$.

(b) Deduce that $V = \ker(T - \mathrm{id}) \oplus \mathrm{im}(T - \mathrm{id})$, with the two subspaces orthogonal.

(c) Show that for every $\mathbf{v} \in V$, the averages $A_N\mathbf{v} = (1/N)\sum_{k=0}^{N-1} T^k\mathbf{v}$ converge to $\mathrm{Pr}_{\ker(T-\mathrm{id})}(\mathbf{v})$ as $N \to \infty$, in the sense that $\left\|A_N\mathbf{v} - \mathrm{Pr}_{\ker(T-\mathrm{id})}(\mathbf{v})\right\| \to 0$.

Solution on page 1533

6.7 Proof patterns

An inner product adds one number $\langle \mathbf{v}, \mathbf{w}\rangle$ to every pair of vectors, and the chapter turned that number into lengths, angles, projections, least-squares solutions, and adjoints. Most of its proofs use a small number of moves to get from a hypothesis about inner products to a conclusion about vectors or maps: expand a squared norm and use positivity, test a vector against everything, move an operator to the other slot, choose an orthonormal basis, or split a vector into orthogonal pieces. This section sorts the chapter's arguments by the question they answer, solves a few problems in several ways to compare the moves, and lists the errors that the chapter's exercises warn against. The exercises at the end practice one move at a time, and the problem set (Problems for Chapter 6) mixes them with earlier chapters.

6.7.1 Proving an inequality

Almost every inequality in the chapter is a squared norm that was dropped because it is nonnegative. The work is in choosing the vector whose norm to drop.

- *Expand a squared norm.* Write $0 \le \|\mathbf{x} - t\mathbf{y}\|^2$, or the norm of some other combination, and expand with Proposition 6.1.29. All three proofs of Theorem 6.2.1 start this way and differ only in the choice of t (Section 6.2.1). The same move gives the reverse triangle inequality (Corollary 6.2.14) and identities such as Exercise 6.1.57.
- *Drop an orthogonal piece* (Technique 6.2.4, Technique 6.4.5). If $\mathbf{x} = \mathbf{p} + \mathbf{q}$ with $\mathbf{p} \perp \mathbf{q}$, then $\|\mathbf{p}\| \le \|\mathbf{x}\|$. This is Bessel's inequality (Theorem 6.3.24), the best-approximation property (Theorem 6.4.33), and, with $\mathbf{p}$ the part of $\mathbf{x}$ along $\mathbf{y}$, Cauchy–Schwarz again.
- *Cauchy–Schwarz for a well-chosen inner product* (Technique 6.2.21). For inequalities between sums or integrals, the choice of space, weights, and splitting of each term is the whole proof (Proposition 6.2.22, Corollary 6.2.26, Exercise 6.2.36).
- *A discriminant* (Technique 6.2.7). When the target has the shape $b^2 \le ac$ and there is no picture, look for a quadratic in a real parameter that is nonnegative for an obvious reason.

- *Normalize.* Both sides of a homogeneous inequality scale the same way, so it suffices to check it on unit vectors (third proof of Theorem 6.2.1). Over $\mathbb{C}$, rotating the phase (Technique 6.2.8) makes an inner product real before a real-variable argument is run.

The equality case comes from the same computation: equality holds exactly when the dropped quantity is 0. That is why it pays to keep the dropped term written down, as in (6.2.2), instead of replacing it by "≥ 0" at once. The triangle inequality (Theorem 6.2.11(b)), Bessel's inequality, and best approximation all get their equality cases this way.

Example 6.7.1 (The Cauchy–Schwarz gap, three ways). Let $\mathbf{x}, \mathbf{y}$ be vectors in an inner product space V, and let

$$D = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - |\langle \mathbf{x}, \mathbf{y}\rangle|^2,$$

the amount by which Cauchy–Schwarz fails to be an equality. Here are three formulas for D.

By orthogonal decomposition. If $\mathbf{y} \neq \mathbf{0}$, (6.2.2) says $D = \|\mathbf{y}\|^2\|\mathbf{x} - c\mathbf{y}\|^2$ with $c = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$. Since $c\mathbf{y}$ is the projection of $\mathbf{x}$ onto the line $\operatorname{span}(\mathbf{y})$, this reads $D = \|\mathbf{y}\|^2 \operatorname{dist}(\mathbf{x}, \operatorname{span}(\mathbf{y}))^2$ (Definition 6.4.34).

In orthonormal coordinates. If $\mathbf{x}, \mathbf{y}$ are linearly dependent, $D = 0$ by Lemma 6.2.3. Otherwise $W = \operatorname{span}(\mathbf{x}, \mathbf{y})$ has an orthonormal basis $\mathcal{O} = \mathbf{u}_1, \mathbf{u}_2$ (Theorem 6.4.20). Write $\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$ and $\mathbf{y} = b_1\mathbf{u}_1 + b_2\mathbf{u}_2$. By Theorem 6.3.22, $\langle \mathbf{x}, \mathbf{y}\rangle = a_1\overline{b_1} + a_2\overline{b_2}$, $\|\mathbf{x}\|^2 = |a_1|^2 + |a_2|^2$, and $\|\mathbf{y}\|^2 = |b_1|^2 + |b_2|^2$. Multiplying out,

$$\|\mathbf{x}\|^2\|\mathbf{y}\|^2 = |a_1b_1|^2 + |a_1b_2|^2 + |a_2b_1|^2 + |a_2b_2|^2,$$
$$|\langle \mathbf{x}, \mathbf{y}\rangle|^2 = |a_1b_1|^2 + |a_2b_2|^2 + 2\operatorname{Re}\left(a_1b_2\overline{a_2b_1}\right),$$

and the difference is $|a_1b_2|^2 - 2\operatorname{Re}\left(a_1b_2\overline{a_2b_1}\right) + |a_2b_1|^2$. So

$$D = |a_1b_2 - a_2b_1|^2.$$

As a determinant. By definition (Definition 6.3.28), the Gram matrix of $\mathbf{x}, \mathbf{y}$ is

$$G = \begin{pmatrix} \langle \mathbf{x}, \mathbf{x}\rangle & \langle \mathbf{y}, \mathbf{x}\rangle \\ \langle \mathbf{x}, \mathbf{y}\rangle & \langle \mathbf{y}, \mathbf{y}\rangle \end{pmatrix},$$

and $\det G = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \langle \mathbf{y}, \mathbf{x}\rangle\langle \mathbf{x}, \mathbf{y}\rangle = D$. In the coordinates above, let $B = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, with columns $[\mathbf{x}]_{\mathcal{O}}$ and $[\mathbf{y}]_{\mathcal{O}}$. By Theorem 6.3.22, the (i, j) entry of B^*B is $[\mathbf{v}_i]_{\mathcal{O}}^*[\mathbf{v}_j]_{\mathcal{O}} = \langle \mathbf{v}_j, \mathbf{v}_i\rangle$, where $\mathbf{v}_1 = \mathbf{x}$ and $\mathbf{v}_2 = \mathbf{y}$; so $G = B^*B$. Then $\det G = \det(B^*)\det B = |\det B|^2$ (Theorem 5.4.10, Exercise 6.3.57), which is the second formula again.

Each formula proves Cauchy–Schwarz with its equality case: $D \geq 0$, and $D = 0$ exactly when $\mathbf{x} \in \operatorname{span}(\mathbf{y})$, or $\det B = 0$, or G is singular, all of which say that $\mathbf{x}, \mathbf{y}$ are dependent (Proposition 6.3.30). They lead in different directions. The first measures the gap by a distance, and Section 6.4 replaces the line by any subspace. The second writes D as one squared modulus, a certificate of nonnegativity in the sense of Technique 6.1.9; over $\mathbb{R}^n$ with the dot product it is Lagrange's identity (Exercise 6.2.43), and the orthonormal basis is what reduces every inner product space to $n = 2$. The third gives the equality case without any basis (Proposition 6.3.30), but its sign comes from $\det G = |\det B|^2$, that is, from the coordinates. It extends to k vectors: the Gram determinant of any list is nonnegative and vanishes exactly for dependent lists, by orthonormal coordinates again (Exercise 6.7.11) or by Gram–Schmidt residuals (Exercise 6.4.59). In $\mathbb{R}^2$, $|\det B|$ is the area of the parallelogram spanned by $\mathbf{x}$ and $\mathbf{y}$ (Remark 6.2.5).

The last two proofs could not have been given in Section 6.2, since they use orthonormal bases, but they are not circular: Gram–Schmidt and Parseval's identity never use Cauchy–Schwarz.

The middle formula is an instance of a move that the chapter used several times without naming it.

Technique 6.7.2 (Pass to orthonormal coordinates). A statement about finitely many vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of an inner product space V involves only $W = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, which is finite-dimensional even when V is not. Choose an orthonormal basis $\mathcal{O}$ of W (Theorem 6.4.20). By Theorem 6.3.22, the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{O}}$ preserves inner products, so W behaves exactly like $\mathbb{F}^r$ with the dot product, where $r = \dim W \leq k$. The list becomes the columns of a matrix $B \in \mathbb{F}^{r \times k}$ with $B^*B = G$, the Gram matrix of the list, and the statement can be checked with entries, sums of squares, and determinants.

Reach for it when a statement about a fixed finite list is easy to verify in $\mathbb{F}^r$ but awkward without coordinates, or when you need a matrix whose Gram matrix is prescribed. The canonical example is the second formula of Example 6.7.1. It also shows that Gram determinants are nonnegative (Exercise 6.7.11) and produces the matrix B in Problem 6.17. It does not apply as stated to a map on an infinite-dimensional V, since the map need not preserve W.

Used in. Exercise 6.7.11, Problem 6.17, Theorem 8.1.3, Proposition 8.1.7, Theorem 8.2.18, Exercise 8.2.42, Theorem 8.3.5, Lemma 8.3.8.

Practice: Exercise 6.7.6 and Exercise 6.7.12.

6.7.2 Showing that a vector or a map is zero

An inner product gives several ways to prove $\mathbf{v} = \mathbf{0}$, or $S = T$, which a bare vector space does not have.

- *Its norm* (Technique 6.1.36). Show $\langle \mathbf{v}, \mathbf{v} \rangle = 0$. If $\mathbf{v}$ is a matrix, show $\operatorname{tr}(\mathbf{v}^*\mathbf{v}) = 0$ (Technique 6.1.18). If the norm is a sum of squares, every term vanishes (Technique 6.1.9).
- *Test against everything.* If $\langle \mathbf{v}, \mathbf{x} \rangle = 0$ for every $\mathbf{x}$ in V, or in a spanning list, then $\mathbf{v} = \mathbf{0}$: take $\mathbf{x} = \mathbf{v}$ (Proposition 6.1.37, Exercise 6.1.50). For a finite-dimensional subspace W this is $W^\perp = \{\mathbf{0}\} \Rightarrow W = V$ (Corollary 6.4.27), and for maps it is Corollary 6.1.38.
- *Test on the diagonal.* Over $\mathbb{C}$, the numbers $\langle T\mathbf{v}, \mathbf{v} \rangle$ already determine T (Proposition 6.1.41). Over $\mathbb{R}$ they determine only self-adjoint operators, and the rotation by $\pi/2$ shows why (Example 6.1.39).
- *Move the operator across* (Technique 6.6.20), to turn the hypothesis into one of the forms above. The standard instance is $\langle T^*T\mathbf{v}, \mathbf{v} \rangle = \|T\mathbf{v}\|^2$, which gives $\ker T^*T = \ker T$ (Proposition 6.6.23).
- *Orthogonal pieces.* A vector in $W \cap W^\perp$ is orthogonal to itself, hence $\mathbf{0}$. This is how uniqueness of orthogonal decompositions is proved (Theorem 6.4.24), and how $\ker T^* = (\operatorname{im} T)^\perp$ is used (Theorem 6.6.21).

Example 6.7.3 (A normal matrix with square zero, three certificates). Let $A \in \mathbb{C}^{n \times n}$ satisfy $A^*A = AA^*$ and $A^2 = 0$. (Matrices with $A^*A = AA^*$ are called *normal*; Section 8.1 studies them.) We show that $A = 0$ in three ways.

Norms. With the dot product, A^* is the adjoint of $\mathbf{v} \mapsto A\mathbf{v}$ (Example 6.6.10). For every $\mathbf{v} \in \mathbb{C}^n$,

$$\|A\mathbf{v}\|^2 = \langle A^*A\mathbf{v}, \mathbf{v} \rangle = \langle AA^*\mathbf{v}, \mathbf{v} \rangle = \|A^*\mathbf{v}\|^2.$$

Applied to $A\mathbf{v}$ this gives $\|A^*A\mathbf{v}\| = \|A^2\mathbf{v}\| = 0$, so $A^*A\mathbf{v} = \mathbf{0}$, and then $\|A\mathbf{v}\|^2 = \langle A^*A\mathbf{v}, \mathbf{v} \rangle = 0$. So $A\mathbf{v} = \mathbf{0}$ for every $\mathbf{v}$.

The trace of B^*B. Let $B = A^*A$. Then $B^* = B$, and $B^2 = A^*(AA^*)A = A^*(A^*A)A = (A^*)^2A^2 = 0$. So $\operatorname{tr}(B^*B) = \operatorname{tr}(B^2) = 0$, which forces $B = 0$ (Technique 6.1.18), and $A^*A = 0$ forces $A = 0$ (Corollary 6.1.19).

Subspaces. $A^2 = 0$ says $\operatorname{im} A \subseteq \ker A$. By Proposition 6.5.9(a), applied to A and to A^*, $\ker A = \ker(A^*A) = \ker(AA^*) = \ker A^*$, and $\ker A^* = (\operatorname{im} A)^\perp$ by Lemma 6.5.4. So $\operatorname{im} A \subseteq (\operatorname{im} A)^\perp$: every $\mathbf{w} \in \operatorname{im} A$ is orthogonal to itself, and $\operatorname{im} A = \{\mathbf{0}\}$.

The norm argument is the shortest; it uses nothing but moving A across the inner product twice. The trace argument never mentions a vector, and it is the one to try when all hypotheses are matrix identities: one nonnegative number controls every entry. The subspace argument shows what normality does: it makes

$\ker A = \ker A^*$, so the image of A, which is orthogonal to $\ker A^*$, is also orthogonal to $\ker A$, and an A with $A^2 = 0$ would need its image inside that kernel.

Each argument breaks at the same hypothesis for $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, which has $N^2 = 0$ but is not normal: $\|N\mathbf{e}_2\| = 1$ while $\|N^*\mathbf{e}_2\| = 0$, and $\ker N = \operatorname{span}(\mathbf{e}_1)$ while $\ker N^* = \operatorname{span}(\mathbf{e}_2)$. The norm argument proves more than was asked. It shows that $\ker A^2 = \ker A$ for every normal A, so by induction $\ker A^k = \ker A$ for all $k \geq 1$, and a normal matrix with $A^k = 0$ is 0.

Practice: Exercise 6.7.7, Exercise 6.7.9, and Exercise 6.7.8.

6.7.3 Projections, distances, and least squares

The projection $\operatorname{Pr}_W(\mathbf{v})$, the closest point of W to $\mathbf{v}$, and the fitted values of a least-squares problem are the same vector (Theorem 6.4.33, Proposition 6.5.2). There are several ways to find it, or to use it without finding it.

- *Orthonormal basis first.* Once W has an orthonormal or orthogonal basis, the projection is a sum of one-term projections (Definition 6.4.11, Corollary 6.4.10), and in $\mathbb{F}^n$ its matrix is QQ^* (Theorem 6.4.31). This pays when an orthogonal spanning list is visible, or can be made by symmetry (Exercise 6.4.54).
- *Normal equations* (Theorem 6.5.5, Proposition 6.5.23). For any spanning list with Gram matrix G, solve $G\mathbf{c} = \mathbf{d}$. Nothing has to be orthogonalized.
- *The complement.* $\operatorname{Pr}_W = \mathrm{id} - \operatorname{Pr}_{W^\perp}$ (Proposition 6.4.29(d)), which is the better route when $W^\perp$ is the smaller space, as for hyperplanes (Example 6.4.35).
- *Recognize instead of compute* (Technique 6.4.7). To show that $\mathbf{p} = \operatorname{Pr}_W(\mathbf{v})$, check $\mathbf{p} \in W$ and $\mathbf{v} - \mathbf{p} \perp W$. This is also how projections onto related subspaces are compared.
- *Perturb a minimizer* (Technique 6.5.6). When a vector is defined by minimizing a squared norm, perturbing it in each direction produces the orthogonality condition, which is the second proof of Theorem 6.5.5.

Example 6.7.4 (One least-squares error, four ways). Fit a line $y = c_0 + c_1 x$ to the points $(1,1), (2,3), (3,2), (4,5)$. With $\mathbf{1} = (1,1,1,1)$, $\mathbf{x} = (1,2,3,4)$, $\mathbf{y} = (1,3,2,5)$, and $W = \operatorname{span}(\mathbf{1}, \mathbf{x}) \subseteq \mathbb{R}^4$, the squared least-squares error is $\operatorname{dist}(\mathbf{y}, W)^2$. Note $\|\mathbf{y}\|^2 = 39$.

Normal equations. With $A = (\mathbf{1}\ \mathbf{x})$,

$$A^{\mathsf{T}}A = \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{y} = \begin{pmatrix} 11 \\ 33 \end{pmatrix},$$

so $\mathbf{c} = (0, 11/10)$ and the line is $y = 11x/10$. The residual is $\mathbf{y} - A\mathbf{c} = (-1/10, 4/5, -13/10, 3/5)$, and its squared norm is $(1 + 64 + 169 + 36)/100 = 27/10$.

An orthogonal basis of W. One step of Gram–Schmidt replaces $\mathbf{x}$ by $\mathbf{x} - (10/4)\mathbf{1}$, and doubling gives $\mathbf{w} = (-3, -1, 1, 3) \perp \mathbf{1}$. By Corollary 6.4.10 and Pythagoras (Proposition 6.4.4),

$$\operatorname{dist}(\mathbf{y}, W)^2 = \|\mathbf{y}\|^2 - \frac{\langle \mathbf{y}, \mathbf{1} \rangle^2}{\|\mathbf{1}\|^2} - \frac{\langle \mathbf{y}, \mathbf{w} \rangle^2}{\|\mathbf{w}\|^2} = 39 - \frac{121}{4} - \frac{121}{20} = \frac{27}{10}.$$

The complement. $W^\perp$ has dimension 2 (Corollary 6.4.25). Among vectors of the form (a, b, b, a), the conditions $\perp \mathbf{1}$ and $\perp \mathbf{x}$ leave $\mathbf{g}_1 = (1, -1, -1, 1)$; among vectors of the form $(a, b, -b, -a)$, they leave $\mathbf{g}_2 = (1, -3, 3, -1)$. Vectors of these two forms are always orthogonal, so $\mathbf{g}_1, \mathbf{g}_2$ is an orthogonal basis of $W^\perp$. The residual is $\operatorname{Pr}_{W^\perp}(\mathbf{y}) = (1/4)\mathbf{g}_1 - (7/20)\mathbf{g}_2$ (Proposition 6.4.29(d)), and

$$\operatorname{dist}(\mathbf{y}, W)^2 = \frac{1^2}{4} + \frac{7^2}{20} = \frac{27}{10}.$$

Gram determinants. By Exercise 6.4.59(b), $\operatorname{dist}(\mathbf{y}, W)^2$ is the ratio of the Gram determinants of $\mathbf{1}, \mathbf{x}, \mathbf{y}$ and of $\mathbf{1}, \mathbf{x}$:

$$\det\begin{pmatrix} 4 & 10 & 11 \\ 10 & 30 & 33 \\ 11 & 33 & 39 \end{pmatrix} \Big/ \det\begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix} = \frac{54}{20} = \frac{27}{10}.$$

With two unknowns, the normal equations are the quickest route, and they also produce the line. The orthogonal basis costs one Gram–Schmidt step, after which each coefficient is a single inner product, and fitting a parabola instead would only add a term: the coefficients already found do not change. The complement computes the residual without the fitted values, and it is the efficient route when $W^\perp$ is smaller than W. Gram determinants need no solving, but they give only the distance, not the closest point.

Practice: Exercise 6.7.5 and Exercise 6.7.13.

6.7.4 Choosing an orthonormal basis

In an orthonormal basis, coordinates are inner products (Proposition 6.3.18), inner products are dot products of coordinates (Theorem 6.3.22), the adjoint is the conjugate transpose (Theorem 6.6.15(b)), self-adjoint operators are Hermitian matrices (Corollary 6.6.28), and an orthogonal projection is $\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ when the first basis vectors span its image (Corollary 6.4.25). So when a statement is basis-free but easier in coordinates, choose an orthonormal basis adapted to the problem: one whose first vectors span a given subspace (Theorem 6.4.20(b)), or one of the span of the vectors at hand (Technique 6.7.2). None of these formulas survives in a basis that is not orthonormal. There the Gram matrix G of the basis enters, as in $\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{w}]^* G [\mathbf{v}]$ and $[T^*] = G^{-1}[T]^* G$ (Exercise 6.6.46, Example 6.6.16).

Coordinates are not always the shorter route. Statements about kernels, images, and complements usually have one-line proofs without a basis, as $\ker T^* = (\operatorname{im} T)^\perp$ does (Theorem 6.6.21), and a coordinate proof of them hides the reason. And in an infinite-dimensional space only finite-dimensional subspaces are guaranteed an orthonormal basis, so an argument that picks an orthonormal basis of V has assumed $\dim V < \infty$. A useful check after a coordinate proof is to ask whether the quantity computed depends on the basis; if the statement was basis-free, it cannot, and showing that directly sometimes gives a better proof (Exercise 6.7.10).

Practice: Exercise 6.7.10 and Exercise 6.7.11.

6.7.5 Lengths determine inner products

By polarization (Proposition 6.1.32, Technique 6.1.33), $\langle \mathbf{x}, \mathbf{y} \rangle$ is a combination of the four lengths $\|\mathbf{x} + i^k \mathbf{y}\|$, or of two over $\mathbb{R}$. So a hypothesis about lengths alone is a hypothesis about inner products: a linear map that preserves lengths preserves inner products (Exercise 6.1.51, Exercise 6.6.40), and two inner products with the same norm are equal (Corollary 6.1.34). For operators, the corresponding statement is about the numbers $\langle T\mathbf{v}, \mathbf{v} \rangle$, and here real and complex scalars part ways: over $\mathbb{C}$ these numbers determine T (Proposition 6.1.41), and they are all real exactly when T is self-adjoint (Proposition 6.6.30); over $\mathbb{R}$ they see only the self-adjoint part $(T + T^*)/2$ (Exercise 6.1.52). The general move behind all of this is to apply a hypothesis to $\mathbf{x} + c\mathbf{y}$ for a few scalars c and solve for the cross term (Technique 6.1.40).

A useful habit is to write every proof over $\mathbb{F}$, with Re wherever a cross term appears, and to specialize to $\mathbb{R}$ by erasing the bars. The differences between the two cases are concentrated in a few places: the converse of Pythagoras (Theorem 6.3.3(b), Non-example 6.3.4), the polarization identities, rotating the phase, and tests on the diagonal. Practice: Exercise 6.7.15 and Exercise 6.7.8.

6.7.6 Reading the chapter's proofs

Uniqueness first, then existence. The projection was characterized by two conditions before anything was computed (Theorem 6.4.6), and uniqueness then gave independence of the basis for free (Corollary 6.4.8). Adjoints went the same way: uniqueness (Proposition 6.6.9) needs only that inner products distinguish vectors, while existence (Theorem 6.6.15) needs a finite-dimensional space and either the Riesz representation theorem or a matrix computation. Recognizing an object by a property that determines it (Technique 6.4.7, Technique 6.6.11) is the inner-product version of checking that a candidate is an inverse (Technique 2.5.16); compare Section 5.6.5.

Where finite dimension enters. Every existence statement of the chapter needs an orthonormal basis of something, and that is the only place where finite dimension is used. For Pr_W and $V = W \oplus W^\perp$, only W has to be finite-dimensional (Theorem 6.4.24). The identity $\ker T^* = (\operatorname{im} T)^\perp$ needs no dimension hypothesis, while $\operatorname{im} T^* = (\ker T)^\perp$ does (Theorem 6.6.21). When a hypothesis of this kind is dropped, the conclusion can fail (Exercise 6.4.63, Non-example 6.6.7).

Orthogonality turns sums into sums of squares. Pythagoras (Corollary 6.3.6), Parseval (Theorem 6.3.22), Bessel (Theorem 6.3.24), and the defect formula (6.2.2) all say that the squared length of a sum of orthogonal pieces is the sum of the squared lengths. That is what makes coefficients in an orthonormal basis independent of each other (Technique 6.3.12), and why projections, unlike arbitrary idempotents, shrink vectors (Exercise 6.4.57).

The scalars are $\mathbb{R}$ or $\mathbb{C}$ for a reason. Positivity and definiteness use the order on $\mathbb{R}$. Over $\mathbb{F}_p$ or with a symmetric bilinear form over $\mathbb{C}$, there are nonzero vectors orthogonal to themselves (Exercise 6.1.58, Proposition 6.1.6), and then projections, Gram–Schmidt, and the normal equations can all fail (Exercise 6.5.44). Section 12.4 studies what survives.

6.7.7 Common errors

Most of these appear in the chapter as find-the-error exercises.

- *A conjugate in the wrong place.* The inner product is conjugate-linear in the second slot: $\langle \mathbf{v}, a\mathbf{w}\rangle = \bar{a}\langle \mathbf{v}, \mathbf{w}\rangle$ (Proposition 6.1.4). Coordinates need the vector in the first slot (Remark 6.3.19), and Gram–Schmidt with the slots exchanged produces vectors that are not orthogonal (Exercise 6.4.21).
- *The transpose instead of the conjugate transpose.* Over $\mathbb{C}$, $A^{\mathsf{T}}A$ can be singular when the columns of A are independent, and $\sum_k z_k w_k$ is not an inner product (Remark 6.5.10, Exercise 6.2.30).
- *Real facts used over $\mathbb{C}$, and complex facts used over $\mathbb{R}$.* The converse of Pythagoras and the two-term polarization identity hold only over $\mathbb{R}$, and a discriminant argument with a complex parameter proves nothing (Exercise 6.2.35). In the other direction, $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$ forces $T = 0$ only over $\mathbb{C}$ (Example 6.1.39, Exercise 6.1.59, Problem 6.6).
- *Projecting onto a spanning list that is not orthogonal*, one vector at a time (Exercise 6.4.42), and its relative $\mathrm{Pr}_{U+W} = \mathrm{Pr}_U + \mathrm{Pr}_W$ (Exercise 6.4.50).
- *A one-sided inverse used as a two-sided one.* $A(A^*A)^{-1}A^*$ is not I_m (Exercise 6.5.36), $T^*T = \mathrm{id}$ does not give $TT^* = \mathrm{id}$ (Exercise 6.6.41), and orthogonal columns do not give orthogonal rows (Exercise 6.3.55).
- *The conjugate transpose in a basis that is not orthonormal* (Example 6.6.16, Exercise 6.6.46).
- *Assuming the orthogonality to be proved.* Pythagoras needs orthogonal pieces, and using it for $\mathbf{v} = P\mathbf{v} + (\mathbf{v} - P\mathbf{v})$ assumes that P is an orthogonal projection (Exercise 6.7.13). Swapping TT^* for T^*T assumes normality (Exercise 6.7.14).
- *A hypothesis that is never used.* In both of the previous errors, the argument never touches one of the hypotheses. That is worth checking in every proof, and it is a warning sign whenever it happens.
- *Dropping finite dimension* (Exercise 6.4.63, Non-example 6.6.7), or definiteness (Exercise 6.1.60).

Exercises

Exercise 6.7.5. ROUTINE For each problem, choose a method, say why it fits, and compute.

(a) The distance from $\mathbf{v} = (3, 1, 4, 1)$ to $W = \operatorname{span}((1, 1, 1, 1), (1, -1, 1, -1))$ in $\mathbb{R}^4$.

(b) The distance from $\mathbf{v} = (1, 2, 3, 4)$ to the hyperplane $x_1 - x_2 + x_3 - x_4 = 0$ in $\mathbb{R}^4$.

(c) The least-squares solution of $x_1 + x_2 = 1$, $x_1 + 2x_2 = 2$, $x_1 + 3x_2 = 2$.

Solution on page 1533

Exercise 6.7.6. STANDARD Let $\mathbf{x}, \mathbf{y}$ be vectors in an inner product space.

(a) Show that $\|\mathbf{x} + t\mathbf{y}\| \geq \|\mathbf{x}\|$ for every $t \in \mathbb{F}$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

(b) Over $\mathbb{C}$, show that it is not enough to require the inequality in (a) for real t only.

Solution on page 1534

Exercise 6.7.7. STANDARD Let V be a finite-dimensional inner product space and $T \in \mathrm{L}(V)$, and suppose that $\langle T\mathbf{u}, \mathbf{v} \rangle = 0$ whenever $\mathbf{u} \perp \mathbf{v}$. Show that $T = c\mathrm{id}_V$ for some $c \in \mathbb{F}$.

Solution on page 1534

Exercise 6.7.8. STANDARD Let V be a finite-dimensional inner product space and $T \in \mathrm{L}(V)$. Show that $T^*T = TT^*$ if and only if $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for every $\mathbf{v} \in V$. (Such T are called *normal*; see Section 8.1.) Over $\mathbb{R}$, where does your proof use that $T^*T - TT^*$ is self-adjoint?

Solution on page 1535

Exercise 6.7.9. STANDARD Let V be a finite-dimensional inner product space, and let $T \in \mathrm{L}(V)$ be self-adjoint with $T^3 = T^2$.

(a) Show that $T^2 = T$, so that T is an orthogonal projection.

(b) Show by an example that (a) fails without the hypothesis $T^* = T$.

Solution on page 1535

Exercise 6.7.10. STANDARD Let V, W be finite-dimensional inner product spaces and $T \in \mathrm{L}(V, W)$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis of V and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ one of W.

(a) Show that

$$\sum_{j=1}^{n} \|T\mathbf{e}_j\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} |\langle T\mathbf{e}_j, \mathbf{f}_i \rangle|^2 = \sum_{i=1}^{m} \|T^*\mathbf{f}_i\|^2.$$

(b) Deduce that $\sum_j \|T\mathbf{e}_j\|^2$ does not depend on the choice of the orthonormal basis of V, and that it equals $\operatorname{tr}(A^*A)$ for the matrix A of T with respect to any orthonormal bases of V and W.

Solution on page 1536

Exercise 6.7.11. STANDARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be vectors in an inner product space V, with Gram matrix G (Definition 6.3.28). Show that $\det G$ is a nonnegative real number, and that $\det G = 0$ if and only if the list is linearly dependent. Use an orthonormal basis of $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, not Gram–Schmidt residuals as in Exercise 6.4.59.

Solution on page 1536

Exercise 6.7.12. HARD Let $n \geq 2$ and $x \in \mathbb{R}$. Show that there are unit vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in some real inner product space with $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = x$ for all $i \neq j$ if and only if

$$-\frac{1}{n-1} \leq x \leq 1.$$

Describe the configurations at the two endpoints.

Solution on page 1537

Exercise 6.7.13. STANDARD Find the error in the following argument, and then prove the claim correctly. (The claim is true; it is part of Exercise 6.4.57.)

Claim. Let V be a finite-dimensional inner product space, and let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$ and $\|P\mathbf{v}\| \leq \|\mathbf{v}\|$ for every $\mathbf{v} \in V$. Then $P = \mathrm{Pr}_{\mathrm{im}\, P}$.

Proof. Let $\mathbf{v} \in V$. Since $P(\mathbf{v} - P\mathbf{v}) = P\mathbf{v} - P^2\mathbf{v} = \mathbf{0}$, we have $\mathbf{v} = P\mathbf{v} + (\mathbf{v} - P\mathbf{v})$ with $P\mathbf{v} \in \mathrm{im}\, P$ and $\mathbf{v} - P\mathbf{v} \in \ker P$. By the Pythagorean theorem, $\|\mathbf{v}\|^2 = \|P\mathbf{v}\|^2 + \|\mathbf{v} - P\mathbf{v}\|^2$, so the two pieces are orthogonal. Hence $\mathbf{v} - P\mathbf{v} \perp \mathrm{im}\, P$, and $P\mathbf{v} = \mathrm{Pr}_{\mathrm{im}\, P}(\mathbf{v})$ by Theorem 6.4.6.

Solution on page 1537

Exercise 6.7.14. STANDARD Find the error in the following argument, give an operator for which its first displayed equation fails, and prove the claim correctly.

Claim. Let V be a finite-dimensional inner product space and $T \in \mathrm{L}(V)$ with $\|T\mathbf{v}\| \leq \|\mathbf{v}\|$ for all $\mathbf{v}$. Then $\|T^*\mathbf{v}\| \leq \|\mathbf{v}\|$ for all $\mathbf{v}$.

Proof. For every $\mathbf{v} \in V$,

$$\|T^*\mathbf{v}\|^2 = \langle TT^*\mathbf{v}, \mathbf{v} \rangle = \langle T^*T\mathbf{v}, \mathbf{v} \rangle = \|T\mathbf{v}\|^2,$$

and $\|T\mathbf{v}\|^2 \leq \|\mathbf{v}\|^2$.

Solution on page 1538

Exercise 6.7.15. ROUTINE For each statement, decide whether it holds in every real inner product space, and whether it holds in every complex one. Here $\mathbf{x}, \mathbf{y}, \mathbf{v} \in V$ and $T \in \mathrm{L}(V)$, with V finite-dimensional.

(a) If $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for every $\mathbf{v}$, then $T = 0$.

(b) If $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, then $\mathbf{x} \perp \mathbf{y}$.

(c) $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}\big(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\big)$.

(d) If $\langle T\mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$ for every $\mathbf{v}$, then T is self-adjoint.

Solution on page 1538

Problems for Chapter 6

These problems combine the sections of Chapter 6 with each other and with earlier chapters: linear systems, dimension, affine subspaces, determinants, and traces. The first three are review prompts: state a main result, prove it without looking, and then answer a question that the text did not ask. The rest are ordered roughly by difficulty. Problems of kind "competition" are harder than anything the course expects, and two problems that go beyond the course are marked. Section 6.7 compares the techniques these problems call for.

Review: state and prove

Problem 6.1. STANDARD

(a) State the Cauchy–Schwarz inequality, including its equality case.

(b) Prove it, by any of the three methods of Section 6.2.1, without looking.

(c) Let $\mathbf{x}, \mathbf{y}$ be nonzero vectors in an inner product space. Show that

$$\left\| \frac{\mathbf{x}}{\|\mathbf{x}\|} - \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\| \le \frac{2\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|},$$

and that equality holds if and only if $\|\mathbf{x}\| = \|\mathbf{y}\|$ or $\mathbf{y}$ is a negative real multiple of $\mathbf{x}$.

Solution on page 1539

Problem 6.2. STANDARD Let W be a finite-dimensional subspace of an inner product space V.

(a) State the orthogonal decomposition theorem and the best-approximation theorem for W.

(b) Prove both, using only that W has an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ (Theorem 6.4.20) and the formula $\operatorname{Pr}_W(\mathbf{v}) = \sum_j \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j$.

(c) On $\mathbb{R}^2$, let $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|)$. Find all points of $W = \operatorname{span}(\mathbf{e}_1)$ closest to $\mathbf{e}_2$ for this norm. Which step of your proof in (b) has no analogue for $\|\cdot\|_\infty$?

Solution on page 1540

Problem 6.3. STANDARD

(a) State the theorem on the kernels and images of T and T^* (Section 6.6.4).

(b) Let $T \in \mathrm{L}(V, W)$ have an adjoint. Prove that $\ker T^* = (\operatorname{im} T)^\perp$, and, if V is finite-dimensional, that $\operatorname{im} T^* = (\ker T)^\perp$.

(c) (*Fredholm alternative.*) Let $A \in \mathbb{F}^{m \times n}$ and $\mathbf{b} \in \mathbb{F}^m$. Show that exactly one of the following holds: $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x} \in \mathbb{F}^n$; or there is $\mathbf{y} \in \mathbb{F}^m$ with $A^*\mathbf{y} = \mathbf{0}$ and $\mathbf{y}^*\mathbf{b} \neq 0$. Use this to decide whether the system

$$x_1 + x_2 = 1, \quad x_2 + x_3 = 1, \quad x_1 - x_3 = 1$$

has a solution, and likewise with right-hand side $(1, 1, 0)$.

Solution on page 1540

Synthesis problems

Problem 6.4. STANDARD Let $\mathbf{p}_0, \dots, \mathbf{p}_k \in \mathbb{R}^n$ be such that $\mathbf{v}_i = \mathbf{p}_i - \mathbf{p}_0$ $(1 \le i \le k)$ are linearly independent, and let $A = \mathbf{p}_0 + \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, the smallest affine subspace containing the points (Section 4.7).

(a) Show that exactly one point $\mathbf{c} \in A$ is equidistant from $\mathbf{p}_0, \dots, \mathbf{p}_k$. (It is the *circumcenter* of the points.)

(b) Show that $\mathbf{q} \in \mathbb{R}^n$ is equidistant from $\mathbf{p}_0, \dots, \mathbf{p}_k$ if and only if $\mathbf{q} - \mathbf{c} \perp \mathbf{v}_i$ for every i. Deduce that $\mathbf{c}$ is the point of A closest to every such $\mathbf{q}$, and that for $k = n$ the circumcenter is the only point of $\mathbb{R}^n$ equidistant from the $n+1$ points.

(c) Find the circumcenter and the common distance for $\mathbf{p}_0 = \mathbf{0}$, $\mathbf{p}_1 = (2, 0, 0)$, and $\mathbf{p}_2 = (1, 2, 2)$.

Solution on page 1541

Problem 6.5. STANDARD Let $x_0, \dots, x_n \in \mathbb{R}$ be distinct, let $V = \mathbb{R}[x]_{\le n}$, and let

$$\langle p, q\rangle = \sum_{k=0}^{n} p(x_k)q(x_k).$$

(a) Show that this is an inner product on V, and that the Lagrange polynomials $L_0, \dots, L_n$ of the nodes (Definition 3.4.26) form an orthonormal basis, in which the coordinates of p are $p(x_0), \dots, p(x_n)$.

(b) Deduce that the matrix with (i, j) entry $\sum_k x_k^{i+j}$ $(0 \le i, j \le n)$ is invertible. (Compare Problem 5.9.)

(c) Let $y_0, \dots, y_n \in \mathbb{R}$, let $p \in V$ satisfy $p(x_k) = y_k$ for all k (Theorem 3.4.24), and let $d < n$. Show that $\mathrm{Pr}_{\mathbb{R}[x]_{\le d}}(p)$ is the least-squares polynomial of degree at most d for the data (x_k, y_k), in the sense of Proposition 6.5.21.

(d) For the nodes $-1, 0, 1$ and the values $0, 1, 3$, find p and the least-squares line by (c).

Solution on page 1542

Problem 6.6. STANDARD Find the error in the following argument, and show that the claim is false.

Claim. Let V be a finite-dimensional real inner product space and $T \in \mathrm{L}(V)$ with $\langle T\mathbf{v}, \mathbf{v}\rangle \ge 0$ for every $\mathbf{v} \in V$. Then T is self-adjoint.

Proof. Let $S = T - T^*$. For every $\mathbf{v}$,

$$\langle T^*\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \langle T\mathbf{v}, \mathbf{v}\rangle,$$

the last step because $\langle T\mathbf{v}, \mathbf{v}\rangle \ge 0$ is real. So $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$, and $S = 0$ by Proposition 6.1.41.

(a) Then show that the same argument is correct for complex inner product spaces.

(b) Over $\mathbb{R}$, show that $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle H\mathbf{v}, \mathbf{v}\rangle$ for every $\mathbf{v}$, where $H = (T + T^*)/2$.

Solution on page 1543

Problem 6.7. STANDARD Prove or disprove each statement.

(a) If $A \in \mathbb{R}^{m\times n}$ has orthonormal columns, then $AA^{\mathsf{T}} = I_m$.

(b) For every $A \in \mathbb{C}^{n\times n}$, $\det(A^*A)$ is a nonnegative real number.

(c) If $P \in \mathbb{R}^{n\times n}$ satisfies $P^2 = P$ and $P^{\mathsf{T}}P = PP^{\mathsf{T}}$, then P is the matrix of an orthogonal projection.

(d) If U, W are subspaces of a finite-dimensional inner product space V, then $\dim\big(U \cap W^{\perp}\big) \ge \dim U - \dim W$.

(e) For $A \in \mathbb{R}^{n\times n}$: $\|A\mathbf{x}\| \le \|\mathbf{x}\|$ for every $\mathbf{x} \in \mathbb{R}^n$ if and only if $|\det A| \le 1$.

Solution on page 1544

Problem 6.8. HARD Let V be a real inner product space and $f : V \to V$ a function (not assumed linear) with $f(\mathbf{0}) = \mathbf{0}$ and

$$\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

(a) Show that $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ and $\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ for all $\mathbf{x}, \mathbf{y}$.

(b) Show that f is linear. Give a proof that works for every V, and, if V is finite-dimensional, a second proof with an orthonormal basis.

(c) Deduce that every distance-preserving $g : \mathbb{R}^n \to \mathbb{R}^n$ has the form $g(\mathbf{x}) = Q\mathbf{x} + \mathbf{b}$ with $Q^{\mathsf{T}}Q = I_n$.

(d) Now let V be a complex inner product space and $f : V \to V$ satisfy the same two conditions. Show that $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ and $f(t\mathbf{x}) = tf(\mathbf{x})$ for real t, but that f need not be linear over $\mathbb{C}$.

Solution on page 1544

Problem 6.9. HARD Let V, W be inner product spaces over $\mathbb{F}$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and $\mathbf{w}_1, \ldots, \mathbf{w}_k \in W$ have the same Gram matrix: $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = \langle \mathbf{w}_j, \mathbf{w}_i \rangle$ for all i, j.

(a) Show that $\|c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k\| = \|c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k\|$ for all $c_1, \ldots, c_k \in \mathbb{F}$. Deduce that $\sum_i c_i\mathbf{v}_i = \mathbf{0}$ if and only if $\sum_i c_i\mathbf{w}_i = \mathbf{0}$.

(b) Show that there is exactly one linear map $S : \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \to \operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_k)$ with $S\mathbf{v}_i = \mathbf{w}_i$ for all i, and that S is bijective with $\langle S\mathbf{x}, S\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y}$.

(c) Let $A, B \in \mathbb{F}^{m\times n}$. Show that $A^*A = B^*B$ if and only if $B = UA$ for some unitary $U \in \mathbb{F}^{m\times m}$.

Solution on page 1545

Problem 6.10. HARD Let V be a finite-dimensional inner product space, and let U, W be subspaces with $\dim U = \dim W$. Give U and W the restricted inner products, and define $A : U \to W$ and $B : W \to U$ by $A\mathbf{u} = \operatorname{Pr}_W(\mathbf{u})$ and $B\mathbf{w} = \operatorname{Pr}_U(\mathbf{w})$.

(a) Show that $\langle A\mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, B\mathbf{w} \rangle$ for all $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Conclude that $B = A^*$.

(b) Show that the following are equivalent:
- (i) $U \cap W^\perp = \{\mathbf{0}\}$;
- (ii) $W \cap U^\perp = \{\mathbf{0}\}$;
- (iii) A is bijective;
- (iv) B is bijective.

(c) Show by an example that (i) and (ii) need not be equivalent when $\dim U \neq \dim W$.

Solution on page 1546

Problem 6.11. HARD Give $\mathbb{F}^{n\times n}$ the Frobenius inner product $\langle X, Y \rangle = \operatorname{tr}(Y^*X)$ (Proposition 6.1.16). Let C be the span of all commutators $XY - YX$ with $X, Y \in \mathbb{F}^{n\times n}$, and let $Z = \{A \in \mathbb{F}^{n\times n} \mid \operatorname{tr} A = 0\}$.

(a) Show that $C \subseteq Z$ and that $Z = \operatorname{span}(I_n)^\perp$.

(b) Show that $\langle XY - YX, M \rangle = \langle Y, X^*M - MX^* \rangle$ for all X, Y, M. Deduce that every $M \in C^\perp$ commutes with every matrix.

(c) Deduce that $C = Z$: every matrix of trace 0 is a sum of commutators.

(d) Give a second proof of $C = Z$, valid over every field F, by writing each matrix in the spanning list of Z found in Example 2.6.28 as a commutator of matrix units.

Solution on page 1547

Problem 6.12. HARD Let $A \in \mathbb{F}^{m\times n}$.

(a) Show that $I_n + A^*A$ is the Gram matrix of the columns $\mathbf{c}_1, \ldots, \mathbf{c}_n$ of the matrix $C \in \mathbb{F}^{(n+m)\times n}$ that has I_n on top of A.

(b) Show that $\det(I_n + A^*A) \geq 1$, with equality if and only if $A = 0$.

Solution on page 1547

Solution on page 1548

Problem 6.13. HARD Let V be a real inner product space of dimension $n \geq 1$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ be nonzero vectors with $\langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq 0$ for all $i \neq j$. Show that $m \leq 2n$, and give an example with $m = 2n$.

Solution on page 1548

Problem 6.14. HARD BEYOND 21-242 (*The kernel trick.*) Let $\Phi \in \mathbb{R}^{m\times N}$ have rows $\boldsymbol{\varphi}_1^{\mathsf{T}}, \dots, \boldsymbol{\varphi}_m^{\mathsf{T}}$ (the *feature vectors* of m data points), let $\mathbf{y} \in \mathbb{R}^m$, and let $K = \Phi\Phi^{\mathsf{T}}$, the matrix with entries $\boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_j$.

(a) Show that the least-squares solution of minimal norm $\mathbf{w}^+$ of $\Phi\mathbf{w} = \mathbf{y}$ (Proposition 6.5.27) lies in $\operatorname{im}\Phi^{\mathsf{T}}$.

(b) Show that for $\boldsymbol{\alpha} \in \mathbb{R}^m$, $\Phi^{\mathsf{T}}\boldsymbol{\alpha}$ is a least-squares solution of $\Phi\mathbf{w} = \mathbf{y}$ if and only if $\boldsymbol{\alpha}$ is a least-squares solution of $K\boldsymbol{\alpha} = \mathbf{y}$. Deduce that $\mathbf{w}^+ = \Phi^{\mathsf{T}}\boldsymbol{\alpha}$ for every least-squares solution $\boldsymbol{\alpha}$ of $K\boldsymbol{\alpha} = \mathbf{y}$, and that for any $\boldsymbol{\varphi} \in \mathbb{R}^N$,

$$\boldsymbol{\varphi} \cdot \mathbf{w}^+ = \sum_{i=1}^{m} \alpha_i \, \boldsymbol{\varphi} \cdot \boldsymbol{\varphi}_i.$$

(c) Data points $t = -1, 0, 1$ have values $\mathbf{y} = (0, 1, 3)$, and each point t has the feature vector $\boldsymbol{\varphi}(t) = \left(1, \sqrt{2}t, t^2\right)$. Using only the kernel $k(s,t) = \boldsymbol{\varphi}(s) \cdot \boldsymbol{\varphi}(t)$, compute K, $\boldsymbol{\alpha}$, and the fitted function $t \mapsto \boldsymbol{\varphi}(t) \cdot \mathbf{w}^+$.

Solution on page 1549

Problem 6.15. HARD For a continuous $f : [0,1] \to \mathbb{R}$, let

$$I(f) = \int_0^1 x^2 f(x)\,\mathrm{d}x \quad \text{and} \quad J(f) = \int_0^1 x f(x)^2\,\mathrm{d}x.$$

Find the largest value of $I(f) - J(f)$ over all such f, and the functions attaining it.

Solution on page 1550

Problem 6.16. HARD Let a, b, c be positive real numbers with $abc = 1$. Show that

$$\frac{1}{a^3(b+c)} + \frac{1}{b^3(c+a)} + \frac{1}{c^3(a+b)} \geq \frac{3}{2},$$

and find when equality holds.

Solution on page 1550

Problem 6.17. CHALLENGE BEYOND 21-242 Let $G \in \mathbb{F}^{k\times k}$ satisfy $G^* = G$ and $\mathbf{c}^* G\mathbf{c} \geq 0$ for every $\mathbf{c} \in \mathbb{F}^k$. For $\mathbf{c}, \mathbf{d} \in \mathbb{F}^k$, let $\beta(\mathbf{c}, \mathbf{d}) = \mathbf{d}^* G\mathbf{c}$.

(a) Show that $|\beta(\mathbf{c}, \mathbf{d})|^2 \leq \beta(\mathbf{c}, \mathbf{c})\beta(\mathbf{d}, \mathbf{d})$ for all $\mathbf{c}, \mathbf{d}$. (Decide which proof of Theorem 6.2.1 survives without definiteness.)

(b) Show that $\beta(\mathbf{c}, \mathbf{c}) = 0$ if and only if $G\mathbf{c} = \mathbf{0}$.

(c) On the quotient space $V = \mathbb{F}^k / \ker G$ (Definition 4.6.20), show that $\langle [\mathbf{c}], [\mathbf{d}] \rangle = \mathbf{d}^* G\mathbf{c}$ is well defined and is an inner product, and that G is the Gram matrix of $[\mathbf{e}_1], \dots, [\mathbf{e}_k]$.

(d) Deduce that a matrix G with $G^* = G$ is the Gram matrix of a list in some inner product space if and only if $\mathbf{c}^* G\mathbf{c} \geq 0$ for all $\mathbf{c}$, if and only if $G = B^* B$ for some $B \in \mathbb{F}^{r\times k}$, where one can take $r = \operatorname{rank} G$.

Solution on page 1551

Problem 6.18. CHALLENGE Let $\mathcal{G}$ be a finite set of invertible matrices in $\mathbb{R}^{n\times n}$ that contains I_n and is closed under multiplication and inverses, and let $|\mathcal{G}|$ be the number of its elements.

(a) Show that $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}} = \sum_{M \in \mathcal{G}} (M\mathbf{x}) \cdot (M\mathbf{y})$ is an inner product on $\mathbb{R}^n$, and that $\langle N\mathbf{x}, N\mathbf{y} \rangle_{\mathcal{G}} = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}}$ for every $N \in \mathcal{G}$.

(b) Let $P = |\mathcal{G}|^{-1} \sum_{M \in \mathcal{G}} M$. Show that $NP = P$ for every $N \in \mathcal{G}$, that $P^2 = P$, and that P is self-adjoint for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$.

(c) Deduce that $\operatorname{tr} P = \operatorname{rank} P$. Conclude that if $\sum_{M \in \mathcal{G}} \operatorname{tr} M = 0$, then $\sum_{M \in \mathcal{G}} M = 0$.

(d) Show that there is an invertible $S \in \mathbb{R}^{n \times n}$ such that $S^{-1}MS$ is orthogonal for every $M \in \mathcal{G}$.

Solution on page 1552

Chapter 7

Eigenvalues and diagonalization

Polynomials, eigenvalues and invariant subspaces, the characteristic polynomial, diagonalization, and Schur's theorem.

7.1 Polynomials

The eigenvalues of a square matrix will turn out to be the roots of a polynomial attached to it, its characteristic polynomial (Section 7.3). So we need to know how many roots a polynomial can have, when it has any at all, and how its roots show up in its factors. All of this rests on multiplying and dividing polynomials. So far (Definition 1.2.13), $F[x]$ has only been a vector space. This section gives it a multiplication and develops the algebra of $F[x]$ that the rest of the book uses. Throughout, F is a field.

7.1.1 Multiplying polynomials

To multiply $a_0 + a_1x + a_2x^2$ by $b_0 + b_1x$, we expand with the distributive law, use $x^ix^j = x^{i+j}$, and collect equal powers of x:

$$\begin{aligned}\left(a_0 + a_1x + a_2x^2\right)(b_0 + b_1x) = a_0b_0 + (a_0b_1 + a_1b_0)x \\ +(a_1b_1 + a_2b_0)x^2 + a_2b_1x^3.\end{aligned}$$

The coefficient of x^k collects the products a_ib_j with $i + j = k$. A polynomial is its sequence of coefficients (Definition 1.2.13), so we take this rule as the definition.

Definition 7.1.1 (Product of polynomials). Let $p = (a_0, a_1, a_2, \ldots)$ and $q = (b_0, b_1, b_2, \ldots)$ be polynomials in $F[x]$. Their **product** is $pq = (c_0, c_1, c_2, \ldots)$, where

$$c_k = \sum_{i=0}^{k} a_ib_{k-i} = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0.$$

We also write $c_k = \sum_{i+j=k} a_ib_j$, a sum over the pairs (i, j) of nonnegative integers with $i + j = k$.

The product is again a polynomial. If $a_i = 0$ for $i > m$ and $b_j = 0$ for $j > n$, then for $k > m + n$ every pair with $i + j = k$ has $i > m$ or $j > n$, so every term of c_k is 0 (Proposition 1.1.15).

Two special cases connect the definition with the notation of Notation 1.2.15. First, $x^jx^k = x^{j+k}$, because the only nonzero term of the product is $1 \cdot 1$, in position $j + k$. In particular x^k is the product of k copies of x, so the notation x^k means what it says. Second, a constant $c = (c, 0, 0, \ldots)$ times q is $(cb_0, cb_1, \ldots)$, which is the scalar multiple cq of Definition 1.2.13. So writing cq is unambiguous.

Example 7.1.2 (Products over three fields).

(a) Over $\mathbb{Q}$, the product $(1+x)\big(1-x+x^2\big)$ has coefficients $c_0 = 1$, $c_1 = 1\cdot(-1)+1\cdot 1 = 0$, $c_2 = 1\cdot 1+1\cdot(-1) = 0$, and $c_3 = 1\cdot 1 = 1$. So $(1+x)\big(1-x+x^2\big) = 1+x^3$.

(b) Over $\mathbb{C}$, $(x-i)(x+i) = x^2+(i-i)x-i^2 = x^2+1$.

(c) Over $\mathbb{F}_2$, $(1+x)^2 = 1+(1+1)x+x^2 = 1+x^2$. So x^2+1 is a square in $\mathbb{F}_2[x]$, while in $\mathbb{R}[x]$ it is not even a product of two polynomials of degree 1 (Corollary 7.1.42).

Proposition 7.1.3 ($F[x]$ is a commutative ring). For all $p, q, r \in F[x]$ and $c \in F$:

(a) $pq = qp$;

(b) $(pq)r = p(qr)$;

(c) $p(q+r) = pq+pr$;

(d) $1p = p$;

(e) $c(pq) = (cp)q = p(cq)$.

In particular, $F[x]$ with its addition and this multiplication is a commutative ring (Definition 1.1.20).

Proof. ROUTINE

Uses. Definition 7.1.1 (Product of polynomials), Proposition 1.2.14 ($F[x]$ is a vector space), Definition 1.1.20 (Commutative ring), Definition 1.1.1 (Field)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Compare the k-th coefficients of the two sides of each identity, the way one compares entries of matrices.
Hint 2. For (b), write the k-th coefficient of each side as a single sum over the triples (i,j,l) of nonnegative integers with $i+j+l = k$.
Sketch. Each identity holds coefficient by coefficient. The k-th coefficient of pq is $\sum_{i+j=k} a_i b_j$, which is symmetric in p and q. Both sides of (b) have k-th coefficient $\sum_{i+j+l=k} a_i b_j d_l$. Parts (c), (d), and (e) follow from distributivity in F, and the additive ring axioms are those of the vector space $F[x]$.

Write $p = (a_i)$, $q = \big(b_j\big)$, and $r = (d_l)$. All sums below are finite, so they can be reordered and regrouped freely.

(a) The k-th coefficient of pq is $\sum_{i+j=k} a_i b_j$, and that of qp is $\sum_{j+i=k} b_j a_i$. These are sums over the same pairs, and $a_i b_j = b_j a_i$.

(b) By distributivity, the k-th coefficient of $(pq)r$ is

$$\sum_{m+l=k}\left(\sum_{i+j=m} a_i b_j\right) d_l = \sum_{i+j+l=k} a_i b_j d_l,$$

since each triple (i,j,l) with $i+j+l=k$ occurs exactly once on the left, with $m = i+j$. In the same way, the k-th coefficient of $p(qr)$ is $\sum_{i+m=k} a_i\big(\sum_{j+l=m} b_j d_l\big) = \sum_{i+j+l=k} a_i b_j d_l$. By associativity in F, the two agree.

(c) The k-th coefficient of $p(q+r)$ is $\sum_{i+j=k} a_i\big(b_j+d_j\big) = \sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i d_j$, which is the k-th coefficient of $pq+pr$.

(d) Since $1 = (1,0,0,\ldots)$, the only possibly nonzero term in the k-th coefficient of $1p$ is $1\cdot a_k = a_k$.

(e) $\sum_{i+j=k}(ca_i)b_j = c\sum_{i+j=k} a_i b_j = \sum_{i+j=k} a_i\big(cb_j\big)$.

Finally, the axioms of a commutative ring that involve only addition (commutativity, associativity, a zero element, and negatives) hold because $F[x]$ is a vector space (Proposition 1.2.14). Together with (a)–(d), this makes $F[x]$ a commutative ring. □

By (c), (e), and $x^i x^j = x^{i+j}$, polynomials are multiplied by expanding and collecting terms:

$$\left(\sum_i a_i x^i\right)\left(\sum_j b_j x^j\right) = \sum_{i,j} a_i b_j x^{i+j}.$$

From now on we compute with polynomials as in school. The identities of Section 1.1.2 that do not use multiplicative inverses, such as $\mathbf{0}p = \mathbf{0}$, $(-1)p = -p$, and the sign rules, have the same proofs in every commutative ring, so they hold in $F[x]$.

Intuition. In a product of two polynomials, the highest power of x comes only from the two leading terms, $(ax^m)(bx^n) = abx^{m+n}$. Every other pair of terms lands lower. Over a field $ab \neq 0$ when $a, b \neq 0$, so nothing can cancel the top term. Over $\mathbb{Z}/4\mathbb{Z}$, where $2 \cdot 2 = 0$, the same definition would give $(2x+1)^2 = 4x^2 + 4x + 1 = 1$, and the degree would drop.

Proposition 7.1.4 (Degree of a product). Let $p, q \in F[x]$ be nonzero, with leading coefficients a and b. Then $pq \neq \mathbf{0}$, $\deg(pq) = \deg p + \deg q$, and the leading coefficient of pq is ab. With the conventions $-\infty + n = n + (-\infty) = -\infty + (-\infty) = -\infty$, the formula $\deg(pq) = \deg p + \deg q$ holds for all $p, q \in F[x]$.

Proof. ROUTINE

Uses. Definition 7.1.1 (Product of polynomials), Definition 1.2.16 (Degree), Proposition 1.1.17 (No zero divisors), Proposition 1.1.15 (Multiplying by zero)
Hint 1. Let $m = \deg p$ and $n = \deg q$. Which pairs (i, j) can give a nonzero term $a_i b_j$?
Hint 2. Only pairs with $i \leq m$ and $j \leq n$, so $i + j \leq m + n$. For $k = m + n$ only the pair (m, n) survives.
Sketch. A term $a_i b_j$ is nonzero only if $i \leq m$ and $j \leq n$. So the coefficients of pq vanish beyond position $m + n$, and the coefficient of x^{m+n} is $a_m b_n = ab$, which is nonzero because F has no zero divisors. If p or q is $\mathbf{0}$, every coefficient of pq is 0.

Write $p = (a_i)$ and $q = (b_j)$, and let $m = \deg p$ and $n = \deg q$, so that $a = a_m$ and $b = b_n$. A term $a_i b_j$ can be nonzero only if $a_i \neq 0$ and $b_j \neq 0$, that is, only if $i \leq m$ and $j \leq n$, and then $i + j \leq m + n$. Hence the k-th coefficient of pq is 0 for $k > m + n$, and for $k = m + n$ the only pair that can contribute is (m, n). So the coefficient of x^{m+n} in pq is ab, which is nonzero by Proposition 1.1.17. Therefore $pq \neq \mathbf{0}$, $\deg(pq) = m + n$, and ab is the leading coefficient of pq.

If $p = \mathbf{0}$ or $q = \mathbf{0}$, every term $a_i b_j$ has a factor 0, so $pq = \mathbf{0}$ (Proposition 1.1.15), and both sides of the formula are $-\infty$. □

Definition 7.1.5 (Monic polynomial). A nonzero polynomial is **monic** if its leading coefficient is 1.

For example, $x^2 - 3x + 2$ and every product $(x - c_1)(x - c_2)\cdots(x - c_n)$ are monic, by Proposition 7.1.4, while $2x + 1$ is not. Every nonzero p is a times a monic polynomial for exactly one $a \in F$, namely its leading coefficient a: indeed $p = a(a^{-1}p)$, and cu has leading coefficient c when u is monic.

The proposition gives an argument that recurs throughout the book.

Technique 7.1.6 (Compare degrees and leading coefficients). To show that a polynomial is zero, or that an identity between polynomials cannot hold, compare the degrees of the two sides, and when the degrees agree, their leading coefficients. Over a field the degree of a product is the sum of the degrees (Proposition 7.1.4), and a sum has degree at most the larger of the two degrees, with equality when they differ (Proposition 1.2.18). The typical use: if $sg = h$ with $s \neq \mathbf{0}$ and $\deg h < \deg s$, then $g = \mathbf{0}$, because a nonzero multiple of s has degree at least $\deg s$.

Reach for it in uniqueness statements about polynomials (quotients and remainders, greatest common divisors, and later minimal polynomials), and to show that a polynomial is constant or zero. Leading coefficients settle what degrees alone cannot, for instance that two monic polynomials that divide each other are equal. The canonical example is the uniqueness half of Theorem 7.1.14. Proposition 1.5.24 was an earlier instance.

Used in. Corollary 7.1.7, Theorem 7.1.14, Theorem 7.1.45, Exercise 7.1.55, Exercise 7.1.65, Theorem 7.6.2, Problem 7.12, Exercise 7.1.69.

Corollary 7.1.7 (No zero divisors, cancellation, and units). Let $p, q, r \in F[x]$.

(a) If $pq = \mathbf{0}$, then $p = \mathbf{0}$ or $q = \mathbf{0}$.

(b) If $p \neq \mathbf{0}$ and $pq = pr$, then $q = r$.

(c) p has a multiplicative inverse in $F[x]$ if and only if p is a nonzero constant.

Proof. ROUTINE

Uses. Proposition 7.1.4 (Degree of a product), Proposition 7.1.3 ($F[x]$ is a commutative ring)
Techniques. Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. (a) is a restatement of part of Proposition 7.1.4. For (b), move everything to one side.
Hint 2. For (c), if $pq = 1$, take degrees of both sides.
Sketch. (a) A product of nonzero polynomials is nonzero. (b) $p(q - r) = \mathbf{0}$ with $p \neq \mathbf{0}$ forces $q - r = \mathbf{0}$. (c) If $pq = 1$, then $\deg p + \deg q = 0$ with both degrees nonnegative, so $\deg p = 0$; nonzero constants c have the inverse c^{-1}.

(a) If p and q are both nonzero, then $pq \neq \mathbf{0}$ by Proposition 7.1.4.

(b) By Proposition 7.1.3, $p(q - r) = pq - pr = \mathbf{0}$. Since $p \neq \mathbf{0}$, (a) gives $q - r = \mathbf{0}$, that is, $q = r$.

(c) A nonzero constant c has the inverse c^{-1}, since $c \cdot c^{-1} = 1$. Conversely, suppose $pq = 1$. Then p and q are nonzero, and $\deg p + \deg q = \deg 1 = 0$ by Proposition 7.1.4. Both degrees are nonnegative integers, so $\deg p = 0$, and p is a nonzero constant. □

So $F[x]$ shares two features with the integers: a product of nonzero elements is nonzero, and few elements are invertible (in $\mathbb{Z}$, only ± 1). In particular x has no inverse, and $F[x]$ is not a field. The analogy with $\mathbb{Z}$ continues in Section 7.1.2 with division with remainder.

Exercise 7.1.8. ROUTINE

(a) Over $\mathbb{Q}$, compute $\left(x^2 - x + 1\right)\left(x^2 + x + 1\right)$ and, for $n \geq 1$, $(1 - x)\left(1 + x + x^2 + \cdots + x^{n-1}\right)$.

(b) Over $\mathbb{F}_2$, compute $\left(1 + x + x^2\right)^2$. Over $\mathbb{F}_3$, compute $(1 + x)^3$.

Solution on page 1553

Evaluation and substitution

Evaluating a polynomial at a scalar (Definition 1.2.17) uses only that scalars can be added, multiplied, and raised to powers. The same formula makes sense with a polynomial in place of the scalar.

Definition 7.1.9 (Substitution). Let $p = \sum_{k=0}^{d} a_k x^k \in F[x]$ and $r \in F[x]$. The polynomial obtained by **substituting** r into p is

$$p(r) = \sum_{k=0}^{d} a_k r^k \in F[x],$$

where $r^0 = 1$ and r^k is the product of k copies of r.

For instance $p(x) = p$, since x^k is the k-th power of x, and if $p = x^2 - 3$, then $p(x + 1) = (x + 1)^2 - 3 = x^2 + 2x - 2$. If $r = c$ is a constant, then $p(c)$ is the constant polynomial whose value is the scalar $p(c)$ of Definition 1.2.17, so the two notations agree.

Proposition 7.1.10 (Evaluation respects sums and products). Let $p, q \in F[x]$ and $c \in F$, and let t be either a scalar in F or a polynomial in $F[x]$. Then

$$(p + q)(t) = p(t) + q(t), \quad (cp)(t) = cp(t),$$

$$(pq)(t) = p(t)q(t).$$

Proof. ROUTINE

Uses. Definition 1.2.17 (Evaluation), Definition 7.1.9 (Substitution), Definition 7.1.1 (Product of polynomials), Proposition 7.1.3 ($F[x]$ is a commutative ring), Proposition 1.2.18 (Degree and evaluation versus the operations)
Hint 1. Write $p = \sum_i a_i x^i$ and $q = \sum_j b_j x^j$, and substitute t into the formula for the coefficients of pq.
Hint 2. Regroup the double sum $\sum_k \sum_{i+j=k}$ as a sum over all pairs (i, j), and use $t^{i+j} = t^i t^j$.
Sketch. The first two identities are the computation of Proposition 1.2.18(c). For the product, $(pq)(t) = \sum_k \bigl(\sum_{i+j=k} a_i b_j\bigr) t^k = \sum_{i,j} (a_i t^i)(b_j t^j) = p(t)q(t)$, by distributivity, $t^{i+j} = t^i t^j$, and commutativity, all of which hold in F and in the commutative ring $F[x]$.

In both cases t lies in a commutative ring, F or $F[x]$ (Proposition 7.1.3), in which scalars multiply as in F (for $F[x]$ this is Proposition 7.1.3(e)) and $t^i t^j = t^{i+j}$. These are the only facts we use.

Choose d with $p = \sum_{i=0}^{d} a_i x^i$ and $q = \sum_{j=0}^{d} b_j x^j$, and put $a_i = b_i = 0$ for $i > d$. Then $p + q = \sum_i (a_i + b_i) x^i$ and $cp = \sum_i (c a_i) x^i$, and distributivity gives $(p + q)(t) = p(t) + q(t)$ and $(cp)(t) = cp(t)$, as in Proposition 1.2.18(c). For the product, $pq = \sum_{k=0}^{2d} \bigl(\sum_{i+j=k} a_i b_j\bigr) x^k$, so

$$\begin{aligned}(pq)(t) &= \sum_{k=0}^{2d} \sum_{i+j=k} a_i b_j t^k = \sum_{i=0}^{d} \sum_{j=0}^{d} (a_i t^i)(b_j t^j) \\ &= \left(\sum_{i=0}^{d} a_i t^i\right)\left(\sum_{j=0}^{d} b_j t^j\right) = p(t)q(t).\end{aligned}$$

The second equality regroups the terms by pairs (i, j) instead of by $k = i + j$ (the pairs with $i > d$ or $j > d$ contribute 0) and uses $t^{i+j} = t^i t^j$ and commutativity. The third is distributivity. □

Remark 7.1.11 (Polynomials in a matrix). For $A \in F^{n \times n}$ and $p = \sum_{k=0}^{d} a_k x^k$, write $p(A) = a_0 I_n + a_1 A + \cdots + a_d A^d$, as in Proposition 4.1.31. The proof above used only that powers of t satisfy $t^i t^j = t^{i+j}$, that products distribute over sums, and that scalars commute with powers of t. It never multiplied t by anything except its own powers and scalars. All of this holds for A (Notation 2.4.16, Proposition 2.4.14), so

$$(p + q)(A) = p(A) + q(A), \quad (cp)(A) = cp(A),$$
$$(pq)(A) = p(A)q(A),$$

even though matrix multiplication is not commutative. In particular $p(A)q(A) = (pq)(A) = (qp)(A) = q(A)p(A)$: any two polynomials in the same matrix commute. Section 7.2 evaluates polynomials at operators in the same way.

Corollary 7.1.12 (Substituting twice). Let $p, r \in F[x]$, and let t be a scalar in F or a polynomial in $F[x]$. Then $(p(r))(t) = p(r(t))$.

Proof. ROUTINE

Uses. Proposition 7.1.10 (Evaluation respects sums and products), Definition 7.1.9 (Substitution), induction on k
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. Both sides respect sums and scalar multiples in p. So it is enough to check the identity for $p = x^k$.
Hint 2. For $p = x^k$ the left side is $\bigl(r^k\bigr)(t)$. Apply the product rule of Proposition 7.1.10 k times.
Sketch. By Proposition 7.1.10, $p \mapsto (p(r))(t)$ and $p \mapsto p(r(t))$ both respect sums and scalar multiples, and every polynomial is a finite combination of the x^k. For $p = x^k$, $\bigl(r^k\bigr)(t) = r(t)^k$ by Proposition 7.1.10 and induction on k, and this is $x^k(r(t))$.

Let $s = r(t)$. By Proposition 7.1.10, applied first with r and then with t, the map $p \mapsto (p(r))(t)$ respects sums and scalar multiples, and so does $p \mapsto p(s)$. Every polynomial is a finite linear combination of $1, x, x^2, \ldots$, so it suffices to prove the identity for $p = x^k$, $k \geq 0$.

For $p = x^k$ we have $x^k(r) = r^k$ and $x^k(s) = s^k$, so we must show $\big(r^k\big)(t) = s^k$. For $k = 0$ both sides are 1. If $\big(r^k\big)(t) = s^k$, then by Proposition 7.1.10, $\big(r^{k+1}\big)(t) = \big(r^k r\big)(t) = \big(r^k\big)(t)\, r(t) = s^k s = s^{k+1}$. □

For example, substituting $x + 1$ into $p(x + 1)$ gives $p((x + 1) + 1) = p(x + 2)$: take $r = t = x + 1$. This is the fact used in Exercise 2.3.48.

7.1.2 Division with remainder

For integers a and $b > 0$ there are integers q and r with $a = bq + r$ and $0 \le r < b$. Polynomials can be divided the same way, with the degree in place of size. Long division, familiar from school, computes the quotient: at each step, subtract a multiple of the divisor that cancels the leading term of what is left.

Example 7.1.13 (Long division). Over $\mathbb{Q}$, divide $p = 2x^4 - x^3 + 3x^2 + x - 1$ by $s = 2x^2 + 1$. The leading term $2x^4$ of p is x^2 times the leading term $2x^2$ of s, so we subtract $x^2 s$. The leading term of what is left is $-x^3 = \left(-\frac{1}{2}x\right) \cdot 2x^2$, so we subtract $-\frac{1}{2}xs$, and so on:

$$\begin{aligned} p - x^2 s &= -x^3 + 2x^2 + x - 1, \\ p - \left(x^2 - \frac{1}{2}x\right)s &= 2x^2 + \frac{3}{2}x - 1, \\ p - \left(x^2 - \frac{1}{2}x + 1\right)s &= \frac{3}{2}x - 2. \end{aligned}$$

The last line has degree $1 < 2 = \deg s$, so we stop. With $q = x^2 - \frac{1}{2}x + 1$ and $r = \frac{3}{2}x - 2$,

$$p = \left(2x^2 + 1\right)\left(x^2 - \frac{1}{2}x + 1\right) + \left(\frac{3}{2}x - 2\right).$$

The coefficient $-\frac{1}{2}$ appeared because we divided by the leading coefficient 2 of s. Over $\mathbb{Z}$ this step would be impossible: dividing by the leading coefficient of the divisor is the one place where long division needs a field. Over $\mathbb{F}_3$, where $2^{-1} = 2$, the same steps give $q = x^2 + x + 1$ and $r = 1$.

Theorem 7.1.14 (Division with remainder). Let $p, s \in F[x]$ with $s \neq \mathbf{0}$. There are unique polynomials $q, r \in F[x]$ such that

$$p = sq + r \quad \text{and} \quad \deg r < \deg s.$$

We call q the **quotient** and r the **remainder** of p on division by s.

The theorem is often called the **division algorithm**, after the procedure that proves it. If $\deg s = 0$, the condition $\deg r < 0$ means $r = \mathbf{0}$: division by a nonzero constant b is exact, with $q = b^{-1}p$.

Intuition. Uniqueness is a degree count. Two divisions $p = sq + r = sq' + r'$ give $s(q - q') = r' - r$. The right side has degree less than $\deg s$, and a nonzero multiple of s cannot, so $q = q'$. For existence, one step of long division replaces p by a polynomial of smaller degree, which suggests strong induction on $\deg p$. A second proof avoids division altogether and counts dimensions instead.

Proof (by induction on the degree). STANDARD

Uses. Proposition 7.1.4 (Degree of a product), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 7.1.3 ($F[x]$ is a commutative ring), Definition 1.2.16 (Degree)
Techniques. Compare degrees and leading coefficients (Technique 7.1.6), Uniqueness via two candidates (Technique 1.1.8), Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Prove uniqueness and existence separately. For uniqueness, subtract two divisions $p = sq + r = sq' + r'$ and compare degrees.
Hint 2. For existence, use strong induction on $\deg p$. If $\deg p < \deg s$ there is nothing to do. Otherwise subtract from p a multiple of s that cancels its leading term.

Hint 3. With a and b the leading coefficients of p and s, the polynomial $p - (a/b)x^{n-m}s$ has degree less than $n = \deg p$. Divide it by the induction hypothesis, and add $(a/b)x^{n-m}$ to the quotient.
Sketch. If $sq + r = sq' + r'$, then $s(q - q') = r' - r$ has degree less than $\deg s$, which forces $q = q'$, and then $r = r'$. For existence, induct on $n = \deg p$. If $n < m = \deg s$, take $q = \mathbf{0}$ and $r = p$. Otherwise the leading terms of p and $(a/b)x^{n-m}s$ agree, so $p_1 = p - (a/b)x^{n-m}s$ has degree less than n. If $p_1 = sq_1 + r_1$, then $p = s(q_1 + (a/b)x^{n-m}) + r_1$.

Let $m = \deg s$, which is a nonnegative integer, and let $b \neq 0$ be the leading coefficient of s.

Uniqueness. Suppose $p = sq + r = sq' + r'$ with $\deg r < m$ and $\deg r' < m$. By Proposition 7.1.3, $s(q - q') = r' - r$, and $\deg(r' - r) < m$ by Proposition 1.2.18(a) and (b). If $q \neq q'$, then by Proposition 7.1.4

$$\deg(s(q - q')) = m + \deg(q - q') \geq m,$$

which is impossible. So $q = q'$, and then $r' - r = s \cdot \mathbf{0} = \mathbf{0}$, so $r = r'$.

Existence. If $p = \mathbf{0}$, take $q = r = \mathbf{0}$. For nonzero p we use strong induction on $n = \deg p$: let $n \geq 0$, and assume that every polynomial of degree less than n (including $\mathbf{0}$) has a quotient and remainder on division by s. Let $\deg p = n$, with leading coefficient a.

If $n < m$, take $q = \mathbf{0}$ and $r = p$. Suppose $n \geq m$. By Proposition 7.1.4, the polynomial $(a/b)x^{n-m}s$ has degree $(n - m) + m = n$ and leading coefficient $(a/b)b = a$. So in

$$p_1 = p - (a/b)x^{n-m}s,$$

the coefficient of x^n is $a - a = 0$, and all coefficients beyond x^n are 0. Hence $\deg p_1 < n$, and by the induction hypothesis $p_1 = sq_1 + r_1$ for some q_1, r_1 with $\deg r_1 < m$. Then

$$p = p_1 + (a/b)x^{n-m}s = s(q_1 + (a/b)x^{n-m}) + r_1,$$

as desired. □

Unwound, the induction is long division. The equation $p = sq + r$ is an invariant, and $\deg r$ is a quantity that drops at every step, so this is an instance of Technique 3.3.5.

Algorithm 7.1.15 (Long division)

Input Polynomials $p, s \in F[x]$ with $s \neq \mathbf{0}$. Let $m = \deg s$ and let b be the leading coefficient of s.
Output The quotient q and the remainder r of p on division by s.

1. Set $q = \mathbf{0}$ and $r = p$.
2. If $\deg r < m$, output q and r and stop.
3. Otherwise, let $d = \deg r$ and let c be the leading coefficient of r. Replace q by $q + (c/b)x^{d-m}$ and r by $r - (c/b)x^{d-m}s$, and return to step 2.

At the start $p = sq + r$, and step 3 adds $(c/b)x^{d-m}s$ to sq and subtracts it from r, so $p = sq + r$ remains true. The same step lowers $\deg r$, as in the proof. So the algorithm stops after at most $\deg p - m + 1$ steps, and then $\deg r < m$.

The second proof is the one in LADR. It proves existence without constructing anything.

Proof of Theorem 7.1.14 (by linear algebra). STANDARD

Uses. Proposition 1.5.24 (Distinct degrees), Example 4.1.7 (Dimensions of the standard spaces), Theorem 4.1.10 (Independent lists are short), Proposition 7.1.4 (Degree of a product), Proposition 7.1.3 ($F[x]$ is a commutative ring)
Techniques. Uniqueness gives existence (Technique 3.4.21), Dimension counting (Technique 4.2.24)
Hint 1. Uniqueness is the degree count from the first proof. For existence, let $n = \deg p \geq m = \deg s$ and work in $F[x]_{\leq n}$. Which polynomials have the form $sq + r$ with $\deg q \leq n - m$ and $\deg r < m$?
Hint 2. They are the linear combinations of $1, x, \ldots, x^{m-1}, s, xs, \ldots, x^{n-m}s$. Show that this list is a basis of $F[x]_{\leq n}$.
Hint 3. The list consists of $n + 1$ nonzero polynomials with the distinct degrees $0, 1, \ldots, n$.

Sketch. Let $n = \deg p \geq m$ (otherwise take $q = \mathbf{0}$, $r = p$). The $n+1$ polynomials $1, x, \ldots, x^{m-1}, s, xs, \ldots, x^{n-m}s$ have the distinct degrees $0, \ldots, n$, so they are independent, and since $\dim F[x]_{\leq n} = n+1$, they form a basis. Expanding p in this basis writes it as $r + sq$, with r a combination of $1, \ldots, x^{m-1}$ and q a combination of $1, \ldots, x^{n-m}$. Uniqueness is as in the first proof.

Uniqueness was proved in the first proof without induction, so we prove existence. Let $m = \deg s$. If $p = \mathbf{0}$ or $\deg p < m$, take $q = \mathbf{0}$ and $r = p$. So let $n = \deg p \geq m$, and consider the list

$$\mathcal{B} = 1, x, \ldots, x^{m-1}, s, xs, \ldots, x^{n-m}s$$

in $F[x]_{\leq n}$. By Proposition 7.1.4, $\deg(x^j s) = j + m$, so the $n+1$ polynomials in $\mathcal{B}$ are nonzero with degrees $0, 1, \ldots, m-1, m, m+1, \ldots, n$, which are distinct. By Proposition 1.5.24, $\mathcal{B}$ is linearly independent. Its length $n+1$ equals $\dim F[x]_{\leq n}$ (Example 4.1.7), so $\mathcal{B}$ is a basis of $F[x]_{\leq n}$ by Theorem 4.1.10(b). Hence there are scalars with

$$\begin{aligned} p = {} & a_0 + a_1 x + \cdots + a_{m-1}x^{m-1} \\ & + b_0 s + b_1 xs + \cdots + b_{n-m}x^{n-m}s. \end{aligned}$$

Let $r = \sum_{i=0}^{m-1} a_i x^i$ and $q = \sum_{j=0}^{n-m} b_j x^j$. By Proposition 7.1.3, $p = sq + r$, and $\deg r \leq m - 1 < m$. □

Intuition (Two proofs compared). The first proof constructs q and r and unwinds into long division, so it is the one to use for computing. The second never divides. Independence of the list $\mathcal{B}$ is the uniqueness statement for $p = \mathbf{0}$, and a dimension count upgrades it to spanning, which is existence. This is Technique 3.4.21, which proved polynomial interpolation in Section 3.4.5. In the language of Section 4.2, the map $F^{n+1} \to F[x]_{\leq n}$ sending a list of coefficients to the corresponding combination of $\mathcal{B}$ is an injective linear map between spaces of the same dimension, hence bijective (Corollary 4.2.18). The count works because the conditions $\deg q \leq n - m$ and $\deg r \leq m - 1$ leave exactly $(n - m + 1) + m = n + 1$ unknown coefficients, one for each coefficient of p.

Definition 7.1.16 (Divisibility). Let $p, s \in F[x]$. We say that s **divides** p, and write $s \mid p$, if $p = sq$ for some $q \in F[x]$. Then s is a **factor** (or **divisor**) of p, and p is a **multiple** of s.

If $s \neq \mathbf{0}$, then $s \mid p$ exactly when the remainder of p on division by s is $\mathbf{0}$, by the uniqueness in Theorem 7.1.14. Some consequences of the definition and of Proposition 7.1.4:

- every s divides $\mathbf{0}$, and $\mathbf{0}$ divides only $\mathbf{0}$;
- a nonzero constant c divides every p, since $p = c(c^{-1}p)$;
- if $s \mid p$ and $p \neq \mathbf{0}$, then $\deg s \leq \deg p$;
- if $s \mid p$ and $p \mid s$ with $p, s \neq \mathbf{0}$, then $p = cs$ for a nonzero constant c. (If $p = sg$ and $s = ph$, then $s = sgh$, so $gh = 1$ by Corollary 7.1.7(b), and g is a nonzero constant by Corollary 7.1.7(c).) In particular, two monic polynomials that divide each other are equal.

For example, $x + 1 \mid x^3 + 1$ over every field, since $x^3 + 1 = (x + 1)(x^2 - x + 1)$.

The quotient and remainder do not depend on the field in which we compute them. This is useful when a polynomial with real coefficients is factored over $\mathbb{C}$.

Corollary 7.1.17 (Division does not depend on the field). Let F be a subfield of a field K (Definition 1.1.32), and let $p, s \in F[x]$ with $s \neq \mathbf{0}$. The quotient and remainder of p on division by s in $K[x]$ are the quotient and remainder in $F[x]$. In particular, s divides p in $K[x]$ if and only if s divides p in $F[x]$.

Proof. ROUTINE

Uses. Definition 1.1.32 (Subfield), Theorem 7.1.14 (Division with remainder), Definition 7.1.1 (Product of polynomials)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Divide in $F[x]$ first. Does the result still satisfy the conditions of Theorem 7.1.14 when it is read in $K[x]$?
Hint 2. Use the uniqueness half of Theorem 7.1.14 over K.

Sketch. $F[x] \subseteq K[x]$, and sums, products, and degrees of polynomials with coefficients in F are the same whether computed in F or in K. So the quotient and remainder over F also satisfy $p = sq + r$ and $\deg r < \deg s$ in $K[x]$, and by uniqueness over K they are the quotient and remainder there.

A polynomial with coefficients in F is also one with coefficients in K, so $F[x] \subseteq K[x]$. The operations of F are those of K restricted to F (Definition 1.1.32), and sums and products of polynomials are computed from their coefficients by the formulas of Definition 1.2.13 and Definition 7.1.1. So for $p, s \in F[x]$, the sum and product are the same whether computed in $F[x]$ or in $K[x]$, and so is the degree.

Let $q, r \in F[x]$ be the quotient and remainder of p on division by s in $F[x]$. Then $p = sq + r$ and $\deg r < \deg s$ also hold in $K[x]$. By the uniqueness in Theorem 7.1.14 over K, q and r are the quotient and remainder in $K[x]$. In particular the remainder in $K[x]$ is $\mathbf{0}$ if and only if the remainder in $F[x]$ is $\mathbf{0}$. □

Exercise 7.1.18. ROUTINE Let $p = x^4 + x^3 + x + 2$ and $s = x^2 + 1$. Find the quotient and remainder of p on division by s over $\mathbb{Q}$, over $\mathbb{F}_2$, and over $\mathbb{F}_3$. Over which of these fields does s divide p?

Solution on page 1553

7.1.3 Roots and linear factors

Definition 7.1.19 (Root). Let $p \in F[x]$. A scalar $c \in F$ is a **root** (or **zero**) of p if $p(c) = 0$.

Which roots exist depends on the field. The polynomial $x^2 + 1$ has no root in $\mathbb{R}$, since $t^2 + 1 \geq 1$ for real t; in $\mathbb{C}$ it has the roots i and $-i$; in $\mathbb{F}_2$ it has the root 1; and in $\mathbb{F}_5$ it has the roots 2 and 3, as $2^2 + 1 = 5$ and $3^2 + 1 = 10$. The polynomial $x^2 - 2$ has no root in $\mathbb{Q}$, because $\sqrt{2}$ is irrational, and two roots in $\mathbb{R}$.

The theorem below is the link between roots and factors. It is the case $s = x - c$ of division with remainder.

Theorem 7.1.20 (Factor theorem). Let $p \in F[x]$ and $c \in F$. There is a unique $q \in F[x]$ with

$$p = (x - c)q + p(c),$$

and if $\deg p \geq 1$, then $\deg q = \deg p - 1$. In particular, c is a root of p if and only if $x - c$ divides p.

Proof (by division with remainder). ROUTINE

Uses. Theorem 7.1.14 (Division with remainder), Proposition 7.1.10 (Evaluation respects sums and products), Proposition 7.1.4 (Degree of a product), Definition 7.1.16 (Divisibility)
Hint 1. Divide p by $x - c$. What can the remainder be?
Hint 2. The remainder is a constant. Evaluate both sides of $p = (x - c)q + r$ at c.
Sketch. Division by $x - c$, which has degree 1, leaves a constant remainder r. Evaluating $p = (x - c)q + r$ at c with Proposition 7.1.10 gives $p(c) = 0 \cdot q(c) + r = r$. The quotient is unique by Theorem 7.1.14, and degrees give $\deg q = \deg p - 1$. If $p(c) = 0$, then $p = (x - c)q$; conversely, $p = (x - c)q'$ gives $p(c) = 0$.

By Theorem 7.1.14 with $s = x - c$, which has degree 1, there are unique q, r with $p = (x - c)q + r$ and $\deg r < 1$, so r is a constant. By Proposition 7.1.10, $p(c) = (c - c)q(c) + r = r$. So $p = (x - c)q + p(c)$. If also $p = (x - c)q' + p(c)$, this is a division with remainder $p(c)$, of degree less than 1, so $q' = q$ by uniqueness.

If $\deg p \geq 1$, then $(x - c)q = p - p(c)$ has the same degree as p, because subtracting a constant does not change any coefficient beyond the constant term. So $q \neq \mathbf{0}$ and $1 + \deg q = \deg p$ (Proposition 7.1.4).

Finally, if $p(c) = 0$, then $p = (x - c)q$, so $x - c \mid p$. Conversely, if $p = (x - c)q'$ for some q', then $p(c) = (c - c)q'(c) = 0$ by Proposition 7.1.10. □

The quotient can also be written down directly. This is how Section 3.4.5 split off a root before products were available, and it gives a fast way to compute q by hand.

Proof (explicitly, by Horner's scheme). STANDARD

Uses. Lemma 3.4.23 (Too many roots), Proposition 7.1.3 ($F[x]$ is a commutative ring), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 7.1.10 (Evaluation respects sums and products)
Hint 1. The identity $x^j - c^j = (x-c)\big(x^{j-1} + cx^{j-2} + \cdots + c^{j-1}\big)$ holds in $F[x]$. Prove it by expanding the right side.
Hint 2. Multiply it by a_j, sum over j, and collect powers of x. The coefficients of q are the numbers b_i from the proof of Lemma 3.4.23.
Sketch. The right side of $x^j - c^j = (x-c)\sum_{i<j} x^i c^{j-1-i}$ telescopes. Multiplying by a_j and summing gives $p - p(c) = (x-c)q$ with $q = \sum_{i<n} b_i x^i$ and $b_i = \sum_{j>i} a_j c^{j-1-i}$, computed from the top by $b_{n-1} = a_n$ and $b_i = a_{i+1} + cb_{i+1}$. Uniqueness of q is cancellation of $x - c$, and $b_{n-1} = a_n \neq 0$ gives the degree.

If p is constant, then $p - p(c) = \mathbf{0} = (x-c)\mathbf{0}$. Otherwise write $p = \sum_{j=0}^{n} a_j x^j$ with $n \geq 1$ and $a_n \neq 0$. For $j \geq 1$, the following sum telescopes in $F[x]$:

$$(x-c)\sum_{i=0}^{j-1} x^i c^{j-1-i} = \sum_{i=0}^{j-1} x^{i+1} c^{j-1-i} - \sum_{i=0}^{j-1} x^i c^{j-i} = x^j - c^j.$$

This is the computation in the proof of Lemma 3.4.23, with the indeterminate x in place of the scalar t. Multiplying by a_j, summing over $j = 1, \ldots, n$, and collecting the terms with the same power of x,

$$p - p(c) = \sum_{j=1}^{n} a_j\big(x^j - c^j\big) = (x-c)q,$$

where

$$q = \sum_{i=0}^{n-1} b_i x^i, \quad b_i = \sum_{j=i+1}^{n} a_j c^{j-1-i}.$$

The leading coefficient of q is $b_{n-1} = a_n \neq 0$, so $\deg q = n - 1$. In both cases, if also $p - p(c) = (x-c)q'$, then $q' = q$ by Corollary 7.1.7(b). The last sentence of the theorem follows as in the first proof. □

As computed in the proof of Lemma 3.4.23, $b_{n-1} = a_n$ and $b_i = a_{i+1} + cb_{i+1}$. Continuing one more step, $a_0 + cb_0 = p(c)$, by expanding the nested form $p(c) = a_0 + c(a_1 + c(a_2 + \cdots + ca_n))$. This recursion is **Horner's scheme**, or **synthetic division**. It computes the quotient by $x - c$ and the value $p(c)$ at once, with n multiplications.

Example 7.1.21 (Synthetic division). Let $p = x^3 - 6x^2 + 11x - 6$ and $c = 1$. Horner's scheme runs through the coefficients $1, -6, 11, -6$ from the top:

$$b_2 = 1, \quad b_1 = -6 + 1 \cdot 1 = -5,$$
$$b_0 = 11 + 1 \cdot (-5) = 6, \quad p(1) = -6 + 1 \cdot 6 = 0.$$

So 1 is a root and $p = (x-1)\big(x^2 - 5x + 6\big)$. Running the scheme again on $x^2 - 5x + 6$ with $c = 2$ gives 1, $-5 + 2 = -3$, and remainder $6 + 2 \cdot (-3) = 0$, so $x^2 - 5x + 6 = (x-2)(x-3)$, and

$$p = (x-1)(x-2)(x-3).$$

In the language of Example 2.6.29, the factor theorem says $\ker \mathrm{ev}_c = \mathrm{im}(M - c\,\mathrm{id})$: the polynomials with root c are exactly the multiples $(x-c)q = (M - c\,\mathrm{id})(q)$.

Intuition. Each root splits off a linear factor. Distinct roots split off distinct factors: after $p = (x-c_1)q$, a second root $c_2 \neq c_1$ of p satisfies $0 = (c_2 - c_1)q(c_2)$, and since $c_2 - c_1 \neq 0$ and F has no zero divisors, c_2 is a root of q. Each factor costs one degree, so there is room for at most $\deg p$ of them.

Proposition 7.1.22 (Counting roots). Let $p \in F[x]$ be nonzero, and let $c_1, \ldots, c_k \in F$ be distinct roots of p. Then

$$(x-c_1)(x-c_2)\cdots(x-c_k) \mid p,$$

and $k \le \deg p$. In particular, a nonzero polynomial p has at most $\deg p$ roots in F.

Proof. STANDARD

Uses. Theorem 7.1.20 (Factor theorem), Proposition 7.1.10 (Evaluation respects sums and products), Proposition 1.1.17 (No zero divisors), Proposition 7.1.4 (Degree of a product)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Induct on k, the length of the list of roots. For the inductive step, assume $p = (x - c_1)\cdots(x - c_{k-1})q$.
Hint 2. Evaluate this factorization at c_k. Why must $q(c_k) = 0$?
Hint 3. For the bound, compare degrees in $p = (x - c_1)\cdots(x - c_k)q'$, noting that $q' \neq \mathbf{0}$.
Sketch. By induction on k, $p = (x - c_1)\cdots(x - c_{k-1})q$. Evaluating at c_k gives $0 = (c_k - c_1)\cdots(c_k - c_{k-1})q(c_k)$, and the differences are nonzero, so $q(c_k) = 0$ and $x - c_k$ divides q by the factor theorem. Then $p = (x - c_1)\cdots(x - c_k)q'$ with $q' \neq \mathbf{0}$, and $\deg p = k + \deg q' \ge k$.

We prove the divisibility by induction on k. For $k = 0$ the product is empty, that is, equal to 1, which divides p. Let $k \ge 1$, and suppose $p = (x - c_1)\cdots(x - c_{k-1})q$ for some $q \in F[x]$. By Proposition 7.1.10,

$$0 = p(c_k) = (c_k - c_1)\cdots(c_k - c_{k-1})\, q(c_k).$$

The roots are distinct, so $c_k - c_i \neq 0$ for each $i < k$, and the product of these differences is nonzero by Proposition 1.1.17. Hence $q(c_k) = 0$, and $q = (x - c_k)q'$ for some q' by Theorem 7.1.20. So $p = (x - c_1)\cdots(x - c_k)q'$, which proves the divisibility.

Since $p \neq \mathbf{0}$, also $q' \neq \mathbf{0}$. Each $x - c_i$ has degree 1, so Proposition 7.1.4 gives $\deg p = k + \deg q' \ge k$. □

The bound $k \le \deg p$ is Lemma 3.4.23 in another form: a polynomial in $F[x]_{\le n}$ with $n + 1$ distinct roots is zero. What is new is the divisibility, which says where the roots sit in p. The proof used that F has no zero divisors, to conclude $q(c_k) = 0$. With coefficients in $\mathbb{Z}/8\mathbb{Z}$, which has zero divisors, $x^2 - 1$ has four roots (Exercise 7.1.52).

Corollary 7.1.23 (Polynomials and their values).

(a) Let $p, q \in F[x]_{\le n}$. If $p(c) = q(c)$ for $n + 1$ distinct $c \in F$, then $p = q$.

(b) A nonzero polynomial has only finitely many roots. If F is infinite, two polynomials with the same polynomial function $F \to F$ are equal.

Proof. ROUTINE

Uses. Proposition 7.1.22 (Counting roots), Proposition 1.2.18 (Degree and evaluation versus the operations)
Techniques. Contrapositive (Technique 1.5.4)
Hint. Apply Proposition 7.1.22 to $p - q$.
Sketch. In (a), $p - q \in F[x]_{\le n}$ has $n + 1$ roots, more than its degree if it were nonzero. In (b), if p and q have the same function, $p - q$ has every element of F as a root, infinitely many.

(a) By Proposition 1.2.18, $p - q \in F[x]_{\le n}$ and $(p - q)(c) = p(c) - q(c) = 0$ for the $n + 1$ given points. If $p - q$ were nonzero, it would have at least $n + 1 > \deg(p - q)$ roots, contradicting Proposition 7.1.22. So $p = q$.

(b) The first claim is Proposition 7.1.22. If $p(c) = q(c)$ for every $c \in F$ and F is infinite, then $p - q$ has infinitely many roots, so $p - q = \mathbf{0}$. □

Part (b) is the fact promised in Aside (Polynomials versus polynomial functions): over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, a polynomial is determined by its polynomial function. Over a finite field it is not. Section 7.1.4 finds all polynomials with a given function.

Multiplicity

A root can split off more than one linear factor, as in $x^2 - 2x + 1 = (x - 1)^2$. Counting roots with these repetitions gives a sharper version of Proposition 7.1.22, and over $\mathbb{C}$ an exact count.

Definition 7.1.24 (Multiplicity of a root). Let $p \in F[x]$ be nonzero and $c \in F$. The **multiplicity** of c as a root of p is the largest integer $m \geq 0$ such that $(x-c)^m \mid p$. A root of multiplicity 1 is **simple**, and a root of multiplicity at least 2 is **repeated** (or **multiple**).

The largest such m exists, because $(x-c)^0 = 1$ divides p, and $(x-c)^m \mid p$ forces $m \leq \deg p$ (degrees, as $p \neq \mathbf{0}$). By the factor theorem, c is a root of p exactly when its multiplicity is at least 1. We do not define multiplicities for the zero polynomial, which every power of $x-c$ divides.

Lemma 7.1.25 (Reading off a multiplicity). Let $p \in F[x]$ be nonzero, $c \in F$, and $m \geq 0$. Then c has multiplicity exactly m as a root of p if and only if $p = (x-c)^m q$ for some $q \in F[x]$ with $q(c) \neq 0$.

Proof. ROUTINE

Uses. Definition 7.1.24 (Multiplicity of a root), Corollary 7.1.7 (No zero divisors, cancellation, and units), Theorem 7.1.20 (Factor theorem)
Hint 1. For "if", suppose also $(x-c)^{m+1} \mid p$, and cancel $(x-c)^m$.
Hint 2. For "only if", if $q(c) = 0$, the factor theorem splits one more factor $x-c$ off q.
Sketch. If $p = (x-c)^m q$ with $q(c) \neq 0$ and also $p = (x-c)^{m+1} q'$, cancelling $(x-c)^m$ gives $q = (x-c)q'$, so $q(c) = 0$, a contradiction; so m is the largest exponent. Conversely, if m is the multiplicity and $p = (x-c)^m q$ with $q(c) = 0$, then $q = (x-c)q''$ and $(x-c)^{m+1}$ divides p.

Suppose $p = (x-c)^m q$ with $q(c) \neq 0$. Then $(x-c)^m \mid p$. If $(x-c)^{m'} \mid p$ for some $m' > m$, then so does $(x-c)^{m+1}$, say $p = (x-c)^{m+1} q'$. Then $(x-c)^m q = (x-c)^m (x-c) q'$, and cancelling the nonzero factor $(x-c)^m$ (Corollary 7.1.7(b)) gives $q = (x-c)q'$, so $q(c) = 0$, a contradiction. Hence m is the multiplicity.

Conversely, let m be the multiplicity, and write $p = (x-c)^m q$. If $q(c) = 0$, then $q = (x-c)q''$ by Theorem 7.1.20, and $(x-c)^{m+1} \mid p$, contradicting the maximality of m. So $q(c) \neq 0$. □

Example 7.1.26 (Multiplicity depends on the field). Let $p = x^3 - x^2 - x + 1$. Over $\mathbb{Q}$, $p = (x-1)^2(x+1)$, and $q = x+1$ has $q(1) = 2 \neq 0$. So 1 is a root of multiplicity 2, and -1 is a simple root. Over $\mathbb{F}_2$ the same polynomial is $x^3 + x^2 + x + 1$, and since $x + 1 = x - 1$ there, $p = (x-1)^3$: the root 1 has multiplicity 3. In the factorization $(x-1)^2(x+1)$, the cofactor now has $q(1) = 1 + 1 = 0$, so Lemma 7.1.25 does not stop at $m = 2$.

Proposition 7.1.27 (Multiplicities add up to at most the degree). Let $p \in F[x]$ be nonzero, and let $c_1, \ldots, c_k \in F$ be distinct, with multiplicities $m_1, \ldots, m_k$ as roots of p. Then

$$(x-c_1)^{m_1}(x-c_2)^{m_2}\cdots(x-c_k)^{m_k} \mid p,$$

and in particular $m_1 + m_2 + \cdots + m_k \leq \deg p$.

Proof. STANDARD

Uses. Lemma 7.1.25 (Reading off a multiplicity), Proposition 7.1.10 (Evaluation respects sums and products), Proposition 1.1.17 (No zero divisors), Proposition 7.1.4 (Degree of a product)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Induct on k, as in the proof of Proposition 7.1.22, now splitting off $(x-c_k)^{m_k}$ at once.
Hint 2. Suppose $p = (x-c_1)^{m_1}\cdots(x-c_{k-1})^{m_{k-1}} q$. Let μ be the multiplicity of c_k as a root of q, and write $q = (x-c_k)^\mu h$ with $h(c_k) \neq 0$.
Hint 3. Then $p = (x-c_k)^\mu g$ for a g with $g(c_k) \neq 0$. Apply Lemma 7.1.25 to p.
Sketch. Induct on k. If $p = \prod_{i<k}(x-c_i)^{m_i} q$ and $q = (x-c_k)^\mu h$ with $h(c_k) \neq 0$, then $p = (x-c_k)^\mu g$ with $g = \prod_{i<k}(x-c_i)^{m_i} h$ and $g(c_k) = \prod_{i<k}(c_k - c_i)^{m_i} h(c_k) \neq 0$. By Lemma 7.1.25 $\mu = m_k$, so $(x-c_k)^{m_k}$ divides q. Degrees give $\sum m_i \leq \deg p$.

We induct on k; for $k = 0$ there is nothing to prove. Let $k \geq 1$ and suppose $p = (x-c_1)^{m_1}\cdots(x-c_{k-1})^{m_{k-1}} q$. Since $p \neq \mathbf{0}, q \neq \mathbf{0}$. Let μ be the multiplicity of c_k as a root of q. By Lemma 7.1.25, $q = (x-c_k)^\mu h$ with $h(c_k) \neq 0$. Reordering the factors,

$$p = (x-c_k)^\mu g, \quad \text{where } g = (x-c_1)^{m_1}\cdots(x-c_{k-1})^{m_{k-1}} h.$$

By Proposition 7.1.10, $g(c_k) = (c_k - c_1)^{m_1} \cdots (c_k - c_{k-1})^{m_{k-1}} h(c_k)$, which is nonzero by Proposition 1.1.17, as every factor is nonzero. By Lemma 7.1.25 applied to p, the multiplicity of c_k as a root of p is μ, so $\mu = m_k$. Hence $p = (x - c_1)^{m_1} \cdots (x - c_k)^{m_k} h$, which proves the divisibility. Comparing degrees with Proposition 7.1.4, $\deg p = m_1 + \cdots + m_k + \deg h \geq m_1 + \cdots + m_k$. □

With all $m_i \geq 1$, this sharpens Proposition 7.1.22: counted with multiplicity, a nonzero polynomial has at most $\deg p$ roots. Over $\mathbb{C}$ the count is exactly $\deg p$ (Corollary 7.1.35).

Repeated roots can be detected without factoring, with the formal derivative p' of Example 2.1.9.

> **Exercise 7.1.28.** STANDARD Let $p \in F[x]$ be nonzero and $c \in F$. Show that c is a repeated root of p if and only if $p(c) = 0$ and $p'(c) = 0$.

Solution on page 1554

7.1.4 Interpolation, revisited

In Section 3.4.5, the existence of interpolating polynomials came from their uniqueness, through a square linear system, and the Lagrange polynomials L_i were *defined* as solutions of that system. Example 3.4.28 previewed a closed form that needed products. We can now prove it.

> **Proposition 7.1.29** (Lagrange polynomials in closed form). Let $c_0, \ldots, c_n \in F$ be distinct. For $i = 0, \ldots, n$,
>
> $$L_i = \prod_{j \neq i} \frac{x - c_j}{c_i - c_j}.$$
>
> Consequently, for $d_0, \ldots, d_n \in F$, the unique $p \in F[x]_{\leq n}$ with $p(c_i) = d_i$ for all i is
>
> $$p = \sum_{i=0}^{n} d_i \prod_{j \neq i} \frac{x - c_j}{c_i - c_j}.$$

Proof. ROUTINE

Uses. Definition 3.4.26 (Lagrange polynomials), Theorem 3.4.24 (Polynomial interpolation), Proposition 3.4.27 (Lagrange's formula), Proposition 7.1.10 (Evaluation respects sums and products), Proposition 7.1.4 (Degree of a product)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Check that the product has degree at most n and takes the values that define L_i.
Hint 2. Evaluate the product at c_k with Proposition 7.1.10. Which factor vanishes when $k \neq i$?
Sketch. The product P_i is a nonzero scalar times a product of n linear factors, so it has degree n. At c_i every factor is 1; at c_k with $k \neq i$, the factor with $j = k$ is 0. So P_i has the defining values of L_i, and $P_i = L_i$ by uniqueness in Theorem 3.4.24. The formula for p is then Proposition 3.4.27.

Let P_i be the product on the right. The scalar $\prod_{j \neq i}(c_i - c_j)$ is nonzero, since the nodes are distinct and F has no zero divisors, so

$$P_i = \left(\prod_{j \neq i} (c_i - c_j) \right)^{-1} \prod_{j \neq i} (x - c_j)$$

is a nonzero scalar times a product of n polynomials of degree 1, and $\deg P_i = n$ by Proposition 7.1.4. By Proposition 7.1.10,

$$P_i(c_k) = \prod_{j \neq i} \frac{c_k - c_j}{c_i - c_j}.$$

For $k = i$ every factor is 1, so $P_i(c_i) = 1$. For $k \neq i$, the factor with $j = k$ is 0, so $P_i(c_k) = 0$. Thus $P_i \in F[x]_{\leq n}$ has the values that define L_i (Definition 3.4.26), and $P_i = L_i$ by the uniqueness in Theorem 3.4.24. The formula for p follows from Proposition 3.4.27. □

Intuition (Two routes to interpolation). With the closed form in hand, Theorem 3.4.24 has a second proof that reverses the roles of existence and uniqueness. Existence is now a construction: the polynomial $\sum_i d_i P_i$ has degree at most n, and the computation above shows that it takes the value d_k at c_k. Uniqueness is Corollary 7.1.23(a), a count of roots. In Section 3.4.5, by contrast, uniqueness was the root count and existence came for free from the square system. The first route needs no formula, which helps when one is hard to guess, as in Exercise 7.1.57. The second gives the answer explicitly and explains the shape of L_i: it is the simplest polynomial vanishing at the other nodes, rescaled to be 1 at c_i.

The node polynomial $\omega = (x - c_0)(x - c_1)\cdots(x - c_n)$, of degree $n + 1$, describes what happens without the degree bound.

Corollary 7.1.30 (All interpolating polynomials). Let $c_0, \ldots, c_n \in F$ be distinct, let $d_0, \ldots, d_n \in F$, let $p_0 \in F[x]_{\leq n}$ be the polynomial with $p_0(c_i) = d_i$ for all i, and let $\omega = (x - c_0)\cdots(x - c_n)$. A polynomial $p \in F[x]$ satisfies $p(c_i) = d_i$ for all i if and only if

$$p = p_0 + \omega h \quad \text{for some } h \in F[x].$$

In particular, p_0 is the remainder of every such p on division by ω.

Proof. ROUTINE

Uses. Proposition 7.1.22 (Counting roots), Proposition 7.1.10 (Evaluation respects sums and products), Theorem 7.1.14 (Division with remainder), Theorem 3.4.24 (Polynomial interpolation)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. The conditions are linear in p, and p_0 is one solution. Which p satisfy the homogeneous conditions $p(c_i) = 0$?
Hint 2. Use the divisibility in Proposition 7.1.22.
Sketch. If $p = p_0 + \omega h$, then $p(c_i) = d_i + 0$. Conversely, $p - p_0$ has the $n + 1$ distinct roots c_i, so ω divides it by Proposition 7.1.22 (or $p - p_0 = \mathbf{0}$). Since $\deg p_0 < \deg \omega$, the equation $p = \omega h + p_0$ is a division with remainder.

If $p = p_0 + \omega h$, then $p(c_i) = p_0(c_i) + \omega(c_i)h(c_i) = d_i + 0 = d_i$ by Proposition 7.1.10, since c_i is a root of ω. Conversely, suppose $p(c_i) = d_i$ for all i. Then $p - p_0$ has the $n + 1$ distinct roots $c_0, \ldots, c_n$. If $p - p_0 = \mathbf{0}$, take $h = \mathbf{0}$. Otherwise $\omega \mid p - p_0$ by Proposition 7.1.22, so $p - p_0 = \omega h$ for some h.

Finally, $p = \omega h + p_0$ with $\deg p_0 \leq n < n + 1 = \deg \omega$, so by the uniqueness in Theorem 7.1.14, p_0 is the remainder of p on division by ω. □

This is Technique 3.1.18 once more: p_0 is a particular solution, and the polynomials vanishing at all nodes are the multiples of ω. The quartic q of Figure 3.4.1 is an example. It passes through the same four points as the cubic p, and indeed $q - p = \frac{1}{2}\big(x^4 - 2x^3 - x^2 + 2x\big) = \frac{1}{2}(x + 1)x(x - 1)(x - 2)$ is a multiple of ω.

The figure can be explored with the least-squares component of Section 6.5. With four points and degree 3, the least-squares cubic passes through all four points, so it is the interpolating polynomial.

$$p = 2L_0 + L_1 + 0L_2 + 5L_3 = 1 - 2x + x^3,$$
$$L_3 = \frac{(x + 1)x(x - 1)}{6}$$

With four points at distinct x-values and degree 3, the residuals are 0 and the curve is the interpolating cubic $1 - 2x + x^3$ of Example 3.4.25. Drag one point up or down while the others stay fixed: the cubic changes by a multiple of the Lagrange polynomial of that node, which vanishes at the other three nodes. Add a fifth point and the cubic no longer passes through all of them, unless the new point happens to lie on it. (Interactive in the web version.)

Remark 7.1.31 (Finite fields). Let F be a finite field with q elements, and take all q elements of F as the nodes. Corollary 7.1.30 with all values $d_i = 0$ shows that the polynomials whose polynomial function is zero are exactly the multiples of

$$\omega_F = \prod_{c \in F} (x - c),$$

a monic polynomial of degree q. Two polynomials have the same polynomial function if and only if their difference is a multiple of ω_F, and each function $F \to F$ is the polynomial function of exactly one polynomial of degree less than q (Theorem 3.4.24 with all q elements as nodes; compare Problem 1.18). For $F = \mathbb{F}_p$, $\omega_F = x^p - x$ (Exercise 7.1.55), which explains the example $x^p - x$ of Aside (Polynomials versus polynomial functions).

7.1.5 The Fundamental Theorem of Algebra

Over $\mathbb{R}$, the polynomial $x^2 + 1$ has no root. Correspondingly, the rotation of the plane by a right angle has no eigenvalue (Section 7.2): no line through $\mathbf{0}$ is mapped to itself. The complex numbers were built to give $x^2 + 1$ a root (Section 1.1.4). The surprise is that nothing more needs to be added: every nonconstant polynomial with complex coefficients, even one with nonreal coefficients such as $x^2 - i$ or $x^5 + ix + 1$, already has a root in $\mathbb{C}$.

Theorem 7.1.32 (Fundamental Theorem of Algebra). Every nonconstant polynomial $p \in \mathbb{C}[x]$ has a root in $\mathbb{C}$.

Intuition (Why there must be a root). Let p have degree $n \geq 1$, and watch the values $p(z)$ as z goes once around the circle $|z| = R$. For large R, the leading term dominates, $p(z) \approx a_n z^n$, and since z^n goes around 0 exactly n times when z goes around once, so does the curve traced by $p(z)$. For tiny R, $p(z) \approx p(0)$, and if $p(0) \neq 0$ (otherwise 0 is a root), the curve is a small loop near $p(0)$ that does not go around 0 at all. As R grows from small to large, the curve moves continuously, and the number of times it winds around 0 changes from 0 to n. It can only change when the curve passes through 0, and a point z with $p(z) = 0$ is a root. Figure 7.1.1 shows this for $p = z^3 + z + 1$.

Making "winds around" precise takes topology. The proof below avoids it. It finds a root as the point where $|p|$ is smallest: at a point where $p \neq 0$, a small step in a well-chosen direction always makes $|p|$ smaller.

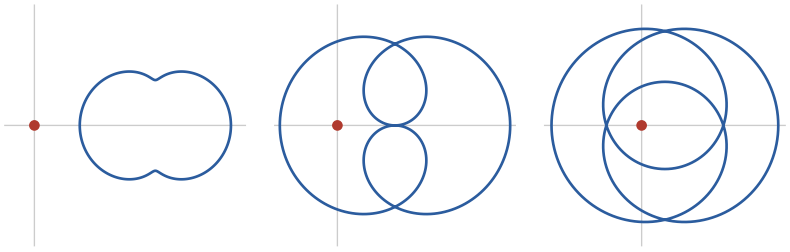


Figure 7.1.1. The images of the circles $|z| = R$ under $p(z) = z^3 + z + 1$, with 0 marked in red. For $R = 1/2$ the curve stays near $p(0) = 1$; for $R = 3/2$ it winds three times around 0. In between it sweeps across 0 three times: at $R \approx 0.68$, the modulus of the real root $-0.68\ldots$, and at $R \approx 1.21$, the common modulus of the two nonreal roots $0.34\ldots \pm 1.16\ldots i$.

Proof of Theorem 7.1.32 (by minimizing $|p|$). CHALLENGE BEYOND 21-242

Uses. Definition B.1.7 (Modulus), Section B.1, the extreme value theorem: a continuous real function on a closed disk attains a minimum, every complex number has a k-th root, for every $k \geq 1$

Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Show first that $|p(z)| > |p(0)|$ outside a large disk. Then $|p|$ has a global minimum at some z_0, by the extreme value theorem.
Hint 2. Suppose $p(z_0) \neq 0$. Write $p(z_0 + w)/p(z_0)$ as $1 + b_k w^k$ plus higher powers of w, with $b_k \neq 0$.
Hint 3. Choose β with $\beta^k = -1/b_k$. For small real $t > 0$, the value at $w = t\beta$ has modulus about $1 - t^k < 1$.
Sketch. For $|z|$ large, $|p(z)| \geq |a_n|\,|z|^n - \sum_{k<n}|a_k|\,|z|^k > |p(0)|$, so the minimum of $|p|$ on a large closed disk is a global minimum, attained at some z_0. If $p(z_0) \neq 0$, then $g(w) = p(z_0 + w)/p(z_0) = 1 + b_k w^k + \cdots$ with $b_k \neq 0$ has $|g| \geq 1$. With $\beta^k = -1/b_k$ and small $t > 0$, $|g(t\beta)| \leq 1 - t^k + Ct^{k+1} < 1$, a contradiction.

Let $p = a_0 + a_1 x + \cdots + a_n x^n$ with $n \geq 1$ and $a_n \neq 0$. We use the modulus $|z|$ (Definition B.1.7), with $|zw| = |z|\,|w|$ and $|z + w| \leq |z| + |w|$, and two facts from outside algebra: the extreme value theorem, and the existence of k-th roots in $\mathbb{C}$ (polar form, Section B.1).

$|p|$ attains a global minimum. Let $A = |a_0| + \cdots + |a_{n-1}|$. For $|z| = \rho \geq 1$, we have $|a_k z^k| \leq |a_k|\,\rho^{n-1}$ for $k < n$, so

$$|p(z)| \geq |a_n|\,\rho^n - A\rho^{n-1} = \rho^{n-1}(|a_n|\,\rho - A).$$

Let $R = \max(1, (A + |a_0| + 1)/|a_n|)$. For $|z| > R$ the right side is at least $|a_0| + 1 > |p(0)|$. The function $z \mapsto |p(z)|$ is continuous, so it attains a minimum on the closed disk $|z| \leq R$, at some z_0. Then $|p(z_0)| \leq |p(0)| < |p(z)|$ for $|z| > R$, so $|p(z_0)| \leq |p(z)|$ for every $z \in \mathbb{C}$.

The minimum is 0. Suppose $p(z_0) \neq 0$. Expanding each $(z_0 + w)^j$ shows that $g(w) = p(z_0 + w)/p(z_0)$ is a polynomial in w of degree n (its coefficient of w^n is $a_n/p(z_0)$), with $g(0) = 1$ and $|g(w)| \geq 1$ for all w. Write

$$g(w) = 1 + b_k w^k + b_{k+1} w^{k+1} + \cdots + b_n w^n,$$

where $1 \leq k \leq n$ and $b_k \neq 0$. Choose $\beta \in \mathbb{C}$ with $\beta^k = -1/b_k$, and let $C = \sum_{j=k+1}^{n} |b_j|\,|\beta|^j$. For real t with $0 < t \leq 1$, we have $b_k(t\beta)^k = -t^k$ and $t^j \leq t^{k+1}$ for $j > k$, so

$$\begin{aligned} |g(t\beta)| &\leq |1 - t^k| + \sum_{j=k+1}^{n} |b_j|\,t^j\,|\beta|^j \\ &\leq 1 - t^k + Ct^{k+1} = 1 - t^k(1 - Ct). \end{aligned}$$

For $0 < t < 1$ with $Ct < 1$, this is less than 1, contradicting $|g| \geq 1$. Hence $p(z_0) = 0$.

The details of both steps, and a second proof by linear algebra due to Derksen, are in Section C.1. □

Some analysis is unavoidable in any proof. The theorem fails for the field $\mathbb{Q}(i) = \{a + bi \mid a, b \in \mathbb{Q}\}$, which satisfies all the field axioms and contains a square root of -1, but in which $x^2 - 2$ has no root. So no argument from the field axioms alone can prove it; every proof uses a property of $\mathbb{R}$ such as completeness.

Splitting into linear factors

Definition 7.1.33 (Splitting). A nonzero polynomial $p \in F[x]$ **splits** over F if

$$p = a(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$$

for some $a, \lambda_1, \ldots, \lambda_n \in F$.

Then $n = \deg p$ and a is the leading coefficient, by Proposition 7.1.4. The polynomial $x^2 + 1$ does not split over $\mathbb{R}$, since a factor $x - \lambda$ would give a real root; it splits over $\mathbb{C}$ as $(x - i)(x + i)$. The polynomial $x^2 - 2$ splits over $\mathbb{R}$ but not over $\mathbb{Q}$. When a polynomial splits, its factors record its roots and their multiplicities.

Proposition 7.1.34 (Roots of a split polynomial). Let $p = a(x - \lambda_1)\cdots(x - \lambda_n)$ with $a \in F$ nonzero and $\lambda_1, \ldots, \lambda_n \in F$. The roots of p are exactly $\lambda_1, \ldots, \lambda_n$, and each root of p occurs in the list $\lambda_1, \ldots, \lambda_n$ exactly as many times as its multiplicity. In particular, the list is determined by p up to order, and the multiplicities of the distinct roots of p add up to $\deg p$.

Proof. ROUTINE

Uses. Lemma 7.1.25 (Reading off a multiplicity), Proposition 7.1.10 (Evaluation respects sums and products), Proposition 1.1.17 (No zero divisors), Proposition 7.1.4 (Degree of a product)
Hint 1. Fix $\lambda \in F$ and let m be the number of indices i with $\lambda_i = \lambda$. Collect these factors in front.
Hint 2. What is the value at λ of the product of the remaining factors? Apply Lemma 7.1.25.
Sketch. For $\lambda \in F$ occurring m times in the list, $p = (x-\lambda)^m q$ with $q = a\prod_{\lambda_i \neq \lambda}(x-\lambda_i)$, and $q(\lambda) = a\prod_{\lambda_i \neq \lambda}(\lambda - \lambda_i) \neq 0$. By Lemma 7.1.25, λ has multiplicity m. So the number of occurrences of each scalar is its multiplicity, which depends only on p; summing over the distinct values in the list gives $n = \deg p$.

Fix $\lambda \in F$, and let m be the number of indices i with $\lambda_i = \lambda$ (possibly $m = 0$). Reordering the factors,

$$p = (x-\lambda)^m q, \quad \text{where } q = a\prod_{i:\lambda_i \neq \lambda}(x-\lambda_i).$$

By Proposition 7.1.10, $q(\lambda) = a\prod_{i:\lambda_i \neq \lambda}(\lambda - \lambda_i)$, which is nonzero by Proposition 1.1.17, as each factor is. By Lemma 7.1.25, the multiplicity of λ as a root of p is m. So λ is a root exactly when it occurs in the list, and it occurs as many times as its multiplicity. The number of occurrences of each scalar is therefore determined by p, which means the list is determined up to order. Summing the numbers of occurrences over the distinct values in the list counts every index once, which gives $n = \deg p$ (Proposition 7.1.4). □

Corollary 7.1.35 (Factorization over $\mathbb{C}$). Let $p \in \mathbb{C}[x]$ be nonzero, of degree n and with leading coefficient a. Then

$$p = a(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n),$$

where $\lambda_1, \ldots, \lambda_n$ are the roots of p, each listed as many times as its multiplicity. The factorization is unique up to the order of the factors. In particular, counted with multiplicity, p has exactly $\deg p$ roots in $\mathbb{C}$.

Proof. ROUTINE

Uses. Theorem 7.1.32 (Fundamental Theorem of Algebra), Theorem 7.1.20 (Factor theorem), Proposition 7.1.4 (Degree of a product), Proposition 7.1.34 (Roots of a split polynomial), induction on n
Hint 1. Induct on n. Use Theorem 7.1.32 to find one root, and Theorem 7.1.20 to split it off.
Hint 2. Uniqueness and the statement about multiplicities are Proposition 7.1.34.
Sketch. For $n = 0$, $p = a$. For $n \geq 1$, Theorem 7.1.32 gives a root λ_n, the factor theorem gives $p = (x-\lambda_n)q$ with $\deg q = n-1$ and leading coefficient a, and induction factors q. The rest is Proposition 7.1.34.

We show by induction on n that p splits. For $n = 0$, $p = a$ is the empty product times a. Let $n \geq 1$. By Theorem 7.1.32, p has a root $\lambda_n \in \mathbb{C}$, and by Theorem 7.1.20, $p = (x-\lambda_n)q$ with $\deg q = n-1$. By Proposition 7.1.4, the leading coefficient of q is that of p, namely a. By the induction hypothesis $q = a(x-\lambda_1)\cdots(x-\lambda_{n-1})$, and so $p = a(x-\lambda_1)\cdots(x-\lambda_n)$. By Proposition 7.1.34, the λ_i are the roots of p, each occurring as often as its multiplicity, and the list is determined by p up to order. □

So over $\mathbb{C}$ the bound of Proposition 7.1.27 is an equality. This is the count behind "an $n \times n$ complex matrix has n eigenvalues, counted with multiplicity" in Section 7.4. The coefficients of a split polynomial are also expressed through its roots. We record the two that matter most later.

Proposition 7.1.36 (Sum and product of the roots). Let $n \geq 1$ and $\lambda_1, \ldots, \lambda_n \in F$. In $(x-\lambda_1)\cdots(x-\lambda_n)$, the coefficient of x^{n-1} is $-(\lambda_1 + \cdots + \lambda_n)$ and the constant term is $(-1)^n \lambda_1 \lambda_2 \cdots \lambda_n$. So for $n \geq 2$,

$$\begin{aligned}(x-\lambda_1)\cdots(x-\lambda_n) = x^n - (\lambda_1 + \cdots + \lambda_n)x^{n-1} \\ + \cdots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n,\end{aligned}$$

where the omitted terms have degrees from 1 to $n-2$.

Proof. ROUTINE

Uses. Proposition 7.1.10 (Evaluation respects sums and products), Proposition 7.1.3 ($F[x]$ is a commutative ring), induction on n

Hint 1. The constant term of a polynomial is its value at 0.
Hint 2. For the coefficient of x^{n-1}, induct on n and multiply by $x - \lambda_{n+1}$.
Sketch. The constant term is the value at 0, which is $\prod(0-\lambda_i) = (-1)^n \prod \lambda_i$. If $(x-\lambda_1)\cdots(x-\lambda_n) = x^n - ex^{n-1} + r$ with $\deg r \le n-2$, then multiplying by $x - \lambda_{n+1}$ gives $x^{n+1} - (e + \lambda_{n+1})x^n$ plus terms of degree at most $n-1$.

Let $P_n = (x-\lambda_1)\cdots(x-\lambda_n)$. Its constant term is $P_n(0) = (0-\lambda_1)\cdots(0-\lambda_n) = (-1)^n\lambda_1\cdots\lambda_n$ by Proposition 7.1.10.

For the coefficient of x^{n-1} we induct on n. For $n = 1$, $P_1 = x - \lambda_1$. Suppose $P_n = x^n - e_n x^{n-1} + r$ with $e_n = \lambda_1 + \cdots + \lambda_n$ and $\deg r \le n-2$. Then

$$\begin{aligned} P_{n+1} &= (x - \lambda_{n+1})P_n \\ &= x^{n+1} - (e_n + \lambda_{n+1})x^n + \left(\lambda_{n+1}e_n x^{n-1} + xr - \lambda_{n+1}r\right), \end{aligned}$$

and the last group has degree at most $n-1$. So the coefficient of x^n in P_{n+1} is $-(\lambda_1 + \cdots + \lambda_{n+1})$. □

By Corollary 7.1.35, for a monic $p = x^n + c_{n-1}x^{n-1} + \cdots + c_0 \in \mathbb{C}[x]$, the sum of its roots counted with multiplicity is $-c_{n-1}$, and their product is $(-1)^n c_0$. Applied to characteristic polynomials, this says that the trace of a complex matrix is the sum of its eigenvalues and the determinant is their product (compare the coefficients in Proposition 5.3.19; see Section 7.5). Exercise 7.1.59 computes the remaining coefficient of a cubic.

7.1.6 Factorization over $\mathbb{R}$

A polynomial with real coefficients is also a polynomial with complex coefficients, so Corollary 7.1.35 factors it over $\mathbb{C}$. The factors need not be real, but they come in conjugate pairs, and each pair multiplies to a real quadratic.

For $f = \sum_k c_k x^k \in \mathbb{C}[x]$, write $\overline{f} = \sum_k \overline{c_k} x^k$ for the polynomial with conjugated coefficients (Definition B.1.1).

Lemma 7.1.37 (Conjugating polynomials). For $f, g \in \mathbb{C}[x]$ and $\lambda \in \mathbb{C}$:

(a) $\overline{f+g} = \overline{f} + \overline{g}$ and $\overline{fg} = \overline{f}\,\overline{g}$;

(b) $\overline{f}\left(\overline{\lambda}\right) = \overline{f(\lambda)}$;

(c) $\overline{f} = f$ if and only if $f \in \mathbb{R}[x]$.

Proof. ROUTINE

Uses. Notation 6.1.1 (Real and complex scalars), Definition 7.1.1 (Product of polynomials), Definition 1.2.17 (Evaluation)
Hint. Conjugation of complex numbers respects sums and products (Notation 6.1.1). Apply it to the formulas for the coefficients of $f+g$ and fg, and to the formula for $f(\lambda)$.
Sketch. The k-th coefficient of $\overline{fg}$ is $\overline{\sum_{i+j=k} a_i b_j} = \sum_{i+j=k} \overline{a_i}\,\overline{b_j}$, that of $\overline{f}\,\overline{g}$; sums are similar. Likewise $\overline{f(\lambda)} = \sum \overline{c_k}\,\overline{\lambda}^k$. A complex number equals its conjugate exactly when it is real.

Write $f = \sum_k a_k x^k$ and $g = \sum_k b_k x^k$. We use $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\,\overline{w}$, and $\overline{z} = z$ if and only if $z \in \mathbb{R}$ (Notation 6.1.1).

(a) The k-th coefficient of $\overline{f+g}$ is $\overline{a_k + b_k} = \overline{a_k} + \overline{b_k}$, that of $\overline{f} + \overline{g}$. The k-th coefficient of $\overline{fg}$ is $\overline{\sum_{i+j=k} a_i b_j} = \sum_{i+j=k} \overline{a_i}\,\overline{b_j}$, that of $\overline{f}\,\overline{g}$.
(b) $\overline{f(\lambda)} = \overline{\sum_k a_k \lambda^k} = \sum_k \overline{a_k}\,\overline{\lambda}^k = \overline{f}\left(\overline{\lambda}\right)$.
(c) $\overline{f} = f$ means $\overline{a_k} = a_k$ for every k, that is, $a_k \in \mathbb{R}$ for every k. □

Intuition. Conjugation is a symmetry of $\mathbb{C}$: it respects sums and products and fixes every real number. So it carries any equation with real coefficients to another true equation. Applied to $p(\lambda) = 0$ for a real polynomial p, it gives $p\left(\overline{\lambda}\right) = 0$, and applied to a factorization, it shows that λ and $\overline{\lambda}$ occur equally often.

Proposition 7.1.38 (Conjugate roots). Let $p \in \mathbb{R}[x]$ be nonzero, regarded as an element of $\mathbb{C}[x]$, and let $\lambda \in \mathbb{C}$. Then λ and $\overline{\lambda}$ have the same multiplicity as roots of p. In particular, the nonreal roots of p come in pairs $\lambda \neq \overline{\lambda}$.

Proof. STANDARD

Uses. Lemma 7.1.37 (Conjugating polynomials), Lemma 7.1.25 (Reading off a multiplicity), Notation 6.1.1 (Real and complex scalars)
Hint 1. First show that $\overline{\lambda}$ is a root whenever λ is: conjugate the equation $p(\lambda) = 0$.
Hint 2. For multiplicities, conjugate a factorization $p = (x-\lambda)^m q$ with $q(\lambda) \neq 0$, and apply Lemma 7.1.25 at $\overline{\lambda}$.
Sketch. Let m be the multiplicity of λ, so $p = (x-\lambda)^m q$ with $q(\lambda) \neq 0$. Conjugating with Lemma 7.1.37, and using $\overline{p} = p$ and $\overline{x-\lambda} = x - \overline{\lambda}$, gives $p = \big(x-\overline{\lambda}\big)^m \overline{q}$ with $\overline{q}\big(\overline{\lambda}\big) = \overline{q(\lambda)} \neq 0$. So $\overline{\lambda}$ has multiplicity m.

Let m be the multiplicity of λ, and write $p = (x-\lambda)^m q$ with $q(\lambda) \neq 0$ (Lemma 7.1.25). Since p has real coefficients, $\overline{p} = p$, and $\overline{x-\lambda} = x - \overline{\lambda}$. By Lemma 7.1.37(a), applied to each factor,

$$p = \overline{p} = \big(x-\overline{\lambda}\big)^m \overline{q},$$

and by Lemma 7.1.37(b), $\overline{q}\big(\overline{\lambda}\big) = \overline{q(\lambda)} \neq 0$. By Lemma 7.1.25, $\overline{\lambda}$ has multiplicity m as a root of p. If λ is not real, then $\lambda \neq \overline{\lambda}$, so the nonreal roots of p come in pairs $\lambda, \overline{\lambda}$ of equal multiplicity. □

Theorem 7.1.39 (Factorization over $\mathbb{R}$). Every nonzero $p \in \mathbb{R}[x]$ can be written as

$$p = a(x-\lambda_1)\cdots(x-\lambda_k)\big(x^2+b_1x+c_1\big)\cdots\big(x^2+b_lx+c_l\big),$$

where $a, \lambda_i, b_j, c_j \in \mathbb{R}$ and $b_j^2 < 4c_j$ for each j (so $k+2l = \deg p$). Here $\lambda_1, \ldots, \lambda_k$ are the real roots of p, each listed as many times as its multiplicity. The factorization is unique up to the order of the factors.

Proof. STANDARD

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Proposition 7.1.38 (Conjugate roots), Proposition 7.1.34 (Roots of a split polynomial), Notation 6.1.1 (Real and complex scalars), Corollary 7.1.17 (Division does not depend on the field)
Hint 1. Factor p over $\mathbb{C}$. The real roots give real linear factors. What is $(x-\mu)(x-\overline{\mu})$ for a nonreal μ?
Hint 2. $(x-\mu)(x-\overline{\mu}) = x^2 - 2(\operatorname{Re}\mu)x + |\mu|^2$. Use Proposition 7.1.38 to pair up all nonreal factors.
Hint 3. For uniqueness, split each real quadratic factor over $\mathbb{C}$ into a conjugate pair, and compare with the complex factorization, which is unique by Proposition 7.1.34.
Sketch. Over $\mathbb{C}$, $p = a\prod_i (x-\lambda_i)$ with a real. By Proposition 7.1.38, the nonreal roots with negative imaginary part are the conjugates of those with positive imaginary part, with the same multiplicities, so the nonreal factors pair up as $(x-\mu)(x-\overline{\mu}) = x^2 - 2(\operatorname{Re}\mu)x + |\mu|^2$, a real quadratic with discriminant $-4(\operatorname{Im}\mu)^2 < 0$. Conversely, each real quadratic with $b^2 < 4c$ is $(x-\nu)(x-\overline{\nu})$ for a nonreal ν, so any real factorization gives the complex one; the real roots are the λ_i, and the roots with positive imaginary part determine the quadratics.

Existence. By Corollary 7.1.35, $p = a(x-\lambda_1)\cdots(x-\lambda_n)$ in $\mathbb{C}[x]$, where a is the leading coefficient of p, which is real, and the λ_i are the roots of p listed with multiplicity. Reorder them so that $\lambda_1, \ldots, \lambda_k$ are the real ones. Let $\mu_1, \ldots, \mu_l$ be the roots of p with positive imaginary part, each listed as many times as its multiplicity. By Proposition 7.1.38, conjugation maps the roots with positive imaginary part bijectively onto the roots with negative imaginary part, preserving multiplicities. So $\overline{\mu_1}, \ldots, \overline{\mu_l}$ are the roots with negative imaginary part, listed with multiplicity. Hence, reordering factors,

$$p = a(x-\lambda_1)\cdots(x-\lambda_k)\prod_{j=1}^{l}\big(x-\mu_j\big)\big(x-\overline{\mu_j}\big).$$

For each j, using $\mu + \overline{\mu} = 2\operatorname{Re}\mu$ and $\mu\overline{\mu} = |\mu|^2$ (Notation 6.1.1),

$$\big(x-\mu_j\big)\big(x-\overline{\mu_j}\big) = x^2 - 2\big(\operatorname{Re}\mu_j\big)x + |\mu_j|^2 = x^2 + b_jx + c_j,$$

with $b_j = -2\operatorname{Re}\mu_j$ and $c_j = |\mu_j|^2$ real, and $b_j^2 - 4c_j = 4\big(\operatorname{Re}\mu_j\big)^2 - 4\Big(\big(\operatorname{Re}\mu_j\big)^2 + \big(\operatorname{Im}\mu_j\big)^2\Big) = -4\big(\operatorname{Im}\mu_j\big)^2 < 0$. All factors now lie in $\mathbb{R}[x]$, and products in $\mathbb{R}[x]$ are computed by the same formula as in $\mathbb{C}[x]$, so this is a

factorization of the required form in $\mathbb{R}[x]$. The λ_i are the real roots listed with their multiplicities in $\mathbb{C}[x]$; for a real root λ this is also its multiplicity in $\mathbb{R}[x]$, because whether $(x-\lambda)^m$ divides p does not depend on the field (Corollary 7.1.17).

Uniqueness. Take any factorization of the stated form, with factors $x-\lambda_i'$ ($i \le k'$) and $x^2+b_j'x+c_j'$ ($j \le l'$) and constant a'. Comparing leading coefficients (Proposition 7.1.4), $a'=a$. For each j, since $4c_j'-b_j'^2 > 0$, the number

$$v_j = \frac{-b_j' + i\sqrt{4c_j' - b_j'^2}}{2}$$

is not real, and $v_j + \overline{v_j} = -b_j'$ and $v_j\overline{v_j} = \left(b_j'^2 + \left(4c_j' - b_j'^2\right)\right)/4 = c_j'$, so $x^2+b_j'x+c_j' = \left(x-v_j\right)\left(x-\overline{v_j}\right)$. Substituting, p splits over $\mathbb{C}$ with the list of roots $\lambda_1', \ldots, \lambda_{k'}', v_1, \overline{v_1}, \ldots, v_{l'}, \overline{v_{l'}}$. By Proposition 7.1.34 this list is, up to order, the list of roots of p with multiplicity. Its real members are the λ_i', so these are the real roots of p listed with multiplicity. Its members with positive imaginary part are the v_j, so these are the roots of p with positive imaginary part, listed with multiplicity. Both lists are determined by p, and each quadratic factor is determined by its v_j, as $\left(x-v_j\right)\left(x-\overline{v_j}\right)$. Hence the factors are determined by p up to order. □

Example 7.1.40 (x^4+4). Since $(1+i)^2 = 2i$, we get $(1+i)^4 = (2i)^2 = -4$, and in the same way $(\pm 1 \pm i)^4 = -4$ for all four choices of signs. These are four distinct roots of x^4+4, so by Corollary 7.1.35

$$x^4+4 = (x-1-i)(x-1+i)(x+1-i)(x+1+i).$$

Pairing conjugates as in the proof, $(x-1-i)(x-1+i) = (x-1)^2+1$ and $(x+1-i)(x+1+i) = (x+1)^2+1$, so

$$x^4+4 = \left(x^2-2x+2\right)\left(x^2+2x+2\right)$$

over $\mathbb{R}$, and even over $\mathbb{Q}$, although x^4+4 has no real root. As Proposition 7.1.36 predicts, the roots add up to 0, the negative of the coefficient of x^3, and their product is $(1+1)(1+1) = 4 = (-1)^4 \cdot 4$.

The real quadratic factors cannot be split further over $\mathbb{R}$. This is the real counterpart of the fact that linear factors cannot be split further over $\mathbb{C}$.

Definition 7.1.41 (Irreducible polynomial). A polynomial $p \in F[x]$ of degree at least 1 is **irreducible** over F if it cannot be written as $p = gh$ with $g, h \in F[x]$ both of degree at least 1. Otherwise, and if $\deg p \ge 1$, it is **reducible**.

Every polynomial of degree 1 is irreducible, since degrees add (Proposition 7.1.4). Irreducibility depends on the field: x^2+1 is irreducible over $\mathbb{R}$ but equals $(x-i)(x+i)$ over $\mathbb{C}$; x^2+x+1 is irreducible over $\mathbb{F}_2$, since it has no root there (Exercise 7.1.61), but equals $(x-1)^2$ over $\mathbb{F}_3$. Irreducible polynomials play the role of prime numbers (Exercise 7.1.68).

Corollary 7.1.42 (Irreducible polynomials over $\mathbb{C}$ and $\mathbb{R}$).

(a) A polynomial in $\mathbb{C}[x]$ is irreducible over $\mathbb{C}$ if and only if it has degree 1.

(b) A polynomial in $\mathbb{R}[x]$ is irreducible over $\mathbb{R}$ if and only if it has degree 1, or it is ax^2+bx+c with $a \ne 0$ and $b^2 < 4ac$.

Proof. ROUTINE

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Theorem 7.1.39 (Factorization over $\mathbb{R}$), Theorem 7.1.20 (Factor theorem), Exercise 1.1.50, Proposition 7.1.4 (Degree of a product)

Hint 1. A polynomial of degree 1 is always irreducible. For larger degrees, read off a factorization from Corollary 7.1.35 or Theorem 7.1.39.

Hint 2. For a real quadratic, a factorization into two linear factors gives a real root. When does $ax^2 + bx + c$ have a real root?
Sketch. Degree 1 is irreducible because degrees add. Over $\mathbb{C}$, degree $n \geq 2$ splits into $n \geq 2$ linear factors. Over $\mathbb{R}$, degree at least 3 gives at least two factors in Theorem 7.1.39, as $k + 2l \geq 3$. A real quadratic is reducible if and only if it has a real root, and by the quadratic formula (Exercise 1.1.50) that happens if and only if $b^2 - 4ac \geq 0$.

Polynomials of degree 1 are irreducible over every field, because in a product gh of two polynomials of degree at least 1, the degree is at least 2.

(a) If $p \in \mathbb{C}[x]$ has degree $n \geq 2$, then by Corollary 7.1.35 $p = (a(x - \lambda_1)) \cdot ((x - \lambda_2) \cdots (x - \lambda_n))$, a product of two polynomials of degrees 1 and $n - 1 \geq 1$.

(b) Let $p \in \mathbb{R}[x]$. If $\deg p \geq 3$, then in Theorem 7.1.39 $k + 2l = \deg p \geq 3$, so there are at least two non-constant factors. Splitting off one of them (and absorbing a into it) writes p as a product of two polynomials of degree at least 1, so p is reducible.

Now let $p = ax^2 + bx + c$ with $a \neq 0$. If p has a real root t, then $p = (x - t)q$ with $\deg q = 1$ (Theorem 7.1.20), so p is reducible. Conversely, if $p = gh$ with $\deg g = \deg h = 1$, say $g = ux + v$ with $u \neq 0$, then $-v/u$ is a root of g, hence of p. So p is irreducible if and only if it has no real root. By Exercise 1.1.50, p has a real root if and only if $b^2 - 4ac = d^2$ for some $d \in \mathbb{R}$, that is, if and only if $b^2 - 4ac \geq 0$. So p is irreducible if and only if $b^2 < 4ac$. □

The last consequence concerns odd degree. It has a proof from the factorization and a proof that avoids $\mathbb{C}$ altogether.

Corollary 7.1.43 (Odd degree gives a real root). Every polynomial $p \in \mathbb{R}[x]$ of odd degree has a real root.

Proof (by conjugate pairs). ROUTINE

Uses. Theorem 7.1.39 (Factorization over $\mathbb{R}$), Proposition 7.1.10 (Evaluation respects sums and products)
Hint. In Theorem 7.1.39, $k + 2l = \deg p$. What does odd degree say about k?
Sketch. $k + 2l = \deg p$ is odd, so $k \geq 1$, and λ_1 is a real root.

In the factorization of Theorem 7.1.39, $k + 2l = \deg p$ is odd, so k is odd and in particular $k \geq 1$. Then λ_1 is a real root of p, by Proposition 7.1.10. □

Proof (by the intermediate value theorem). STANDARD

Uses. the intermediate value theorem
Hint 1. Assume the leading coefficient is positive. Show that $p(t) > 0$ for large positive t and $p(t) < 0$ for large negative t.
Hint 2. For $|t| \geq 1$, bound the lower terms by $A\,|t|^{n-1}$, where A is the sum of the absolute values of the lower coefficients.
Sketch. With $a_n > 0$ and $A = \sum_{k<n} |a_k|$, for $|t| \geq 1$ the lower terms are at most $A\,|t|^{n-1}$ in absolute value. At $T = 1 + A/a_n$ this gives $p(T) > 0 > p(-T)$, because n is odd. The intermediate value theorem gives a root in $(-T, T)$.

Let $p = \sum_{k=0}^{n} a_k x^k$ with n odd and $a_n \neq 0$. Replacing p by $-p$, which has the same roots, we may assume $a_n > 0$. Let $A = |a_0| + \cdots + |a_{n-1}|$. For real t with $|t| \geq 1$, we have $|a_k t^k| \leq |a_k|\,|t|^{n-1}$ for $k < n$, so

$$|p(t) - a_n t^n| \leq A\,|t|^{n-1}.$$

Let $T = 1 + A/a_n \geq 1$, so that $a_n T - A = a_n > 0$. Then $p(T) \geq a_n T^n - A T^{n-1} = T^{n-1} a_n > 0$, and since n is odd, $(-T)^n = -T^n$ and

$$p(-T) \leq -a_n T^n + A T^{n-1} = -T^{n-1} a_n < 0.$$

The function $t \mapsto p(t)$ is continuous on $[-T, T]$, so by the intermediate value theorem it vanishes somewhere in $(-T, T)$. □

The second proof is the more elementary one: it uses the completeness of $\mathbb{R}$ directly, through the intermediate value theorem, while the first relies on the Fundamental Theorem of Algebra, whose proof uses completeness in a much more elaborate way. On the other hand, the first proof gives more: the number of

real roots, counted with multiplicity, has the same parity as $\deg p$. In Section 7.3, this corollary shows that every operator on a real vector space of odd dimension has an eigenvalue.

7.1.7 Greatest common divisors BEYOND 21-242

Greatest common divisors of polynomials are not covered in lecture. They are used in Section 7.6, Section 10.2, and Section 14.4.

The integers have greatest common divisors, computed by the Euclidean algorithm, and $\gcd(a, b)$ is an integer combination $ua + vb$ (Bézout's identity). All of this carries over to $F[x]$, because the only tool it needs is division with remainder. The application to linear algebra is this: if p and s have no common factor, then $up + vs = 1$ for some polynomials u and v, and for a square matrix A, Remark 7.1.11 turns this into $u(A)p(A) + v(A)s(A) = I_n$. Section 7.6 and Section 10.2 use such identities to split a space into the kernels of $p(A)$ and $s(A)$.

Definition 7.1.44 (Greatest common divisor). Let $p, s \in F[x]$, not both $\mathbf{0}$. A **greatest common divisor** of p and s is a monic polynomial d such that $d \mid p$, $d \mid s$, and every common divisor of p and s divides d. If 1 is a greatest common divisor of p and s, they are **coprime**.

"Greatest" refers to divisibility, but it implies the largest degree: every common divisor divides d, so its degree is at most $\deg d$.

Theorem 7.1.45 (Bézout's identity). Let $p, s \in F[x]$, not both $\mathbf{0}$. Then p and s have exactly one greatest common divisor d, written $\gcd(p, s)$, and there are $u, v \in F[x]$ with

$$d = up + vs.$$

Moreover, d is the monic polynomial of smallest degree of the form $up + vs$.

Intuition. Consider all combinations $up + vs$. Every common divisor of p and s divides each of them, so a combination that itself divides p and s is automatically a greatest common divisor. The smallest nonzero combination is the natural candidate: dividing p by it leaves a remainder that is again a combination, and smaller, so it must be $\mathbf{0}$.

Proof (by a combination of smallest degree). HARD

Uses. Theorem 7.1.14 (Division with remainder), Definition 7.1.44 (Greatest common divisor), Definition 7.1.16 (Divisibility), Corollary 7.1.7 (No zero divisors, cancellation, and units)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Uniqueness via two candidates (Technique 1.1.8), Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. Let $I = \{up + vs \mid u, v \in F[x]\}$. Choose a nonzero element of I of smallest degree, and scale it to be monic.
Hint 2. Divide p by this element d. Show that the remainder lies in I, and conclude from minimality that it is $\mathbf{0}$.
Hint 3. For uniqueness, two greatest common divisors divide each other, and both are monic.
Sketch. The set I of combinations $up + vs$ is closed under sums and under multiplication by polynomials, and contains a nonzero element. Let $d \in I$ be monic of smallest degree. The remainder of p on division by d is $p - qd \in I$ and has degree less than $\deg d$, so it is $\mathbf{0}$; thus $d \mid p$, and likewise $d \mid s$. A common divisor of p and s divides $up + vs = d$. Two greatest common divisors divide each other and are monic, so they are equal.

Let $I = \{up + vs \mid u, v \in F[x]\}$. It contains $p = 1p + \mathbf{0}s$ and s, so it contains a nonzero polynomial. It is closed under sums, since $(up + vs) + (u'p + v's) = (u + u')p + (v + v')s$, and under multiplication by any $w \in F[x]$, since $w(up + vs) = (wu)p + (wv)s$.

Existence. Among the nonzero elements of I, choose one of smallest degree. Multiplying it by the inverse of its leading coefficient keeps it in I and makes it monic; call the result $d = up + vs$. Divide p by d: $p = dq + r$ with $\deg r < \deg d$ (Theorem 7.1.14). Then

$$r = p - qd = (1 - qu)p + (-qv)s \in I.$$

By the choice of d, no nonzero element of I has degree less than $\deg d$, so $r = \mathbf{0}$ and $d \mid p$. In the same way $d \mid s$. If e divides both p and s, say $p = ep'$ and $s = es'$, then $d = up + vs = e(up' + vs')$, so $e \mid d$. Hence d is a greatest common divisor, and it is a combination of p and s.

Uniqueness. If d and d' are both greatest common divisors, then each divides the other, by the definition applied twice. As noted after Definition 7.1.16, two monic polynomials that divide each other are equal.

Smallest degree. We showed that the gcd is a monic element of I of smallest degree. Conversely, any monic $d' \in I$ of smallest degree is, by the existence argument applied to d', a greatest common divisor, so it equals d. □

The second proof computes d, and u and v as well, by repeated division.

Algorithm 7.1.46 (Euclidean algorithm)

Input Polynomials $p, s \in F[x]$ with $s \neq \mathbf{0}$.
Output $\gcd(p, s)$.

1. Set $r_0 = p$ and $r_1 = s$, and $i = 1$.
2. If $r_i = \mathbf{0}$, output r_{i-1} divided by its leading coefficient, and stop.
3. Otherwise let r_{i+1} be the remainder of r_{i-1} on division by r_i, increase i by 1, and return to step 2.

Proof of Theorem 7.1.45 (by the Euclidean algorithm). HARD

> *Uses.* Theorem 7.1.14 (Division with remainder), Definition 7.1.44 (Greatest common divisor), Definition 7.1.16 (Divisibility)
> *Techniques.* Algorithms as proofs: invariants and termination (Technique 3.3.5)
> *Hint 1.* If $s = \mathbf{0}$, the gcd is p made monic. Otherwise run Algorithm 7.1.46. Why does it stop?
> *Hint 2.* Find an invariant: compare the common divisors of r_{i-1} and r_i with those of r_i and r_{i+1}.
> *Hint 3.* Show also by induction that every r_i is a combination $u_i p + v_i s$.
> *Sketch.* The degrees $\deg r_1 > \deg r_2 > \cdots$ strictly decrease, so some $r_{N+1} = \mathbf{0}$. Since $r_{i+1} = r_{i-1} - q_i r_i$ and $r_{i-1} = q_i r_i + r_{i+1}$, the pairs (r_{i-1}, r_i) and (r_i, r_{i+1}) have the same common divisors. So the common divisors of p and s are those of r_N and $\mathbf{0}$, the divisors of r_N. Each r_i is a combination of p and s by induction. Making r_N monic gives d; uniqueness is as in the first proof.

If $s = \mathbf{0}$, then $p \neq \mathbf{0}$, and the common divisors of p and $\mathbf{0}$ are the divisors of p. So p divided by its leading coefficient a is a greatest common divisor, and it equals $a^{-1}p + \mathbf{0}s$. Now let $s \neq \mathbf{0}$ and run Algorithm 7.1.46. Each r_{i+1} is a remainder on division by r_i, so $\deg r_{i+1} < \deg r_i$. A strictly decreasing sequence of nonnegative integers is finite, so some $r_{N+1} = \mathbf{0}$ with $r_N \neq \mathbf{0}$, and the algorithm stops.

Invariant. Write $r_{i-1} = q_i r_i + r_{i+1}$. If e divides r_{i-1} and r_i, it divides $r_{i+1} = r_{i-1} - q_i r_i$; if e divides r_i and r_{i+1}, it divides r_{i-1}. So the pairs (r_{i-1}, r_i) and (r_i, r_{i+1}) have the same common divisors. Following the chain from $(p, s) = (r_0, r_1)$ to $(r_N, r_{N+1}) = (r_N, \mathbf{0})$, the common divisors of p and s are exactly the divisors of r_N. In particular, with a the leading coefficient of r_N, the monic polynomial $d = a^{-1}r_N$ divides p and s, and every common divisor divides d. So d is a greatest common divisor.

Combination. We show by induction on i that $r_i = u_i p + v_i s$ for some u_i, v_i. This holds for $r_0 = 1p + \mathbf{0}s$ and $r_1 = \mathbf{0}p + 1s$, and if it holds for r_{i-1} and r_i, then $r_{i+1} = r_{i-1} - q_i r_i = (u_{i-1} - q_i u_i)p + (v_{i-1} - q_i v_i)s$. So $d = a^{-1}u_N p + a^{-1}v_N s$. Every nonzero combination $up + vs$ is divisible by d, which divides p and s, so it has degree at least $\deg d$; and a monic combination d' of degree $\deg d$ is divisible by d, so $d' = d$ (compare leading coefficients). Uniqueness of the gcd is proved as in the first proof. □

Example 7.1.47 (A greatest common divisor over $\mathbb{Q}$). Let $p = x^4 - 1$ and $s = x^3 + 2x^2 + 2x + 1$. The Euclidean algorithm, with $r_0 = p$ and $r_1 = s$, gives

$$p = (x - 2)s + r_2, \qquad r_2 = 2x^2 + 3x + 1,$$

$$s = \left(\frac{1}{2}x + \frac{1}{4}\right)r_2 + r_3, r_3 = \frac{3}{4}(x+1),$$
$$r_2 = \left(\frac{8}{3}x + \frac{4}{3}\right)r_3.$$

The last nonzero remainder is $r_3 = \frac{3}{4}(x+1)$, so $\gcd(p,s) = x+1$. Running the substitutions backwards, with $r_2 = p - (x-2)s$ and $r_3 = s - \left(\frac{1}{2}x + \frac{1}{4}\right)r_2$,

$$x + 1 = \frac{4}{3}r_3 = -\left(\frac{2}{3}x + \frac{1}{3}\right)p + \left(\frac{2}{3}x^2 - x + \frac{2}{3}\right)s.$$

The factorizations $p = (x-1)(x+1)\left(x^2+1\right)$ and $s = (x+1)\left(x^2+x+1\right)$ confirm the answer.

Corollary 7.1.48 (Coprime polynomials). Let $p, s \in F[x]$, not both $\mathbf{0}$.

(a) p and s are coprime if and only if $up + vs = 1$ for some $u, v \in F[x]$.

(b) If $F = \mathbb{C}$, then p and s are coprime if and only if they have no common root.

Proof. STANDARD

Uses. Theorem 7.1.45 (Bézout's identity), Corollary 7.1.7 (No zero divisors, cancellation, and units), Theorem 7.1.20 (Factor theorem), Theorem 7.1.32 (Fundamental Theorem of Algebra)
Hint 1. For (a), "only if" is Theorem 7.1.45. For "if", what can a common divisor of p and s be, if it divides 1?
Hint 2. For (b), a common root λ gives the common factor $x - \lambda$. Conversely, apply Theorem 7.1.32 to the gcd.
Sketch. (a) If p, s are coprime, Bézout gives $up + vs = 1$. Conversely, every common divisor divides $up + vs = 1$, so it is a nonzero constant, and the monic gcd is 1. (b) A common root λ makes $x - \lambda$ a common divisor of degree 1, so the gcd is not 1. If the gcd d is not 1, it is monic of degree at least 1 and has a root by Theorem 7.1.32, which is a root of p and of s.

(a) If p and s are coprime, Theorem 7.1.45 gives $up + vs = \gcd(p,s) = 1$. Conversely, suppose $up + vs = 1$, and let $d = \gcd(p,s)$. Then d divides $up + vs = 1$, so $\deg d \leq 0$, and d, being monic, is 1.

(b) If λ is a common root, then $x - \lambda$ divides p and s (Theorem 7.1.20), so it divides $\gcd(p,s)$, which therefore has degree at least 1 and is not 1. Conversely, if $d = \gcd(p,s) \neq 1$, then d is monic of degree at least 1, since the only monic constant is 1. By Theorem 7.1.32, d has a root $\lambda \in \mathbb{C}$, and since d divides p and s, λ is a root of both. □

Part (b) fails over $\mathbb{R}$: $x^2 + 1$ and $x^4 - 1$ have no common real root, but their gcd is $x^2 + 1$.

Exercises

Exercise 7.1.49. ROUTINE Let $p = x^4 - 2x^3 + 2x - 1$. Factor p into linear factors over $\mathbb{Q}$, over $\mathbb{F}_3$, and over $\mathbb{F}_2$, and in each case list the roots of p with their multiplicities.

Solution on page 1554

Exercise 7.1.50. STANDARD Over $\mathbb{Q}$, find the remainder of x^{100} on division by (a) $x^2 - 1$; (b) $x^2 + 1$; (c) $(x-1)^2$.

Solution on page 1554

Exercise 7.1.51. STANDARD Let $s \in F[x]$ have degree $m \geq 1$, let $n \geq m$, and for $p \in F[x]_{\leq n}$ let $R(p)$ be the remainder of p on division by s.

(a) Show that $R : F[x]_{\leq n} \to F[x]_{\leq m-1}$ is linear.

(b) Show that R is surjective and that $\ker R = \{sq \mid q \in F[x]_{\leq n-m}\}$.

(c) Verify the rank–nullity formula (Theorem 4.2.11) for R.

Solution on page 1555

Exercise 7.1.52. STANDARD The definitions of this section make sense with coefficients in $\mathbb{Z}/8\mathbb{Z}$, which is a commutative ring but not a field (Lemma 1.1.21).

(a) Show that $x^2 - 1$ has four roots in $\mathbb{Z}/8\mathbb{Z}$, and that $x^2 - 1 = (x-1)(x+1) = (x-3)(x+3)$ in $(\mathbb{Z}/8\mathbb{Z})[x]$.

(b) The proof of Proposition 7.1.22, run over $\mathbb{Z}/8\mathbb{Z}$, would show that $x^2 - 1$ has at most two roots. Find the first step that fails.

Solution on page 1555

Exercise 7.1.53. STANDARD Prove or disprove, for $p, q, r \in F[x]$.

(a) If $p(c) = q(c)$ for every $c \in F$, then $p = q$.

(b) If $p \mid q$ and $q \mid p$, then $p = q$.

(c) If $p \mid qr$, then $p \mid q$ or $p \mid r$.

(d) If $p \in \mathbb{R}[x]$ has no real root, then $\deg p$ is even.

(e) If $p \in \mathbb{R}[x]$ is nonzero, then the number of real roots of p, counted with multiplicity, has the same parity as $\deg p$.

Solution on page 1556

Exercise 7.1.54. STANDARD

(a) Let $n \geq 1$. Show that $x^n - 1$ has n distinct roots in $\mathbb{C}$.

(b) Let p be a prime. Show that $x^p - 1 = (x-1)^p$ in $\mathbb{F}_p[x]$, so that 1 is the only root of $x^p - 1$ in $\mathbb{F}_p$, with multiplicity p.

Solution on page 1556

Exercise 7.1.55. HARD Let p be a prime.

(a) Show that $x^p - x = \prod_{a \in \mathbb{F}_p} (x - a)$ in $\mathbb{F}_p[x]$.

(b) Deduce Wilson's theorem: $(p-1)! \equiv -1 \pmod{p}$.

Solution on page 1557

Exercise 7.1.56. STANDARD Consider $f \in F[x]$ with $f(x+1) = f$.

(a) Show that if F has characteristic 0 (Definition 1.1.28), then f is constant.

(b) Let p be a prime. Find a nonconstant $f \in \mathbb{F}_p[x]$ with $f(x+1) = f$.

Solution on page 1557

Exercise 7.1.57. STANDARD Let $a, b, c, d \in F$.

(a) Show that there is exactly one $f \in F[x]_{\leq 3}$ with $f(0) = a$, $f'(0) = b$, $f(1) = c$, and $f'(1) = d$, where f' is the formal derivative.

(b) Find f over $\mathbb{Q}$ for $(a, b, c, d) = (1, 0, 0, 0)$.

Solution on page 1558

Exercise 7.1.58. STANDARD Let $n \geq 1$, let $c_0, \ldots, c_n \in F$ be distinct, let $L_0, \ldots, L_n$ be the Lagrange polynomials (Definition 3.4.26), and let $\omega = (x - c_0) \cdots (x - c_n)$.

(a) Show that $L_0 + L_1 + \cdots + L_n = 1$ and $c_0 L_0 + c_1 L_1 + \cdots + c_n L_n = x$.

(b) Show that $\omega'(c_i) = \prod_{j \neq i} (c_i - c_j)$, and deduce that $L_i = \omega_i / \omega'(c_i)$, where $\omega_i = \prod_{j \neq i} (x - c_j)$.

Solution on page 1558

Exercise 7.1.59. ROUTINE Let $\lambda_1, \lambda_2, \lambda_3$ be the roots of $x^3 - 2x + 5$ in $\mathbb{C}$, listed with multiplicity.

(a) Expand $(x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$.

(b) Compute $\lambda_1 + \lambda_2 + \lambda_3$, $\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$, $\lambda_1\lambda_2\lambda_3$, and $\lambda_1^2 + \lambda_2^2 + \lambda_3^2$.

Solution on page 1558

Exercise 7.1.60. STANDARD Factor $x^4 + 1$ and $x^6 - 1$ into irreducible polynomials over $\mathbb{C}$ and over $\mathbb{R}$.

Solution on page 1559

Exercise 7.1.61. ROUTINE

(a) Let $p \in F[x]$ have degree 2 or 3. Show that p is irreducible over F if and only if p has no root in F.

(b) Show that (a) fails in degree 4: find a polynomial of degree 4 over $\mathbb{Q}$, and one over $\mathbb{F}_2$, that has no root but is reducible.

(c) Which of $x^2 + x + 1$, $x^3 + x + 1$, and $x^3 + 1$ are irreducible over $\mathbb{F}_2$?

Solution on page 1559

Exercise 7.1.62. STANDARD Let p be a prime. Show that there are exactly $p(p-1)/2$ monic irreducible polynomials of degree 2 in $\mathbb{F}_p[x]$, and list them for $p = 3$.

Solution on page 1560

Exercise 7.1.63. STANDARD Let $f \in \mathbb{C}[x]_{\leq n}$, and suppose that $f(t_k) \in \mathbb{R}$ for $n+1$ distinct real numbers $t_0, \ldots, t_n$. Show that f has real coefficients. Show by an example that the hypothesis $\deg f \leq n$ cannot be dropped.

Solution on page 1560

Exercise 7.1.64. HARD Let $p \in \mathbb{R}[x]$ satisfy $p(t) \geq 0$ for every $t \in \mathbb{R}$. Show that $p = f^2 + g^2$ for some $f, g \in \mathbb{R}[x]$.

Solution on page 1560

Exercise 7.1.65. HARD

(a) Let F be a field with $1 + 1 \neq 0$. Find all $p \in F[x]$ with $p(x^2) = p^2$.

(b) Show that every $p \in \mathbb{F}_2[x]$ satisfies $p(x^2) = p^2$.

Solution on page 1561

Exercise 7.1.66. CHALLENGE BEYOND 21-242 (Gauss–Lucas) Let $p \in \mathbb{C}[x]$ have degree $n \geq 2$ and roots $\lambda_1, \ldots, \lambda_n$, listed with multiplicity. Show that every root w of p' is a convex combination of the roots of p: $w = t_1\lambda_1 + \cdots + t_n\lambda_n$ for some real $t_i \geq 0$ with $t_1 + \cdots + t_n = 1$.

Solution on page 1562

Exercise 7.1.67. STANDARD BEYOND 21-242 Let $p = x^4 + x^3 + 2x^2 + x + 1$ and $s = x^3 - 1$.

(a) Over $\mathbb{Q}$, compute $\gcd(p, s)$ with the Euclidean algorithm, and find $u, v \in \mathbb{Q}[x]$ with $up + vs = \gcd(p, s)$.

(b) Compute $\gcd(p, s)$ over $\mathbb{F}_2$. Why is the answer different?

Solution on page 1562

Exercise 7.1.68. HARD BEYOND 21-242

(a) (Euclid's lemma) Let $p \in F[x]$ be irreducible and $g, h \in F[x]$. Show that if $p \mid gh$, then $p \mid g$ or $p \mid h$.

(b) Show that every polynomial of degree at least 1 is a product of irreducible polynomials, and that the factorization is unique in the following sense: if $ap_1 \cdots p_k = bq_1 \cdots q_l$ with $a, b \in F$ nonzero and all p_i, q_j monic irreducible, then $a = b$, $k = l$, and the q_j are the p_i in some order.

Solution on page 1563

Exercise 7.1.69. HARD BEYOND 21-242 (Chinese remainder theorem) Let $m_1, m_2 \in F[x]$ be coprime, of degrees $d_1, d_2 \geq 1$, and let $r_1, r_2 \in F[x]$.

(a) Show that there is exactly one $p \in F[x]$ with $\deg p < d_1 + d_2$ such that $m_1 \mid p - r_1$ and $m_2 \mid p - r_2$.

(b) Explain why polynomial interpolation at two nodes (Theorem 3.4.24 with $n = 1$) is the special case $m_i = x - c_i$.

Solution on page 1563

7.2 Eigenvalues, eigenvectors, and invariant subspaces

A diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ is the simplest kind of operator: it multiplies the i-th coordinate by d_i, so it sends each standard basis vector $\mathbf{e}_i$ to the multiple $d_i\mathbf{e}_i$ of itself. A general operator mixes directions, but it may still have some directions that it only stretches, that is, nonzero vectors $\mathbf{v}$ with $T\mathbf{v} = \lambda\mathbf{v}$. If V has a basis of such vectors, the matrix of T in that basis is diagonal, and questions about T (its powers, its kernel, its long-run behavior) become questions about the scalars λ. This section introduces these vectors, the eigenvectors, together with the more general subspaces that an operator maps into themselves. Computing eigenvalues is the subject of Section 7.3, and the question of when a basis of eigenvectors exists is the subject of Section 7.4.

Throughout, V is a vector space over a field F, possibly infinite-dimensional unless we say otherwise, and $T \in \mathrm{L}(V)$ is an operator on V (Notation 2.1.2). We write $T\mathbf{v}$ for $T(\mathbf{v})$ when no confusion can arise.

7.2.1 Invariant subspaces

To understand an operator, we would like to cut V into pieces that T does not mix. The minimal requirement on a piece U is that T does not move vectors of U out of U. Then T restricts to an operator on U, which we can study on its own.

Definition 7.2.1 (Invariant subspace). Let $T \in \mathrm{L}(V)$. A subspace U of V is **invariant under** T (or T-**invariant**) if $T\mathbf{u} \in U$ for every $\mathbf{u} \in U$, that is, if $\operatorname{im}(T|_U) \subseteq U$. In that case we regard $T|_U$ as an operator on U, $T|_U \in \mathrm{L}(U)$.

Example 7.2.2 (Invariant subspaces).

(a) The subspaces $\{\mathbf{0}\}$ and V are invariant under every operator on V.

(b) $\ker T$ and $\operatorname{im} T$ are T-invariant. If $\mathbf{u} \in \ker T$, then $T\mathbf{u} = \mathbf{0} \in \ker T$; and $T\mathbf{u} \in \operatorname{im} T$ for every $\mathbf{u} \in V$, in particular for every $\mathbf{u} \in \operatorname{im} T$.

(c) Every subspace is invariant under a scaling map $\lambda \mathrm{id}_V$, since $\lambda\mathbf{u} \in U$ whenever $\mathbf{u} \in U$.

(d) Differentiation D on $F[x]$ (Example 2.1.9) lowers degrees, so for each $n \geq 0$ the subspace $F[x]_{\leq n}$ is D-invariant. Multiplication by x, $M(p) = xp$, raises degrees, and $F[x]_{\leq n}$ is not M-invariant, as $M(x^n) = x^{n+1}$. The subspace $\{p \mid p(0) = 0\}$ is M-invariant, since $(xp)(0) = 0$ for every p.

(e) The cyclic permutation $Z(x, y, z) = (z, x, y)$ of Example 4.5.37 leaves the line spanned by $(1, 1, 1)$ and the plane $x + y + z = 0$ invariant.

Non-example 7.2.3 (A line that moves). Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, the shear $(x, y) \mapsto (x + y, y)$ of F^2. Then $A\mathbf{e}_1 = \mathbf{e}_1$, so the line $\operatorname{span}(\mathbf{e}_1)$ is invariant. But $A\mathbf{e}_2 = (1, 1) \notin \operatorname{span}(\mathbf{e}_2)$, so the line $\operatorname{span}(\mathbf{e}_2)$ is not, although it is a complement of an invariant line. Invariance is a property of a subspace relative to T, and complements of invariant subspaces need not inherit it.

In the last two examples we checked invariance only on spanning vectors. That is always enough.

Proposition 7.2.4 (Invariance is checked on a spanning list). Let $T \in \mathrm{L}(V)$ and $U = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$. Then U is T-invariant if and only if $T\mathbf{u}_i \in U$ for $i = 1, \ldots, k$.

Proof. ROUTINE

Uses. Definition 7.2.1 (Invariant subspace), Definition 1.3.10 (Span), Definition 2.1.1 (Linear map), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. One direction is immediate. For the other, write an arbitrary $\mathbf{u} \in U$ as a combination of the $\mathbf{u}_i$.
Hint 2. Apply T to the combination and use that U is closed under linear combinations.
Sketch. If U is invariant, then in particular $T\mathbf{u}_i \in U$. Conversely, $T(\sum a_i\mathbf{u}_i) = \sum a_i T\mathbf{u}_i$ is a linear combination of vectors of the subspace U, so it lies in U.

If U is T-invariant, then $T\mathbf{u}_i \in U$ because $\mathbf{u}_i \in U$. Conversely, suppose $T\mathbf{u}_i \in U$ for each i, and let $\mathbf{u} \in U$. By Definition 1.3.10, $\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$ for some $a_i \in F$, and by linearity

$$T\mathbf{u} = a_1 T\mathbf{u}_1 + \cdots + a_k T\mathbf{u}_k.$$

This is a linear combination of vectors of the subspace U, so it lies in U (Proposition 1.4.14). □

Remark 7.2.5 (Invariant subspaces and block matrices). Suppose V is finite-dimensional. By Proposition 4.5.35, a subspace U is T-invariant exactly when, in a basis of V that starts with a basis of U, the matrix of T has the block form

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

and then A is the matrix of $T|_U$. If $V = U \oplus W$ with both U and W invariant, then $B = 0$ as well, and T splits into the two operators $T|_U$ and $T|_W$ (Technique 4.5.36). So invariant subspaces are what make matrices block triangular, and invariant complements are what make them block diagonal. The smallest nonzero blocks have size 1×1, and they come from invariant lines.

7.2.2 Eigenvalues and eigenvectors

Intuition (Invariant lines). A line $\operatorname{span}(\mathbf{v})$, with $\mathbf{v} \neq \mathbf{0}$, is T-invariant exactly when $T\mathbf{v} \in \operatorname{span}(\mathbf{v})$ (Proposition 7.2.4), that is, when $T\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ. On such a line T acts as multiplication by λ:

$T(c\mathbf{v}) = c\lambda\mathbf{v} = \lambda(c\mathbf{v})$. So the one-dimensional invariant subspaces are the directions that T stretches (by λ, possibly 0 or negative) without turning.

Definition 7.2.6 (Eigenvalue, eigenvector). Let $T \in \mathrm{L}(V)$. A scalar $\lambda \in F$ is an **eigenvalue** of T if there is a vector $\mathbf{v} \in V$ with

$$\mathbf{v} \neq \mathbf{0} \quad \text{and} \quad T\mathbf{v} = \lambda\mathbf{v}.$$

Such a vector $\mathbf{v}$ is an **eigenvector** of T with eigenvalue λ (or an eigenvector for λ). For $A \in F^{n\times n}$, the eigenvalues and eigenvectors of A are those of T_A: nonzero $\mathbf{v} \in F^n$ with $A\mathbf{v} = \lambda\mathbf{v}$.

Three remarks on the definition. First, eigenvectors are nonzero by definition: $T\mathbf{0} = \lambda\mathbf{0}$ holds for every λ, so allowing $\mathbf{0}$ would make every scalar an eigenvalue. The eigenvalue 0 is allowed, and it occurs exactly when $\ker T \neq \{\mathbf{0}\}$. Second, an eigenvector determines its eigenvalue: if $\lambda\mathbf{v} = \mu\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $(\lambda - \mu)\mathbf{v} = \mathbf{0}$, so $\lambda = \mu$ by Proposition 1.2.35. Third, eigenvalues are elements of F. The same matrix can have eigenvalues over $\mathbb{C}$ and none over $\mathbb{R}$ (Example 7.2.18), so the field is part of the question.

Proposition 7.2.7 (Invariant lines are spanned by eigenvectors). Let $T \in \mathrm{L}(V)$ and $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$. Then $\operatorname{span}(\mathbf{v})$ is T-invariant if and only if $\mathbf{v}$ is an eigenvector of T. In that case every nonzero vector of $\operatorname{span}(\mathbf{v})$ is an eigenvector of T with the same eigenvalue.

Proof. ROUTINE

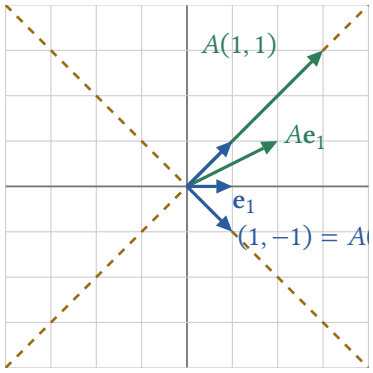
Uses. Proposition 7.2.4 (Invariance is checked on a spanning list), Definition 7.2.6 (Eigenvalue, eigenvector), Definition 1.3.10 (Span)
Hint 1. Apply Proposition 7.2.4 to the list consisting of $\mathbf{v}$ alone.
Hint 2. For the last claim, compute $T(c\mathbf{v})$.
Sketch. By Proposition 7.2.4, $\operatorname{span}(\mathbf{v})$ is invariant iff $T\mathbf{v} \in \operatorname{span}(\mathbf{v}) = \{c\mathbf{v} \mid c \in F\}$, which says $T\mathbf{v} = \lambda\mathbf{v}$ for some λ. Then $T(c\mathbf{v}) = \lambda(c\mathbf{v})$.

By Proposition 7.2.4, $\operatorname{span}(\mathbf{v})$ is T-invariant if and only if $T\mathbf{v} \in \operatorname{span}(\mathbf{v}) = \{c\mathbf{v} \mid c \in F\}$, that is, if and only if $T\mathbf{v} = \lambda\mathbf{v}$ for some $\lambda \in F$. Since $\mathbf{v} \neq \mathbf{0}$, this says that $\mathbf{v}$ is an eigenvector. If $T\mathbf{v} = \lambda\mathbf{v}$ and $c \neq 0$, then $c\mathbf{v} \neq \mathbf{0}$ and $T(c\mathbf{v}) = c\lambda\mathbf{v} = \lambda(c\mathbf{v})$. □

Example 7.2.8 (Eigenvectors of a symmetric 2×2 matrix). Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Then

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

So $(1, 1)$ is an eigenvector with eigenvalue 3, $(1, -1)$ is an eigenvector with eigenvalue 1, and $\mathbf{e}_1$ is not an eigenvector, since $(2, 1)$ is not a multiple of $(1, 0)$. The lines $y = x$ and $y = -x$ are invariant. In the basis $(1, 1), (1, -1)$ the matrix of T_A is $\operatorname{diag}(3, 1)$: A stretches the first diagonal by 3 and fixes the second pointwise. We will see in Corollary 7.2.22 that A has no other eigenvalues.



The map of $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. The dashed lines $y = x$ and $y = -x$ are invariant: $A(1,1) = (3,3)$, while $(1,-1)$ is fixed. The vector $\mathbf{e}_1$ is turned toward the diagonal. In the interactive version, drag $A\mathbf{e}_1$ and $A\mathbf{e}_2$ and look for lines through $\mathbf{0}$ that are mapped to themselves. The far corner of the image of the unit square is $A(1,1) = A\mathbf{e}_1 + A\mathbf{e}_2$, so the diagonal $y = x$ is invariant exactly when that corner stays on it, that is, when the two rows of A have the same sum. The x-axis is invariant exactly when $A\mathbf{e}_1$ lies on it. Try a rotation by an angle that is not a multiple of π, which leaves no line in place, and a shear, which keeps exactly one. (Interactive in the web version.)

To find all eigenvalues, rewrite $T\mathbf{v} = \lambda\mathbf{v}$ as $(T - \lambda\,\mathrm{id})\mathbf{v} = \mathbf{0}$, which turns the search for eigenvalues into a question about kernels.

Theorem 7.2.9 (Eigenvalues and invertibility). Let $T \in \mathrm{L}(V)$ and $\lambda \in F$, and consider the conditions

(I) λ is an eigenvalue of T;

(II) $\ker(T - \lambda\,\mathrm{id}) \neq \{\mathbf{0}\}$, that is, $T - \lambda\,\mathrm{id}$ is not injective;

(III) $T - \lambda\,\mathrm{id}$ is not surjective;

(IV) $T - \lambda\,\mathrm{id}$ is not invertible.

Then (I) and (II) are equivalent. If V is finite-dimensional, all four are equivalent. In particular, λ is an eigenvalue of $A \in F^{n\times n}$ if and only if $A - \lambda I_n$ is not invertible, if and only if $\operatorname{rank}(A - \lambda I_n) < n$, if and only if $\det(A - \lambda I_n) = 0$.

Proof. ROUTINE

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Definition 2.1.21 (Pointwise operations), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Corollary 4.2.18 (Injective if and only if surjective), Proposition 2.7.7 (Bijective linear maps are invertible), Corollary 3.5.25 (The invertible matrix theorem), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Rewrite $T\mathbf{v} = \lambda\mathbf{v}$ using the operator $T - \lambda\,\mathrm{id}$.
Hint 2. For (I) $\Leftrightarrow$ (II), use Theorem 2.6.11. In finite dimensions, Corollary 4.2.18 compares (II), (III), and (IV).
Hint 3. For matrices, $T_{A-\lambda I_n} = T_A - \lambda\,\mathrm{id}$. Translate "not injective" with Corollary 3.5.25 and Theorem 5.4.7.
Sketch. $(T - \lambda\,\mathrm{id})\mathbf{v} = T\mathbf{v} - \lambda\mathbf{v}$, so the eigenvectors for λ are exactly the nonzero vectors of $\ker(T - \lambda\,\mathrm{id})$, and such vectors exist iff $T - \lambda\,\mathrm{id}$ is not injective. If V is finite-dimensional, an operator on V is injective iff surjective iff bijective iff invertible. For A, apply this to $T_A - \lambda\,\mathrm{id} = T_{A-\lambda I_n}$ and use the invertible matrix theorem and det.

By Definition 2.1.21, $(T - \lambda\,\mathrm{id})\mathbf{v} = T\mathbf{v} - \lambda\mathbf{v}$, so $T\mathbf{v} = \lambda\mathbf{v}$ if and only if $\mathbf{v} \in \ker(T - \lambda\,\mathrm{id})$. Hence λ is an eigenvalue if and only if $\ker(T - \lambda\,\mathrm{id})$ contains a nonzero vector, which proves (I) $\Leftrightarrow$ (II). The two forms of (II) are equivalent by Theorem 2.6.11, applied to the linear map $T - \lambda\,\mathrm{id}$.

Now let V be finite-dimensional. By Corollary 4.2.18, the operator $T - \lambda\,\mathrm{id}$ on V is injective if and only if it is surjective, if and only if it is bijective, and by Proposition 2.7.7 a bijective linear map is invertible (conversely, an invertible map is bijective). Negating, (II), (III), and (IV) are equivalent.

For $A \in F^{n\times n}$, $(A - \lambda I_n)\mathbf{x} = A\mathbf{x} - \lambda\mathbf{x}$ for every $\mathbf{x}$, so $T_{A-\lambda I_n} = T_A - \lambda\,\mathrm{id}$. Thus λ is an eigenvalue of A if and only if $T_{A-\lambda I_n}$ is not injective. By Theorem 3.5.12 and Corollary 3.5.25, this happens if and only if $\operatorname{rank}(A - \lambda I_n) < n$, if and only if $A - \lambda I_n$ is not invertible, and by Theorem 5.4.7 if and only if $\det(A - \lambda I_n) = 0$. □

The determinant condition is the practical one: $\det(xI_n - A)$ is a polynomial in x whose roots are the eigenvalues (Section 7.3). Until then we find eigenvalues by solving $A\mathbf{v} = \lambda\mathbf{v}$ directly, as in the examples below.

Remark 7.2.10 (Infinite dimensions). Without finite dimension, (III) and (IV) are not equivalent to (I). Multiplication by x on $F[x]$, $M(p) = xp$, is injective but not surjective (the constant 1 is not of the form xp). So $M - 0\,\mathrm{id} = M$ is not invertible, yet 0 is not an eigenvalue of M, since $xp = \mathbf{0}$ forces $p = \mathbf{0}$. In fact M has no eigenvalues at all (Example 7.2.20).

Exercise 7.2.11. ROUTINE Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Decide which of the vectors

$$(1,1,0), \quad (1,-1,0), \quad (0,0,1), \quad (1,0,0), \quad (1,-1,1), \quad (1,1,1)$$

are eigenvectors of A, and give the eigenvalue of each one that is.

Solution on page 1564

Exercise 7.2.12. ROUTINE Let $T \in \mathrm{L}(V)$ be invertible. Show that 0 is not an eigenvalue of T, and that for $\lambda \neq 0$, a vector $\mathbf{v}$ is an eigenvector of T with eigenvalue λ if and only if it is an eigenvector of T^{-1} with eigenvalue λ^{-1}.

Solution on page 1564

The eigenvectors for a fixed λ, together with $\mathbf{0}$, form the kernel in Theorem 7.2.9(II). That kernel gets a name.

Definition 7.2.13 (Eigenspace). Let $T \in \mathrm{L}(V)$ and $\lambda \in F$. The **eigenspace** of T for λ is

$$E_\lambda = E_\lambda(T) = \ker(T - \lambda\,\mathrm{id}) = \{\mathbf{v} \in V \mid T\mathbf{v} = \lambda\mathbf{v}\}.$$

For $A \in F^{n\times n}$, $E_\lambda(A) = \ker(A - \lambda I_n)$ is the null space of $A - \lambda I_n$.

The eigenspace is defined for every $\lambda \in F$, and it is a subspace, being a kernel (Proposition 2.6.8). By Theorem 7.2.9, $E_\lambda \neq \{\mathbf{0}\}$ exactly when λ is an eigenvalue, and then E_λ consists of $\mathbf{0}$ and the eigenvectors for λ. Each eigenspace is T-invariant, since T acts on it as the scaling $\lambda\,\mathrm{id}$. The eigenvectors of T for all eigenvalues together, with $\mathbf{0}$ added, do not form a subspace in general, as Exercise 7.2.11 showed. Its dimension, $\dim E_\lambda$, is the geometric multiplicity of λ (Section 7.4).

Example 7.2.14 (Computing eigenspaces). Let $J \in \mathbb{R}^{3\times 3}$ be the matrix whose entries are all 1, and let

$$A = 2I_3 + J = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

Since $J(x,y,z) = (s,s,s)$ with $s = x + y + z$, we have $A\mathbf{v} = 2\mathbf{v} + s(1,1,1)$. So

$$E_2 = \ker J = \{(x,y,z) \mid x + y + z = 0\} = \operatorname{span}((1,-1,0),(0,1,-1)),$$

a plane, and $A(1,1,1) = (5,5,5)$, so 5 is an eigenvalue. To find all of E_5, row reduce $A - 5I_3$:

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

so $E_5 = \{(x,y,z) \mid x = z, y = z\} = \operatorname{span}((1,1,1))$. We will see right after Theorem 7.2.21 that 2 and 5 are the only eigenvalues of A.

7.2.3 A gallery of examples

The following operators will serve as examples and counterexamples for the rest of the chapter. In each case we solve $T\mathbf{v} = \lambda\mathbf{v}$ by hand.

Example 7.2.15 (Diagonal matrices). Let $D = \operatorname{diag}(d_1, \ldots, d_n)$. Then $D\mathbf{e}_i = d_i\mathbf{e}_i$, so each d_i is an eigenvalue, with eigenvector $\mathbf{e}_i$. There are no others: if $\lambda \neq d_i$ for every i, then $D - \lambda I_n = \operatorname{diag}(d_1 - \lambda, \ldots, d_n - \lambda)$ has inverse $\operatorname{diag}\big((d_1-\lambda)^{-1}, \ldots, (d_n-\lambda)^{-1}\big)$, so λ is not an eigenvalue by Theorem 7.2.9. The eigenspace E_λ is spanned by the $\mathbf{e}_i$ with $d_i = \lambda$, because $(D - \lambda I_n)\mathbf{x}$ has coordinates $(d_i - \lambda)x_i$, which all vanish exactly when $x_i = 0$ for every i with $d_i \neq \lambda$.

Example 7.2.16 (Projections). Let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$, for instance the projection Pr_ℓ of $\mathbb{R}^2$ onto a line ℓ (Example 2.1.11). If $P\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then

$$\lambda\mathbf{v} = P\mathbf{v} = P(P\mathbf{v}) = P(\lambda\mathbf{v}) = \lambda^2\mathbf{v},$$

so $\big(\lambda^2 - \lambda\big)\mathbf{v} = \mathbf{0}$ and $\lambda(\lambda - 1) = 0$. So the only possible eigenvalues are 0 and 1. Their eigenspaces are

$$E_1 = \operatorname{im} P, \quad E_0 = \ker P.$$

The second is the definition. For the first, if $\mathbf{v} = P\mathbf{w}$, then $P\mathbf{v} = P^2\mathbf{w} = P\mathbf{w} = \mathbf{v}$, and conversely $P\mathbf{v} = \mathbf{v}$ exhibits $\mathbf{v}$ as a value of P. For Pr_ℓ, $E_1 = \ell$ and E_0 is the line through $\mathbf{0}$ perpendicular to ℓ. Every vector is the sum of its components in these two eigenspaces, $\mathbf{v} = P\mathbf{v} + (\mathbf{v} - P\mathbf{v})$.

Example 7.2.17 (Reflections, and the field $\mathbb{F}_2$). Let $S \in \mathrm{L}(V)$ satisfy $S^2 = \mathrm{id}$, for instance the reflection S_ℓ of $\mathbb{R}^2$ across a line ℓ. If $S\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{v} = S^2\mathbf{v} = \lambda^2\mathbf{v}$, so $\lambda^2 = 1$, that is, $(\lambda - 1)(\lambda + 1) = 0$, and $\lambda = \pm 1$. For S_ℓ, $E_1 = \ell$ and E_{-1} is the perpendicular line.

Now take the swap $\sigma(x,y) = (y,x)$ on F^2, the reflection across the diagonal. Then $\sigma(1,1) = (1,1)$ and $\sigma(1,-1) = (-1,1) = -(1,-1)$. If $1 + 1 \neq 0$ in F, then $1 \neq -1$, and $(1,1)$, $(1,-1)$ are eigenvectors for the two different eigenvalues 1 and -1; they form a basis, in which σ has matrix $\operatorname{diag}(1,-1)$. Over $\mathbb{F}_2$, however, $-1 = 1$, and 1 is the only possible eigenvalue. Its eigenspace is $\ker(\sigma - \mathrm{id})$, and $(\sigma - \mathrm{id})(x,y) = (y - x, x - y)$, so $E_1 = \operatorname{span}((1,1))$ is a line. Over $\mathbb{F}_2$ the swap has no basis of eigenvectors, although it still satisfies $\sigma^2 = \mathrm{id}$.

Example 7.2.18 (Rotations). Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so that $A(x,y) = (-y,x)$; over $\mathbb{R}$ this is the rotation $R_{\pi/2}$ by a right angle. Suppose $A(x,y) = \lambda(x,y)$, that is, $-y = \lambda x$ and $x = \lambda y$. Substituting each equation into the other gives

$$x = \lambda y = -\lambda^2 x, \quad y = -\lambda x = -\lambda^2 y, \quad \text{so} \quad \big(1 + \lambda^2\big)x = \big(1 + \lambda^2\big)y = 0.$$

So if $\lambda^2 \neq -1$, then $(x,y) = \mathbf{0}$, and λ is not an eigenvalue. If $\lambda^2 = -1$, then $(1,-\lambda)$ is an eigenvector: $A(1,-\lambda) = (\lambda,1) = \lambda(1,-\lambda)$, since $-\lambda^2 = 1$. So the eigenvalues of A are exactly the square roots of -1 in F:

- Over $\mathbb{R}$: none, since $\lambda^2 \geq 0 > -1$. A rotation by a right angle turns every line, so its only invariant subspaces are $\{\mathbf{0}\}$ and $\mathbb{R}^2$ (by Proposition 7.2.7 and Proposition 4.1.21, an invariant subspace other than these would be an invariant line).
- Over $\mathbb{C}$: $\lambda = i$ with eigenvector $(1,-i)$, and $\lambda = -i$ with eigenvector $(1,i)$.
- Over $\mathbb{F}_5$: $2^2 = 4 = -1$ and $3^2 = 9 = -1$, so the eigenvalues are 2 and 3, with eigenvectors $(1,-2) = (1,3)$ and $(1,-3) = (1,2)$.
- Over $\mathbb{F}_3$: the squares are $0, 1, 1$, and $-1 = 2$ is not among them, so there are no eigenvalues.
- Over $\mathbb{F}_2$: $-1 = 1$, so $\lambda = 1$, and A is the swap of Example 7.2.17, with the single eigenvector line $\operatorname{span}((1,1))$.

This is the polynomial x^2+1 of Section 7.1.5 in disguise: it has no real root, and accordingly the rotation has no eigenvalue.

Example 7.2.19 (Differentiation). Let D be differentiation on $F[x]$, over any field. We claim that 0 is its only eigenvalue. Suppose $Dp = \lambda p$ with $p \neq \mathbf{0}$. If p is a nonzero constant, then $Dp = \mathbf{0}$, so $\lambda p = \mathbf{0}$ and $\lambda = 0$. If $\deg p \geq 1$, then Dp is either $\mathbf{0}$ or of degree less than $\deg p$ (Example 2.1.9), while λp has degree $\deg p$ unless $\lambda = 0$; so again $\lambda = 0$. And 0 is an eigenvalue, as $D1 = \mathbf{0}$. Its eigenspace $E_0 = \ker D$ is the constants over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$; over $\mathbb{F}_p$ it is larger, since $D(x^p) = px^{p-1} = \mathbf{0}$ (Example 4.1.28).

Contrast this with differentiation on the real vector space of infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$. There, for every $\lambda \in \mathbb{R}$, the function $f_\lambda(t) = e^{\lambda t}$ satisfies $f_\lambda{}' = \lambda f_\lambda$, so every real number is an eigenvalue. For $\lambda \neq 0$ these eigenvectors are not polynomials.

Example 7.2.20 (Shifts). Write $p \in F[x]$ by its coefficient sequence $(a_0, a_1, a_2, \ldots)$ (Definition 1.2.13). Multiplication by x, M, shifts the sequence to the right, and the map L of Exercise 2.5.41,

$$L(a_0, a_1, a_2, \ldots) = (a_1, a_2, a_3, \ldots),$$

shifts it to the left. M has no eigenvalues: if $xp = \lambda p$ with $p \neq \mathbf{0}$, the left side has degree $\deg p + 1$ and the right side has degree at most $\deg p$. L lowers degrees like D, and the argument of Example 7.2.19 shows that 0 is its only eigenvalue, with $E_0 = \ker L$ the constants.

On the space F^X of all sequences $(x_0, x_1, \ldots)$, where $X = \{0, 1, 2, \ldots\}$ (Remark 1.2.11), the left shift L given by the same formula behaves differently. $L\mathbf{x} = \lambda\mathbf{x}$ says $x_{n+1} = \lambda x_n$ for all n, that is, $x_n = \lambda^n x_0$. So every $\lambda \in F$ is an eigenvalue, with

$$E_\lambda = \operatorname{span}\big((1, \lambda, \lambda^2, \lambda^3, \ldots)\big),$$

a line of geometric sequences. Only for $\lambda = 0$ is the eigenvector a polynomial.

Finally, on F^n the shift $N(x_1, \ldots, x_n) = (x_2, \ldots, x_n, 0)$ of Section 4.1.7 satisfies $N^n = 0$. If $N\mathbf{v} = \lambda\mathbf{v}$, then applying N repeatedly gives $\mathbf{0} = N^n\mathbf{v} = \lambda^n\mathbf{v}$, so $\lambda = 0$; and $E_0 = \ker N = \operatorname{span}(\mathbf{e}_1)$.

So an operator may have no eigenvalues (a rotation of $\mathbb{R}^2$, multiplication by x), a few, or every scalar (the left shift on sequences). In finite dimensions there are at most $\dim V$ eigenvalues, as we now show; over an infinite field, this rules out the last behavior.

7.2.4 Eigenvectors for distinct eigenvalues

In every example so far, eigenvectors for different eigenvalues pointed in different directions. They are in fact always linearly independent. The reason is that T treats them differently, and that difference can be used to separate them.

Intuition (Separating eigenvectors). Let $T\mathbf{v}_l = \lambda_l\mathbf{v}_l$. The operator $T - \mu\,\mathrm{id}$ multiplies each $\mathbf{v}_l$ by $\lambda_l - \mu$:

$$(T - \mu\,\mathrm{id})\mathbf{v}_l = T\mathbf{v}_l - \mu\mathbf{v}_l = (\lambda_l - \mu)\mathbf{v}_l.$$

Choosing $\mu = \lambda_j$ kills $\mathbf{v}_j$ and only rescales the others, by nonzero factors when the eigenvalues are distinct. So, given a linear relation among eigenvectors, we can apply such operators to destroy chosen terms of the relation while keeping control of the rest. There are two ways to organize this. We can kill all terms but one at once, using a product of such operators, or we can kill one term at a time and argue that a shortest relation cannot exist.

Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent). Let $T \in \mathrm{L}(V)$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be eigenvectors of T with eigenvalues $\lambda_1, \ldots, \lambda_m$, which are pairwise distinct. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is linearly independent.

Proof (kill all terms but one). STANDARD

> *Uses.* Definition 7.2.6 (Eigenvalue, eigenvector), Definition 2.1.21 (Pointwise operations), Proposition 1.1.17 (No zero divisors), Proposition 1.2.35 (Zero products), Theorem 1.5.5 (Independence is decided at zero)
> *Techniques.* Apply an operator to a relation to kill terms (Technique 4.1.25)
> *Hint 1.* Take a relation $a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$ and fix an index i. Find an operator that sends $\mathbf{v}_l$ to $\mathbf{0}$ for every $l \neq i$, but not $\mathbf{v}_i$.
> *Hint 2.* Try the composite P_i of the operators $T - \lambda_j \operatorname{id}$ for all $j \neq i$. Each factor multiplies $\mathbf{v}_l$ by $\lambda_l - \lambda_j$.
> *Hint 3.* So $P_i\mathbf{v}_l = \prod_{j\neq i}(\lambda_l - \lambda_j)\mathbf{v}_l$. For $l \neq i$ the product contains the factor $\lambda_l - \lambda_l = 0$; for $l = i$ every factor is nonzero.
> *Sketch.* Let P_i be the composite of the $T - \lambda_j \operatorname{id}$, $j \neq i$. Since $(T - \lambda_j \operatorname{id})\mathbf{v}_l = (\lambda_l - \lambda_j)\mathbf{v}_l$, applying the factors one at a time gives $P_i\mathbf{v}_l = \prod_{j\neq i}(\lambda_l - \lambda_j)\mathbf{v}_l$, which is $\mathbf{0}$ for $l \neq i$ and a nonzero multiple of $\mathbf{v}_i$ for $l = i$. Applying P_i to a relation $\sum a_l\mathbf{v}_l = \mathbf{0}$ leaves $a_ic_i\mathbf{v}_i = \mathbf{0}$ with $c_i \neq 0$, so $a_i = 0$.

First, for every l and every $\mu \in F$,

$$(T - \mu \operatorname{id})\mathbf{v}_l = T\mathbf{v}_l - \mu\mathbf{v}_l = (\lambda_l - \mu)\mathbf{v}_l. \tag{7.2.1}$$

Suppose $a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$, and fix $i \in \{1, \ldots, m\}$. Let P_i be the composite, in increasing order of j, of the operators $T - \lambda_j \operatorname{id}$ with $j \neq i$. It is linear, as a composite of linear maps. Applying its factors to $\mathbf{v}_l$ one at a time, (7.2.1) shows that each factor multiplies $\mathbf{v}_l$ by a scalar, so

$$P_i\mathbf{v}_l = c_{il}\mathbf{v}_l, \quad \text{where} \quad c_{il} = \prod_{j\neq i}(\lambda_l - \lambda_j).$$

If $l \neq i$, the product c_{il} contains the factor $\lambda_l - \lambda_l = 0$ (the factor with $j = l$), so $c_{il} = 0$. If $l = i$, every factor $\lambda_i - \lambda_j$ is nonzero because the eigenvalues are distinct, so $c_{ii} \neq 0$ by Proposition 1.1.17. Applying P_i to the relation therefore gives

$$\mathbf{0} = P_i\left(\sum_{l=1}^{m} a_l\mathbf{v}_l\right) = \sum_{l=1}^{m} a_lc_{il}\mathbf{v}_l = a_ic_{ii}\mathbf{v}_i.$$

Since $\mathbf{v}_i \neq \mathbf{0}$, Proposition 1.2.35 gives $a_ic_{ii} = 0$, and since $c_{ii} \neq 0$, $a_i = 0$. This holds for every i, so the list is independent by Theorem 1.5.5. □

Proof (a shortest relation). STANDARD

> *Uses.* Definition 7.2.6 (Eigenvalue, eigenvector), Definition 1.5.3 (Linear relation), Proposition 1.2.35 (Zero products), Theorem 1.5.5 (Independence is decided at zero)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* Argue by contradiction. Among all nontrivial relations $\sum a_l\mathbf{v}_l = \mathbf{0}$, choose one with as few nonzero coefficients as possible.
> *Hint 2.* Pick k with $a_k \neq 0$ and apply $T - \lambda_k \operatorname{id}$ to the relation. Which coefficients of the new relation are nonzero?
> *Hint 3.* The new relation has coefficients $a_l(\lambda_l - \lambda_k)$: zero at $l = k$, and nonzero exactly where $a_l \neq 0$ for $l \neq k$. It is shorter, so it must be trivial.
> *Sketch.* Take a nontrivial relation with the fewest nonzero coefficients, and k with $a_k \neq 0$. Applying $T - \lambda_k \operatorname{id}$ gives the relation $\sum a_l(\lambda_l - \lambda_k)\mathbf{v}_l = \mathbf{0}$, which has strictly fewer nonzero coefficients, hence is trivial. So $a_l = 0$ for all $l \neq k$, and the original relation reads $a_k\mathbf{v}_k = \mathbf{0}$ with $a_k \neq 0$, contradicting $\mathbf{v}_k \neq \mathbf{0}$.

Suppose, for contradiction, that the list is linearly dependent. Then it has nontrivial linear relations (Theorem 1.5.5). Choose one, $a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$, in which the number of nonzero coefficients is as small as possible, and choose k with $a_k \neq 0$. Applying $T - \lambda_k \operatorname{id}$ and using (7.2.1),

$$\sum_{l=1}^{m} a_l(\lambda_l - \lambda_k)\mathbf{v}_l = \mathbf{0}.$$

In this relation the coefficient of $\mathbf{v}_k$ is 0, and for $l \neq k$ the coefficient $a_l(\lambda_l - \lambda_k)$ is nonzero exactly when $a_l \neq 0$, since $\lambda_l \neq \lambda_k$. So it has exactly one nonzero coefficient fewer than the chosen relation. By minimality it is not nontrivial: all its coefficients are 0, so $a_l = 0$ for every $l \neq k$. The chosen relation then reads $a_k\mathbf{v}_k = \mathbf{0}$ with $a_k \neq 0$, so $\mathbf{v}_k = \mathbf{0}$ by Proposition 1.2.35. This contradicts that $\mathbf{v}_k$ is an eigenvector. □

The first proof is constructive: the operator P_i isolates the i-th term of any combination of the $\mathbf{v}_l$, and after dividing by c_{ii} it extracts the coefficient a_i. In the language of Section 7.2.5 below, $c_{ii}^{-1}P_i$ is $L_i(T)$, where L_i is the Lagrange polynomial (Definition 3.4.26) for the nodes $\lambda_1, \dots, \lambda_m$ with $L_i(\lambda_l) = 0$ for $l \neq i$ and $L_i(\lambda_i) = 1$. The second proof needs only one operator at a time, and it is the argument the lecture gives as an induction on m: assuming the result for $m-1$ eigenvectors, apply $T - \lambda_m \mathrm{id}$. Choosing a shortest relation does that induction in one step.

Corollary 7.2.22 (At most $\dim V$ eigenvalues). If V is finite-dimensional, an operator $T \in \mathrm{L}(V)$ has at most $\dim V$ distinct eigenvalues. In particular, $A \in F^{n\times n}$ has at most n eigenvalues.

Proof. ROUTINE

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 4.1.10 (Independent lists are short)
Hint 1. Choose one eigenvector for each of m distinct eigenvalues.
Hint 2. Compare the length of that list with $\dim V$.
Sketch. Eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_m$ for m distinct eigenvalues form an independent list by Theorem 7.2.21, so $m \leq \dim V$.

Let $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of T, and choose an eigenvector $\mathbf{v}_i$ for each λ_i. By Theorem 7.2.21 the list $\mathbf{v}_1, \dots, \mathbf{v}_m$ is independent, so $m \leq \dim V$ by Theorem 4.1.10(a). The statement for A is the case $T = T_A$, $V = F^n$. □

So the symmetric matrix of Example 7.2.8 has no eigenvalues besides 3 and 1. The theorem also says something about whole eigenspaces.

Corollary 7.2.23 (Eigenspaces are independent). Let $\lambda_1, \dots, \lambda_m \in F$ be distinct. Then the sum $E_{\lambda_1} + \cdots + E_{\lambda_m}$ of the eigenspaces of T is direct. If V is finite-dimensional, then

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_m} \leq \dim V.$$

Proof. STANDARD

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 4.4.7 (Criteria for direct sums), Definition 4.4.4 (Direct sum), Theorem 4.1.10 (Independent lists are short), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. By Theorem 4.4.7, it suffices to show: if $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$ with $\mathbf{u}_i \in E_{\lambda_i}$, then every $\mathbf{u}_i = \mathbf{0}$. Which of the $\mathbf{u}_i$ are eigenvectors?
Hint 2. For the dimension bound, concatenate bases of the eigenspaces and show that the long list is independent.
Sketch. The nonzero $\mathbf{u}_i$ in $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$ are eigenvectors for distinct eigenvalues, and the equation is a relation among them with all coefficients 1, so by Theorem 7.2.21 there are none. For the bound, a relation among the concatenated bases groups into $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$ with $\mathbf{u}_i \in E_{\lambda_i}$, so each $\mathbf{u}_i = \mathbf{0}$, and independence of each basis kills the coefficients.

Suppose $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$ with $\mathbf{u}_i \in E_{\lambda_i}$, and let $I = \{i \mid \mathbf{u}_i \neq \mathbf{0}\}$. For $i \in I$, $\mathbf{u}_i$ is an eigenvector with eigenvalue λ_i, and these eigenvalues are distinct. If I were nonempty, $\sum_{i\in I} 1 \cdot \mathbf{u}_i = \mathbf{0}$ would be a nontrivial relation among them, contradicting Theorem 7.2.21. So I is empty, every $\mathbf{u}_i$ is $\mathbf{0}$, and the sum is direct by the implication (II) ⇒ (I) of Theorem 4.4.7.

Now let V be finite-dimensional, let $d_i = \dim E_{\lambda_i}$, and choose a basis $\mathbf{b}_{i1}, \dots, \mathbf{b}_{id_i}$ of each E_{λ_i}. Suppose $\sum_{i=1}^{m}\sum_{j=1}^{d_i} a_{ij}\mathbf{b}_{ij} = \mathbf{0}$. With $\mathbf{u}_i = \sum_j a_{ij}\mathbf{b}_{ij} \in E_{\lambda_i}$ this reads $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$, so every $\mathbf{u}_i = \mathbf{0}$ by the first part, and then every $a_{ij} = 0$ because the $\mathbf{b}_{ij}$ with fixed i are independent. So the concatenated list, of length $d_1 + \cdots + d_m$, is independent, and $d_1 + \cdots + d_m \leq \dim V$ by Theorem 4.1.10(a). □

For the matrix $A = 2I_3 + J$ of Example 7.2.14, $\dim E_2 + \dim E_5 = 2 + 1 = 3$. If A had a third eigenvalue λ, its eigenspace would add at least 1 to this sum, which would then exceed $\dim \mathbb{R}^3 = 3$. So 2 and 5 are the only eigenvalues. When the dimensions of the eigenspaces add up to $\dim V$, concatenating their bases gives a basis of eigenvectors; this is the starting point of Section 7.4.

Example 7.2.24 (Independence for free). Two independence results from earlier chapters are instances of the theorem.

(a) The geometric sequences $\left(1, \lambda_i, \lambda_i^2, \ldots\right)$ for distinct $\lambda_1, \ldots, \lambda_k \in F$ are eigenvectors of the left shift on F^X with the distinct eigenvalues λ_i (Example 7.2.20). So they are independent, which is Exercise 1.5.44.

(b) The functions $t \mapsto e^{\lambda_i t}$ for distinct real λ_i are eigenvectors of differentiation with distinct eigenvalues (Example 7.2.19), so they are independent, which is Example 1.5.27 and its generalization.

Exercise 7.2.25. ROUTINE Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be eigenvectors of T with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, where $m \geq 2$. Show that $\mathbf{v}_1 + \cdots + \mathbf{v}_m$ is not an eigenvector of T.

Solution on page 1565

7.2.5 Polynomials in an operator

The proofs above composed operators of the form $T - \lambda\,\mathrm{id}$. Such composites are polynomials in T, and polynomials in T are the natural source of operators that interact well with T.

Definition 7.2.26 (Polynomials in an operator). Let $T \in \mathrm{L}(V)$ and $p = a_0 + a_1 x + \cdots + a_d x^d \in F[x]$. Then

$$p(T) = a_0 \mathrm{id}_V + a_1 T + a_2 T^2 + \cdots + a_d T^d \in \mathrm{L}(V),$$

where T^k is the k-fold composite (Notation 2.1.18) and $T^0 = \mathrm{id}_V$. For a matrix $A \in F^{n\times n}$, $p(A) = a_0 I_n + a_1 A + \cdots + a_d A^d$, as in Remark 7.1.11.

For example, if $p = x^2 - 3x + 2 = (x-1)(x-2)$, then $p(T) = T^2 - 3T + 2\,\mathrm{id}$, and the next proposition shows that this equals $(T - \mathrm{id})(T - 2\,\mathrm{id})$, and also $(T - 2\,\mathrm{id})(T - \mathrm{id})$.

Proposition 7.2.27 (Evaluating polynomials at an operator). Let $T \in \mathrm{L}(V)$, $p, q \in F[x]$, and $c \in F$.

(a) $(p+q)(T) = p(T) + q(T)$, $(cp)(T) = cp(T)$, and $(pq)(T) = p(T)q(T)$, where the product of operators is composition.

(b) In particular $p(T)q(T) = q(T)p(T)$, and $p(T)$ commutes with T.

(c) If V is finite-dimensional and $\mathcal{B}$ is a basis of V, then $[p(T)]_{\mathcal{B}\mathcal{B}} = p([T]_{\mathcal{B}\mathcal{B}})$.

(d) Every T-invariant subspace is $p(T)$-invariant.

Proof. ROUTINE

Uses. Proposition 7.1.10 (Evaluation respects sums and products), Remark 7.1.11 (Polynomials in a matrix), Proposition 2.1.24 (Algebra of composition), Notation 2.1.18, Theorem 2.4.2 (The matrix of a composite), Proposition 7.2.4 (Invariance is checked on a spanning list)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. For (a), copy the proof of Proposition 7.1.10. Which facts about multiplication did it use, and do they hold for composition of operators?
Hint 2. It used $t^i t^j = t^{i+j}$, distributivity, and that scalars commute with powers of t. For operators these are Notation 2.1.18 and Proposition 2.1.24. Then (b) is (a) with $pq = qp$.
Hint 3. For (d), show by induction that $T^k \mathbf{u} \in U$ for $\mathbf{u} \in U$, then take a combination.
Sketch. The proof of Proposition 7.1.10 used only $t^i t^j = t^{i+j}$, that multiplication distributes over sums, and that scalars commute with powers of t; all three hold for composition of operators (Proposition 2.1.24), so (a) follows, as for matrices in Remark 7.1.11. (b) is $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$ with $q = x$. (c): $T \mapsto [T]_{\mathcal{B}\mathcal{B}}$ respects sums, scalar multiples, and composites. (d): T^k maps U into U by induction, and $p(T)\mathbf{u}$ is a combination of the $T^k\mathbf{u}$.

(a) The proof of Proposition 7.1.10, with $t = T$, used only three facts: $t^i t^j = t^{i+j}$; multiplication distributes over sums on both sides; and scalars can be moved past powers of t, $\left(at^i\right)\left(bt^j\right) = (ab)t^{i+j}$. For composition of operators, the first is Notation 2.1.18 and the other two are Proposition 2.1.24. Commutativity of t with

itself was never needed beyond $t^i t^j = t^{i+j}$. So the same computation proves (a), exactly as for matrices in Remark 7.1.11.

(b) Since $F[x]$ is commutative, $pq = qp$, so by (a), $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$. Taking $q = x$, for which $q(T) = T$, shows that $p(T)T = Tp(T)$.

(c) Write $A = [T]_{\mathcal{BB}}$. By Theorem 2.4.2 and induction on k, $\left[T^k\right]_{\mathcal{BB}} = A^k$, and $[\mathrm{id}_V]_{\mathcal{BB}} = I_n$ with $n = \dim V$. The map $S \mapsto [S]_{\mathcal{BB}}$ respects sums and scalar multiples, so $\left[\sum_k a_k T^k\right]_{\mathcal{BB}} = \sum_k a_k A^k = p(A)$.

(d) Let U be T-invariant and $\mathbf{u} \in U$. Then $T^0\mathbf{u} = \mathbf{u} \in U$, and if $T^k\mathbf{u} \in U$, then $T^{k+1}\mathbf{u} = T\left(T^k\mathbf{u}\right) \in U$. So every $T^k\mathbf{u}$ lies in U, and so does the linear combination $p(T)\mathbf{u} = \sum_k a_k T^k\mathbf{u}$. □

Proposition 7.2.28 (Polynomials act on eigenvectors by their values). Let $T \in \mathrm{L}(V)$, and let $\mathbf{v}$ be an eigenvector of T with eigenvalue λ. Then for every $p \in F[x]$,

$$p(T)\mathbf{v} = p(\lambda)\mathbf{v}.$$

Consequently, if $p(T) = 0$, then every eigenvalue of T is a root of p.

Proof. ROUTINE

Uses. Definition 7.2.26 (Polynomials in an operator), Definition 1.2.17 (Evaluation), Proposition 1.2.35 (Zero products)
Hint 1. First show $T^k\mathbf{v} = \lambda^k\mathbf{v}$ by induction on k.
Hint 2. For the consequence, apply $p(T) = 0$ to an eigenvector.
Sketch. By induction, $T^k\mathbf{v} = \lambda^k\mathbf{v}$, so $p(T)\mathbf{v} = \sum a_k\lambda^k\mathbf{v} = p(\lambda)\mathbf{v}$. If $p(T) = 0$ and $\mathbf{v}$ is an eigenvector for λ, then $p(\lambda)\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$, so $p(\lambda) = 0$.

We have $T^0\mathbf{v} = \mathbf{v} = \lambda^0\mathbf{v}$, and if $T^k\mathbf{v} = \lambda^k\mathbf{v}$, then $T^{k+1}\mathbf{v} = T\left(\lambda^k\mathbf{v}\right) = \lambda^k T\mathbf{v} = \lambda^{k+1}\mathbf{v}$. So for $p = \sum_k a_k x^k$,

$$p(T)\mathbf{v} = \sum_k a_k T^k\mathbf{v} = \sum_k a_k\lambda^k\mathbf{v} = p(\lambda)\mathbf{v}.$$

If $p(T) = 0$ and λ is an eigenvalue with eigenvector $\mathbf{v}$, then $\mathbf{0} = p(T)\mathbf{v} = p(\lambda)\mathbf{v}$, so $p(\lambda) = 0$ by Proposition 1.2.35. □

This explains the computations in Example 7.2.16, Example 7.2.17, and Example 7.2.20 at once. A projection satisfies $p(P) = 0$ for $p = x^2 - x$, so its eigenvalues lie among the roots $0, 1$. An involution satisfies $x^2 - 1$, with roots ± 1. A nilpotent operator satisfies x^m, whose only root is 0. The cyclic permutation Z of Example 7.2.2 satisfies $Z^3 = \mathrm{id}$, so its eigenvalues are cube roots of 1: over $\mathbb{R}$ the only candidate is 1, and over $\mathbb{C}$ the candidates are $1, \omega, \omega^2$, where $\omega = e^{2\pi i/3}$. The converse fails: the identity satisfies $x^2 - x$, but 0 is not one of its eigenvalues. A polynomial with $p(T) = 0$ only bounds the eigenvalues; Section 7.6 finds the polynomial that pins them down exactly.

Exercise 7.2.29. ROUTINE Let $T \in \mathrm{L}(V)$ satisfy $T^2 = 5T - 6\,\mathrm{id}$.

(a) Show that every eigenvalue of T is 2 or 3.

(b) Assume $1 + 1 \neq 0$ in F. Show that $T - 4\,\mathrm{id}$ is invertible, and write its inverse as a polynomial in T.

Solution on page 1565

The commutativity in Proposition 7.2.27(b) has a geometric consequence, which is the main source of invariant subspaces in this book.

Proposition 7.2.30 (Kernels and images of commuting operators are invariant). Let $S, T \in \mathrm{L}(V)$ with $ST = TS$. Then $\ker S$ and $\operatorname{im} S$ are T-invariant. In particular:

(a) for every $p \in F[x]$, $\ker p(T)$ and $\operatorname{im} p(T)$ are T-invariant;

(b) every eigenspace $E_\lambda(S)$ of S is T-invariant.

Proof. ROUTINE

Uses. Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 7.2.13 (Eigenspace), Proposition 2.1.24 (Algebra of composition)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Take $\mathbf{v} \in \ker S$ and compute $S(T\mathbf{v})$, moving T to the left with $ST = TS$.
Hint 2. For $\operatorname{im} S$, write the vector as $S\mathbf{w}$ and rewrite $TS\mathbf{w}$ as a value of S.
Hint 3. For (a) and (b), find an operator that commutes with T and has the given kernel or image.
Sketch. If $S\mathbf{v} = \mathbf{0}$, then $ST\mathbf{v} = TS\mathbf{v} = \mathbf{0}$. And $T(S\mathbf{w}) = S(T\mathbf{w}) \in \operatorname{im} S$. Then (a) is the case $S = p(T)$, which commutes with T by Proposition 7.2.27(b), and (b) is the case $S - \lambda \operatorname{id}$, which commutes with T because S and id do.

Let $\mathbf{v} \in \ker S$. Then $S(T\mathbf{v}) = T(S\mathbf{v}) = T\mathbf{0} = \mathbf{0}$, so $T\mathbf{v} \in \ker S$. Let $\mathbf{u} \in \operatorname{im} S$, say $\mathbf{u} = S\mathbf{w}$. Then $T\mathbf{u} = TS\mathbf{w} = S(T\mathbf{w}) \in \operatorname{im} S$. (This is part (a) of Exercise 2.6.52.)

(a) $p(T)$ commutes with T by Proposition 7.2.27(b), so the first part applies with $S = p(T)$.

(b) By Proposition 2.1.24, $(S - \lambda \operatorname{id})T = ST - \lambda T = TS - \lambda T = T(S - \lambda \operatorname{id})$. Apply the first part to $S - \lambda \operatorname{id}$, whose kernel is $E_\lambda(S)$. □

Technique 7.2.31 (Look for invariant subspaces among kernels and images of commuting operators). To find subspaces that T maps into themselves, look for operators S with $ST = TS$: the kernel and the image of each such S, and each eigenspace of S, are T-invariant. Polynomials in T always commute with T (Proposition 7.2.27), so $\ker(T - \lambda \operatorname{id})^k$, $\operatorname{im}(T - \lambda \operatorname{id})^k$, and $\ker p(T)$ are always available. Conversely, if the operators at hand commute, an eigenspace of one is a space on which the others act, and the problem can often be reduced to that smaller space.

Reach for it when a problem involves two operators that commute, or when you need an invariant subspace of T and have a polynomial identity for T. The canonical example is Proposition 7.2.30. Generalized eigenspaces $\ker(T - \lambda \operatorname{id})^k$ (Section 10.1) and common eigenvectors of commuting operators (Section 7.6) rest on the same observation.
Used in. Exercise 7.2.47, Exercise 7.2.48, Exercise 7.2.55, Exercise 7.4.50, Exercise 7.4.51, Proposition 7.6.12, Theorem 7.6.28, Theorem 7.6.31, Theorem 7.6.34, Exercise 7.7.6, Exercise 7.7.12, Exercise 7.7.13, Problem 7.4, Exercise 8.2.36, Exercise 7.4.36.

Example 7.2.32 (What commutes with a diagonal matrix). Let $D = \operatorname{diag}(d_1, \ldots, d_n)$ with distinct d_i, and suppose $SD = DS$ for some $S \in F^{n \times n}$. Each eigenspace $E_{d_i}(D) = \operatorname{span}(\mathbf{e}_i)$ (Example 7.2.15) is S-invariant by Proposition 7.2.30(b), so $\mathbf{e}_i$ is an eigenvector of S by Proposition 7.2.7: $S\mathbf{e}_i = s_i\mathbf{e}_i$. That is, the i-th column of S is $s_i\mathbf{e}_i$, and S is diagonal. This recovers Exercise 2.4.42 without computing a single entry of $SD - DS$. If two of the d_i coincide, the eigenspace is a plane, and S only has to preserve the plane, so non-diagonal S appear.

7.2.6 Krylov subspaces and chains

Given a single vector $\mathbf{v}$, what is the smallest T-invariant subspace containing it? It must contain $T\mathbf{v}$, then $T^2\mathbf{v}$, and so on, and all their linear combinations. These combinations are exactly the vectors $p(T)\mathbf{v}$.

Definition 7.2.33 (Krylov subspace). Let $T \in \mathrm{L}(V)$ and $\mathbf{v} \in V$. The **Krylov subspace** (or **cyclic subspace**) of $\mathbf{v}$ under T is

$$K_T(\mathbf{v}) = \{p(T)\mathbf{v} \mid p \in F[x]\},$$

the set of all finite linear combinations of $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$.

Proposition 7.2.34 (The smallest invariant subspace containing a vector). Let $T \in \mathrm{L}(V)$ and $\mathbf{v} \in V$. Then $K_T(\mathbf{v})$ is a T-invariant subspace containing $\mathbf{v}$, and it is contained in every T-invariant subspace that contains $\mathbf{v}$.

Proof. ROUTINE

Uses. Definition 7.2.33 (Krylov subspace), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 1.4.1 (Subspace), Definition 7.2.1 (Invariant subspace)
Hint 1. Use Proposition 7.2.27(a): $p(T)\mathbf{v} + cq(T)\mathbf{v} = (p+cq)(T)\mathbf{v}$ and $Tp(T)\mathbf{v} = (xp)(T)\mathbf{v}$.
Hint 2. For minimality, take an invariant U containing $\mathbf{v}$, and use Proposition 7.2.27(d).
Sketch. $K_T(\mathbf{v})$ contains $\mathbf{0} = \mathbf{0}(T)\mathbf{v}$ and $\mathbf{v} = 1(T)\mathbf{v}$, is closed under combinations by $(p+cq)(T) = p(T) + cq(T)$, and is invariant because $Tp(T)\mathbf{v} = (xp)(T)\mathbf{v}$. If U is invariant and $\mathbf{v} \in U$, then $p(T)\mathbf{v} \in U$ by Proposition 7.2.27(d).

The constant polynomials $\mathbf{0}$ and 1 give $\mathbf{0} = \mathbf{0}(T)\mathbf{v}$ and $\mathbf{v} = \mathrm{id}\,\mathbf{v} = 1(T)\mathbf{v}$, so $\mathbf{0}, \mathbf{v} \in K_T(\mathbf{v})$. For $p, q \in F[x]$ and $c \in F$, Proposition 7.2.27(a) gives

$$p(T)\mathbf{v} + cq(T)\mathbf{v} = (p+cq)(T)\mathbf{v} \in K_T(\mathbf{v}), \quad T(p(T)\mathbf{v}) = (xp)(T)\mathbf{v} \in K_T(\mathbf{v}).$$

So $K_T(\mathbf{v})$ is a subspace (Definition 1.4.1), and it is T-invariant. If U is a T-invariant subspace with $\mathbf{v} \in U$, then U is $p(T)$-invariant for every p by Proposition 7.2.27(d), so $p(T)\mathbf{v} \in U$, and $K_T(\mathbf{v}) \subseteq U$. □

In particular, $K_T(\mathbf{v})$ is a line exactly when $\mathbf{v}$ is an eigenvector, and $\{\mathbf{0}\}$ exactly when $\mathbf{v} = \mathbf{0}$. In finite dimensions the sequence $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ must eventually become dependent, and the first dependence tells us everything about $K_T(\mathbf{v})$.

Proposition 7.2.35 (A basis of a Krylov subspace). Let $T \in \mathrm{L}(V)$ and $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$, and suppose $K_T(\mathbf{v})$ is finite-dimensional (for instance, V is). Then there is a smallest integer $d \geq 1$ with

$$T^d\mathbf{v} \in \operatorname{span}\left(\mathbf{v}, T\mathbf{v}, \ldots, T^{d-1}\mathbf{v}\right),$$

and for this d:

(a) $\mathbf{v}, T\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$ is a basis of $K_T(\mathbf{v})$, so $\dim K_T(\mathbf{v}) = d$;

(b) writing $T^d\mathbf{v} = -\left(c_0\mathbf{v} + c_1T\mathbf{v} + \cdots + c_{d-1}T^{d-1}\mathbf{v}\right)$, the monic polynomial $p = x^d + c_{d-1}x^{d-1} + \cdots + c_0$ satisfies $p(T)\mathbf{v} = \mathbf{0}$, and no nonzero polynomial of degree less than d sends $\mathbf{v}$ to $\mathbf{0}$.

Proof. STANDARD

Uses. Lemma 1.5.10 (Linear dependence lemma), Theorem 4.1.10 (Independent lists are short), Proposition 7.2.4 (Invariance is checked on a spanning list), Proposition 7.2.34 (The smallest invariant subspace containing a vector), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Too many vectors force a dependence (Technique 4.1.30), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For existence of d: if $\dim K_T(\mathbf{v}) = m$, the $m+1$ vectors $\mathbf{v}, \ldots, T^m\mathbf{v}$ lie in $K_T(\mathbf{v})$. Apply the linear dependence lemma.
Hint 2. Independence of $\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$: if they were dependent, the linear dependence lemma would produce a smaller d.
Hint 3. Spanning: show that $W = \operatorname{span}\left(\mathbf{v}, \ldots, T^{d-1}\mathbf{v}\right)$ is T-invariant by checking the spanning vectors, then use Proposition 7.2.34.
Sketch. The $m+1$ vectors $\mathbf{v}, \ldots, T^m\mathbf{v}$ in the m-dimensional $K_T(\mathbf{v})$ are dependent, and the linear dependence lemma gives some $T^j\mathbf{v}$ in the span of the earlier ones, with $j \geq 1$ as $\mathbf{v} \neq \mathbf{0}$; so d exists. The same lemma shows that $\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$ is independent, by minimality of d. Their span W is invariant, since T sends each $T^i\mathbf{v}$ to the next one and $T^d\mathbf{v} \in W$; so $K_T(\mathbf{v}) \subseteq W \subseteq K_T(\mathbf{v})$. In (b), a polynomial of degree less than d killing $\mathbf{v}$ is a relation among the basis vectors.

Let $m = \dim K_T(\mathbf{v})$. The vectors $\mathbf{v}, T\mathbf{v}, \ldots, T^m\mathbf{v}$ lie in $K_T(\mathbf{v})$, and there are $m+1$ of them, so they are dependent (Theorem 4.1.10(a)). By Lemma 1.5.10 some $T^j\mathbf{v}$, $0 \leq j \leq m$, lies in the span of $\mathbf{v}, \ldots, T^{j-1}\mathbf{v}$. Here $j \neq 0$, because the span of the empty list is $\{\mathbf{0}\}$ and $\mathbf{v} \neq \mathbf{0}$. So the set of integers $d \geq 1$ with $T^d\mathbf{v} \in \operatorname{span}\left(\mathbf{v}, \ldots, T^{d-1}\mathbf{v}\right)$ is nonempty, and it has a smallest element d.

(a) If $\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$ were dependent, Lemma 1.5.10 would give $j \leq d-1$ with $T^j\mathbf{v} \in \operatorname{span}\left(\mathbf{v}, \ldots, T^{j-1}\mathbf{v}\right)$, and as before $j \geq 1$; this contradicts the minimality of d. So the list is independent. Let W be its span. For $i < d-1$, $T\left(T^i\mathbf{v}\right) = T^{i+1}\mathbf{v} \in W$, and $T\left(T^{d-1}\mathbf{v}\right) = T^d\mathbf{v} \in W$ by the choice of d. So W is T-invariant by Proposition 7.2.4, and it contains $\mathbf{v}$, so $K_T(\mathbf{v}) \subseteq W$ by Proposition 7.2.34. Conversely, each $T^i\mathbf{v}$ lies in $K_T(\mathbf{v})$, so $W \subseteq K_T(\mathbf{v})$. Hence $W = K_T(\mathbf{v})$, and the list is a basis.

(b) By the choice of $c_0, \dots, c_{d-1}$,

$$p(T)\mathbf{v} = T^d\mathbf{v} + c_{d-1}T^{d-1}\mathbf{v} + \cdots + c_0\mathbf{v} = \mathbf{0}.$$

If $q = b_0 + b_1x + \cdots + b_{d-1}x^{d-1}$ satisfies $q(T)\mathbf{v} = \mathbf{0}$, then $b_0\mathbf{v} + \cdots + b_{d-1}T^{d-1}\mathbf{v} = \mathbf{0}$ is a linear relation among the basis vectors of (a), so every $b_i = 0$ and $q = \mathbf{0}$ (Theorem 1.5.5). □

In the basis of (a), T shifts each basis vector to the next, except that the last one goes to $T^d\mathbf{v} = -\sum c_iT^i\mathbf{v}$. So the matrix of $T\restriction_{K_T(\mathbf{v})}$ has 1's just below the diagonal, the numbers $-c_0, \dots, -c_{d-1}$ in its last column, and 0's elsewhere. This is the companion matrix of p, which reappears in Section 7.3.

Example 7.2.36 (Krylov subspaces of the cyclic permutation). Let $Z(x, y, z) = (z, x, y)$ on F^3.

(a) For $\mathbf{v} = (1, 1, 1)$, $Z\mathbf{v} = \mathbf{v}$, so $d = 1$, $p = x - 1$, and $K_Z(\mathbf{v})$ is the line $\operatorname{span}(\mathbf{v})$.

(b) For $\mathbf{v} = (1, -1, 0)$, $Z\mathbf{v} = (0, 1, -1)$, which is not a multiple of $\mathbf{v}$, and $Z^2\mathbf{v} = (-1, 0, 1) = -\mathbf{v} - Z\mathbf{v}$. So $d = 2$, $p = x^2 + x + 1$, and $K_Z(\mathbf{v}) = \operatorname{span}((1, -1, 0), (0, 1, -1))$ is the plane $x + y + z = 0$.

(c) For $\mathbf{v} = \mathbf{e}_1$, the vectors $\mathbf{e}_1, Z\mathbf{e}_1 = \mathbf{e}_2, Z^2\mathbf{e}_1 = \mathbf{e}_3$ are independent and $Z^3\mathbf{e}_1 = \mathbf{e}_1$. So $d = 3$, $p = x^3 - 1$, and $K_Z(\mathbf{e}_1) = F^3$.

The polynomials fit together: $x^3 - 1 = (x - 1)\big(x^2 + x + 1\big)$, the line in (a) and the plane in (b) are the two invariant subspaces of Example 7.2.2, and $\mathbf{e}_1 = \frac{1}{3}(1, 1, 1) + \frac{1}{3}(2, -1, -1)$ splits along them (when $3 \neq 0$ in F).

Chains

The Krylov sequence of $\mathbf{v}$ can also end by reaching $\mathbf{0}$. In Section 4.1.7 we proved that if $T^k\mathbf{v} \neq \mathbf{0}$ and $T^{k+1}\mathbf{v} = \mathbf{0}$, then $\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}$ is independent (Proposition 4.1.26). The same holds with T replaced by $T - \lambda\,\mathrm{id}$, and in that form the chain carries information about eigenvalues: its last nonzero vector is an eigenvector.

Proposition 7.2.37 (Chains are independent). Let $T \in \mathrm{L}(V)$, $\lambda \in F$, and $N = T - \lambda\,\mathrm{id}$. Let $\mathbf{v} \in V$ and $k \geq 0$ satisfy

$$N^k\mathbf{v} \neq \mathbf{0} \quad \text{and} \quad N^{k+1}\mathbf{v} = \mathbf{0}.$$

Then $\mathbf{v}, N\mathbf{v}, N^2\mathbf{v}, \dots, N^k\mathbf{v}$ is linearly independent.

For $\lambda = 0$ this is Proposition 4.1.26. Every power N^j with $j \geq k + 1$ kills $\mathbf{v}$, since $N^j\mathbf{v} = N^{j-k-1}\big(N^{k+1}\mathbf{v}\big) = \mathbf{0}$. So applying a power of N to a combination of the chain pushes every term toward the end of the chain, and terms that are pushed past $N^k\mathbf{v}$ disappear. The two proofs below use this in different ways.

Proof (peel off the coefficients). STANDARD

Uses. Theorem 1.5.5 (Independence is decided at zero), Notation 2.1.18, Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Proposition 1.2.35 (Zero products)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Take a relation $a_0\mathbf{v} + a_1N\mathbf{v} + \cdots + a_kN^k\mathbf{v} = \mathbf{0}$ and apply N^k. Which terms survive?
Hint 2. Only $a_0N^k\mathbf{v}$ survives, so $a_0 = 0$. Now apply N^{k-1} to what is left.
Hint 3. Prove $a_i = 0$ by induction on i: if $a_0 = \cdots = a_{i-1} = 0$, apply N^{k-i}.
Sketch. Let $\sum_{j=0}^{k} a_jN^j\mathbf{v} = \mathbf{0}$. Assume $a_j = 0$ for $j < i$ and apply N^{k-i}: the term $j = i$ becomes $a_iN^k\mathbf{v}$, and every term with $j > i$ becomes $N^{k+(j-i)}\mathbf{v} = \mathbf{0}$. So $a_iN^k\mathbf{v} = \mathbf{0}$, and $a_i = 0$ because $N^k\mathbf{v} \neq \mathbf{0}$. By induction all $a_i = 0$.

Suppose that

$$a_0\mathbf{v} + a_1N\mathbf{v} + \cdots + a_kN^k\mathbf{v} = \mathbf{0}. \tag{7.2.2}$$

We show $a_i = 0$ by induction on $i \in \{0, \dots, k\}$. Let i be given and assume $a_j = 0$ for all $j < i$ (no assumption when $i = 0$). Then (7.2.2) reads $\sum_{j=i}^{k} a_jN^j\mathbf{v} = \mathbf{0}$. Applying the linear map N^{k-i},

$$\mathbf{0} = \sum_{j=i}^{k} a_j N^{k-i+j}\mathbf{v} = a_i N^k \mathbf{v} + \sum_{j=i+1}^{k} a_j N^{k+(j-i)}\mathbf{v} = a_i N^k \mathbf{v},$$

because each exponent $k + (j - i)$ in the last sum is at least $k + 1$. As $N^k\mathbf{v} \neq \mathbf{0}$, Proposition 1.2.35 gives $a_i = 0$. So all $a_i = 0$, and the list is independent by Theorem 1.5.5. □

Proof (lowest nonzero coefficient). STANDARD

Uses. Theorem 1.5.5 (Independence is decided at zero), Notation 2.1.18, Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Proposition 1.2.35 (Zero products)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Suppose there is a nontrivial relation, and let m be the smallest index with $a_m \neq 0$.
Hint 2. Apply a single power of N that moves $N^m\mathbf{v}$ exactly to the end of the chain. Where do the terms with $j > m$ go?
Hint 3. Applying N^{k-m} leaves only $a_m N^k\mathbf{v}$.
Sketch. If $\sum_{j=0}^{k} a_j N^j\mathbf{v} = \mathbf{0}$ is nontrivial, let m be the least index with $a_m \neq 0$. Applying N^{k-m} to $\sum_{j\geq m} a_j N^j\mathbf{v} = \mathbf{0}$ sends the term $j = m$ to $a_m N^k\mathbf{v}$ and every term with $j > m$ past the end of the chain, to $\mathbf{0}$. So $a_m N^k\mathbf{v} = \mathbf{0}$, forcing $a_m = 0$, a contradiction.

Suppose, for contradiction, that (7.2.2) holds with not all a_j zero, and let m be the smallest index with $a_m \neq 0$. Then $\sum_{j=m}^{k} a_j N^j\mathbf{v} = \mathbf{0}$. Applying N^{k-m},

$$\mathbf{0} = a_m N^k\mathbf{v} + \sum_{j=m+1}^{k} a_j N^{k+(j-m)}\mathbf{v} = a_m N^k\mathbf{v},$$

since $k + (j - m) \geq k + 1$ for $j > m$. As $N^k\mathbf{v} \neq \mathbf{0}$, this forces $a_m = 0$ (Proposition 1.2.35), contradicting the choice of m. So the relation is trivial, and the list is independent by Theorem 1.5.5. □

Intuition (Why the lowest index). The first proof peels off the coefficients from the front of the chain, one application at a time, and needs an induction to organize the peeling. The second proof asks which single term of a relation is hardest to kill, and kills everything else. Applying N^s shifts every term s places toward the end, and the term that reaches the end last is the one with the smallest index. So the lowest index with a nonzero coefficient is the right extremal choice, and one application of N^{k-m} isolates it. Choosing the highest index instead does not work: applying $N^{k-m'}$ for the highest index m' sends every lower term somewhere strictly inside the chain, so nothing is isolated.

The pattern "choose the extremal term, then apply the operator that kills everything except that term" is exactly the shortest-relation proof of Theorem 7.2.21, with "smallest index" in place of "fewest terms". It solves Exercise 7.2.51, where several chains are mixed together.

Corollary 7.2.38 (The span of a chain). Under the hypotheses of Proposition 7.2.37:

(a) $N^k\mathbf{v}$ is an eigenvector of T with eigenvalue λ;

(b) $\operatorname{span}\big(\mathbf{v}, N\mathbf{v}, \ldots, N^k\mathbf{v}\big) = K_T(\mathbf{v})$, so $\dim K_T(\mathbf{v}) = k + 1$;

(c) in the basis $N^k\mathbf{v}, N^{k-1}\mathbf{v}, \ldots, N\mathbf{v}, \mathbf{v}$ of $K_T(\mathbf{v})$, the matrix of $T\restriction_{K_T(\mathbf{v})}$ has every diagonal entry equal to λ, every entry just above the diagonal equal to 1, and all other entries 0.

Proof. ROUTINE

Uses. Proposition 7.2.37 (Chains are independent), Proposition 7.2.34 (The smallest invariant subspace containing a vector), Proposition 7.2.4 (Invariance is checked on a spanning list), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 2.3.1 (Matrix of a linear map)
Hint 1. For (a), compute $(T - \lambda\,\mathrm{id})\big(N^k\mathbf{v}\big)$.
Hint 2. For (b), each $N^j\mathbf{v} = (x - \lambda)^j(T)\mathbf{v}$ lies in $K_T(\mathbf{v})$. For the other inclusion, show the span of the chain is T-invariant, using $T = N + \lambda\,\mathrm{id}$.
Hint 3. For (c), compute T on each basis vector: $T\big(N^j\mathbf{v}\big) = N^{j+1}\mathbf{v} + \lambda N^j\mathbf{v}$.

Sketch. (a) $N\big(N^k\mathbf{v}\big) = N^{k+1}\mathbf{v} = \mathbf{0}$ and $N^k\mathbf{v} \neq \mathbf{0}$. (b) $N^j\mathbf{v} = \big((x-\lambda)^j\big)(T)\mathbf{v} \in K_T(\mathbf{v})$, and $T\big(N^j\mathbf{v}\big) = N^{j+1}\mathbf{v} + \lambda N^j\mathbf{v}$ lies in the span W of the chain, so W is invariant and contains $K_T(\mathbf{v})$. The chain is independent, so $\dim = k+1$. (c) Read off the coordinates of $T\big(N^j\mathbf{v}\big)$.

(a) $(T-\lambda\,\mathrm{id})\big(N^k\mathbf{v}\big) = N^{k+1}\mathbf{v} = \mathbf{0}$, so $T\big(N^k\mathbf{v}\big) = \lambda N^k\mathbf{v}$, and $N^k\mathbf{v} \neq \mathbf{0}$ by hypothesis.

(b) Let $W = \operatorname{span}\big(\mathbf{v}, N\mathbf{v}, \ldots, N^k\mathbf{v}\big)$. Each $N^j = (T-\lambda\,\mathrm{id})^j$ equals $q_j(T)$ for $q_j = (x-\lambda)^j$ (Proposition 7.2.27(a)), so $N^j\mathbf{v} \in K_T(\mathbf{v})$ and $W \subseteq K_T(\mathbf{v})$. Since $T = N + \lambda\,\mathrm{id}$,

$$T\big(N^j\mathbf{v}\big) = N^{j+1}\mathbf{v} + \lambda N^j\mathbf{v}, \tag{7.2.3}$$

which lies in W for $j < k$, and equals $\lambda N^k\mathbf{v} \in W$ for $j = k$. So W is T-invariant (Proposition 7.2.4) and contains $\mathbf{v}$, hence $K_T(\mathbf{v}) \subseteq W$ (Proposition 7.2.34). Thus $W = K_T(\mathbf{v})$, and by Proposition 7.2.37 the chain is a basis of it, of length $k+1$.

(c) Number the basis as $\mathbf{b}_1 = N^k\mathbf{v}, \mathbf{b}_2 = N^{k-1}\mathbf{v}, \ldots, \mathbf{b}_{k+1} = \mathbf{v}$, so $\mathbf{b}_i = N^{k+1-i}\mathbf{v}$. By (7.2.3), $T\mathbf{b}_1 = \lambda\mathbf{b}_1$ and $T\mathbf{b}_i = \mathbf{b}_{i-1} + \lambda\mathbf{b}_i$ for $i \geq 2$. So the i-th column of the matrix has λ in row i, 1 in row $i-1$ when $i \geq 2$, and 0 elsewhere (Definition 2.3.1). □

For $k = 2$ the matrix in (c) is

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

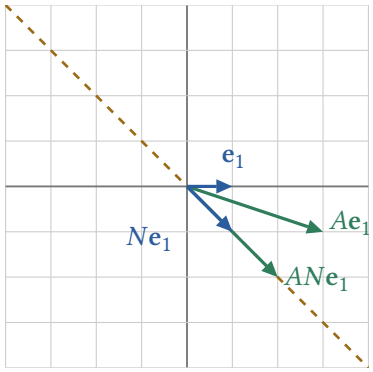
Matrices of this shape are Jordan blocks, and Section 10.3 shows that every operator on a finite-dimensional complex vector space is built out of chains like these. Here is the smallest example that is not diagonal in any basis.

Example 7.2.39 (A matrix with a single invariant line). Let $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, $\lambda = 2$, and

$$N = A - 2I_2 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 1-1 & 1-1 \\ -1+1 & -1+1 \end{pmatrix} = 0.$$

Take $\mathbf{v} = \mathbf{e}_1$: $N\mathbf{e}_1 = (1,-1) \neq \mathbf{0}$ and $N^2\mathbf{e}_1 = \mathbf{0}$, so Proposition 7.2.37 applies with $k = 1$. By Corollary 7.2.38, $(1,-1)$ is an eigenvector with eigenvalue 2 (indeed $A(1,-1) = (2,-2)$), and in the basis $(1,-1), \mathbf{e}_1$ the matrix of T_A is $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

The only eigenvalue is 2: if $A\mathbf{w} = \mu\mathbf{w}$, then $N\mathbf{w} = (\mu-2)\mathbf{w}$ and $\mathbf{0} = N^2\mathbf{w} = (\mu-2)^2\mathbf{w}$, so $\mu = 2$ when $\mathbf{w} \neq \mathbf{0}$. Its eigenspace $E_2 = \ker N = \operatorname{span}((1,-1))$ is a line, since $N(x,y) = (x+y)(1,-1)$. So A has exactly one invariant line, and there is no basis of $\mathbb{R}^2$ consisting of eigenvectors of A.



The map of $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$ from Example 7.2.39. The dashed line $y = -x$ is stretched by 2 and is the only invariant line. The vector $\mathbf{e}_1$ goes to $A\mathbf{e}_1 = 2\mathbf{e}_1 + N\mathbf{e}_1$: twice itself plus a vector on the invariant line. In the interactive version, change the entries slightly and the single invariant line either splits into two or disappears. For $\begin{pmatrix} 3 & b \\ c & 1 \end{pmatrix}$ there are two invariant lines when $bc > -1$, one when $bc = -1$, and none when $bc < -1$. (Interactive in the web version.)

Exercise 7.2.40. ROUTINE Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in F^{3\times 3}$ and $N = A - I_3$.

(a) Compute $N\mathbf{e}_3$, $N^2\mathbf{e}_3$, and $N^3\mathbf{e}_3$, and deduce that $\mathbf{e}_3, N\mathbf{e}_3, N^2\mathbf{e}_3$ is a basis of F^3.

(b) Show that 1 is the only eigenvalue of A and that $E_1(A) = \operatorname{span}(\mathbf{e}_1)$.

Solution on page 1565

Exercises

Exercise 7.2.41. ROUTINE Let $B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$. Find all eigenvalues of B, with an eigenvector for each, over $\mathbb{R}$, $\mathbb{F}_2$, $\mathbb{F}_3$, $\mathbb{F}_5$, and $\mathbb{F}_7$.

Solution on page 1566

Exercise 7.2.42. STANDARD For $\theta \in \mathbb{R}$, let $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, the rotation by θ (Example 2.1.11).

(a) For which θ does R_θ have a real eigenvalue? Find the eigenspaces in those cases.

(b) Regard R_θ as a matrix in $\mathbb{C}^{2\times 2}$. Show that $(1, -i)$ and $(1, i)$ are eigenvectors, with eigenvalues $\cos\theta + i\sin\theta$ and $\cos\theta - i\sin\theta$.

Solution on page 1566

Exercise 7.2.43. STANDARD Let $S, T \in \mathrm{L}(V)$. Prove or disprove:

(a) If λ is an eigenvalue of S and μ is an eigenvalue of T, then $\lambda + \mu$ is an eigenvalue of $S + T$.

(b) If λ is an eigenvalue of S and μ is an eigenvalue of T, then $\lambda\mu$ is an eigenvalue of ST.

(c) If $\mathbf{v}$ is an eigenvector of both S and T, then it is an eigenvector of $S + T$ and of ST.

(d) If $\mathbf{v}$ is an eigenvector of ST, then it is an eigenvector of TS.

(e) If $ST = TS$, then every eigenvector of S is an eigenvector of T.

Solution on page 1566

Exercise 7.2.44. STANDARD Let $A \in F^{n\times n}$.

(a) Suppose every row of A has the same sum s. Show that s is an eigenvalue of A.

(b) Show that A and A^{T} have the same eigenvalues.

(c) Suppose every column of A has the same sum s. Show that s is an eigenvalue of A. Must $(1, \ldots, 1)$ be an eigenvector?

Solution on page 1567

Exercise 7.2.45. STANDARD

(a) Show that an operator $T \in \mathrm{L}(\mathbb{R}^2)$ has 0, 1, 2, or infinitely many invariant lines, and give an example of each.

(b) Over $\mathbb{F}_2$, show that every $T \in \mathrm{L}(\mathbb{F}_2^2)$ has 0, 1, 2, or 3 invariant lines, and give an example of each.

Solution on page 1567

Exercise 7.2.46. STANDARD Let D be differentiation on $\mathbb{R}[x]$.

(a) Show that the D-invariant subspaces of $\mathbb{R}[x]$ are exactly $\{\mathbf{0}\}$, the spaces $\mathbb{R}[x]_{\leq n}$ for $n \geq 0$, and $\mathbb{R}[x]$ itself.

(b) Over $\mathbb{F}_p$, find a D-invariant subspace of $\mathbb{F}_p[x]$ that is not of this form.

Solution on page 1568

Exercise 7.2.47. STANDARD Let $\dim V = n$, and let $T \in \mathrm{L}(V)$ have n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Let $S \in \mathrm{L}(V)$ commute with T.

(a) Show that each $\mathbf{v}_i$ is an eigenvector of S.

(b) Show that $S = p(T)$ for some $p \in F[x]$ of degree less than n.

Solution on page 1568

Exercise 7.2.48. HARD Let $\dim V = n$, and let $T \in \mathrm{L}(V)$ have n distinct eigenvalues. Show that T has exactly 2^n invariant subspaces.

Solution on page 1569

Exercise 7.2.49. STANDARD Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors of $T \in \mathrm{L}(V)$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, and let $\mathbf{w} = \mathbf{v}_1 + \cdots + \mathbf{v}_n$. Show that $\mathbf{w}, T\mathbf{w}, \ldots, T^{n-1}\mathbf{w}$ is linearly independent, and deduce that $K_T(\mathbf{w}) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$.

Solution on page 1569

Exercise 7.2.50. HARD Let $\dim V = n \geq 1$, and let $N \in \mathrm{L}(V)$ satisfy $N^n = 0$ and $N^{n-1} \neq 0$.

(a) Show that $\dim \ker N^j = j$ for $j = 0, 1, \ldots, n$.

(b) Show that the N-invariant subspaces of V are exactly $\ker N^0, \ker N^1, \ldots, \ker N^n$. In particular there are exactly $n + 1$ of them.

Solution on page 1570

Exercise 7.2.51. HARD Let $N \in \mathrm{L}(V)$, let $\mathbf{u}_1, \ldots, \mathbf{u}_r \in V$, and let $k_1, \ldots, k_r \geq 0$ satisfy $N^{k_i+1}\mathbf{u}_i = \mathbf{0}$ for each i. Suppose the "tops" $N^{k_1}\mathbf{u}_1, \ldots, N^{k_r}\mathbf{u}_r$ are linearly independent. Show that the list of all vectors

$$N^j \mathbf{u}_i \quad (1 \le i \le r, 0 \le j \le k_i)$$

is linearly independent.

Solution on page 1570

Exercise 7.2.52. STANDARD Let $n \ge 1$ and define $T : F[x]_{\le n} \to F[x]_{\le n}$ by $T(p) = xp'$, where p' is the formal derivative.

(a) Over $\mathbb{R}$, find the eigenvalues and eigenspaces of T.

(b) Over $\mathbb{F}_3$ with $n = 5$, find the eigenvalues and eigenspaces of T.

Solution on page 1571

Exercise 7.2.53. STANDARD Find the error in the following argument, and decide whether the claim is true.

Claim. Let $T \in \mathrm{L}(V)$ with V a complex vector space. Every eigenvector of T^2 is an eigenvector of T.

"Proof." Let $T^2\mathbf{v} = \mu\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and choose $\lambda \in \mathbb{C}$ with $\lambda^2 = \mu$. Then

$$(T - \lambda \,\mathrm{id})(T + \lambda \,\mathrm{id})\mathbf{v} = \big(T^2 - \lambda^2 \,\mathrm{id}\big)\mathbf{v} = \mathbf{0}.$$

So either $(T + \lambda \,\mathrm{id})\mathbf{v} = \mathbf{0}$ or $(T - \lambda \,\mathrm{id})\mathbf{v} = \mathbf{0}$. In the first case $\mathbf{v}$ is an eigenvector of T for $-\lambda$, and in the second for λ.

Solution on page 1571

Exercise 7.2.54. STANDARD Let $T \in \mathrm{L}(V)$, and let U, W be subspaces of V. Prove or disprove:

(a) If U and W are T-invariant, then so are $U + W$ and $U \cap W$.

(b) If U is T^2-invariant, then U is T-invariant.

(c) If V is finite-dimensional, T is invertible, and U is T-invariant, then U is T^{-1}-invariant.

(d) Statement (c) holds without assuming that V is finite-dimensional.

Solution on page 1572

Exercise 7.2.55. HARD BEYOND 21-242 Let $T \in \mathrm{L}(V)$ and let $p, q \in F[x]$ be coprime, with $p(T)q(T) = 0$. Show that

$$V = \ker p(T) \oplus \ker q(T),$$

and that both summands are T-invariant. Deduce that if $1 + 1 \neq 0$ in F and $T^2 = \mathrm{id}$, then $V = E_1 \oplus E_{-1}$.

Solution on page 1572

Exercise 7.2.56. STANDARD Let $Z(x, y, z) = (z, x, y)$ be the cyclic permutation of coordinates on F^3. Find all eigenvalues of Z, and an eigenvector for each, over $\mathbb{C}$, over $\mathbb{R}$, over $\mathbb{F}_7$, and over $\mathbb{F}_3$.

Solution on page 1573

Exercise 7.2.57. HARD Let $\dim V = n$ and $1 \le k < n$. Suppose $T \in \mathrm{L}(V)$ leaves every k-dimensional subspace of V invariant. Show that $T = \lambda \,\mathrm{id}$ for some $\lambda \in F$.

Solution on page 1573

Exercise 7.2.58. ROUTINE Find $A \in \mathbb{Q}^{2\times 2}$ with $A(1, 2) = 3(1, 2)$ and $A(1, 1) = -(1, 1)$. Is A unique?

Solution on page 1574

7.3 The characteristic polynomial

In Section 7.2 we found eigenvalues by solving $T\mathbf{v} = \lambda\mathbf{v}$ by hand, with a separate argument for each operator. Theorem 7.2.9 points to a systematic method: λ is an eigenvalue of $A \in F^{n\times n}$ exactly when $\det(A - \lambda I_n) = 0$. If we leave λ unknown and call it x, the Leibniz formula turns $\det(xI_n - A)$ into a polynomial of degree n in x, and the eigenvalues of A are its roots. This section makes that precise, reads the trace and the determinant of A off the coefficients, and shows that every monic polynomial occurs. It also settles the existence question raised by the rotation in Example 7.2.18. Over $\mathbb{C}$ every operator on a nonzero finite-dimensional space has an eigenvalue, and we prove this twice, once with the characteristic polynomial and once without determinants.

7.3.1 Determinants with polynomial entries

The matrix $xI_n - A$ has entries in $F[x]$, and $F[x]$ is not a field: x has no multiplicative inverse in $F[x]$ (Corollary 7.1.7). Chapter 5 defined the determinant only for matrices with entries in a field. So before we can write $\det(xI_n - A)$, we have to say what it means and which rules it obeys.

Intuition (What the Leibniz formula needs). The Leibniz formula (Theorem 5.3.3) builds $\det A$ out of the entries of A with additions, subtractions, and multiplications, and never divides. So it makes sense for entries from any commutative ring (Definition 1.1.20), and for $F[x]$ in particular (Proposition 7.1.3). The real question is which theorems of Chapter 5 survive. Most facts there were proved twice: once by row reduction, which divides by pivots, and once from the Leibniz formula or from the uniqueness of multilinear alternating functions, which do not divide. The second kind of proof works over a commutative ring word for word. The one statement that changes is the invertibility criterion, because in a ring a nonzero element need not be invertible.

Definition 7.3.1 (Determinant over a commutative ring). Let R be a commutative ring and $n \geq 1$. We write $R^{n\times n}$ for the set of $n \times n$ arrays $M = (m_{ij})$ with entries in R, and add, multiply, and transpose such arrays by the same formulas as matrices over a field. The **determinant** of $M \in R^{n\times n}$ is

$$\det M = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) m_{\sigma(1)1} m_{\sigma(2)2} \cdots m_{\sigma(n)n} \in R,$$

where $\operatorname{sgn}(\sigma) \in \{1, -1\}$ is read in R. When $R = F$ is a field, this is the determinant of Definition 5.3.6, by Theorem 5.3.3.

For example, with $R = F[x]$,

$$\det\begin{pmatrix} x-a & -b \\ -c & x-d \end{pmatrix} = (x-a)(x-d) - (-b)(-c) = x^2 - (a+d)x + (ad - bc).$$

Proposition 7.3.2 (Determinants over a commutative ring). Let R and R' be commutative rings, $n \geq 1$, and $M, N \in R^{n\times n}$.

(a) Multiplication in $R^{n\times n}$ is associative and distributes over addition, $(rM)N = r(MN) = M(rN)$ for $r \in R$, and the array I_n with 1 on the diagonal and 0 elsewhere is a two-sided identity. In particular the scalar arrays rI_n commute with every array.

(b) det is linear in each column (with coefficients in R), it vanishes when two columns are equal, and $\det I_n = 1$. Every function $D : R^{n\times n} \to R$ with these first two properties satisfies $D(M) = D(I_n)\det M$.

(c) $\det(MN) = \det M \det N$.

(d) $\det M^{\mathsf{T}} = \det M$. So (b) holds for rows as well as columns.

(e) If $M = \begin{pmatrix} P & Q \\ 0 & S \end{pmatrix}$ with square blocks P and S, then $\det M = \det P \det S$. In particular, if M is upper or lower triangular, then $\det M = m_{11}m_{22}\cdots m_{nn}$.

(f) For $n \geq 2$, cofactor expansion along any row or column computes $\det M$, and $M\,\mathrm{adj}(M) = \mathrm{adj}(M)M = (\det M)I_n$.

(g) Let $\varphi : R \to R'$ satisfy $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$, and $\varphi(1) = 1$ for all $a, b \in R$, and write $\varphi(M) = \big(\varphi(m_{ij})\big)$. Then $\det \varphi(M) = \varphi(\det M)$.

Proof. STANDARD

Uses. Definition 7.3.1 (Determinant over a commutative ring), Proposition 2.4.14 (Algebra of matrices), Theorem 5.3.4 (Existence and uniqueness of the determinant), Theorem 5.3.3 (The Leibniz formula), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Theorem 5.4.26 (Block-triangular matrices), Proposition 5.3.15 (Triangular matrices), Theorem 5.5.3 (Cofactor expansion), Theorem 5.5.9 (The adjugate and the inverse), Definition 1.1.20 (Commutative ring)
Techniques. Cancel terms in pairs (Technique 5.3.5), Recognize the determinant by its axioms (Technique 5.4.11)
Hint 1. For (a)–(f), go back to Chapter 5 and, for each fact, find a proof that never divides. Which proofs used row reduction, and which used only the Leibniz formula or the uniqueness of multilinear alternating functions?
Hint 2. Alternation is the delicate point: the shortcut "$L(A) = -L(A)$" divides by $1+1$. The proof of Theorem 5.3.4 cancelled terms in pairs instead.
Hint 3. For (g), apply φ to the Leibniz sum term by term. You need $\varphi(0) = 0$ and $\varphi(-a) = -\varphi(a)$; derive them from additivity.
Sketch. Each of (a)–(f) has a proof in Chapter 5 that uses only the commutative ring axioms: entrywise computation for (a); the Leibniz formula, pairing σ with $\sigma\tau$, for (b); uniqueness for (c), since $\mathbf{b} \mapsto \det(M\mathbf{b})$ is multilinear and alternating; reindexing by σ^{-1} for (d); the Leibniz formula for (e) and (f). For (g), $\varphi(0) = 0$ and $\varphi(-a) = -\varphi(a)$ follow from additivity, so φ maps each Leibniz term of M to the corresponding term of $\varphi(M)$.

We go through Chapter 5 and point, for each statement, to an argument that uses only the commutative ring axioms, that is, never divides and never uses $1 + 1 \neq 0$.

(a) The entrywise proof of Proposition 2.4.14 compares the entries of both sides, each of which is a sum of products of entries, using associativity, commutativity, and distributivity in the scalars. These hold in R.

(b) The proof of Theorem 5.3.4 (parts 2 and 3) shows that the Leibniz formula is normalized, linear in each column, and alternating. Normalization and multilinearity are checked term by term, since each term contains exactly one entry from each column. Alternation is proved by pairing each even σ with $\sigma\tau$, where τ swaps the two equal columns, and the two terms cancel (Technique 5.3.5). None of this divides. The last statement is Theorem 5.3.3. Its proof expands $D(M)$ by multilinearity into a sum over functions $[n] \to [n]$, discards the non-injective ones because D vanishes on arrays with two equal columns, and evaluates D on permuted identity columns with Theorem 5.2.31; these steps use multilinearity, alternation, and the ring axioms only.

(c) We use the second proof of Theorem 5.4.10 (Technique 5.4.11). Fix M and let $D(N) = \det(MN)$. Column j of MN is M times column j of N, so D is linear in each column of N by (a) and (b), and two equal columns of N give two equal columns of MN, where det vanishes. By (b), $D(N) = D(I_n)\det N = \det M \det N$.

(d) The second proof of Theorem 5.4.14 reindexes the Leibniz sum by $\sigma \mapsto \sigma^{-1}$ and uses $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$; this is ring arithmetic.

(e) In the Leibniz formula for $\begin{pmatrix} P & Q \\ 0 & S \end{pmatrix}$, with P of size $k \times k$, a term can be nonzero only if σ maps $\{1, \ldots, k\}$ into itself, and then the term is a product of a Leibniz term of P and one of S. This is the proof of Theorem 5.4.26 via the Leibniz formula, and the triangular case is the Leibniz proof of Proposition 5.3.15. Lower block-triangular arrays follow by (d).

(f) The second proof of Theorem 5.5.3 derives the expansion from the Leibniz formula by regrouping its terms, using the signs of cycles and the multiplicativity of sgn, and gets the row version from (d). The identity $M \operatorname{adj}(M) = (\det M) I_n$ is proved in Theorem 5.5.9 by expanding along a row and noting that expansion along the "wrong" row computes the determinant of an array with two equal rows, which is 0 by (b) and (d). Neither step divides.

(g) First, $\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$, so $\varphi(0) = 0$; and $\varphi(a) + \varphi(-a) = \varphi(0) = 0$, so $\varphi(-a) = -\varphi(a)$. Hence $\varphi(\operatorname{sgn}(\sigma)) = \operatorname{sgn}(\sigma)$, and since φ respects sums and products of any finite length,

$$\varphi(\det M) = \sum_{\sigma \in S_n} \varphi(\operatorname{sgn}(\sigma)) \varphi\left(m_{\sigma(1)1}\right) \cdots \varphi\left(m_{\sigma(n)n}\right) = \det \varphi(M).$$

□

The most important instance of (g) is evaluation. For $M \in F[x]^{n \times n}$ and $\lambda \in F$, let $M(\lambda) \in F^{n \times n}$ be the matrix obtained by evaluating every entry at λ. Evaluation respects sums, products, and 1 (Proposition 7.1.10), so

$$(\det M)(\lambda) = \det M(\lambda). \tag{7.3.1}$$

In words, it does not matter whether we first compute the determinant as a polynomial and then plug in λ, or first plug in λ and then compute a determinant of scalars. The same holds for substituting a polynomial $r \in F[x]$ for x in every entry (also Proposition 7.1.10), and for complex conjugation of the coefficients of polynomials in $\mathbb{C}[x]$ (Lemma 7.1.37).

Remark 7.3.3 (What does not carry over). Over a field, M is invertible if and only if $\det M \neq 0$ (Theorem 5.4.7). Over $F[x]$ this fails: $\det(xI_n) = x^n \neq \mathbf{0}$, but xI_n has no inverse in $F[x]^{n \times n}$: if $(xI_n)N = I_n$, then $xn_{11} = 1$, which is impossible by Corollary 7.1.7(c). The proof of Theorem 5.4.7 used row reduction, which divides by pivots. What survives is (f): if $\det M$ has an inverse in R, then $(\det M)^{-1} \operatorname{adj}(M)$ is an inverse of M. The converse also holds, by (c): $MN = I_n$ gives $\det M \det N = 1$. So M is invertible in $R^{n \times n}$ exactly when $\det M$ is invertible in R, which for $R = F[x]$ means that $\det M$ is a nonzero constant. We will not need this, but it explains why nothing like "$\det(xI_n - A) \neq \mathbf{0}$, so $xI_n - A$ is invertible" can be used.

ASIDE: *Rational functions*

There is a second way to make sense of $\det(xI_n - A)$. Fractions p/q of polynomials with $q \neq \mathbf{0}$, with the usual rules for fractions, form a field $F(x)$ that contains $F[x]$, in the same way that $\mathbb{Q}$ contains $\mathbb{Z}$. All of Chapter 5 applies to matrices over $F(x)$, and the Leibniz formula shows that a matrix with polynomial entries has a polynomial determinant. This route gives everything in Proposition 7.3.2 for $R = F[x]$ at once, at the cost of constructing $F(x)$. We stay with commutative rings, which we need anyway for (g).

7.3.2 The characteristic polynomial

Definition 7.3.4 (Characteristic polynomial). Let $n \geq 1$ and $A \in F^{n \times n}$. The **characteristic polynomial** of A is

$$p_A = \det(xI_n - A) \in F[x],$$

the determinant (Definition 7.3.1) of the matrix $xI_n - A \in F[x]^{n \times n}$, whose diagonal entries are $x - a_{ii}$ and whose off-diagonal entries are the constants $-a_{ij}$. When the polynomial is evaluated, we write $p_A(\lambda)$ for its value at $\lambda \in F$.

Some texts, including the lecture notes, use $\det(A - xI_n)$ instead. Multiplying each of the n columns by -1 shows $\det(A - xI_n) = (-1)^n p_A$ (Proposition 7.3.2(b)), so the two conventions differ by a sign and have the same roots. With our sign, p_A is monic (Proposition 7.3.18).

Example 7.3.5 (Small cases).

(a) For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the computation after Definition 7.3.1 gives

$$p_A = x^2 - (a+d)x + (ad - bc) = x^2 - (\operatorname{tr} A)x + \det A.$$

(b) For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ of Example 7.2.8, $p_A = x^2 - 4x + 3 = (x-1)(x-3)$, and its roots 1 and 3 are the eigenvalues found there.

(c) For the rotation $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ of Example 7.2.18, $p_A = x^2 + 1$. It has no root in $\mathbb{R}$ or $\mathbb{F}_3$, the roots $\pm i$ in $\mathbb{C}$, and the roots $2, 3$ in $\mathbb{F}_5$, matching the eigenvalues found there field by field.

(d) If A is upper triangular, so is $xI_n - A$, with diagonal entries $x - a_{ii}$. By Proposition 7.3.2(e),

$$p_A = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}).$$

For instance, $p_A = (x-2)^2$ for $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, the matrix of Example 7.2.39 in its chain basis.

Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial). Let $A \in F^{n \times n}$ and $\lambda \in F$. Then λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$.

Proof. ROUTINE

Uses. Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.1.10 (Evaluation respects sums and products), Definition 7.3.4 (Characteristic polynomial)
Hint 1. Theorem 7.2.9 already says that λ is an eigenvalue if and only if $\det(A - \lambda I_n) = 0$. How is that number related to the polynomial p_A?
Hint 2. Evaluate $xI_n - A$ entry by entry at λ and use (7.3.1).
Sketch. By (7.3.1), $p_A(\lambda) = \det(\lambda I_n - A) = (-1)^n \det(A - \lambda I_n)$, which vanishes exactly when λ is an eigenvalue, by Theorem 7.2.9.

Evaluating every entry of $xI_n - A$ at λ gives $\lambda I_n - A$, so by (7.3.1), $p_A(\lambda) = \det(\lambda I_n - A)$. Multiplying each column by -1 gives $\det(\lambda I_n - A) = (-1)^n \det(A - \lambda I_n)$. So $p_A(\lambda) = 0$ if and only if $\det(A - \lambda I_n) = 0$, which by Theorem 7.2.9 holds if and only if λ is an eigenvalue of A. □

The theorem turns the search for eigenvalues into a search for roots of one polynomial of degree n. It reproves Corollary 7.2.22 for matrices: a nonzero polynomial of degree n has at most n roots (Proposition 7.1.22).

Algorithm 7.3.7 (Eigenvalues by the characteristic polynomial)

Input A matrix $A \in F^{n \times n}$.
Output The eigenvalues of A in F, and a basis of each eigenspace.

1. Compute $p_A = \det(xI_n - A)$, by cofactor expansion along a sparse row or column, or by the Leibniz formula for small n.
2. Find the roots $\lambda_1, \ldots, \lambda_k$ of p_A in F. These are the eigenvalues.
3. For each λ_j, row reduce $A - \lambda_j I_n$ and read off a basis of its null space E_{λ_j} from the free columns.

The algorithm is correct by Theorem 7.3.6 and Definition 7.2.13. In step 3 the matrix $A - \lambda_j I_n$ must have rank less than n; if row reduction finds n pivots, then λ_j was not a root, and the arithmetic in step 1 or step 2 should be checked. The hard step is step 2: for $n \geq 5$ there is no formula in radicals for the roots in

general, and in practice eigenvalues are computed by other methods. For hand computations in this book, the roots will be findable.

Example 7.3.8 (A 3×3 computation). Let $A = \begin{pmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. Then

$$xI_3 - A = \begin{pmatrix} x-4 & 0 & -1 \\ 2 & x-1 & 0 \\ 2 & 0 & x-1 \end{pmatrix}.$$

The middle column has a single nonzero entry, so expanding along it (Proposition 7.3.2(f)),

$$p_A = (x-1)\det\begin{pmatrix} x-4 & -1 \\ 2 & x-1 \end{pmatrix} = (x-1)((x-4)(x-1)+2) = (x-1)\left(x^2-5x+6\right) = (x-1)(x-2)(x-3).$$

So the eigenvalues are 1, 2, 3. For $\lambda = 2$, row reduction gives

$$A - 2I_3 = \begin{pmatrix} 2 & 0 & 1 \\ -2 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

so $E_2 = \operatorname{span}((1,-2,-2))$; check: $A(1,-2,-2) = (2,-4,-4)$. In the same way, $E_1 = \operatorname{span}((0,1,0))$ and $E_3 = \operatorname{span}((1,-1,-1))$. Realize that $\operatorname{tr} A = 6 = 1+2+3$ and $\det A = 6 = 1\cdot 2\cdot 3$; Remark 7.3.19 explains this.

$$\begin{pmatrix} 2 & 0 & 1 \\ -2 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Row reduction of $A - 2I_3$ from Example 7.3.8. The third column is free, so E_2 is a line. Use "Edit" to try $A - \lambda I_3$ for other values of λ: you get three pivots, and only the zero vector in the null space, unless λ is 1, 2, or 3. (Interactive in the web version.)

Exercise 7.3.9. ROUTINE Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Find p_A, and find the eigenvalues of A together with an eigenvector for each, over $\mathbb{R}$, over $\mathbb{F}_2$, over $\mathbb{F}_5$, and over $\mathbb{F}_{11}$.

Solution on page 1574

Remark 7.3.10 (A polynomial, not a function). Over a finite field, the function $\lambda \mapsto \det(\lambda I_n - A)$ on F carries less information than the polynomial p_A. Over $\mathbb{F}_2$, take $A = 0_{3\times 3}$ and $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Then $p_A = x^3$, and by Proposition 7.3.2(e) and Example 7.3.5(a), $p_B = x\left(x^2+x+1\right) = x^3+x^2+x$. Both polynomials take the value 0 at 0 and 1 at 1, so they define the same function on $\mathbb{F}_2$. But they differ, and so do $\operatorname{tr} A = 0$ and $\operatorname{tr} B = 1$, which Proposition 7.3.18 reads off the coefficients. This is why we defined p_A through determinants over $F[x]$, instead of as the function $\lambda \mapsto \det(\lambda I_n - A)$.

7.3.3 Similarity and operators

The characteristic polynomial of a matrix should depend only on the operator the matrix represents, as the trace and the determinant do (Definition 4.5.30, Definition 5.4.18).

Proposition 7.3.11 (Similar matrices have the same characteristic polynomial). If $A, B \in F^{n\times n}$ are similar, then $p_A = p_B$.

Proof (by multiplicativity over $F[x]$). ROUTINE

Uses. Definition 4.5.23 (Similar matrices), Proposition 7.3.2 (Determinants over a commutative ring), Definition 7.3.4 (Characteristic polynomial)
Hint 1. Write $B = S^{-1}AS$. Can you write $xI_n - B$ as a product of three matrices with entries in $F[x]$?
Hint 2. $xI_n = S^{-1}(xI_n)S$. Then take determinants in $F[x]$, using Proposition 7.3.2(c).
Sketch. $xI_n - S^{-1}AS = S^{-1}(xI_n - A)S$ in $F[x]^{n\times n}$. Taking determinants with Proposition 7.3.2(c) gives $p_B = \det S^{-1}\, p_A \det S = p_A$, since $\det S^{-1} \det S = \det I_n = 1$.

Let $B = S^{-1}AS$ with $S \in F^{n\times n}$ invertible. In $F[x]^{n\times n}$, by Proposition 7.3.2(a),

$$S^{-1}(xI_n - A)S = xS^{-1}S - S^{-1}AS = xI_n - B.$$

Taking determinants and using Proposition 7.3.2(c) twice,

$$p_B = \det\bigl(S^{-1}\bigr)\det(xI_n - A)\det(S) = \det\bigl(S^{-1}\bigr)\det(S)p_A = \det\bigl(S^{-1}S\bigr)p_A = p_A,$$

where we used that $F[x]$ is commutative and $\det I_n = 1$. □

Proof (pointwise, when F is infinite). ROUTINE

Uses. Theorem 5.4.10 (Multiplicativity), Corollary 7.1.23 (Polynomials and their values), (7.3.1)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. Compare the values $p_A(\lambda)$ and $p_B(\lambda)$ for a scalar λ. Only determinants over F are involved.
Hint 2. Two polynomials that agree at infinitely many points are equal.
Sketch. For $\lambda \in F$, $p_B(\lambda) = \det\bigl(S^{-1}(\lambda I_n - A)S\bigr) = p_A(\lambda)$ by multiplicativity over F. If F is infinite, polynomials that agree at every point are equal.

Let $B = S^{-1}AS$ and $\lambda \in F$. By (7.3.1) and multiplicativity over the field F (Theorem 5.4.10),

$$p_B(\lambda) = \det\bigl(S^{-1}(\lambda I_n - A)S\bigr) = \det\bigl(S^{-1}\bigr)\det(\lambda I_n - A)\det(S) = p_A(\lambda).$$

So p_A and p_B agree at every point of F. If F is infinite, Corollary 7.1.23 gives $p_A = p_B$. □

The second proof avoids determinants over $F[x]$ but only proves equality of functions, which is weaker than equality of polynomials over a finite field (Remark 7.3.10). Over an infinite field this is no loss. The idea of proving a polynomial identity by checking it at many points returns in Proposition 7.3.22.

Definition 7.3.12 (Characteristic polynomial of an operator). Let V be finite-dimensional with $\dim V \geq 1$, and $T \in \mathrm{L}(V)$. The **characteristic polynomial** of T is

$$p_T = p_{[T]_{\mathcal{B}\mathcal{B}}},$$

where $\mathcal{B}$ is any basis of V. Matrices of T in two bases are similar (Theorem 4.5.25), so by Proposition 7.3.11 this does not depend on $\mathcal{B}$.

Corollary 7.3.13 (Eigenvalues of an operator). Let V be finite-dimensional with $\dim V \geq 1$, $T \in \mathrm{L}(V)$, and $\lambda \in F$. Then λ is an eigenvalue of T if and only if $p_T(\lambda) = 0$.

Proof. ROUTINE

Uses. Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Definition 7.3.12 (Characteristic polynomial of an operator), Definition 2.3.1 (Matrix of a linear map), Proposition 1.6.20 (Coordinates translate questions about V into F^n)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Fix a basis $\mathcal{B}$ and let $A = [T]_{\mathcal{B}\mathcal{B}}$. Compare the eigenvectors of T with those of A.
Hint 2. $[T\mathbf{v}]_{\mathcal{B}} = A[\mathbf{v}]_{\mathcal{B}}$, and the coordinate map is a linear bijection.
Sketch. With $A = [T]_{\mathcal{B}\mathcal{B}}$, $T\mathbf{v} = \lambda\mathbf{v}$ if and only if $A[\mathbf{v}]_{\mathcal{B}} = \lambda[\mathbf{v}]_{\mathcal{B}}$, and $\mathbf{v} \neq \mathbf{0}$ if and only if $[\mathbf{v}]_{\mathcal{B}} \neq \mathbf{0}$. So T and A have the same eigenvalues, which are the roots of $p_A = p_T$.

Let $\mathcal{B}$ be a basis of V and $A = [T]_{\mathcal{B}\mathcal{B}}$, so $p_T = p_A$. For $\mathbf{v} \in V$, $[T\mathbf{v}]_{\mathcal{B}} = A[\mathbf{v}]_{\mathcal{B}}$ (Definition 2.3.1), and the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$ is a linear bijection (Proposition 1.6.20). So $T\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$ if and only if $A\mathbf{w} = \lambda\mathbf{w}$ with $\mathbf{w} = [\mathbf{v}]_{\mathcal{B}} \neq \mathbf{0}$. Hence T and A have the same eigenvalues, and Theorem 7.3.6 finishes the proof. □

Example 7.3.14 (Characteristic polynomials of operators).

(a) Differentiation D on $F[x]_{\leq n}$ has, in the basis $1, x, \ldots, x^n$, a strictly upper triangular matrix (it lowers degrees). So $p_D = x^{n+1}$ by Example 7.3.5(d), and 0 is its only eigenvalue, as found in Example 7.2.19.

(b) On the real vector space $W = \operatorname{span}(\cos, \sin)$ of functions $\mathbb{R} \to \mathbb{R}$, differentiation maps $\cos \mapsto -\sin$ and $\sin \mapsto \cos$. Its matrix in the basis cos, sin is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, with characteristic polynomial $x^2 + 1$. So it has no real eigenvalue: no nonzero $a\cos + b\sin$ is proportional to its own derivative.

(c) The cyclic permutation $Z(x, y, z) = (z, x, y)$ of Example 7.2.36 has matrix $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ in the standard basis. Expanding along the first row,

$$p_Z = \det\begin{pmatrix} x & 0 & -1 \\ -1 & x & 0 \\ 0 & -1 & x \end{pmatrix} = x \cdot x^2 + (-1) \cdot (-1)^{1+3} \cdot ((-1)(-1) - x \cdot 0) = x^3 - 1.$$

The converse of Proposition 7.3.11 is false. The matrices I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ both have characteristic polynomial $(x-1)^2$, but the only matrix similar to I_2 is $S^{-1} I_2 S = I_2$ itself. So the characteristic polynomial is one more similarity invariant to try (Technique 4.5.29), stronger than trace and determinant together, but not a complete one. Chapter 10 finds the complete invariant over $\mathbb{C}$ (Section 10.3).

Block-triangular matrices and invariant subspaces

Proposition 7.3.15 (Block-triangular matrices). Let $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in F^{n \times n}$, where $A \in F^{k \times k}$ and $D \in F^{(n-k)\times(n-k)}$ with $1 \leq k < n$. Then $p_M = p_A p_D$. The same holds for $\begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$.

Proof. ROUTINE

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Definition 7.3.4 (Characteristic polynomial)
Hint 1. Write $xI_n - M$ in block form. Is it still block triangular?
Hint 2. Apply Proposition 7.3.2(e) over the ring $F[x]$.
Sketch. $xI_n - M = \begin{pmatrix} xI_k - A & -B \\ 0 & xI_{n-k} - D \end{pmatrix}$ is block upper triangular in $F[x]^{n\times n}$, so its determinant is $p_A p_D$ by Proposition 7.3.2(e). The lower case is the same, or follows by transposing.

In $F[x]^{n\times n}$,

$$xI_n - M = \begin{pmatrix} xI_k - A & -B \\ 0 & xI_{n-k} - D \end{pmatrix},$$

which is block upper triangular with diagonal blocks $xI_k - A$ and $xI_{n-k} - D$. By Proposition 7.3.2(e), $p_M = \det(xI_k - A)\det(xI_{n-k} - D) = p_A p_D$. For $\begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$, the matrix $xI_n - M$ is block lower triangular, and Proposition 7.3.2(e) applies again. □

Combined with the block form of an operator with an invariant subspace (Remark 7.2.5), this says that an invariant subspace splits off a factor of the characteristic polynomial.

Corollary 7.3.16 (Invariant subspaces give factors). Let V be finite-dimensional, $T \in \mathrm{L}(V)$, and U a T-invariant subspace with $\dim U \geq 1$. Then $p_{T|_U}$ divides p_T.

Proof. ROUTINE

Uses. Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Proposition 7.3.15 (Block-triangular matrices), Theorem 4.1.16 (Extending to a basis), Definition 7.3.12 (Characteristic polynomial of an operator)
Techniques. Choose a basis adapted to a preserved subspace (Technique 4.5.36)
Hint 1. Choose a basis of V that starts with a basis of U. What does the matrix of T look like?
Hint 2. Proposition 4.5.35 gives $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ with $A = [T|_U]$. Now use Proposition 7.3.15.

Sketch. If $U = V$ there is nothing to prove. Otherwise extend a basis of U to a basis $\mathcal{B}$ of V; then $[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ with A the matrix of $T{\restriction_U}$ (Proposition 4.5.35), and $p_T = p_A p_D = p_{T{\restriction_U}} p_D$.

If $U = V$, then $T{\restriction_U} = T$ and there is nothing to prove. Otherwise let $k = \dim U$ and $n = \dim V$, so $1 \le k < n$. Extend a basis $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of U to a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of V (Theorem 4.1.16). By Proposition 4.5.35,

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \quad \text{with } A = [T{\restriction_U}]_{\mathcal{B}_U \mathcal{B}_U}.$$

By Proposition 7.3.15, $p_T = p_A p_D = p_{T{\restriction_U}} p_D$. □

Exercise 7.3.17. ROUTINE Let $A \in F^{n\times n}$.

(a) Show that $p_{A^{\mathsf{T}}} = p_A$.

(b) For $F = \mathbb{C}$, let $\overline{A}$ be the entrywise conjugate of A, and $\overline{f}$ the polynomial whose coefficients are the conjugates of those of $f \in \mathbb{C}[x]$. Show that $p_{\overline{A}} = \overline{p_A}$, and deduce that the eigenvalues of A^{H} are the conjugates of the eigenvalues of A.

Solution on page 1574

7.3.4 The coefficients: trace and determinant

Proposition 7.3.18 (The first and last coefficients). Let $A \in F^{n\times n}$. Then p_A is monic of degree n, the coefficient of x^{n-1} in p_A is $-\operatorname{tr} A$, and its constant term is $(-1)^n \det A$. That is, for $n \ge 3$,

$$p_A = x^n - (\operatorname{tr} A)x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_1 x + (-1)^n \det A$$

for some $c_1, \ldots, c_{n-2} \in F$, while $p_A = x^2 - (\operatorname{tr} A)x + \det A$ for $n = 2$ and $p_A = x - a_{11}$ for $n = 1$.

Proof. STANDARD

Uses. Definition 7.3.1 (Determinant over a commutative ring), Proposition 7.3.2 (Determinants over a commutative ring), Definition 2.4.26 (Trace), Proposition 7.1.4 (Degree of a product), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t)

Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)

Hint 1. Each entry of $xI_n - A$ has degree at most 1, and only the diagonal entries involve x. In the Leibniz term of σ, which factors have degree 1?

Hint 2. Only the factors $x - a_{jj}$ with $\sigma(j) = j$. A permutation $\sigma \neq \mathrm{id}$ moves at least two points. What does that say about the degree of its term?

Hint 3. So x^n and x^{n-1} come only from $(x - a_{11})\cdots(x - a_{nn})$. For the constant term, evaluate at 0 with (7.3.1).

Sketch. The Leibniz term of σ is a product of n factors of degree at most 1, of which only those with $\sigma(j) = j$ have degree 1. A permutation other than id fixes at most $n - 2$ points, so its term has degree at most $n - 2$. Thus the coefficients of x^n and x^{n-1} are those of $\prod_j (x - a_{jj}) = x^n - \left(\sum_j a_{jj}\right)x^{n-1} + \cdots$. The constant term is $p_A(0) = \det(-A) = (-1)^n \det A$.

Let $M = xI_n - A$, so $m_{jj} = x - a_{jj}$ and $m_{ij} = -a_{ij}$ for $i \neq j$. The Leibniz term of $\sigma \in S_n$ is

$$t_\sigma = \operatorname{sgn}(\sigma) \prod_{j=1}^{n} m_{\sigma(j)j}.$$

Each factor has degree at most 1, with degree 1 exactly when $\sigma(j) = j$. So t_σ is $\mathbf{0}$ or has degree at most the number of fixed points of σ (Proposition 7.1.4). If $\sigma \neq \mathrm{id}$, pick j with $\sigma(j) \neq j$; then $i = \sigma(j)$ is also moved, since $\sigma(i) = i = \sigma(j)$ would contradict injectivity. So σ fixes at most $n - 2$ points, and t_σ contributes nothing to the coefficients of x^n and x^{n-1}.

These two coefficients therefore come from

$$t_{\mathrm{id}} = (x - a_{11})(x - a_{22})\cdots(x - a_{nn}).$$

Expanding by distributivity, a product of n factors $x - a_{jj}$ is a sum over the ways to choose x or $-a_{jj}$ from each factor. The only way to get x^n is to choose x every time, and the ways to get x^{n-1} are to choose $-a_{jj}$ from exactly one factor j. So the coefficient of x^n is 1 and that of x^{n-1} is $-(a_{11} + \cdots + a_{nn}) = -\operatorname{tr} A$. In particular p_A is monic of degree n.

The constant term of a polynomial is its value at 0. By (7.3.1), $p_A(0) = \det(-A)$, and multiplying each of the n columns of A by -1 multiplies det by $(-1)^n$ (Proposition 7.3.2(b)). So the constant term is $(-1)^n \det A$. □

This is the computation in the proof of Proposition 5.3.19, done in $F[x]$ instead of for each scalar t. The coefficients in between are sums of minors (Exercise 7.3.53).

Remark 7.3.19 (Trace and determinant from the roots). Suppose p_A splits over F, say $p_A = (x - \lambda_1)\cdots(x - \lambda_n)$ with $\lambda_i \in F$. By Proposition 7.1.36, the coefficient of x^{n-1} in this product is $-(\lambda_1 + \cdots + \lambda_n)$ and its constant term is $(-1)^n \lambda_1 \cdots \lambda_n$. Comparing with Proposition 7.3.18 gives

$$\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n, \quad \det A = \lambda_1 \lambda_2 \cdots \lambda_n.$$

The λ_i are the eigenvalues of A (Theorem 7.3.6), each repeated as often as its multiplicity as a root of p_A, its algebraic multiplicity (Definition 7.4.11). Repetitions matter: $\left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ has the single eigenvalue 2 but trace $4 = 2 + 2$. Over $\mathbb{C}$, p_A always splits (Corollary 7.1.35), and Theorem 7.5.22 proves the same statement a second way, by triangularizing A.

For 2×2 matrices the trace and determinant determine $p_A = x^2 - (\operatorname{tr} A)x + \det A$, and with it the eigenvalues. Over $\mathbb{R}$, the quadratic formula says that the number of real eigenvalues is governed by the discriminant

$$\Delta = (\operatorname{tr} A)^2 - 4 \det A :$$

two real eigenvalues if $\Delta > 0$, one repeated real eigenvalue if $\Delta = 0$, and a pair of conjugate nonreal eigenvalues if $\Delta < 0$ (Theorem B.1.30).

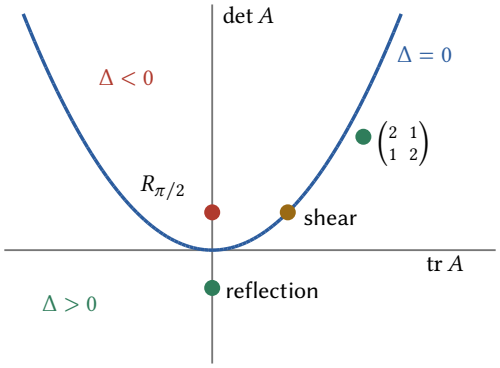


Figure 7.3.1. Real 2×2 matrices placed by $(\operatorname{tr} A, \det A)$, with $\Delta = (\operatorname{tr} A)^2 - 4 \det A$. Below the parabola $\det A = (\operatorname{tr} A)^2/4$ there are two real eigenvalues, and above it none. The rotation $R_{\pi/2}$ sits at $(0, 1)$, above the parabola. The matrix $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ of Example 7.2.8 sits at $(4, 3)$ and the reflection $\operatorname{diag}(1, -1)$ at $(0, -1)$, both below. The shear $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ sits on the parabola at $(2, 1)$, as do I_2 and the matrix of Example 7.2.39 (at $(4, 4)$). A point on the parabola can be a multiple of I_2 or not: the plane cannot tell them apart.

Exercise 7.3.20. ROUTINE Let $A \in \mathbb{R}^{3\times 3}$ have $\operatorname{tr} A = 2$ and $\det A = -2$, and suppose $A - I_3$ is not invertible and p_A splits over $\mathbb{R}$. Find p_A and the eigenvalues of A.

Solution on page 1575

Products in either order

For square matrices, AB and BA are usually different, and they need not be similar. But their characteristic polynomials always agree.

Technique 7.3.21 (Compare AB and BA). The products AB and BA are close relatives. If A is invertible, they are similar: $BA = A^{-1}(AB)A$. In general they are not necessarily similar, but they still share the characteristic polynomial (when square) and the nonzero eigenvalues (always). Three ways to transfer information between them are useful. Conjugate by A when A is invertible, and reduce the general case to it by a polynomial identity in a perturbation parameter. Or embed both products in block matrices of twice the size that are similar. Or transport eigenvectors directly: if $AB\mathbf{v} = \lambda\mathbf{v}$ with $\lambda \neq 0$, then $B\mathbf{v} \neq \mathbf{0}$ and $BA(B\mathbf{v}) = \lambda B\mathbf{v}$.

Reach for it whenever a question involves a product in both orders, such as $A^{\mathsf{T}}A$ and AA^{T}, or $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ and $\mathbf{v}^{\mathsf{T}}\mathbf{u}$. The canonical example is Proposition 7.3.22.

Used in. Proposition 7.3.22, Exercise 7.3.40, Exercise 7.3.41, Exercise 7.7.10, Proposition 8.3.20.

Proposition 7.3.22 (AB and BA). For $A, B \in F^{n\times n}$, $p_{AB} = p_{BA}$. In particular AB and BA have the same eigenvalues.

Intuition. If A is invertible, then $BA = A^{-1}(AB)A$ is similar to AB, and we are done by Proposition 7.3.11. If A is not invertible, AB and BA need not be similar: for $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $AB = B \neq 0$ while $BA = 0$. Two ideas get around this. The first replaces A by the invertible matrices $A + tI_n$, which differ from A by a parameter t that we then set to 0; this needs the identity to be polynomial in t, and enough values of t to conclude. The second finds a pair of larger matrices, built from AB and BA, that are similar for every A.

Proof (by block matrices). STANDARD

Uses. Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Proposition 7.3.15 (Block-triangular matrices), Theorem 2.4.34 (Block multiplication), Corollary 7.1.7 (No zero divisors, cancellation, and units)
Techniques. Compare AB and BA (Technique 7.3.21), Block decomposition (Technique 2.4.35)
Hint 1. Look for a $2n \times 2n$ matrix S with $\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} S = S \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$. Try S block upper triangular with identity blocks on the diagonal.
Hint 2. $S = \begin{pmatrix} I_n & A \\ 0 & I_n \end{pmatrix}$ works. Both block matrices are block triangular; compute their characteristic polynomials.
Hint 3. You get $x^n p_{AB} = x^n p_{BA}$. Cancel x^n in $F[x]$.
Sketch. With $X = \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$, and $S = \begin{pmatrix} I_n & A \\ 0 & I_n \end{pmatrix}$, block multiplication gives $XS = \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix} = SY$, and S is invertible, so X and Y are similar. By Proposition 7.3.15, $p_X = p_{AB}x^n$ and $p_Y = x^n p_{BA}$. Cancel x^n (Corollary 7.1.7).

Consider the $2n \times 2n$ matrices

$$X = \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}, \quad S = \begin{pmatrix} I_n & A \\ 0 & I_n \end{pmatrix}.$$

By block multiplication (Theorem 2.4.34),

$$XS = \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix} = SY.$$

The matrix S is invertible, with inverse $\begin{pmatrix} I_n & -A \\ 0 & I_n \end{pmatrix}$ (multiply out). So $Y = S^{-1}XS$, and $p_X = p_Y$ by Proposition 7.3.11. Both matrices are block triangular, so by Proposition 7.3.15,

$$p_X = p_{AB}p_{0_{n\times n}} = p_{AB}x^n, \quad p_Y = p_{0_{n\times n}}p_{BA} = x^n p_{BA}.$$

So $x^n p_{AB} = x^n p_{BA}$, and since $x^n \neq \mathbf{0}$, cancellation in $F[x]$ (Corollary 7.1.7(b)) gives $p_{AB} = p_{BA}$. The statement about eigenvalues follows from Theorem 7.3.6. □

Proof (by perturbation, when F is infinite). HARD

> *Uses.* Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Proposition 5.3.19 ($\det(A+tI_n)$ as a function of t), Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.1.22 (Counting roots), Corollary 7.1.23 (Polynomials and their values), Theorem 5.4.7 (Invertibility and the determinant)
> *Techniques.* Compare AB and BA (Technique 7.3.21), Polynomial identity arguments (Technique 5.5.24), det is a polynomial in the entries (Technique 5.3.18)
> *Hint 1.* Fix $\lambda \in F$. For a scalar s with $A+sI_n$ invertible, compare $\det(\lambda I_n - (A+sI_n)B)$ with $\det(\lambda I_n - B(A+sI_n))$.
> *Hint 2.* They are equal, by similarity. Now regard s as an indeterminate t: both sides are values of polynomials in t, and their difference has infinitely many roots.
> *Hint 3.* Set $t=0$ to get $p_{AB}(\lambda) = p_{BA}(\lambda)$ for every λ, and use that F is infinite once more.
> *Sketch.* Fix λ. In $F[t]$ let $g = \det(\lambda I_n - (A+tI_n)B) - \det(\lambda I_n - B(A+tI_n))$. For each $s \in F$ with $A+sI_n$ invertible, $g(s) = 0$ by similarity; this excludes at most n values of s. So g has infinitely many roots and $g = \mathbf{0}$. Then $g(0) = p_{AB}(\lambda) - p_{BA}(\lambda) = 0$ for every $\lambda \in F$, and Corollary 7.1.23 finishes.

Assume F is infinite, and fix $\lambda \in F$. Work in $F[t]$, polynomials in a new indeterminate t, and let

$$g = \det(\lambda I_n - (A+tI_n)B) - \det(\lambda I_n - B(A+tI_n)) \in F[t].$$

For $s \in F$, evaluating every entry at $t = s$ (Proposition 7.3.2(g)) gives

$$g(s) = \det(\lambda I_n - (A+sI_n)B) - \det(\lambda I_n - B(A+sI_n)).$$

If $A+sI_n$ is invertible, then $B(A+sI_n) = (A+sI_n)^{-1}((A+sI_n)B)(A+sI_n)$, so the two matrices $(A+sI_n)B$ and $B(A+sI_n)$ are similar, and their characteristic polynomials agree at λ (Proposition 7.3.11). So $g(s) = 0$.

By Proposition 5.3.19 and Theorem 5.4.7, $A+sI_n$ fails to be invertible only for roots s of a monic polynomial of degree n, that is, for at most n values of s (Proposition 7.1.22). Since F is infinite, g has infinitely many roots, so $g = \mathbf{0}$, again by Proposition 7.1.22. In particular $g(0) = 0$, which says $p_{AB}(\lambda) = p_{BA}(\lambda)$ by (7.3.1). This holds for every $\lambda \in F$, so $p_{AB} = p_{BA}$ by Corollary 7.1.23. □

The first proof is shorter and works over every field. The second shows a pattern that recurs: prove an identity for invertible matrices, where it is easy, and extend it to all matrices because both sides are polynomials. Over a finite field there are too few values of t, and this pattern alone does not suffice.

7.3.5 Companion matrices

Which polynomials occur as characteristic polynomials? Every characteristic polynomial of an $n \times n$ matrix is monic of degree n. Conversely, every such polynomial occurs, and the matrix realizing it already appeared twice: in Proposition 5.5.20 as a determinant computation, and after Proposition 7.2.35 as the matrix of an operator on a Krylov subspace.

Definition 7.3.23 (Companion matrix). Let $p = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$ be monic of degree $n \geq 1$. The **companion matrix** of p is

$$C_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in F^{n\times n},$$

whose column k is $\mathbf{e}_{k+1}$ for $k < n$ and whose last column is $-(a_0, a_1, \ldots, a_{n-1})$.

For example, $C_{x-c} = (c)$, the companion matrix of x^2+1 is the rotation $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, and the companion matrix of $x^3 - 1$ is the matrix of the cyclic permutation Z in Example 7.3.14. In general $C_p\mathbf{e}_k = \mathbf{e}_{k+1}$ for $k < n$: C_p shifts the standard basis forward, and the last basis vector goes to $C_p\mathbf{e}_n = -(a_0\mathbf{e}_1 + \cdots + a_{n-1}\mathbf{e}_n)$. Some books use the transpose C_p^{T} instead; it has the same characteristic polynomial (Exercise 7.3.17).

Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial). Let $p \in F[x]$ be monic of degree $n \geq 1$. Then $p_{C_p} = p$.

Proof (by rerunning Proposition 5.5.20). STANDARD

Uses. Proposition 5.5.20 (Companion matrices), Proposition 7.3.2 (Determinants over a commutative ring)
Hint 1. Proposition 5.5.20 proves $\det(tI_n - C) = p(t)$ for every scalar $t \in F$. Over a finite field this is weaker than an identity in $F[x]$. Which tools did its first proof use?
Hint 2. It used cofactor expansion along the first row, a triangular determinant, and induction on n. All of these are available over $F[x]$ by Proposition 7.3.2.
Sketch. The first proof of Proposition 5.5.20 expands $\det\big(tI_n - C_p\big)$ along the first row, recognizes one minor as the same kind of matrix for $q = (p - a_0)/x$ and the other as triangular with -1's on the diagonal, and inducts. Each step is valid with t replaced by the indeterminate x, by Proposition 7.3.2(e) and (f). So $\det\big(xI_n - C_p\big) = xq + a_0 = p$ in $F[x]$.

The first proof of Proposition 5.5.20 computes $\det\big(tI_n - C_p\big)$ by induction on n. It expands along the first row, identifies the minor at position $(1,1)$ as $tI_{n-1} - C_q$ for $q = a_1 + a_2x + \cdots + a_{n-1}x^{n-2} + x^{n-1}$, and the minor at position $(1,n)$ as an upper triangular matrix with diagonal entries -1. It then concludes $\det\big(tI_n - C_p\big) = tq(t) + a_0 = p(t)$. Every step uses only cofactor expansion, the determinant of a triangular matrix, and ring arithmetic. By Proposition 7.3.2(e) and (f) these hold in $F[x]^{n\times n}$, so the same induction, with the indeterminate x in place of the scalar t, proves $\det\big(xI_n - C_p\big) = xq + a_0 = p$ in $F[x]$. □

Proof (by expanding along the last column). STANDARD

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Definition 7.3.23 (Companion matrix)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. The last column of $xI_n - C_p$ is $(a_0, a_1, \ldots, a_{n-2}, x + a_{n-1})$. Expand along it.
Hint 2. Deleting row k and column n leaves a block-diagonal matrix: a lower triangular block of size $k-1$ with x on the diagonal, and an upper triangular block of size $n-k$ with -1 on the diagonal.
Hint 3. So the minor is $x^{k-1}(-1)^{n-k}$, and the sign $(-1)^{k+n}$ cancels the $(-1)^{n-k}$.
Sketch. Let $M = xI_n - C_p$. The minor $M_{\widehat{kn}}$ is block diagonal, with a lower triangular $(k-1)\times(k-1)$ block with diagonal x and an upper triangular $(n-k)\times(n-k)$ block with diagonal -1, so $\det M_{\widehat{kn}} = x^{k-1}(-1)^{n-k}$. Expanding along column n: $\det M = \sum_k (-1)^{k+n} m_{kn} x^{k-1}(-1)^{n-k} = \sum_k m_{kn}x^{k-1} = a_0 + a_1x + \cdots + a_{n-2}x^{n-2} + (x + a_{n-1})x^{n-1} = p$.

For $n = 1$, $C_p = (-a_0)$ and $\det\big(xI_1 - C_p\big) = x + a_0 = p$. Let $n \geq 2$ and $M = xI_n - C_p$. Row k of M has -1 in column $k-1$ (if $k \geq 2$), x in column k (if $k \leq n-1$), and $m_{kn} = a_{k-1}$ in column n, except that $m_{nn} = x + a_{n-1}$; all other entries are 0. Fix k and delete row k and column n.

- Rows $1, \ldots, k-1$ have their nonzero entries (outside column n) in columns $1, \ldots, k-1$. They form a $(k-1)\times(k-1)$ lower triangular block L with diagonal entries x.
- Rows $k+1, \ldots, n$ have their nonzero entries (outside column n) in columns $k, \ldots, n-1$: row i has -1 in column $i-1$ and x in column i (if $i \leq n-1$). They form an $(n-k)\times(n-k)$ upper triangular block U with diagonal entries -1.

So $M_{\widehat{kn}} = \begin{pmatrix} L & 0 \\ 0 & U \end{pmatrix}$, and by Proposition 7.3.2(e), $\det M_{\widehat{kn}} = x^{k-1}(-1)^{n-k}$. (For $k = 1$ there is no block L, and for $k = n$ no block U; the formula still holds.) Expanding along column n (Proposition 7.3.2(f)),

$$\begin{aligned} \det M &= \sum_{k=1}^{n} (-1)^{k+n} m_{kn} x^{k-1}(-1)^{n-k} = \sum_{k=1}^{n} m_{kn}x^{k-1} \\ &= a_0 + a_1x + \cdots + a_{n-2}x^{n-2} + (x + a_{n-1})x^{n-1} = p, \end{aligned}$$

since $(-1)^{k+n}(-1)^{n-k} = (-1)^{2n} = 1$. □

The first proof builds p from the inside out, as in Horner's scheme. The second reads all coefficients off at once: the cofactor of the entry in row k of the last column is exactly x^{k-1}, so the last column of $xI_n - C_p$ is the list of coefficients of p and the expansion reassembles p.

Example 7.3.25 (Matrices with prescribed eigenvalues).

(a) The polynomial $p = x^3 - 2$ has no root in $\mathbb{Q}$ (a rational root would be an integer dividing 2, by the argument of Exercise 7.3.42, and none of $\pm 1, \pm 2$ works). So $C_p = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$ has no eigenvalues in $\mathbb{Q}$, although over $\mathbb{R}$ it has the eigenvalue $\sqrt[3]{2}$.

(b) To get a real 2×2 matrix with eigenvalues 3 and -1 that is neither diagonal nor triangular, take $p = (x-3)(x+1) = x^2 - 2x - 3$ and $C_p = \begin{pmatrix} 0 & 3 \\ 1 & 2 \end{pmatrix}$.

The connection with Krylov subspaces gives the companion matrix its meaning. By Proposition 7.2.35, if $\mathbf{v} \neq \mathbf{0}$ and $K_T(\mathbf{v})$ is finite-dimensional, then $\mathbf{v}, T\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$ is a basis of $K_T(\mathbf{v})$, and

$$T^d\mathbf{v} = -\big(c_0\mathbf{v} + c_1 T\mathbf{v} + \cdots + c_{d-1}T^{d-1}\mathbf{v}\big).$$

In this basis, $T\restriction_{K_T(\mathbf{v})}$ moves each basis vector to the next and sends the last one to the vector with coordinates $-(c_0, \ldots, c_{d-1})$. That is the description of C_p for $p = x^d + c_{d-1}x^{d-1} + \cdots + c_0$.

Proposition 7.3.26 (Krylov subspaces and the characteristic polynomial). Let V be finite-dimensional, $T \in \mathrm{L}(V)$, and $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$. Let d and the monic polynomial p of degree d be as in Proposition 7.2.35, so that $p(T)\mathbf{v} = \mathbf{0}$. Then

(a) the matrix of $T\restriction_{K_T(\mathbf{v})}$ in the basis $\mathbf{v}, T\mathbf{v}, \ldots, T^{d-1}\mathbf{v}$ is C_p;

(b) $p_{T\restriction_{K_T(\mathbf{v})}} = p$;

(c) p divides p_T.

Proof. ROUTINE

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Proposition 7.2.34 (The smallest invariant subspace containing a vector), Definition 2.3.1 (Matrix of a linear map), Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial), Corollary 7.3.16 (Invariant subspaces give factors)
Hint 1. Compute T on each basis vector $T^i\mathbf{v}$ and write the result in the basis.
Hint 2. For (b) use Theorem 7.3.24, and for (c) use that $K_T(\mathbf{v})$ is T-invariant.
Sketch. $T\big(T^i\mathbf{v}\big) = T^{i+1}\mathbf{v}$ for $i < d-1$, and $T\big(T^{d-1}\mathbf{v}\big) = -\sum_i c_i T^i\mathbf{v}$, so the columns are $\mathbf{e}_2, \ldots, \mathbf{e}_d$ and $-(c_0, \ldots, c_{d-1})$, which is C_p. Then the characteristic polynomial is $p_{C_p} = p$, and p divides p_T because $K_T(\mathbf{v})$ is invariant (Corollary 7.3.16).

Write $p = x^d + c_{d-1}x^{d-1} + \cdots + c_0$ and $\mathbf{b}_{i+1} = T^i\mathbf{v}$ for $0 \leq i \leq d-1$. For $i < d-1$, $T\mathbf{b}_{i+1} = T^{i+1}\mathbf{v} = \mathbf{b}_{i+2}$, and $T\mathbf{b}_d = T^d\mathbf{v} = -(c_0\mathbf{b}_1 + \cdots + c_{d-1}\mathbf{b}_d)$ because $p(T)\mathbf{v} = \mathbf{0}$. So column k of the matrix of $T\restriction_{K_T(\mathbf{v})}$ is $\mathbf{e}_{k+1}$ for $k < d$, and the last column is $-(c_0, \ldots, c_{d-1})$ (Definition 2.3.1). This is C_p, proving (a). Then (b) follows from (a), Definition 7.3.12, and Theorem 7.3.24. Since $K_T(\mathbf{v})$ is T-invariant (Proposition 7.2.34) and has dimension $d \geq 1$, (c) follows from (b) and Corollary 7.3.16. □

So every nonzero vector produces a factor of p_T, namely the polynomial that records the first dependence among $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$. In Example 7.2.36, the vectors $(1,1,1)$, $(1,-1,0)$, and $\mathbf{e}_1$ produced $x-1$, x^2+x+1, and x^3-1, and each divides $p_Z = x^3 - 1$. When $K_T(\mathbf{v}) = V$, part (b) says $p_T = p$ (Exercise 7.3.50).

7.3.6 Eigenvalues exist over $\mathbb{C}$

Example 7.2.18 showed that a real operator need not have an eigenvalue, and Example 7.2.20 that an operator on an infinite-dimensional space need not either. Neither obstruction survives over $\mathbb{C}$ in finite dimensions.

Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$). Let V be a finite-dimensional complex vector space with $V \neq \{\mathbf{0}\}$, and let $T \in \mathrm{L}(V)$. Then T has an eigenvalue. In particular, every $A \in \mathbb{C}^{n\times n}$ with $n \geq 1$ has an eigenvalue.

The two proofs below locate the eigenvalue in two different polynomials. The first uses p_T, which collects all eigenvalues at once and needs determinants. The second uses the polynomial of a single Krylov sequence, which only needs linear dependence.

Proof (via the characteristic polynomial). ROUTINE

> *Uses.* Proposition 7.3.18 (The first and last coefficients), Theorem 7.1.32 (Fundamental Theorem of Algebra), Corollary 7.3.13 (Eigenvalues of an operator)
> *Hint 1.* What is the degree of p_T?
> *Hint 2.* A polynomial of degree at least 1 over $\mathbb{C}$ has a root (Theorem 7.1.32).
> *Sketch.* p_T has degree $\dim V \geq 1$, so it has a root $\lambda \in \mathbb{C}$ by the FTA, and λ is an eigenvalue by Corollary 7.3.13.

Let $n = \dim V \geq 1$. By Proposition 7.3.18, p_T is monic of degree n, in particular nonconstant. By Theorem 7.1.32 it has a root $\lambda \in \mathbb{C}$, and λ is an eigenvalue of T by Corollary 7.3.13. □

> *Intuition (A proof without determinants).* Take any $\mathbf{v} \neq \mathbf{0}$ and look at $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$. In a space of dimension n, the $n+1$ vectors $\mathbf{v}, \ldots, T^n\mathbf{v}$ are dependent, so some nonzero polynomial q has $q(T)\mathbf{v} = \mathbf{0}$. Over $\mathbb{C}$ we can factor $q = c(x-\lambda_1)\cdots(x-\lambda_m)$, so
>
> $$c(T - \lambda_1 \operatorname{id})(T - \lambda_2 \operatorname{id})\cdots(T - \lambda_m \operatorname{id})\mathbf{v} = \mathbf{0}$$
>
> with $\mathbf{v} \neq \mathbf{0}$. A composite of injective maps is injective, so some factor $T - \lambda_j \operatorname{id}$ is not injective, and λ_j is an eigenvalue. If we use the *first* dependence, the polynomial p of Proposition 7.2.35, we can even say which vector is the eigenvector: split off one linear factor, $p = (x-\lambda)q$, and apply $q(T)$ to $\mathbf{v}$. The result is nonzero because q has smaller degree than p, and $T - \lambda \operatorname{id}$ kills it.

> **Technique 7.3.28** (Krylov sequences must become dependent). In a finite-dimensional space, the sequence $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ eventually becomes dependent, and the first dependence is a monic polynomial p with $p(T)\mathbf{v} = \mathbf{0}$ and no nonzero polynomial of smaller degree killing $\mathbf{v}$ (Proposition 7.2.35). This converts a question about T into a question about one polynomial: factor p, and use the factors to build vectors. If $p = fq$ with $\deg f \geq 1$, then $\mathbf{w} = q(T)\mathbf{v}$ is nonzero and $f(T)\mathbf{w} = \mathbf{0}$, so a factor $f = x - \lambda$ produces an eigenvector, and a quadratic factor produces a two-dimensional invariant subspace (Theorem 7.3.31).
>
> Reach for it when you need an eigenvector or a small invariant subspace and want to avoid determinants, or when V is infinite-dimensional but some Krylov subspace is not. The canonical example is the second proof of Theorem 7.3.27.
> *Used in.* Theorem 7.3.27, Theorem 7.3.31, Exercise 7.3.45, Exercise 7.3.46, Exercise 7.3.51, Lemma 7.6.21, Exercise 7.6.47, Problem 7.9, Problem 7.15, Theorem 8.1.3.

Proof (via a Krylov sequence, without determinants). STANDARD

> *Uses.* Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.1.32 (Fundamental Theorem of Algebra), Theorem 7.1.20 (Factor theorem), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 7.2.6 (Eigenvalue, eigenvector)
> *Techniques.* Krylov sequences must become dependent (Technique 7.3.28), Too many vectors force a dependence (Technique 4.1.30)
> *Hint 1.* Fix $\mathbf{v} \neq \mathbf{0}$. The sequence $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ cannot stay independent. Use the first dependence, from Proposition 7.2.35.
> *Hint 2.* This gives a monic p of degree $d \geq 1$ with $p(T)\mathbf{v} = \mathbf{0}$, and no nonzero polynomial of degree less than d kills $\mathbf{v}$. Use the FTA to split off a linear factor $x - \lambda$.
> *Hint 3.* Write $p = (x-\lambda)q$. Then $q(T)\mathbf{v} \neq \mathbf{0}$ (why?) and $(T - \lambda \operatorname{id})q(T)\mathbf{v} = p(T)\mathbf{v}$.
> *Sketch.* Let $\mathbf{v} \neq \mathbf{0}$ and let p, of degree $d \geq 1$, record the first dependence among $\mathbf{v}, T\mathbf{v}, \ldots$ (Proposition 7.2.35). By the FTA and the factor theorem, $p = (x-\lambda)q$ with q monic of degree $d-1$. Then $\mathbf{w} = q(T)\mathbf{v} \neq \mathbf{0}$ by the minimality of d, and $(T - \lambda \operatorname{id})\mathbf{w} = p(T)\mathbf{v} = \mathbf{0}$. So $\mathbf{w}$ is an eigenvector for λ.

Choose $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$. Since V is finite-dimensional, so is $K_T(\mathbf{v})$, and Proposition 7.2.35 gives $d \geq 1$ and a monic polynomial $p \in \mathbb{C}[x]$ of degree d with $p(T)\mathbf{v} = \mathbf{0}$, such that no nonzero polynomial of degree less than d sends $\mathbf{v}$ to $\mathbf{0}$.

As $\deg p = d \geq 1$, the FTA (Theorem 7.1.32) gives a root $\lambda \in \mathbb{C}$ of p, and the factor theorem (Theorem 7.1.20) gives $p = (x - \lambda)q$ for some $q \in \mathbb{C}[x]$. Comparing leading coefficients and degrees, q is monic of degree $d - 1$; in particular $q \neq \mathbf{0}$. Let $\mathbf{w} = q(T)\mathbf{v}$. Since q is nonzero of degree less than d, $\mathbf{w} \neq \mathbf{0}$. By Proposition 7.2.27(a),

$$(T - \lambda\,\mathrm{id})\mathbf{w} = ((x - \lambda)q)(T)\mathbf{v} = p(T)\mathbf{v} = \mathbf{0}.$$

So $T\mathbf{w} = \lambda\mathbf{w}$ with $\mathbf{w} \neq \mathbf{0}$, and λ is an eigenvalue of T (Definition 7.2.6). □

Example 7.3.29 (Finding an eigenvector from a Krylov sequence). Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ and $\mathbf{v} = \mathbf{e}_1$. Then $A\mathbf{v} = \mathbf{e}_2$, which is not a multiple of $\mathbf{e}_1$, and $A^2\mathbf{v} = -\mathbf{e}_1 = -\mathbf{v}$. So the first dependence is $A^2\mathbf{v} + \mathbf{v} = \mathbf{0}$, and $p = x^2 + 1 = (x - i)(x + i)$. Splitting off $x - i$ leaves $q = x + i$, and

$$\mathbf{w} = (A + iI_2)\mathbf{e}_1 = (i, 1), \quad A\mathbf{w} = (-1, i) = i(i, 1).$$

Splitting off $x + i$ instead gives $(A - iI_2)\mathbf{e}_1 = (-i, 1)$, an eigenvector for $-i$. For the cyclic permutation Z and $\mathbf{v} = \mathbf{e}_1$, the first dependence is $x^3 - 1 = (x - 1)\big(x^2 + x + 1\big)$ (Example 7.2.36), and $\big(Z^2 + Z + \mathrm{id}\big)\mathbf{e}_1 = (1, 1, 1)$ is the eigenvector for 1.

Intuition (Comparing the two proofs). The first proof is shorter, but it rests on determinants over $F[x]$, and it finds an eigenvalue without an eigenvector. The second proof needs only linear dependence and polynomial arithmetic. It also uses less: it never needs all of V, only the finite-dimensional subspace $K_T(\mathbf{v})$. So it shows, for instance, that an operator on an infinite-dimensional complex space has an eigenvalue as soon as some nonzero vector has a finite-dimensional Krylov subspace. On the other hand, the first proof locates every eigenvalue at once, with multiplicities, which the second does not.

Both proofs use exactly one property of $\mathbb{C}$: every nonconstant polynomial over $\mathbb{C}$ has a root. So both prove the theorem over any field with this property, and over $\mathbb{R}$ both fail at the same step, with the rotation as the witness: its characteristic polynomial is $x^2 + 1$, and so is the Krylov polynomial of every $\mathbf{v} \neq \mathbf{0}$. Over a finite field the theorem always fails (Exercise 7.3.44).

7.3.7 Real operators BEYOND 21-242

Real operators and complexification are covered in LADR and LADW (5.8), but not in 21-242 lecture.

Over $\mathbb{R}$ an operator may have no eigenvalue, but it cannot avoid small invariant subspaces altogether.

Proposition 7.3.30 (Odd dimension). Let V be a real vector space of odd dimension. Then every $T \in \mathrm{L}(V)$ has an eigenvalue.

Proof. ROUTINE

Uses. Proposition 7.3.18 (The first and last coefficients), Corollary 7.1.43 (Odd degree gives a real root), Corollary 7.3.13 (Eigenvalues of an operator)
Hint. What is the degree of p_T, and what do you know about real polynomials of that degree?
Sketch. $p_T \in \mathbb{R}[x]$ has odd degree $\dim V$, so it has a real root (Corollary 7.1.43), which is an eigenvalue.

By Proposition 7.3.18, $p_T \in \mathbb{R}[x]$ has degree $\dim V$, which is odd. By Corollary 7.1.43 it has a real root λ, and λ is an eigenvalue of T by Corollary 7.3.13. □

For example, every operator on $\mathbb{R}^3$ maps some line to itself; for a rotation this is its axis. In even dimension there may be no eigenvalue: the operator on $\mathbb{R}^4$ with matrix $\begin{pmatrix} R_{\pi/2} & 0 \\ 0 & R_{\pi/2} \end{pmatrix}$ has characteristic polynomial $\big(x^2 + 1\big)^2$ by Proposition 7.3.15. It still has invariant planes, such as $\mathrm{span}(\mathbf{e}_1, \mathbf{e}_2)$.

Theorem 7.3.31 (Small invariant subspaces). Let V be a finite-dimensional real vector space with $V \neq \{\mathbf{0}\}$, and let $T \in \mathrm{L}(V)$. Then T has an invariant subspace of dimension 1 or 2.

Proof (via a Krylov sequence). STANDARD

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.1.39 (Factorization over $\mathbb{R}$), Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 7.2.4 (Invariance is checked on a spanning list), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. Imitate the second proof of Theorem 7.3.27. What replaces the FTA over $\mathbb{R}$?
Hint 2. By Theorem 7.1.39, the polynomial p of the first dependence has a monic factor f of degree 1 or 2. Write $p = fq$ and set $\mathbf{w} = q(T)\mathbf{v}$.
Hint 3. If $f = x^2 + bx + c$, then $T^2\mathbf{w} = -bT\mathbf{w} - c\mathbf{w}$. Which subspace is invariant?
Sketch. Let p record the first dependence of a Krylov sequence. Over $\mathbb{R}$, p has a monic factor f of degree 1 or 2 (Theorem 7.1.39); write $p = fq$. Then $\mathbf{w} = q(T)\mathbf{v} \neq \mathbf{0}$ and $f(T)\mathbf{w} = \mathbf{0}$. If $f = x - \lambda$, $\mathbf{w}$ spans an invariant line. If $f = x^2 + bx + c$, then $T(T\mathbf{w}) = -bT\mathbf{w} - c\mathbf{w}$, so $\operatorname{span}(\mathbf{w}, T\mathbf{w})$ is invariant, of dimension 1 or 2.

Choose $\mathbf{v} \neq \mathbf{0}$ and let p, monic of degree $d \geq 1$, be as in Proposition 7.2.35. By Theorem 7.1.39, p is a product of monic factors of degree 1 and 2 (times its leading coefficient 1), so $p = fq$ with f monic of degree 1 or 2 and q monic of degree less than d. Let $\mathbf{w} = q(T)\mathbf{v}$. Then $\mathbf{w} \neq \mathbf{0}$ by the minimality in Proposition 7.2.35, and $f(T)\mathbf{w} = p(T)\mathbf{v} = \mathbf{0}$ by Proposition 7.2.27(a).

If $f = x - \lambda$, then $T\mathbf{w} = \lambda\mathbf{w}$, so $\operatorname{span}(\mathbf{w})$ is an invariant line (Proposition 7.2.7). If $f = x^2 + bx + c$, then $T^2\mathbf{w} + bT\mathbf{w} + c\mathbf{w} = \mathbf{0}$. Let $U = \operatorname{span}(\mathbf{w}, T\mathbf{w})$. Then $T\mathbf{w} \in U$ and $T(T\mathbf{w}) = -bT\mathbf{w} - c\mathbf{w} \in U$, so U is T-invariant by Proposition 7.2.4. It has dimension 1 or 2, since $\mathbf{w} \neq \mathbf{0}$ and it is spanned by two vectors. □

A second proof, through complex eigenvectors, follows in the next subsection.

7.3.8 Complexification BEYOND 21-242

To use complex eigenvalues for a real operator, we need a complex vector space for it to act on. For $\mathbb{R}^n$ the answer is $\mathbb{C}^n$: a real matrix is also a complex matrix. For an abstract real vector space V, we build the complex space by adjoining "imaginary parts", just as $\mathbb{C}$ is built from $\mathbb{R}$ (Definition 1.1.29).

Definition 7.3.32 (Complexification). Let V be a real vector space. Its **complexification** $V_{\mathbb{C}}$ is the set $V \times V$, whose element $(\mathbf{u}, \mathbf{v})$ we write $\mathbf{u} + i\mathbf{v}$, with addition $(\mathbf{u} + i\mathbf{v}) + (\mathbf{u}' + i\mathbf{v}') = (\mathbf{u} + \mathbf{u}') + i(\mathbf{v} + \mathbf{v}')$ and scalar multiplication by $a + bi \in \mathbb{C}$ $(a, b \in \mathbb{R})$ given by

$$(a + bi)(\mathbf{u} + i\mathbf{v}) = (a\mathbf{u} - b\mathbf{v}) + i(b\mathbf{u} + a\mathbf{v}).$$

We identify $\mathbf{u} \in V$ with $\mathbf{u} + i\mathbf{0} \in V_{\mathbb{C}}$. For $T \in \mathrm{L}(V)$, its **complexification** is $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$, $T_{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = T\mathbf{u} + iT\mathbf{v}$.

The formula for scalar multiplication is what we get by expanding $(a + bi)(\mathbf{u} + i\mathbf{v})$ formally with $i^2 = -1$. For $V = \mathbb{R}^n$, the map $\mathbf{u} + i\mathbf{v} \mapsto \mathbf{u} + i\mathbf{v} \in \mathbb{C}^n$ (with the right side computed coordinatewise) identifies $V_{\mathbb{C}}$ with $\mathbb{C}^n$, and $T_{\mathbb{C}}$ with the complex matrix of T.

Proposition 7.3.33 (Basic properties). Let V be a real vector space and $T \in \mathrm{L}(V)$.

(a) $V_{\mathbb{C}}$ is a complex vector space.

(b) If $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is a basis of V over $\mathbb{R}$, it is a basis of $V_{\mathbb{C}}$ over $\mathbb{C}$. In particular $\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim_{\mathbb{R}} V$.

(c) $T_{\mathbb{C}}$ is $\mathbb{C}$-linear, and $[T_{\mathbb{C}}]_{\mathcal{B}\mathcal{B}} = [T]_{\mathcal{B}\mathcal{B}}$ for every basis $\mathcal{B}$ of V. In particular $p_{T_{\mathbb{C}}} = p_T$ when V is finite-dimensional and nonzero.

Proof. STANDARD

Uses. Definition 7.3.32 (Complexification), Definition 1.2.1 (Vector space), Definition 1.6.1 (Basis), Definition 2.3.1 (Matrix of a linear map), Definition 7.3.12 (Characteristic polynomial of an operator)
Hint 1. For (a), most axioms hold coordinatewise. The interesting one is $(\alpha\beta)\mathbf{z} = \alpha(\beta\mathbf{z})$; expand both sides with $\alpha = a + bi$, $\beta = c + di$.
Hint 2. For (b), write $\mathbf{u} + i\mathbf{v}$ with $\mathbf{u} = \sum \alpha_k \mathbf{b}_k$ and $\mathbf{v} = \sum \beta_k \mathbf{b}_k$, and compute $\sum(\alpha_k + i\beta_k)\mathbf{b}_k$.
Hint 3. For (c), check $T_{\mathbb{C}}((a + bi)\mathbf{z}) = (a + bi)T_{\mathbb{C}}(\mathbf{z})$ directly, and compute $T_{\mathbb{C}}\mathbf{b}_j$.

Sketch. (a) Addition is that of $V \times V$; scalar multiplication satisfies $1\mathbf{z} = \mathbf{z}$, both distributive laws (by linearity of the formula in (a, b) and in $(\mathbf{u}, \mathbf{v})$), and associativity, which reduces to the multiplication rule of $\mathbb{C}$. (b) $\sum(\alpha_k + i\beta_k)\mathbf{b}_k = \sum \alpha_k \mathbf{b}_k + i\sum \beta_k \mathbf{b}_k$, so spanning and independence over $\mathbb{C}$ reduce to the same statements over $\mathbb{R}$ for the real and imaginary parts. (c) Expanding gives linearity; $T_{\mathbb{C}}\mathbf{b}_j = T\mathbf{b}_j$ has the same real coordinates.

(a) $V \times V$ is a real vector space under coordinatewise operations, so addition in $V_{\mathbb{C}}$ satisfies the additive axioms. For scalar multiplication, $1(\mathbf{u} + i\mathbf{v}) = \mathbf{u} + i\mathbf{v}$ is the case $a = 1$, $b = 0$. The formula is additive in $a + bi$ and in $\mathbf{u} + i\mathbf{v}$ separately, which gives the two distributive laws. For associativity, let $\alpha = a + bi$ and $\beta = c + di$. Then

$$\begin{aligned}\alpha(\beta(\mathbf{u} + i\mathbf{v})) &= \alpha((c\mathbf{u} - d\mathbf{v}) + i(d\mathbf{u} + c\mathbf{v})) \\ &= (a(c\mathbf{u} - d\mathbf{v}) - b(d\mathbf{u} + c\mathbf{v})) + i(b(c\mathbf{u} - d\mathbf{v}) + a(d\mathbf{u} + c\mathbf{v})) \\ &= ((ac - bd)\mathbf{u} - (ad + bc)\mathbf{v}) + i((ad + bc)\mathbf{u} + (ac - bd)\mathbf{v}),\end{aligned}$$

which is $(\alpha\beta)(\mathbf{u} + i\mathbf{v})$ since $\alpha\beta = (ac - bd) + (ad + bc)i$.

(b) For $\alpha_k, \beta_k \in \mathbb{R}$, $(\alpha_k + i\beta_k)\mathbf{b}_k = (\alpha_k + i\beta_k)(\mathbf{b}_k + i\mathbf{0}) = \alpha_k\mathbf{b}_k + i\beta_k\mathbf{b}_k$, so

$$\sum_k (\alpha_k + i\beta_k)\mathbf{b}_k = \sum_k \alpha_k \mathbf{b}_k + i\sum_k \beta_k \mathbf{b}_k. \tag{7.3.2}$$

Every $\mathbf{u} + i\mathbf{v} \in V_{\mathbb{C}}$ is of this form, by expanding $\mathbf{u}$ and $\mathbf{v}$ in the basis, so the $\mathbf{b}_k$ span $V_{\mathbb{C}}$. If the left side of (7.3.2) is $\mathbf{0}$, then $\sum \alpha_k \mathbf{b}_k = \mathbf{0}$ and $\sum \beta_k \mathbf{b}_k = \mathbf{0}$, so all α_k and β_k vanish. Hence the $\mathbf{b}_k$ are independent over $\mathbb{C}$.

(c) $T_{\mathbb{C}}$ is additive because T is. For $a + bi \in \mathbb{C}$,

$$T_{\mathbb{C}}((a + bi)(\mathbf{u} + i\mathbf{v})) = T(a\mathbf{u} - b\mathbf{v}) + iT(b\mathbf{u} + a\mathbf{v}) = (aT\mathbf{u} - bT\mathbf{v}) + i(bT\mathbf{u} + aT\mathbf{v}),$$

which is $(a + bi)(T\mathbf{u} + iT\mathbf{v})$. Finally, $T_{\mathbb{C}}\mathbf{b}_j = T\mathbf{b}_j + i\mathbf{0} = \sum_k t_{kj}\mathbf{b}_k$, where $[T]_{\mathcal{B}\mathcal{B}} = (t_{kj})$, so $T_{\mathbb{C}}$ has the same matrix (Definition 2.3.1) and the same characteristic polynomial. □

Conjugation $\overline{\mathbf{u} + i\mathbf{v}} = \mathbf{u} - i\mathbf{v}$ on $V_{\mathbb{C}}$ commutes with $T_{\mathbb{C}}$, and this pairs up the eigenvectors of $T_{\mathbb{C}}$.

Proposition 7.3.34 (Eigenvalues of a complexification). Let $V \neq \{\mathbf{0}\}$ be a finite-dimensional real vector space and $T \in \mathrm{L}(V)$.

(a) A real number λ is an eigenvalue of T if and only if it is an eigenvalue of $T_{\mathbb{C}}$.

(b) If $T_{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = \lambda(\mathbf{u} + i\mathbf{v})$, then $T_{\mathbb{C}}(\mathbf{u} - i\mathbf{v}) = \overline{\lambda}(\mathbf{u} - i\mathbf{v})$. So the nonreal eigenvalues of $T_{\mathbb{C}}$ come in conjugate pairs.

(c) If $\lambda = a + bi$ with $b \neq 0$ and $\mathbf{u} + i\mathbf{v}$ is an eigenvector of $T_{\mathbb{C}}$ for λ, then $\mathbf{u}, \mathbf{v}$ is linearly independent, $U = \mathrm{span}(\mathbf{u}, \mathbf{v})$ is T-invariant, and the matrix of $T|_U$ in the basis $\mathbf{u}, \mathbf{v}$ is $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Proof. STANDARD

Uses. Proposition 7.3.33 (Basic properties), Corollary 7.3.13 (Eigenvalues of an operator), Proposition 7.2.4 (Invariance is checked on a spanning list), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Proposition 1.2.35 (Zero products)
Techniques. Complexify, then take real and imaginary parts (Technique 7.3.35)
Hint 1. For (a), compare characteristic polynomials.
Hint 2. For (b) and (c), write out $T_{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = (a + bi)(\mathbf{u} + i\mathbf{v})$ and compare the parts in front of i and outside it.
Hint 3. You should get $T\mathbf{u} = a\mathbf{u} - b\mathbf{v}$ and $T\mathbf{v} = b\mathbf{u} + a\mathbf{v}$. If $\mathbf{u}, \mathbf{v}$ were dependent, they would lie on a T-invariant line.
Sketch. (a) $p_{T_{\mathbb{C}}} = p_T$. (b), (c) Comparing real and imaginary parts, $T\mathbf{u} = a\mathbf{u} - b\mathbf{v}$ and $T\mathbf{v} = b\mathbf{u} + a\mathbf{v}$, which gives both the conjugate eigenvector and the invariant plane with matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. If $\mathbf{u}, \mathbf{v}$ were dependent, they would lie on an invariant line on which T acts by a real μ, forcing $\lambda = \mu$.

(a) By Proposition 7.3.33(c), $p_{T_{\mathbb{C}}} = p_T$. For real λ, both statements say $p_T(\lambda) = 0$, by Corollary 7.3.13 applied to T and to $T_{\mathbb{C}}$.

(b) Let $\lambda = a + bi$. The equation $T\mathbf{u} + iT\mathbf{v} = (a\mathbf{u} - b\mathbf{v}) + i(b\mathbf{u} + a\mathbf{v})$ is an equation in $V \times V$, so

$$T\mathbf{u} = a\mathbf{u} - b\mathbf{v}, \quad T\mathbf{v} = b\mathbf{u} + a\mathbf{v}. \tag{7.3.3}$$

Then $T_{\mathbb{C}}(\mathbf{u} - i\mathbf{v}) = (a\mathbf{u} - b\mathbf{v}) - i(b\mathbf{u} + a\mathbf{v})$, and the definition of scalar multiplication gives the same vector for $(a - bi)(\mathbf{u} + i(-\mathbf{v}))$. Since $\mathbf{u} + i\mathbf{v} \neq \mathbf{0}$, also $\mathbf{u} - i\mathbf{v} \neq \mathbf{0}$, so $\overline{\lambda}$ is an eigenvalue whenever λ is.

(c) By (7.3.3) and Proposition 7.2.4, U is T-invariant, and the coordinates of $T\mathbf{u}$ and $T\mathbf{v}$ in $\mathbf{u}, \mathbf{v}$ are $(a, -b)$ and (b, a), giving the matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, once we know that $\mathbf{u}, \mathbf{v}$ is independent. Suppose it is not. Since $\mathbf{u}$ and $\mathbf{v}$ are not both $\mathbf{0}$, U is then a line $\operatorname{span}(\mathbf{z})$, and it is T-invariant, so $T\mathbf{z} = \mu\mathbf{z}$ for some $\mu \in \mathbb{R}$ (Proposition 7.2.7). Writing $\mathbf{u} = \alpha\mathbf{z}$ and $\mathbf{v} = \beta\mathbf{z}$ with $\alpha, \beta \in \mathbb{R}$, we get $T_{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = \mu(\mathbf{u} + i\mathbf{v})$. Comparing with $T_{\mathbb{C}}(\mathbf{u} + i\mathbf{v}) = \lambda(\mathbf{u} + i\mathbf{v})$ gives $(\lambda - \mu)(\mathbf{u} + i\mathbf{v}) = \mathbf{0}$, so $\lambda = \mu \in \mathbb{R}$ by Proposition 1.2.35, contradicting $b \neq 0$. □

Technique 7.3.35 (Complexify, then take real and imaginary parts). To prove something about an operator T on a real vector space, pass to $T_{\mathbb{C}}$ on $V_{\mathbb{C}}$, where the FTA is available, use a complex result (an eigenvector, later a triangular form), and then split the conclusion into real and imaginary parts. A complex eigenvector $\mathbf{u} + i\mathbf{v}$ for $a + bi$ becomes the pair of real equations (7.3.3), and conjugation exchanges λ with $\overline{\lambda}$.

Reach for it when a statement about real operators is easy over $\mathbb{C}$, and the obstruction over $\mathbb{R}$ is only that p_T may have nonreal roots. The canonical example is the second proof of Theorem 7.3.31.
Used in. Proposition 7.3.34, Theorem 7.3.31, Exercise 7.7.12, Problem 7.17, Exercise 8.2.48, Theorem 8.4.27.

Proof of Theorem 7.3.31 (via complexification). STANDARD

Uses. Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 7.3.33 (Basic properties), Proposition 7.3.34 (Eigenvalues of a complexification), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors)
Techniques. Complexify, then take real and imaginary parts (Technique 7.3.35)
Hint 1. $T_{\mathbb{C}}$ acts on a nonzero finite-dimensional complex space. What does Theorem 7.3.27 give?
Hint 2. Split into the cases $\lambda \in \mathbb{R}$ and $\lambda \notin \mathbb{R}$, and use Proposition 7.3.34.
Sketch. $T_{\mathbb{C}}$ has an eigenvalue λ. If λ is real, it is an eigenvalue of T, and an eigenvector spans an invariant line. Otherwise an eigenvector $\mathbf{u} + i\mathbf{v}$ gives the invariant plane $\operatorname{span}(\mathbf{u}, \mathbf{v})$ of Proposition 7.3.34(c).

By Proposition 7.3.33(b), $V_{\mathbb{C}}$ is a nonzero finite-dimensional complex vector space, so $T_{\mathbb{C}}$ has an eigenvalue λ by Theorem 7.3.27. If $\lambda \in \mathbb{R}$, it is an eigenvalue of T by Proposition 7.3.34(a), and any eigenvector spans a T-invariant line (Proposition 7.2.7). If $\lambda \notin \mathbb{R}$, an eigenvector $\mathbf{u} + i\mathbf{v}$ of $T_{\mathbb{C}}$ gives the two-dimensional T-invariant subspace $\operatorname{span}(\mathbf{u}, \mathbf{v})$ of Proposition 7.3.34(c). □

The Krylov proof stays inside V and uses the real factorization of polynomials. The complexification proof uses only the complex theorem, but has to leave V and come back, and it is the version that generalizes, to the block triangular form of real operators (Section 7.5) and to real normal operators in Chapter 8.

Example 7.3.36 (Decomplexification). The opposite construction forgets structure. A complex vector space is also a real vector space, by restricting scalars to $\mathbb{R}$, and $\mathbb{C}^n$ becomes $\mathbb{R}^{2n}$ with real basis $\mathbf{e}_1, \ldots, \mathbf{e}_n, i\mathbf{e}_1, \ldots, i\mathbf{e}_n$. A complex matrix $A = P + iQ$, with $P, Q \in \mathbb{R}^{n \times n}$, acts by

$$(P + iQ)(\mathbf{x} + i\mathbf{y}) = (P\mathbf{x} - Q\mathbf{y}) + i(Q\mathbf{x} + P\mathbf{y}) \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n),$$

so as a real linear map of $\mathbb{R}^{2n}$ its matrix is $\begin{pmatrix} P & -Q \\ Q & P \end{pmatrix}$. For $n = 1$, multiplication by $a + bi$ on $\mathbb{C} = \mathbb{R}^2$ becomes $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, as in Proposition B.1.18. Conversely, decomplexifying $V_{\mathbb{C}}$ gives $V \times V$, on which $T_{\mathbb{C}}$ acts as $\begin{pmatrix} T & 0 \\ 0 & T \end{pmatrix}$.

Exercises

Exercise 7.3.37. ROUTINE For each matrix, find p_A, the eigenvalues, and a basis of each eigenspace.

(a) $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(b) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(c) $A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, over $\mathbb{R}$ and over $\mathbb{C}$.

Solution on page 1575

Exercise 7.3.38. STANDARD Let $A, B \in F^{n\times n}$. Prove or disprove:

(a) If $p_A = p_B$, then A and B are similar.

(b) If $p_A = p_B$, then $\operatorname{rank} A = \operatorname{rank} B$.

(c) For every $c \in F$, p_{A+cI_n} is obtained from p_A by substituting $x - c$ for x, that is, $p_{A+cI_n} = p_A(x - c)$.

(d) If $p_A = p_B$ and p_A has n distinct roots in F, then A and B are similar.

Solution on page 1576

Exercise 7.3.39. STANDARD

(a) Let $\tau : F^{n\times n} \to F^{n\times n}$ be transposition, $\tau(X) = X^{\mathsf{T}}$. Show that

$$p_\tau = (x-1)^{n(n+1)/2}(x+1)^{n(n-1)/2}.$$

(b) Let W be the real vector space of functions $\mathbb{R} \to \mathbb{R}$ with basis $\sin t, \cos t, t\sin t, t\cos t$. Find the characteristic polynomial of differentiation $D : W \to W$, and conclude that no nonzero function in W satisfies $f' = \lambda f$ with $\lambda \in \mathbb{R}$.

Solution on page 1576

Exercise 7.3.40. STANDARD Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$, where $m, n \geq 1$.

(a) Show that $x^n p_{AB} = x^m p_{BA}$.

(b) Deduce that AB and BA have the same nonzero eigenvalues.

(c) Prove (b) again without characteristic polynomials: if $AB\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$ and $\lambda \neq 0$, show that $B\mathbf{v}$ is an eigenvector of BA.

(d) Give an example in which 0 is an eigenvalue of BA but not of AB.

Solution on page 1577

Exercise 7.3.41. STANDARD Let $n \geq 2$ and $\mathbf{u}, \mathbf{v} \in F^n$.

(a) Show that $p_{\mathbf{u}\mathbf{v}^{\mathsf{T}}} = x^{n-1}\left(x - \mathbf{v}^{\mathsf{T}}\mathbf{u}\right)$.

(b) Let $J \in F^{n\times n}$ be the matrix whose entries are all 1. Find p_J and the eigenvalues of J over $\mathbb{Q}$, and over $\mathbb{F}_p$ for a prime p dividing n.

Solution on page 1577

Exercise 7.3.42. STANDARD Let $A \in \mathbb{Q}^{n\times n}$ have integer entries.

(a) Show that all coefficients of p_A are integers.

(b) Show that every rational eigenvalue of A is an integer that divides $\det A$.

(c) Show that $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ has no rational eigenvalue.

Solution on page 1578

Exercise 7.3.43. STANDARD Let $A \in \mathbb{R}^{n\times n}$ satisfy $A^2 = -I_n$.

(a) Show that n is even, using determinants.

(b) Give a second proof using p_A and the eigenvalues of A regarded as a complex matrix.

(c) Show that, conversely, $A^2 = -I_n$ has a solution in $\mathbb{R}^{n\times n}$ for every even n, and in $\mathbb{C}^{n\times n}$ for every n.

Solution on page 1578

Exercise 7.3.44. STANDARD Let F be a finite field with elements $c_1, \ldots, c_q$.

(a) Show that $p = (x - c_1)(x - c_2)\cdots\left(x - c_q\right) + 1$ has no root in F.

(b) Deduce that some operator on F^q has no eigenvalue. So Theorem 7.3.27 fails over every finite field.

(c) Write down this operator for $F = \mathbb{F}_2$ and check directly that it has no eigenvalue.

Solution on page 1579

Exercise 7.3.45. STANDARD Let $p = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in F[x]$ with $n \geq 2$, and let $C = C_p$.

(a) Show that $\operatorname{rank}(C - \lambda I_n) \geq n - 1$ for every $\lambda \in F$. So every eigenspace of C is a line.

(b) Let λ be a root of p and write $p = (x - \lambda)q$ with $q = x^{n-1} + b_{n-2}x^{n-2} + \cdots + b_0$. Show that $(b_0, b_1, \ldots, b_{n-2}, 1)$ is an eigenvector of C for λ.

(c) Show that F^n has a basis of eigenvectors of C if and only if p has n distinct roots in F.

Solution on page 1579

Exercise 7.3.46. STANDARD Let $A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{C}^{3\times 3}$.

(a) For $\mathbf{v} = \mathbf{e}_1 + \mathbf{e}_3$, compute $\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, A^3\mathbf{v}$, and find the polynomial p of Proposition 7.2.35.

(b) Factor p over $\mathbb{C}$, and use the second proof of Theorem 7.3.27 to produce an eigenvector for each root.

(c) Repeat (a) for $\mathbf{v} = \mathbf{e}_1$. Which eigenvalue does this Krylov sequence miss, and why is that consistent with Proposition 7.3.26?

Solution on page 1579

Exercise 7.3.47. STANDARD Find the error in the following argument.

Claim. For every $A \in F^{n\times n}$, $p_A(A) = 0$.

"Proof." By definition, $p_A(x) = \det(xI_n - A)$. Substituting $x = A$ gives $p_A(A) = \det(AI_n - A) = \det(A - A) = \det 0 = 0$.

Solution on page 1580

Exercise 7.3.48. HARD Let $A \in F^{n\times n}$ with $A^k = 0$ for some $k \geq 1$.

(a) Show that in $F[x]^{n\times n}$,
$$(xI_n - A)\left(x^{k-1}I_n + x^{k-2}A + \cdots + xA^{k-2} + A^{k-1}\right) = x^k I_n.$$
(b) Deduce that $p_A = x^n$. In particular $\operatorname{tr} A = 0$, and 0 is the only eigenvalue of A.
(c) For $F = \mathbb{C}$, give a shorter proof of (b) using eigenvalues. Why does it not work over every field?

Solution on page 1581

Exercise 7.3.49. HARD Let $A \in F^{n\times n}$. For $f \in F[x]$, write $f(x^2)$ and $f(-x)$ for the polynomials obtained by substituting x^2 and $-x$ for x (Definition 7.1.9).
(a) Show that $p_{A^2}(x^2) = (-1)^n p_A(x)p_A(-x)$.
(b) Suppose $p_A = (x-\lambda_1)\cdots(x-\lambda_n)$. Show that $p_{A^2} = \left(x-\lambda_1^2\right)\cdots\left(x-\lambda_n^2\right)$.
(c) Deduce that if $A \in \mathbb{R}^{n\times n}$, p_A splits over $\mathbb{R}$, and $\operatorname{tr}\left(A^2\right) = 0$, then $p_A = x^n$.

Solution on page 1581

Exercise 7.3.50. HARD Let $\dim V = n \geq 1$ and $T \in \mathrm{L}(V)$. A vector $\mathbf{v}$ is a **cyclic vector** for T if $K_T(\mathbf{v}) = V$.
(a) Show that if T has a cyclic vector, then the matrix of T in some basis is C_{p_T}.
(b) Suppose $A, B \in F^{n\times n}$ both have cyclic vectors (for T_A and T_B). Show that A and B are similar if and only if $p_A = p_B$.
(c) Show that I_2 has no cyclic vector, and that C_p and C_p^{T} are similar for every monic p.

Solution on page 1582

Exercise 7.3.51. HARD BEYOND 21-242 Let V be finite-dimensional with $\dim V \geq 1$, and $T \in \mathrm{L}(V)$.
(a) Use Proposition 7.3.26 to show that $p_T(T) = 0$. This is the Cayley–Hamilton theorem (Theorem 7.6.20).
(b) Deduce that every $A \in F^{n\times n}$ satisfies a monic polynomial of degree n, improving the degree n^2 of Proposition 4.1.31.

Solution on page 1582

Exercise 7.3.52. HARD Let q be a prime and $f = x^2 - tx + d \in \mathbb{F}_q[x]$.
(a) Show that the number of matrices $A \in \mathbb{F}_q^{2\times 2}$ with $p_A = f$ is $q^2 + q$ if f has two distinct roots in $\mathbb{F}_q$, q^2 if it has one repeated root, and $q^2 - q$ if it has no root.
(b) Deduce that exactly $\left(q^2 - q\right)^2/2$ matrices in $\mathbb{F}_q^{2\times 2}$ have no eigenvalue in $\mathbb{F}_q$, and check that the counts in (a) add up to q^4.

Solution on page 1582

Exercise 7.3.53. HARD BEYOND 21-242 For $A \in F^{n\times n}$ and $S \subseteq \{1, \ldots, n\}$, let A_S be the principal submatrix of A with rows and columns in S, and let $e_k(A) = \sum_{|S|=k} \det A_S$ for $1 \leq k \leq n$, and $e_0(A) = 1$. Show that
$$p_A = \sum_{k=0}^{n} (-1)^k e_k(A) x^{n-k},$$
and check that this agrees with Proposition 7.3.18. For $n = 3$, write out the coefficient of x.

Solution on page 1583

Exercise 7.3.54. STANDARD BEYOND 21-242 Let V be a finite-dimensional real vector space, and let $T \in \mathrm{L}(V)$ have no eigenvalue.

(a) Show that every T-invariant subspace of V has even dimension.

(b) Give such a T on $\mathbb{R}^4$, and find three different T-invariant planes.

Solution on page 1584

Exercise 7.3.55. CHALLENGE Let $A, B \in \mathbb{C}^{n\times n}$ satisfy $AB - BA = A$.

(a) Show that $A^k B - BA^k = kA^k$ for every $k \geq 1$.

(b) Show that $A^k = 0$ for some k.

(c) Deduce that $p_A = x^n$.

Solution on page 1584

7.4 Diagonalization

Let $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$, and suppose we want A^{100}. Multiplying out 100 copies of A is hopeless by hand, and the entries of the partial products show no obvious pattern. But A has two eigenvectors,

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

and they form a basis of $\mathbb{R}^2$. In this basis A only scales: it doubles the first coordinate and triples the second, so A^{100} multiplies them by 2^{100} and 3^{100}. All that remains is to convert coordinates back. This section finds which operators admit such a basis, by a test in terms of two kinds of multiplicity, and then uses such bases for matrix powers, linear recurrences, and the long-run behavior of linear dynamical systems.

Throughout, V is a finite-dimensional vector space over a field F with $\dim V = n \geq 1$, and $T \in \mathrm{L}(V)$.

7.4.1 Diagonalizable operators

Definition 7.4.1 (Diagonalizable). An operator $T \in \mathrm{L}(V)$ is **diagonalizable** if there is a basis $\mathcal{B}$ of V for which $[T]_{\mathcal{B}\mathcal{B}}$ is a diagonal matrix. A matrix $A \in F^{n\times n}$ is **diagonalizable** if it is similar to a diagonal matrix, that is, if

$$A = SDS^{-1}$$

for some invertible $S \in F^{n\times n}$ and some diagonal $D \in F^{n\times n}$. Such a factorization is a **diagonalization** of A.

The lecture defines a diagonalizable operator as one with a basis of eigenvectors. The next theorem shows that this is the same notion, and that the matrix and operator versions agree. A basis of V consisting of eigenvectors of T is called an **eigenbasis** for T.

Theorem 7.4.2 (Diagonalizable means an eigenbasis). Let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V and $\lambda_1, \ldots, \lambda_n \in F$. Then

$$[T]_{\mathcal{B}\mathcal{B}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \iff T\mathbf{b}_j = \lambda_j \mathbf{b}_j \text{ for } j = 1, \ldots, n.$$

In particular, T is diagonalizable if and only if V has an eigenbasis for T, and then the diagonal entries of $[T]_{\mathcal{B}\mathcal{B}}$ are the eigenvalues of the vectors of the eigenbasis, in order.

For $A \in F^{n\times n}$ the following are equivalent:

(I) A is diagonalizable;

(II) the operator T_A on F^n is diagonalizable;

(III) F^n has a basis of eigenvectors of A.

More precisely, if S is invertible with columns $\mathbf{s}_1, \ldots, \mathbf{s}_n$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then $A = SDS^{-1}$ if and only if $A\mathbf{s}_j = \lambda_j \mathbf{s}_j$ for every j.

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.14 (Coordinates), Definition 7.2.6 (Eigenvalue, eigenvector), Example 2.4.10 (Diagonal matrices), Corollary 3.5.25 (The invertible matrix theorem), Proposition 2.3.16 (The matrix of T_A)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. What is the j-th column of $[T]_{\mathcal{BB}}$? When is it $\lambda_j \mathbf{e}_j$?
Hint 2. For matrices, rewrite $A = SDS^{-1}$ as $AS = SD$ and compare the two sides column by column. Right multiplication by a diagonal matrix scales columns.
Hint 3. Columns of an invertible matrix form a basis of F^n, and conversely (Corollary 3.5.25).
Sketch. The j-th column of $[T]_{\mathcal{BB}}$ is $\left[T\mathbf{b}_j\right]_{\mathcal{B}}$, and it equals $\lambda_j \mathbf{e}_j$ exactly when $T\mathbf{b}_j = \lambda_j \mathbf{b}_j$; basis vectors are nonzero, so these are eigenvectors. For matrices, $A = SDS^{-1}$ iff $AS = SD$, whose j-th columns are $A\mathbf{s}_j$ and $\lambda_j \mathbf{s}_j$. An invertible S is the same as a basis of columns, which gives (I) $\Leftrightarrow$ (III), and (II) $\Leftrightarrow$ (III) is the first part for T_A.

By Definition 2.3.1, the j-th column of $[T]_{\mathcal{BB}}$ is the coordinate vector $\left[T\mathbf{b}_j\right]_{\mathcal{B}}$. Since coordinates are unique (Definition 1.6.14), $\left[T\mathbf{b}_j\right]_{\mathcal{B}} = \lambda_j \mathbf{e}_j$ if and only if $T\mathbf{b}_j = \lambda_j \mathbf{b}_j$. So $[T]_{\mathcal{BB}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ exactly when $T\mathbf{b}_j = \lambda_j \mathbf{b}_j$ for all j. The vectors of a basis are nonzero, so in that case each $\mathbf{b}_j$ is an eigenvector with eigenvalue λ_j (Definition 7.2.6), and $\mathcal{B}$ is an eigenbasis. Conversely, if $\mathcal{B}$ is an eigenbasis with $T\mathbf{b}_j = \lambda_j \mathbf{b}_j$, then $[T]_{\mathcal{BB}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. This proves the first two claims.

For matrices, let S be invertible with columns $\mathbf{s}_j$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Multiplying on the right by S or by S^{-1} shows that $A = SDS^{-1}$ if and only if $AS = SD$. The j-th column of AS is $A\mathbf{s}_j$, and the j-th column of SD is $\lambda_j \mathbf{s}_j$, since right multiplication by a diagonal matrix scales the columns (Example 2.4.10). So $A = SDS^{-1}$ if and only if $A\mathbf{s}_j = \lambda_j \mathbf{s}_j$ for every j.

(I) $\Leftrightarrow$ (III). If $A = SDS^{-1}$, the columns of S form a basis of F^n by Corollary 3.5.25, and by the previous paragraph they are eigenvectors of A. Conversely, if $\mathbf{s}_1, \ldots, \mathbf{s}_n$ is a basis of F^n with $A\mathbf{s}_j = \lambda_j \mathbf{s}_j$, the matrix S with these columns is invertible (Corollary 3.5.25), and $A = SDS^{-1}$ with $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

(II) $\Leftrightarrow$ (III). The eigenvectors of A are by definition those of T_A, so this is the first part of the theorem applied to $T = T_A$. □

Remark 7.4.3 (Two readings of $A = SDS^{-1}$). The proof read $A = SDS^{-1}$ column by column: A scales the columns of S. There is also the reading of Section 4.5. If $\mathring{\mathcal{B}}$ is the list of columns of S, then $S = [\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}}$, and Theorem 4.5.17 gives

$$D = S^{-1}AS = [\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{E}}[T_A]_{\mathcal{EE}}[\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}} = [T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}.$$

Read from right to left, $A = SDS^{-1}$ says: convert standard coordinates into eigenbasis coordinates (S^{-1}), scale each coordinate by its eigenvalue (D), and convert back (S). Diagonalization is the strongest instance of Technique 4.5.33: the operator is as simple as an operator can be, and the basis that achieves this is made of the operator's own eigenvectors.

Example 7.4.4 (Diagonalizing a 2×2 matrix). Let $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, the matrix of the introduction. By Theorem 7.2.9, λ is an eigenvalue exactly when

$$\det(A - \lambda I_2) = (4 - \lambda)(1 - \lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$$

vanishes, so the eigenvalues are 2 and 3. Row reducing,

$$A - 2I_2 = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad A - 3I_2 = \begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix},$$

so $E_2 = \operatorname{span}((1,1))$ and $E_3 = \operatorname{span}((2,1))$. The vectors $(1,1)$ and $(2,1)$ are not parallel, so they form an eigenbasis, and by Theorem 7.4.2

$$A = SDS^{-1}, \quad S = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}.$$

To check the work it is enough to verify $AS = SD$: both sides equal $\begin{pmatrix} 2 & 6 \\ 2 & 3 \end{pmatrix}$. The order of the columns is our choice. Swapping them swaps the diagonal entries of D, and rescaling a column of S by $c \neq 0$ changes nothing in D. So a diagonalization is far from unique, but D is unique up to the order of its entries (Theorem 7.4.16 will show why).

$$[\mathbf{e}_1]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad [A\mathbf{e}_1]_{\mathcal{B}} = \left[\begin{pmatrix} 4 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

The standard grid and the grid of the eigenbasis $\mathcal{B} = (1,1),(2,1)$ of $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$. In $\mathcal{B}$-coordinates A only scales, $[A\mathbf{v}]_{\mathcal{B}} = D[\mathbf{v}]_{\mathcal{B}}$ with $D = \operatorname{diag}(2,3)$. The vector $\mathbf{v} = \mathbf{e}_1$ starts with $[\mathbf{v}]_{\mathcal{B}} = (-1,1)$, so $A\mathbf{e}_1 = -2(1,1) + 3(2,1) = (4,1)$, which is the first column of A. In the interactive version, type $\mathbf{v} = (4,1)$ and confirm that its $\mathcal{B}$-coordinates are $(-2,3)$. Moving $\mathbf{v}$ along a line of the $\mathcal{B}$-grid changes only one coordinate, and A moves the point along the corresponding line of the stretched grid. (Interactive in the web version.)

In Example 7.4.4 the two eigenvalues were different, and that alone guaranteed the eigenbasis.

Corollary 7.4.5 (Distinct eigenvalues give a diagonalization). If T has $n = \dim V$ distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then T is diagonalizable: any choice of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for $\lambda_1, \ldots, \lambda_n$ is an eigenbasis, and in it T has matrix $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. In particular, $A \in F^{n \times n}$ with n distinct eigenvalues is diagonalizable.

Proof. ROUTINE

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 4.1.10 (Independent lists are short), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Hint 1. Pick one eigenvector for each eigenvalue. What do you know about this list?
Hint 2. It is independent and has length $\dim V$.
Sketch. Eigenvectors for distinct eigenvalues are independent (Theorem 7.2.21), and n independent vectors in an n-dimensional space form a basis (Theorem 4.1.10(b)). Apply Theorem 7.4.2.

The list $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is independent by Theorem 7.2.21, and it has length $n = \dim V$, so it is a basis by Theorem 4.1.10(b). It consists of eigenvectors, so Theorem 7.4.2 gives $[T]_{\mathcal{B}\mathcal{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ for $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$. □

The converse fails: for $n \geq 2$, I_n is diagonal, with the single eigenvalue 1. Repeated eigenvalues are allowed as long as each one comes with enough eigenvectors.

Example 7.4.6 (A repeated eigenvalue with enough eigenvectors). The matrix $A = 2I_3 + J = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$ of Example 7.2.14 has only the eigenvalues 2 and 5, with

$$E_2 = \operatorname{span}((1,-1,0),(0,1,-1)), \quad E_5 = \operatorname{span}((1,1,1)).$$

These three vectors form a basis of $\mathbb{R}^3$: the sum $E_2 + E_5$ is direct (Corollary 7.2.23), so concatenating bases gives an independent list (Theorem 4.4.13), of length 3. Hence

$$A = S\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}S^{-1}, \quad S = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Obstructions to diagonalizability

There are two ways an operator can fail to be diagonalizable: it may have too few eigenvalues in F, or an eigenvalue may have too few eigenvectors. The simplest instance of the second failure is an operator with one eigenvalue.

Proposition 7.4.7 (One eigenvalue). Suppose T is diagonalizable and λ is its only eigenvalue. Then $T = \lambda \mathrm{id}_V$. In particular, if $A \in F^{n\times n}$ is diagonalizable with only eigenvalue λ, then $A = \lambda I_n$.

Proof. ROUTINE

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis), Definition 7.2.6 (Eigenvalue, eigenvector)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. Take an eigenbasis. What is the eigenvalue of each of its vectors?
Hint 2. Two linear maps that agree on a basis are equal.
Sketch. Every vector of an eigenbasis is an eigenvector, so its eigenvalue is λ. Then T and λ id agree on a basis, so they are equal.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be an eigenbasis for T. The eigenvalue of each $\mathbf{b}_j$ is an eigenvalue of T, hence equals λ, so $T\mathbf{b}_j = \lambda\mathbf{b}_j = (\lambda \mathrm{id}_V)\mathbf{b}_j$ for every j. Two linear maps that agree on a basis are equal (Technique 2.2.7), so $T = \lambda\mathrm{id}_V$. For A, apply this to T_A: then $T_A = \lambda$ id, and its matrix in the standard basis is $A = \lambda I_n$. □

Example 7.4.8 (Shears and nilpotent matrices). Let $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, a shear, over any field. Then $J - \mu I_2 = \begin{pmatrix} 1-\mu & 1 \\ 0 & 1-\mu \end{pmatrix}$ has determinant $(1-\mu)^2$, so 1 is the only eigenvalue (Theorem 7.2.9). Since $J \neq I_2$, Proposition 7.4.7 shows that J is not diagonalizable. Directly: $E_1 = \ker\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \mathrm{span}(\mathbf{e}_1)$ is a line, so there is no basis of eigenvectors.

More generally, let $N \in F^{n\times n}$ be nilpotent and nonzero, say $N^k = 0$. If $N\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{0} = N^k\mathbf{v} = \lambda^k\mathbf{v}$, so $\lambda = 0$; and 0 is an eigenvalue, because N is not injective (otherwise N^k would be injective too, while $N^k = 0$ on $F^n \neq \{\mathbf{0}\}$). So 0 is the only eigenvalue, and $N \neq 0 = 0 \cdot I_n$ is not diagonalizable. The same argument works for an upper triangular matrix with all diagonal entries equal to λ: its only eigenvalue is λ, because $A - \mu I_n$ is triangular with diagonal entries $\lambda - \mu$, hence invertible for $\mu \neq \lambda$ (Proposition 5.3.15, Theorem 5.4.7). So it is diagonalizable only if it equals λI_n.

These are the minimal counterexamples of Technique 1.4.18 for diagonalizability. In particular, being invertible (J) and being diagonalizable are unrelated: J is invertible but not diagonalizable, and $\mathrm{diag}(1, 0)$ is diagonal but not invertible.

Example 7.4.9 (Rotations: the field matters). Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the rotation by a right angle. By Example 7.2.18 its eigenvalues are the square roots of -1 in F.

- Over $\mathbb{R}$ there are none, so there are no eigenvectors at all, and A is not diagonalizable over $\mathbb{R}$.
- Over $\mathbb{C}$ the eigenvalues i and $-i$ are distinct, with eigenvectors $(1, -i)$ and $(1, i)$, so by Corollary 7.4.5

$$A = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1}.$$

- Over $\mathbb{F}_5$, where $2^2 = 3^2 = -1$, it is diagonalizable with eigenvalues 2 and 3; over $\mathbb{F}_3$ there are no eigenvalues.
- Over $\mathbb{F}_2$, A is the swap $(x, y) \mapsto (y, x)$, whose only eigenvalue is 1 (Example 7.2.17). Since $A \neq I_2$, it is not diagonalizable over $\mathbb{F}_2$, by Proposition 7.4.7. Over every field with $1 + 1 \neq 0$ the swap is diagonalizable, with eigenbasis $(1, 1), (1, -1)$.

So "diagonalizable" is a property of a matrix together with a field. The real matrix A fails for lack of eigenvalues, while over $\mathbb{F}_2$ the matrix fails for lack of eigenvectors.

Exercise 7.4.10. ROUTINE Without computing any eigenvectors, decide which of the following real matrices are diagonalizable over $\mathbb{R}$, and which over $\mathbb{C}$.

$$A_1 = \begin{pmatrix} 1 & 5 \\ 0 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

Solution on page 1584

7.4.2 Algebraic and geometric multiplicity

The obstructions above suggest counting, for each eigenvalue, how many times it "should" occur on the diagonal and how many independent eigenvectors it actually has. The first count comes from the characteristic polynomial, the second from the eigenspace.

Recall from Section 7.3 that the characteristic polynomial of $A \in F^{n\times n}$ is the monic polynomial $p_A = \det(xI_n - A) \in F[x]$ of degree n (Definition 7.3.4), and that its roots in F are exactly the eigenvalues of A (Theorem 7.3.6). Similar matrices have the same characteristic polynomial, so the **characteristic polynomial of** T, $p_T = p_A$ for $A = [T]_{\mathcal{B}\mathcal{B}}$, does not depend on the basis $\mathcal{B}$. Determinants of matrices with entries in $F[x]$ obey the same rules as over F (triangular and block triangular matrices, products), as explained in Section 7.3; we use these rules freely.

Definition 7.4.11 (Algebraic multiplicity). Let $\lambda \in F$. The **algebraic multiplicity** of λ for T, written $\operatorname{am}_T(\lambda)$, is the multiplicity of λ as a root of p_T (Definition 7.1.24): the largest $m \geq 0$ with $(x-\lambda)^m$ dividing p_T. For a matrix A, $\operatorname{am}_A(\lambda)$ is the multiplicity of λ as a root of p_A.

Definition 7.4.12 (Geometric multiplicity). Let $\lambda \in F$. The **geometric multiplicity** of λ for T is

$$\operatorname{gm}_T(\lambda) = \dim E_\lambda(T) = \dim \ker(T - \lambda \mathrm{id}_V),$$

and for a matrix A, $\operatorname{gm}_A(\lambda) = \dim \ker(A - \lambda I_n)$.

We drop the subscript when the operator is clear. Both multiplicities are defined for every $\lambda \in F$, and by Theorem 7.3.6 and Theorem 7.2.9 each is positive exactly when λ is an eigenvalue. The geometric multiplicity is computed by row reduction, through a rank.

Lemma 7.4.13 (Geometric multiplicity through a rank). For every basis $\mathcal{B}$ of V and every $\lambda \in F$,

$$\operatorname{gm}_T(\lambda) = n - \operatorname{rank}([T]_{\mathcal{B}\mathcal{B}} - \lambda I_n).$$

In particular, $\operatorname{gm}_T(\lambda) = \operatorname{gm}_A(\lambda)$ for every matrix A of T, and similar matrices have the same geometric multiplicities.

Proof. ROUTINE

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.2.15 (Rank through any matrix), Definition 2.3.1 (Matrix of a linear map), Theorem 4.5.25 (Similar matrices are one operator in two bases)
Hint 1. $\operatorname{gm}_T(\lambda)$ is the nullity of $T - \lambda$ id. What does rank–nullity say?
Hint 2. What is the matrix of $T - \lambda$ id in the basis $\mathcal{B}$?
Sketch. By rank–nullity, $\dim \ker(T - \lambda \mathrm{id}) = n - \operatorname{rank}(T - \lambda \mathrm{id})$. The matrix of $T - \lambda$ id in $\mathcal{B}$ is $[T]_{\mathcal{B}\mathcal{B}} - \lambda I_n$, and a map has the rank of any of its matrices. For a matrix A this is the same formula with $T = T_A$; similar matrices are matrices of one operator.

By Theorem 4.2.11, $\operatorname{gm}_T(\lambda) = \dim \ker(T - \lambda \mathrm{id}) = n - \operatorname{rank}(T - \lambda \mathrm{id})$. The matrix of a sum and of a scalar multiple of maps is the sum and the scalar multiple of the matrices, and $[\mathrm{id}]_{\mathcal{B}\mathcal{B}} = I_n$, so $[T - \lambda \mathrm{id}]_{\mathcal{B}\mathcal{B}} =$

$[T]_{\mathcal{BB}} - \lambda I_n$. A linear map has the same rank as each of its matrices (Proposition 4.2.15), which gives the formula. For $A \in F^{n\times n}$, $\mathrm{gm}_A(\lambda) = n - \mathrm{rank}(A - \lambda I_n)$ by the same computation with $T = T_A$ and the standard basis. Finally, two similar matrices are matrices of one operator in two bases (Theorem 4.5.25), so they have the same geometric multiplicities, namely those of that operator. □

Example 7.4.14 (The multiplicities of a diagonal matrix). Let $D = \mathrm{diag}(d_1, \ldots, d_n)$ and $\lambda \in F$, and let m be the number of indices i with $d_i = \lambda$. Then $xI_n - D = \mathrm{diag}(x - d_1, \ldots, x - d_n)$ is diagonal, so

$$p_D = (x - d_1)(x - d_2)\cdots(x - d_n),$$

and by Proposition 7.1.34, $\mathrm{am}_D(\lambda) = m$. On the other hand $D - \lambda I_n = \mathrm{diag}(d_1 - \lambda, \ldots, d_n - \lambda)$ has exactly $n - m$ nonzero columns, which are multiples of distinct standard basis vectors, so its rank is $n - m$, and $\mathrm{gm}_D(\lambda) = m$ by Lemma 7.4.13. So for a diagonal matrix the two multiplicities agree: both count how often λ occurs on the diagonal. Consequently, if T is diagonalizable, the number of times λ occurs on the diagonal of a diagonal matrix of T does not depend on the eigenbasis: it is $\mathrm{gm}_T(\lambda)$. This is why D in a diagonalization $A = SDS^{-1}$ is unique up to order.

For the shear $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $p_J = (x-1)^2$, so $\mathrm{am}_J(1) = 2$, while $\mathrm{gm}_J(1) = 2 - \mathrm{rank}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$.

Intuition (Why the geometric multiplicity cannot exceed the algebraic one). Suppose λ has g independent eigenvectors. Put them first in a basis. Then the first g columns of the matrix of T are $\lambda\mathbf{e}_1, \ldots, \lambda\mathbf{e}_g$, so the matrix is block upper triangular with top left block λI_g, and its characteristic polynomial picks up the factor $(x - \lambda)^g$ from that block. The bottom right block may contribute further factors $x - \lambda$, which are not visible as eigenvectors. For the shear, the bottom right block is the 1×1 matrix (1), and it contributes the second factor $x - 1$ without contributing a second eigenvector.

Theorem 7.4.15 (Geometric multiplicity is at most algebraic multiplicity). For every $\lambda \in F$,

$$\mathrm{gm}_T(\lambda) \le \mathrm{am}_T(\lambda).$$

If λ is an eigenvalue of T, then $1 \le \mathrm{gm}_T(\lambda) \le \mathrm{am}_T(\lambda)$.

Proof. STANDARD

Uses. Theorem 4.1.16 (Extending to a basis), Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Definition 7.2.13 (Eigenspace), Theorem 5.4.26 (Block-triangular matrices), Definition 7.3.4 (Characteristic polynomial), Definition 7.1.24 (Multiplicity of a root), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Choose a basis adapted to a preserved subspace (Technique 4.5.36)
Hint 1. Let $g = \mathrm{gm}_T(\lambda)$. Choose a basis of E_λ and extend it to a basis of V. What does the matrix of T look like?
Hint 2. E_λ is T-invariant and T acts on it as λ id, so by Proposition 4.5.35 the matrix is $\begin{pmatrix} \lambda I_g & B \\ 0 & C \end{pmatrix}$.
Hint 3. Compute $\det(xI_n - [T]_{\mathcal{BB}})$ with the block triangular rule. Which power of $x - \lambda$ divides it?
Sketch. Extend a basis of E_λ, of length g, to a basis $\mathcal{B}$ of V. Since T acts on E_λ as λ id, $[T]_{\mathcal{BB}} = \begin{pmatrix} \lambda I_g & B \\ 0 & C \end{pmatrix}$, so $p_T = \det\big((x - \lambda)I_g\big)\det\big(xI_{n-g} - C\big) = (x - \lambda)^g p_C$. Hence $(x - \lambda)^g$ divides p_T, and $g \le \mathrm{am}_T(\lambda)$. If λ is an eigenvalue, $E_\lambda \ne \{\mathbf{0}\}$, so $g \ge 1$.

Let $g = \mathrm{gm}_T(\lambda) = \dim E_\lambda$. If $g = 0$ there is nothing to prove. If $g = n$, then $E_\lambda = V$ (Proposition 4.1.21), so $T = \lambda \mathrm{id}_V$, $[T]_{\mathcal{BB}} = \lambda I_n$ in every basis, and $p_T = (x - \lambda)^n$; so $\mathrm{am}_T(\lambda) = n = g$.

Let $1 \le g < n$. Choose a basis $\mathbf{u}_1, \ldots, \mathbf{u}_g$ of E_λ and extend it to a basis $\mathcal{B} = \mathbf{u}_1, \ldots, \mathbf{u}_g, \mathbf{w}_1, \ldots, \mathbf{w}_{n-g}$ of V (Theorem 4.1.16). For $\mathbf{u} \in E_\lambda$ we have $T\mathbf{u} = \lambda\mathbf{u} \in E_\lambda$, so E_λ is mapped into itself, and the restriction of T to E_λ is $\lambda \mathrm{id}_{E_\lambda}$, whose matrix in $\mathbf{u}_1, \ldots, \mathbf{u}_g$ is λI_g. By Proposition 4.5.35,

$$[T]_{\mathcal{BB}} = \begin{pmatrix} \lambda I_g & B \\ 0 & C \end{pmatrix}$$

for some $B \in F^{g\times(n-g)}$ and $C \in F^{(n-g)\times(n-g)}$. Then

$$xI_n - [T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} (x-\lambda)I_g & -B \\ 0 & xI_{n-g} - C \end{pmatrix}$$

is block upper triangular with entries in $F[x]$, and by Theorem 5.4.26

$$p_T = \det\Big((x-\lambda)I_g\Big) \cdot \det\Big(xI_{n-g} - C\Big) = (x-\lambda)^g p_C.$$

So $(x-\lambda)^g$ divides p_T, and $g \leq \mathrm{am}_T(\lambda)$ by Definition 7.1.24.

If λ is an eigenvalue, then E_λ contains a nonzero vector, so $g \geq 1$. □

The lecture calls Theorem 7.4.15 the "AM–GM inequality", as a joke: the inequality between the arithmetic and geometric means has the same letters in the same order. The proof is Technique 4.5.36 applied to the most obvious invariant subspace of T, an eigenspace.

We can now say exactly when T is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Every eigenvector lies in some E_{λ_i}, the sum of these eigenspaces is direct (Corollary 7.2.23), and p_T has degree n. So

$$\mathrm{gm}(\lambda_1) + \cdots + \mathrm{gm}(\lambda_m) \leq \mathrm{am}(\lambda_1) + \cdots + \mathrm{am}(\lambda_m) \leq \deg p_T = n, \tag{7.4.1}$$

where the first inequality is Theorem 7.4.15 term by term, and the second is Proposition 7.1.27, since the λ_i are distinct roots of p_T. Diagonalizability turns out to mean that the left end equals the right end, which forces both inequalities to be equalities.

Theorem 7.4.16 (Criteria for diagonalizability). Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T ($m \geq 0$). The following are equivalent.

(I) T is diagonalizable.

(II) $V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_m}$.

(III) $\mathrm{gm}(\lambda_1) + \mathrm{gm}(\lambda_2) + \cdots + \mathrm{gm}(\lambda_m) = n$.

(IV) p_T splits over F, and $\mathrm{gm}(\lambda_i) = \mathrm{am}(\lambda_i)$ for $i = 1, \ldots, m$.

In that case, for every eigenbasis $\mathcal{B}$ the matrix $[T]_{\mathcal{B}\mathcal{B}}$ has λ_i on its diagonal exactly $\mathrm{gm}(\lambda_i) = \mathrm{am}(\lambda_i)$ times.

Proof, part 1 of 2: (I) ⇒ (II) ⇒ (III) ⇒ (I). STANDARD

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis), Corollary 7.2.23 (Eigenspaces are independent), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.1.21 (Dimensions of subspaces), Proposition 4.4.2 (Sums are the smallest containing subspaces)
Techniques. Proving TFAE (Technique 2.7.1), Dimension counting (Technique 4.2.24), Glue bases of the summands (Technique 4.4.14)
Hint 1. For (I) ⇒ (II): where does each vector of an eigenbasis live? The sum is direct by Corollary 7.2.23.
Hint 2. For (II) ⇒ (III) and (III) ⇒ (I), use Theorem 4.4.13: the dimension of a direct sum, and bases of a direct sum by concatenation.
Hint 3. In (III) ⇒ (I), the concatenated list is independent of length n. Why is it a basis of V?
Sketch. (I) ⇒ (II): each vector of an eigenbasis lies in some E_{λ_i}, so the sum of the eigenspaces contains a basis and equals V; it is direct by Corollary 7.2.23. (II) ⇒ (III): the dimension of a direct sum is the sum of the dimensions. (III) ⇒ (I): the sum of the eigenspaces is direct, so concatenating their bases gives a basis of the sum, of length n; the sum is then all of V, and the concatenation is an eigenbasis.

Write $U = E_{\lambda_1} + \cdots + E_{\lambda_m}$. By Corollary 7.2.23 the sum is direct, whatever T is.

(I) ⇒ (II). Let $\mathcal{B}$ be an eigenbasis. Each vector of $\mathcal{B}$ is an eigenvector, whose eigenvalue is one of the λ_i, so it lies in $E_{\lambda_i} \subseteq U$. So U, which is a subspace (Proposition 4.4.2), contains a basis of V, hence contains its span V. Thus $V = U$, and the sum is direct.

(II) ⇒ (III). By Theorem 4.4.13(c), $n = \dim V = \dim E_{\lambda_1} + \cdots + \dim E_{\lambda_m}$.

(III) ⇒ (I). Choose a basis of each E_{λ_i} and concatenate them. Since U is a direct sum, the concatenation is a basis of U by Theorem 4.4.13(b), so $\dim U = \sum_i \operatorname{gm}(\lambda_i) = n$, and $U = V$ by Proposition 4.1.21(b). So the concatenation is a basis of V consisting of eigenvectors, and T is diagonalizable by Theorem 7.4.2.

So far nothing used the characteristic polynomial: (I), (II), and (III) are statements about eigenvectors only. Condition (IV) brings in the polynomial, through the squeeze (7.4.1).

Proof, part 2 of 2: (III) ⇔ (IV). STANDARD

> *Uses.* Theorem 7.4.15 (Geometric multiplicity is at most algebraic multiplicity), Proposition 7.1.27 (Multiplicities add up to at most the degree), Proposition 7.1.34 (Roots of a split polynomial), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Proposition 7.1.4 (Degree of a product), Example 7.4.14 (The multiplicities of a diagonal matrix)
> *Hint 1.* Look at the chain (7.4.1). Condition (III) says that its two ends are equal.
> *Hint 2.* A chain $a \le b \le c$ with $a = c$ forces $a = b$ and $b = c$. The first equality is a sum of termwise inequalities.
> *Hint 3.* For $\sum_i \operatorname{am}(\lambda_i) = n$ ⇔ "p_T splits": Proposition 7.1.27 gives a monic divisor of p_T of degree $\sum_i \operatorname{am}(\lambda_i)$; compare degrees.
> *Sketch.* In $\sum \operatorname{gm}(\lambda_i) \le \sum \operatorname{am}(\lambda_i) \le n$, (III) says the ends are equal. That holds iff $\operatorname{gm}(\lambda_i) = \operatorname{am}(\lambda_i)$ for each i (the terms satisfy gm ≤ am) and $\sum \operatorname{am}(\lambda_i) = n$. The latter means that $\prod (x - \lambda_i)^{\operatorname{am}(\lambda_i)}$, a monic divisor of p_T, has degree n, so equals p_T, i.e. p_T splits; conversely, if p_T splits, its roots are the λ_i and their multiplicities add up to n.

Write $g_i = \operatorname{gm}(\lambda_i)$ and $a_i = \operatorname{am}(\lambda_i)$. By Theorem 7.4.15, $g_i \le a_i$ for each i, so $\sum_i g_i \le \sum_i a_i$, with equality if and only if $g_i = a_i$ for every i. By Theorem 7.3.6 the λ_i are distinct roots of p_T, so by Proposition 7.1.27,

$$q = (x - \lambda_1)^{a_1} \cdots (x - \lambda_m)^{a_m} \text{ divides } p_T, \quad \text{and} \quad \sum_i a_i \le n.$$

We claim that $\sum_i a_i = n$ if and only if p_T splits. If $\sum_i a_i = n$, write $p_T = qh$. Comparing degrees (Proposition 7.1.4), $\deg h = 0$, and comparing leading coefficients, which are 1 for p_T and q, gives $h = 1$. So $p_T = q$ splits. Conversely, if p_T splits, then by Proposition 7.1.34 the multiplicities of its distinct roots add up to $\deg p_T = n$, and its roots are exactly the eigenvalues $\lambda_1, \ldots, \lambda_m$ (Theorem 7.3.6). So $\sum_i a_i = n$.

Now (III) says $\sum_i g_i = n$. Since $\sum_i g_i \le \sum_i a_i \le n$, this holds if and only if $\sum_i g_i = \sum_i a_i$ and $\sum_i a_i = n$, that is, if and only if $g_i = a_i$ for every i and p_T splits. This is (IV).

For the last claim, if $\mathcal{B}$ is an eigenbasis, $[T]_{\mathcal{B}\mathcal{B}}$ is a diagonal matrix of T, and by Example 7.4.14 and Lemma 7.4.13, λ_i occurs on its diagonal $\operatorname{gm}(\lambda_i)$ times. □

Remark 7.4.17 (Over $\mathbb{C}$, and over other fields). Over $\mathbb{C}$ every polynomial splits (Corollary 7.1.35), so a complex operator is diagonalizable if and only if $\operatorname{gm}(\lambda) = \operatorname{am}(\lambda)$ for every eigenvalue. Over $\mathbb{R}$, or over $\mathbb{Q}$ or $\mathbb{F}_p$, splitting is a genuine condition: the real rotation by a right angle has $p = x^2 + 1$, both sums in (7.4.1) are 0, and $n = 2$. Condition (III) handles both failures at once, which makes it the practical test: add up the dimensions of the eigenspaces and compare with n. Splitting alone is not enough, as the shear shows.

Algorithm 7.4.18 (Diagonalizing a matrix)

Input $A \in F^{n \times n}$.
Output An invertible S and a diagonal D with $A = SDS^{-1}$, or the conclusion that A is not diagonalizable over F.

1. Compute $p_A = \det(xI_n - A)$ and find its distinct roots $\lambda_1, \ldots, \lambda_m$ in F.
2. For each λ_i, row reduce $A - \lambda_i I_n$ and read off a basis of $E_{\lambda_i} = \ker(A - \lambda_i I_n)$, of length $g_i = n - \operatorname{rank}(A - \lambda_i I_n)$.
3. If $g_1 + \cdots + g_m < n$, stop: A is not diagonalizable over F.
4. Otherwise let S be the matrix whose columns are the basis vectors found in step 2, and D the diagonal matrix whose j-th entry is the eigenvalue of the j-th column of S. Check $AS = SD$.

The algorithm is correct by Theorem 7.4.16, (I) ⇔ (III), and the proof of (III) ⇒ (I) shows that the columns found in step 2 form a basis. In step 1 one can stop early if p_A does not split, and in step 2 as soon as some g_i is less than the multiplicity of λ_i as a root.

Example 7.4.19 (Two matrices that differ in one entry). Let

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad A' = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Both are upper triangular with diagonal $3, 2, 2$, so $p_A = p_{A'} = (x-3)(x-2)^2$: the eigenvalue 3 is simple, and $\operatorname{am}(2) = 2$ for both. For A,

$$A - 2I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

has rank 1, so $\operatorname{gm}_A(2) = 2$, with $E_2 = \{(x, y, z) \mid x + y + z = 0\} = \operatorname{span}((1, -1, 0), (1, 0, -1))$. Since 3 is simple, $\operatorname{gm}_A(3) = 1$, by Theorem 7.4.15, and $E_3 = \operatorname{span}(\mathbf{e}_1)$. The dimensions add up to 3, and

$$A = S\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}S^{-1}, \quad S = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

For A',

$$A' - 2I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

has rank 2, so $\operatorname{gm}_{A'}(2) = 1 < 2 = \operatorname{am}_{A'}(2)$, and A' is not diagonalizable. The single entry $a'_{23} = 1$ destroys one eigenvector for 2. In particular A and A' are not similar, although they have the same characteristic polynomial (Lemma 7.4.13).

Example 7.4.20 (The Fibonacci matrix over several fields). Let $Q = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, so that $p_Q = \det\left(\begin{smallmatrix} x-1 & -1 \\ -1 & x \end{smallmatrix}\right) = x^2 - x - 1$. Each failure mode of Theorem 7.4.16 occurs over some field.

- Over $\mathbb{F}_2$, $p_Q = x^2 + x + 1$ has no roots ($p_Q(0) = p_Q(1) = 1$). So Q has no eigenvalues and is not diagonalizable: p_Q does not split.
- Over $\mathbb{F}_5$, $(x+2)^2 = x^2 + 4x + 4 = x^2 - x - 1$, so p_Q splits with the single root $-2 = 3$, of algebraic multiplicity 2. But $Q - 3I_2 = \left(\begin{smallmatrix} -2 & 1 \\ 1 & -3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ has rank 1 (its second row is 2 times its first in $\mathbb{F}_5$), so $\operatorname{gm}(3) = 1 < 2$. Not diagonalizable: the multiplicities differ.
- Over $\mathbb{F}_{11}$, $(x-4)(x+3) = x^2 - x - 12 = x^2 - x - 1$, so Q has the distinct eigenvalues 4 and $-3 = 8$, and is diagonalizable by Corollary 7.4.5.
- Over $\mathbb{R}$, the roots $\varphi = \left(1 + \sqrt{5}\right)/2$ and $\psi = \left(1 - \sqrt{5}\right)/2$ are distinct, and Q is diagonalizable; Section 7.4.4 uses this to find the Fibonacci numbers.

The discriminant 5 of $x^2 - x - 1$ explains the pattern: the roots are $\left(1 \pm \sqrt{5}\right)/2$ whenever $1 + 1 \neq 0$, they coincide when $5 = 0$, and they lie in $\mathbb{F}_p$ when 5 is a square there, as $4^2 = 16 = 5$ is in $\mathbb{F}_{11}$.

Exercise 7.4.21. ROUTINE Let

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

Find $\mathrm{am}_A(\lambda)$ and $\mathrm{gm}_A(\lambda)$ for every eigenvalue λ, and decide whether A is diagonalizable. Which single entry of A above the diagonal can you change to make A diagonalizable?

Solution on page 1585

7.4.3 Computing with a diagonalization

A diagonalization turns questions about A into questions about the scalars on the diagonal of D, because conjugating by S respects products.

Technique 7.4.22 (Diagonalize, solve, conjugate back). If $A = SDS^{-1}$ with D diagonal, then $A^k = SD^kS^{-1}$, and more generally every expression built from A by sums, products, and scalar multiples is $S(\cdots)S^{-1}$ with D in place of A. Diagonal matrices are computed entry by entry, so a problem about A splits into n problems about scalars. The recipe: diagonalize; solve the problem for D, one diagonal entry at a time; conjugate back by S. In operator language, pass to eigenbasis coordinates, where T is n independent scalings, solve there, and translate back.

Reach for it when a problem involves powers or polynomials of a matrix (closed forms for A^k, square roots, $p(A)$), a linear recurrence $\mathbf{x}_{k+1} = A\mathbf{x}_k$, or a linear differential equation $\mathbf{x}' = A\mathbf{x}$. If A is not diagonalizable, the same recipe works with a triangular or Jordan form in place of D (Section 7.5, Section 10.3), at the price of harder computations with that form. The canonical example is Proposition 7.4.23.

Used in. Proposition 7.4.23, Theorem 7.4.30, Proposition 7.4.33, Proposition 7.4.34, Exercise 7.4.36, Exercise 7.4.44, Exercise 7.4.46, Exercise 7.7.11, Problem 7.8, Problem 7.13, Problem 7.16, Exercise 8.2.16, Exercise 8.2.32, Exercise 7.4.27, Exercise 7.4.52.

Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix). Let $A = SDS^{-1}$ with $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then for every integer $k \geq 0$ and every $p \in F[x]$,

$$A^k = S\begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}S^{-1}, \quad p(A) = S\begin{pmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_n) \end{pmatrix}S^{-1}.$$

If every λ_i is nonzero, then A is invertible, $A^{-1} = S\,\mathrm{diag}\left(\lambda_1^{-1}, \ldots, \lambda_n^{-1}\right)S^{-1}$, and the formula for A^k holds for all $k \in \mathbb{Z}$.

Proof. ROUTINE

Uses. Proposition 2.4.14 (Algebra of matrices), Example 2.4.10 (Diagonal matrices), Definition 7.2.26 (Polynomials in an operator), Proposition 2.5.17 (Algebra of inverses)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. Multiply out $(SDS^{-1})(SDS^{-1})$. What happens in the middle?
Hint 2. Induct on k, then use linearity for $p(A) = \sum_k a_k A^k$. Diagonal matrices multiply entry by entry.
Sketch. The inner factors cancel: $(SDS^{-1})\left(SD^kS^{-1}\right) = SD^{k+1}S^{-1}$, so by induction $A^k = SD^kS^{-1}$, and $D^k = \mathrm{diag}\left(\lambda_i^k\right)$ since diagonal matrices multiply entrywise. Summing, $p(A) = Sp(D)S^{-1}$ with $p(D) = \mathrm{diag}(p(\lambda_i))$. For the inverse, check that $S\,\mathrm{diag}\left(\lambda_i^{-1}\right)S^{-1}$ multiplies A to I_n on both sides.

We show $A^k = SD^kS^{-1}$ by induction on k. For $k = 0$ both sides are I_n. If it holds for k, then by associativity (Proposition 2.4.14)

$$A^{k+1} = AA^k = \left(SDS^{-1}\right)\left(SD^kS^{-1}\right) = SD\left(S^{-1}S\right)D^kS^{-1} = SD^{k+1}S^{-1}.$$

Diagonal matrices multiply entry by entry (Example 2.4.10), so $D^k = \operatorname{diag}\left(\lambda_1^k, \ldots, \lambda_n^k\right)$. For $p = \sum_k a_k x^k$, distributivity gives

$$p(A) = \sum_k a_k S D^k S^{-1} = S\left(\sum_k a_k D^k\right) S^{-1} = S \operatorname{diag}(p(\lambda_1), \ldots, p(\lambda_n)) S^{-1}.$$

If every $\lambda_i \neq 0$, let $D' = \operatorname{diag}\left(\lambda_1^{-1}, \ldots, \lambda_n^{-1}\right)$, so that $DD' = D'D = I_n$. Then $A\left(SD'S^{-1}\right) = SDD'S^{-1} = I_n$ and likewise $\left(SD'S^{-1}\right)A = I_n$, so A is invertible with $A^{-1} = SD'S^{-1}$. Applying the first part to $A^{-1} = SD'S^{-1}$ gives $A^{-k} = SD'^k S^{-1} = S \operatorname{diag}\left(\lambda_i^{-k}\right) S^{-1}$ for $k \geq 0$. □

Example 7.4.24 (A closed form for A^k). For $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} = S \operatorname{diag}(2,3) S^{-1}$ from Example 7.4.4,

$$A^k = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3^k - 2^k & 2^{k+1} - 2 \cdot 3^k \\ 3^k - 2^k & 2^{k+1} - 3^k \end{pmatrix}$$

for every $k \in \mathbb{Z}$. At $k = 1$ this returns A, and at $k = 0$ it returns I_2. Every entry of A^k is a combination $\alpha 2^k + \beta 3^k$, and for large k the term 3^k dominates.

Often only $A^k \mathbf{v}$ is needed, and then S^{-1} is unnecessary. Expand $\mathbf{v}$ in the eigenbasis, $\mathbf{v} = c_1 \mathbf{s}_1 + c_2 \mathbf{s}_2$; then $A^k \mathbf{v} = c_1 2^k \mathbf{s}_1 + c_2 3^k \mathbf{s}_2$, because $A^k \mathbf{s}_j = \lambda_j^k \mathbf{s}_j$ (Proposition 7.2.28). For instance $(0,1) = 2(1,1) - (2,1)$, so $A^k(0,1) = \left(2^{k+1} - 2 \cdot 3^k, 2^{k+1} - 3^k\right)$, the second column above.

The second paragraph of Example 7.4.24 proved a matrix identity by checking it on eigenvectors. For diagonalizable operators this is a general method.

Technique 7.4.25 (Test an identity on eigenvectors). To show that two operators built from a diagonalizable T are equal, such as $p(T) = 0$ or $S = q(T)$, it is enough to show that they agree on each vector of an eigenbasis. On an eigenvector $\mathbf{v}$ with eigenvalue λ, every polynomial in T acts as the scalar $p(\lambda)$ (Proposition 7.2.28), so the check becomes an identity between scalars, one eigenvalue at a time. This is Technique 2.2.7 with the most convenient basis.

Reach for it when the hypothesis says "diagonalizable" and the claim is an identity between polynomials in T, or between T and another operator with the same eigenvectors. The canonical example is Proposition 7.4.26. Without diagonalizability the method fails: the shear J satisfies $p(J)\mathbf{v} = p(1)\mathbf{v}$ on its only eigenvector line, but $(J - I)$ kills that line without being 0.

Used in. Proposition 7.4.26, Exercise 7.4.49, Exercise 7.4.52, Exercise 7.7.7, Proposition 8.2.20, Proposition 8.2.24, Notation 8.3.10, Exercise 8.3.40, Exercise 8.3.43, Exercise 8.2.33.

Proposition 7.4.26 (Polynomials that kill a diagonalizable operator). Let T be diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, and let $p \in F[x]$. Then $p(T) = 0$ if and only if $p(\lambda_i) = 0$ for $i = 1, \ldots, m$. In particular,

$$(T - \lambda_1 \operatorname{id})(T - \lambda_2 \operatorname{id}) \cdots (T - \lambda_m \operatorname{id}) = 0.$$

Proof. ROUTINE

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Proposition 7.2.27 (Evaluating polynomials at an operator)
Techniques. Test an identity on eigenvectors (Technique 7.4.25), Check it on a basis (Technique 2.2.7)
Hint 1. One direction is the last sentence of Proposition 7.2.28.
Hint 2. For the other, evaluate $p(T)$ on each vector of an eigenbasis.
Sketch. If $p(T) = 0$, every eigenvalue is a root of p (Proposition 7.2.28). Conversely, on an eigenbasis vector $\mathbf{b}$ with eigenvalue λ_i, $p(T)\mathbf{b} = p(\lambda_i)\mathbf{b} = \mathbf{0}$; a linear map that vanishes on a basis is 0. The product is $p(T)$ for $p = \prod(x - \lambda_i)$ (Proposition 7.2.27).

If $p(T) = 0$, then every eigenvalue of T is a root of p, by Proposition 7.2.28. Conversely, suppose $p(\lambda_i) = 0$ for every i, and let $\mathcal{B}$ be an eigenbasis. Each $\mathbf{b} \in \mathcal{B}$ has some eigenvalue λ_i, so $p(T)\mathbf{b} = p(\lambda_i)\mathbf{b} = \mathbf{0}$ by Proposition 7.2.28. The linear map $p(T)$ vanishes on a basis, so $p(T) = 0$. For the last claim, take $p = (x - \lambda_1)\cdots(x - \lambda_m)$; by Proposition 7.2.27(a), $p(T)$ is the product of the operators $T - \lambda_i \operatorname{id}$. □

For instance, a diagonalizable T whose eigenvalues lie in $\{0, 1\}$ satisfies $T^2 - T = T(T - \operatorname{id}) = 0$, so it is a projection; one with eigenvalues in $\{1, -1\}$ satisfies $T^2 = \operatorname{id}$. The converse of the last statement, that an operator killed by a product of distinct linear factors is diagonalizable, is also true. Exercise 7.4.49 proves it, and Section 7.6 builds the minimal polynomial on it.

> **Exercise 7.4.27.** ROUTINE Let $A \in F^{n\times n}$ be diagonalizable. Show that A^{T} and $p(A)$, for every $p \in F[x]$, are diagonalizable, and that A^{-1} is diagonalizable if A is invertible. How are the eigenvalues of $p(A)$ related to those of A?

Solution on page 1585

7.4.4 Linear recurrences

The Fibonacci numbers $f_0 = 0$, $f_1 = 1$, $f_{k+2} = f_{k+1} + f_k$ are generated by a matrix: with $\mathbf{u}_k = (f_{k+1}, f_k)$,

$$\mathbf{u}_{k+1} = \begin{pmatrix} f_{k+1} + f_k \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{u}_k = Q\mathbf{u}_k,$$

so $\mathbf{u}_k = Q^k\mathbf{u}_0$ with $\mathbf{u}_0 = (1, 0)$, as in Exercise 2.4.47. Diagonalizing Q gives a closed formula.

Example 7.4.28 (Binet's formula). By Example 7.4.20, Q has the distinct real eigenvalues $\varphi = \left(1 + \sqrt{5}\right)/2$ and $\psi = \left(1 - \sqrt{5}\right)/2$, the roots of $x^2 = x + 1$. Eigenvectors are $(\varphi, 1)$ and $(\psi, 1)$: for instance $Q(\varphi, 1) = (\varphi + 1, \varphi) = (\varphi^2, \varphi) = \varphi(\varphi, 1)$. Since $\varphi - \psi = \sqrt{5}$,

$$\mathbf{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}}\left(\begin{pmatrix} \varphi \\ 1 \end{pmatrix} - \begin{pmatrix} \psi \\ 1 \end{pmatrix}\right), \quad \text{so} \quad \mathbf{u}_k = Q^k\mathbf{u}_0 = \frac{1}{\sqrt{5}}\left(\varphi^k\begin{pmatrix} \varphi \\ 1 \end{pmatrix} - \psi^k\begin{pmatrix} \psi \\ 1 \end{pmatrix}\right).$$

The second coordinate is **Binet's formula**

$$f_k = \frac{\varphi^k - \psi^k}{\sqrt{5}} \quad \text{for all } k \geq 0.$$

Since $|\psi| < 1$, the term $\psi^k/\sqrt{5}$ has absolute value at most $1/\sqrt{5} < 1/2$, so f_k is the integer nearest to $\varphi^k/\sqrt{5}$. For instance $\varphi^{10}/\sqrt{5} \approx 55.004$, and $f_{10} = 55$. The eigenvalue $\varphi \approx 1.618$ of larger absolute value controls the growth.

The same formula appeared in Exercise 1.6.39 without any matrices, by finding a basis of geometric sequences of the space of all solutions. The two arguments are the same argument seen from two sides: the map $\mathbf{s} \mapsto (s_1, s_0)$ identifies the solution space with $\mathbb{R}^2$, and under it, shifting a sequence one step to the left becomes multiplication by Q. The eigenvectors $(\varphi, 1)$ and $(\psi, 1)$ correspond to the geometric sequences $\left(\varphi^k\right)$ and $\left(\psi^k\right)$, which the left shift multiplies by φ and ψ (Example 7.2.20). We now do this for every recurrence of finite order.

Fix $k \geq 1$ and $c_0, \ldots, c_{k-1} \in F$, and consider sequences $\mathbf{s} = (s_0, s_1, s_2, \ldots)$ in F with

$$s_{j+k} = c_{k-1}s_{j+k-1} + \cdots + c_1 s_{j+1} + c_0 s_j \quad \text{for all } j \geq 0. \tag{7.4.2}$$

The **characteristic polynomial of the recurrence** is $q = x^k - c_{k-1}x^{k-1} - \cdots - c_1 x - c_0$; a geometric sequence $\left(r^j\right)_{j\geq 0}$ satisfies (7.4.2) exactly when $r^j q(r) = 0$ for all j, that is, when $q(r) = 0$. Let W be the set of solutions of (7.4.2), a subspace of the sequence space F^X, $X = \{0, 1, 2, \ldots\}$ (the recurrence is linear in $\mathbf{s}$). Let $C \in F^{k\times k}$ be the matrix with

$$C\begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ x_k \end{pmatrix} = \begin{pmatrix} x_2 \\ \vdots \\ x_k \\ c_0x_1 + c_1x_2 + \cdots + c_{k-1}x_k \end{pmatrix}, \quad \text{that is,} \quad C = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ c_0 & c_1 & \cdots & c_{k-1} \end{pmatrix},$$

a companion matrix of q (compare Definition 7.3.23; conventions differ by a transpose and the order of the basis). For $k = 2$, $c_0 = c_1 = 1$, it is Q with the basis reversed.

Proposition 7.4.29 (The matrix of a recurrence). With the notation above:

(a) if $\mathbf{s} \in W$ and $\mathbf{u}_j = (s_j, s_{j+1}, \ldots, s_{j+k-1})$, then $\mathbf{u}_{j+1} = C\mathbf{u}_j$, so $\mathbf{u}_j = C^j\mathbf{u}_0$;

(b) the eigenvalues of C are the roots of q in F, and for each root r, $E_r(C) = \operatorname{span}\left(\left(1, r, r^2, \ldots, r^{k-1}\right)\right)$;

(c) C is diagonalizable if and only if q has k distinct roots in F.

Proof. STANDARD

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Definition 7.2.13 (Eigenspace), Theorem 7.4.16 (Criteria for diagonalizability)
Hint 1. (a) is the definition of C and the recurrence at index j.
Hint 2. For (b), write out $C\mathbf{v} = \lambda\mathbf{v}$ coordinate by coordinate. The first $k-1$ equations determine $\mathbf{v}$ from v_1.
Hint 3. For (c), every eigenspace is a line. Use criterion (III) of Theorem 7.4.16.
Sketch. $C\mathbf{v} = \lambda\mathbf{v}$ says $v_{i+1} = \lambda v_i$ for $i < k$, so $\mathbf{v} = v_1\left(1, \lambda, \ldots, \lambda^{k-1}\right)$, and the last coordinate then says $v_1 q(\lambda) = 0$. So eigenvalues are roots of q, each with a one-dimensional eigenspace. Then $\sum \operatorname{gm}$ = the number of distinct roots of q in F, which is k exactly when q has k distinct roots.

(a) The first $k-1$ coordinates of $C\mathbf{u}_j$ are $s_{j+1}, \ldots, s_{j+k-1}$, and the last is $c_0s_j + \cdots + c_{k-1}s_{j+k-1} = s_{j+k}$ by (7.4.2). So $C\mathbf{u}_j = \mathbf{u}_{j+1}$, and $\mathbf{u}_j = C^j\mathbf{u}_0$ by induction.

(b) Let $\lambda \in F$ and $\mathbf{v} = (v_1, \ldots, v_k) \in F^k$. Comparing coordinates, $C\mathbf{v} = \lambda\mathbf{v}$ says

$$v_2 = \lambda v_1, \quad v_3 = \lambda v_2, \quad \ldots, \quad v_k = \lambda v_{k-1}, \quad c_0v_1 + c_1v_2 + \cdots + c_{k-1}v_k = \lambda v_k.$$

The first $k-1$ equations hold if and only if $v_i = \lambda^{i-1}v_1$ for all i, that is, $\mathbf{v} = v_1\mathbf{w}_\lambda$ with $\mathbf{w}_\lambda = \left(1, \lambda, \ldots, \lambda^{k-1}\right)$. Substituting into the last equation gives $v_1\left(c_0 + c_1\lambda + \cdots + c_{k-1}\lambda^{k-1}\right) = v_1\lambda^k$, that is, $v_1 q(\lambda) = 0$. So if $q(\lambda) \neq 0$, then $v_1 = 0$ and $\mathbf{v} = \mathbf{0}$: λ is not an eigenvalue. If $q(\lambda) = 0$, the solutions are exactly the multiples of $\mathbf{w}_\lambda \neq \mathbf{0}$, so λ is an eigenvalue with $E_\lambda(C) = \operatorname{span}(\mathbf{w}_\lambda)$.

(c) By (b), $\operatorname{gm}_C(r) = 1$ for each root r of q in F, so the sum of the geometric multiplicities is the number of distinct roots of q in F. By Theorem 7.4.16, (I) $\Leftrightarrow$ (III), C is diagonalizable if and only if this number is k. □

So C never has two independent eigenvectors for the same eigenvalue: a repeated root of q always destroys diagonalizability. When the roots are distinct, the solutions are combinations of geometric sequences.

Theorem 7.4.30 (Recurrences with distinct characteristic roots). Suppose q has k distinct roots $r_1, \ldots, r_k \in F$. Then the geometric sequences $\left(r_i^j\right)_{j\geq 0}$, $i = 1, \ldots, k$, form a basis of W. That is, every solution of (7.4.2) is

$$s_j = a_1r_1^j + a_2r_2^j + \cdots + a_kr_k^j \quad (j \geq 0)$$

for unique $a_1, \ldots, a_k \in F$, which are determined by $s_0, \ldots, s_{k-1}$.

Proof (diagonalize the matrix of the recurrence). STANDARD

Uses. Proposition 7.4.29 (The matrix of a recurrence), Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. By Proposition 7.4.29, $\mathbf{u}_j = C^j\mathbf{u}_0$ and C has the eigenbasis $\mathbf{w}_{r_1}, \ldots, \mathbf{w}_{r_k}$. Expand $\mathbf{u}_0$ in it.

Hint 2. Then $C^j\mathbf{u}_0 = \sum_i a_i r_i^j \mathbf{w}_{r_i}$. Read off the first coordinate, and note that each $\mathbf{w}_{r_i}$ has first coordinate 1.
Hint 3. For uniqueness, $s_0, \dots, s_{k-1}$ are the coordinates of $\mathbf{u}_0$, and coordinates in a basis are unique.
Sketch. The vectors $\mathbf{w}_{r_i} = \left(1, r_i, \dots, r_i^{k-1}\right)$ are eigenvectors for distinct eigenvalues, so an eigenbasis. Write $\mathbf{u}_0 = \sum a_i \mathbf{w}_{r_i}$; then $\mathbf{u}_j = C^j\mathbf{u}_0 = \sum a_i r_i^j \mathbf{w}_{r_i}$, whose first coordinate is $s_j = \sum a_i r_i^j$. So the geometric sequences span W; they lie in W since $q(r_i) = 0$. If $\sum a_i r_i^j = 0$ for all j, then $\sum a_i \mathbf{w}_{r_i} = \mathbf{0}$ (the cases $j < k$), so all $a_i = 0$.

Write $\mathbf{w}_i = \left(1, r_i, r_i^2, \dots, r_i^{k-1}\right)$. By Proposition 7.4.29(b), $\mathbf{w}_i$ is an eigenvector of C with eigenvalue r_i, and since the r_i are distinct, $\mathbf{w}_1, \dots, \mathbf{w}_k$ is a basis of F^k (Corollary 7.4.5). Each geometric sequence $\left(r_i^j\right)_j$ lies in W, because $q(r_i) = 0$.

Let $\mathbf{s} \in W$, and write $\mathbf{u}_0 = (s_0, \dots, s_{k-1}) = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k$. By Proposition 7.4.29(a) and Proposition 7.2.28,

$$\mathbf{u}_j = C^j\mathbf{u}_0 = \sum_{i=1}^{k} a_i C^j \mathbf{w}_i = \sum_{i=1}^{k} a_i r_i^j \mathbf{w}_i.$$

The first coordinate of $\mathbf{u}_j$ is s_j, and the first coordinate of each $\mathbf{w}_i$ is 1, so $s_j = \sum_i a_i r_i^j$ for all j. Thus the geometric sequences span W.

They are independent: if $\sum_i a_i r_i^j = 0$ for all $j \geq 0$, then the cases $j = 0, \dots, k-1$ say $\sum_i a_i \mathbf{w}_i = \mathbf{0}$, so all $a_i = 0$. Hence they form a basis of W, the a_i are unique, and they are the coordinates of $\mathbf{u}_0 = (s_0, \dots, s_{k-1})$ in the basis $\mathbf{w}_1, \dots, \mathbf{w}_k$. □

Proof (count dimensions in the solution space). STANDARD

Uses. Theorem 4.1.10 (Independent lists are short), Example 7.2.24 (Independence for free), Definition 4.5.1 (Isomorphism)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Show that $\mathbf{s} \mapsto (s_0, \dots, s_{k-1})$ is an isomorphism $W \to F^k$. So $\dim W = k$.
Hint 2. The k geometric sequences lie in W. Why are they independent? Think of the left shift.
Hint 3. k independent vectors in a k-dimensional space form a basis.
Sketch. A solution is determined by its first k terms, and any first k terms extend (define s_{j+k} by (7.4.2)), so $\mathbf{s} \mapsto (s_0, \dots, s_{k-1})$ is a linear bijection and $\dim W = k$. The geometric sequences lie in W and are eigenvectors of the left shift with distinct eigenvalues r_i, hence independent; k of them form a basis.

Let $\Phi : W \to F^k, \Phi(\mathbf{s}) = (s_0, \dots, s_{k-1})$. It is linear. It is injective: if $s_0 = \cdots = s_{k-1} = 0$, then (7.4.2) and induction on j give $s_j = 0$ for all j. It is surjective: given $(t_0, \dots, t_{k-1})$, set $s_j = t_j$ for $j < k$ and define s_{j+k} for $j = 0, 1, 2, \dots$ by (7.4.2); the resulting sequence lies in W. So Φ is an isomorphism (Definition 4.5.1) and $\dim W = k$.

Since $q(r_i) = 0$, each geometric sequence $\left(r_i^j\right)_j$ lies in W. These k sequences are eigenvectors of the left shift on F^X with the distinct eigenvalues $r_1, \dots, r_k$, so they are independent (Example 7.2.24(a)). By Theorem 4.1.10(b) they form a basis of the k-dimensional space W, and the coefficients a_i of a solution are its coordinates in this basis, which are unique. Applying Φ, they are determined by $s_0, \dots, s_{k-1}$. □

The first proof computes: it finds the a_i by solving one linear system, $\mathbf{u}_0 = \sum a_i \mathbf{w}_i$, whose matrix has rows $(1, \dots, 1), (r_1, \dots, r_k), \dots$ (a Vandermonde matrix, Theorem 5.5.21). The second proof needs no matrix at all, only the dimension of W; it is Technique 4.2.24, and it is the proof that generalizes to linear differential equations (Section 15.2).

Remark 7.4.31 (Repeated roots). If q has a repeated root, C is not diagonalizable (Proposition 7.4.29(c)), and the geometric sequences no longer span W. For $s_{j+2} = 2s_{j+1} - s_j$ we have $q = (x-1)^2$, the only geometric solution is the constant sequence, and W consists of the arithmetic sequences $s_j = a + bj$. In general, a root r of multiplicity m with $r \neq 0$ contributes the solutions $\left(j^i r^j\right)_j$ for $i < m$ (when F has characteristic 0); the source of these is the Jordan form of C (Section 10.3), and Section 15.2 develops them.

7.4.5 Discrete dynamical systems

A **discrete linear dynamical system** is a sequence of states $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ in F^n with $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for a fixed $A \in F^{n\times n}$, so $\mathbf{x}_k = A^k\mathbf{x}_0$. If A has an eigenbasis $\mathbf{s}_1, \ldots, \mathbf{s}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, and $\mathbf{x}_0 = c_1\mathbf{s}_1 + \cdots + c_n\mathbf{s}_n$, then

$$\mathbf{x}_k = c_1\lambda_1^k\mathbf{s}_1 + c_2\lambda_2^k\mathbf{s}_2 + \cdots + c_n\lambda_n^k\mathbf{s}_n. \tag{7.4.3}$$

Each eigen-component evolves on its own, by repeated multiplication by its eigenvalue. Over $\mathbb{R}$ or $\mathbb{C}$, a component with $|\lambda_i| < 1$ dies out, one with $|\lambda_i| > 1$ grows, and the components with the largest $|\lambda_i|$ eventually dominate.

Example 7.4.32 (Migration between two cities). Each year, 20% of the population of city 1 moves to city 2, and 30% of the population of city 2 moves to city 1; nobody else moves, and the total population is constant. With $\mathbf{x}_k = (a_k, b_k)$ the populations in year k,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \quad A = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}.$$

Here $\det(A - \lambda I_2) = \lambda^2 - 1.5\lambda + 0.5 = (\lambda - 1)(\lambda - 0.5)$, with eigenvectors $(3, 2)$ for 1 and $(1, -1)$ for 0.5. Writing $\mathbf{x}_0 = (a, b) = c_1(3, 2) + c_2(1, -1)$ gives $c_1 = (a + b)/5$ and $c_2 = (2a - 3b)/5$, so by (7.4.3)

$$\mathbf{x}_k = \frac{a+b}{5}\begin{pmatrix}3\\2\end{pmatrix} + \frac{2a-3b}{5}\cdot 2^{-k}\begin{pmatrix}1\\-1\end{pmatrix} \quad\to\quad \frac{a+b}{5}\begin{pmatrix}3\\2\end{pmatrix} \quad (k \to \infty).$$

Whatever the initial distribution, the populations approach a $60 : 40$ split of the total $a + b$. The limit is the component of $\mathbf{x}_0$ in E_1, and the deviation from it halves every year. The columns of A add up to 1, which is why 1 is an eigenvalue (the rows of $A - I_2$ add up to zero) and why the total is preserved. Such matrices are the subject of Section 11.1.

Proposition 7.4.33 (Long-run behavior). Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable, with an eigenbasis $\mathbf{s}_1, \ldots, \mathbf{s}_n$ and corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $\mathbf{x}_0 = c_1\mathbf{s}_1 + \cdots + c_n\mathbf{s}_n$ and $\mathbf{x}_k = A^k\mathbf{x}_0$. Limits of vectors are taken coordinate by coordinate.

(a) If $|\lambda_i| < 1$ for every i, then $\mathbf{x}_k \to \mathbf{0}$ for every $\mathbf{x}_0$, and $A^k \to 0$ entry by entry.

(b) Suppose $\mu \neq 0$ is an eigenvalue with $|\lambda_i| < |\mu|$ for every $\lambda_i \neq \mu$. Then

$$\mu^{-k}\mathbf{x}_k \to \sum_{i:\lambda_i=\mu} c_i\mathbf{s}_i,$$

the component of $\mathbf{x}_0$ in E_μ. In particular, if $\mu = 1$, then $\mathbf{x}_k$ converges to the component of $\mathbf{x}_0$ in E_1.

Proof. STANDARD

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.4.16 (Criteria for diagonalizability)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. Start from (7.4.3). Each coordinate of $\mathbf{x}_k$ is a fixed linear combination of the λ_i^k.
Hint 2. For complex z with $|z| < 1$, $|z^k| = |z|^k \to 0$.
Hint 3. For (b), divide (7.4.3) by μ^k and split the sum according to whether $\lambda_i = \mu$.
Sketch. By (7.4.3), $\mathbf{x}_k = \sum c_i\lambda_i^k\mathbf{s}_i$, a combination with fixed vectors and coefficients λ_i^k. If all $|\lambda_i| < 1$, each $\lambda_i^k \to 0$. For (b), $\mu^{-k}\mathbf{x}_k = \sum_{\lambda_i=\mu} c_i\mathbf{s}_i + \sum_{\lambda_i\neq\mu} c_i(\lambda_i/\mu)^k\mathbf{s}_i$, and the second sum tends to $\mathbf{0}$. Apply (a) to the columns $A^k\mathbf{e}_j$ to get $A^k \to 0$.

Since $A^k\mathbf{s}_i = \lambda_i^k\mathbf{s}_i$ (Proposition 7.2.28), (7.4.3) holds: $\mathbf{x}_k = \sum_i c_i\lambda_i^k\mathbf{s}_i$. Each coordinate of $\mathbf{x}_k$ is therefore $\sum_i c_i\lambda_i^k s_{i,l}$, where $s_{i,l}$ is the l-th coordinate of $\mathbf{s}_i$, a sum of n fixed complex numbers times k-th powers.

(a) If $|\lambda_i| < 1$, then $|\lambda_i^k| = |\lambda_i|^k \to 0$, so $\lambda_i^k \to 0$, and each coordinate of $\mathbf{x}_k$ tends to 0. The j-th column of A^k is $A^k\mathbf{e}_j$, which is the case $\mathbf{x}_0 = \mathbf{e}_j$, so every entry of A^k tends to 0.

(b) Dividing by μ^k,

$$\mu^{-k}\mathbf{x}_k = \sum_{i:\lambda_i=\mu} c_i\mathbf{s}_i + \sum_{i:\lambda_i\neq\mu} c_i\left(\frac{\lambda_i}{\mu}\right)^k \mathbf{s}_i.$$

The first sum does not depend on k, and in the second $|\lambda_i/\mu| < 1$, so each term tends to $\mathbf{0}$ as in (a). The first sum lies in E_μ, and the second in the sum of the other eigenspaces; since $\mathbb{C}^n$ is the direct sum of the eigenspaces (Theorem 7.4.16), the first sum is the component of $\mathbf{x}_0$ in E_μ. □

In Example 7.4.32, $\mu = 1$ and the other eigenvalue is 0.5. In Example 7.4.28, $\mu = \varphi$ dominates ψ, so $\varphi^{-k}\mathbf{u}_k \to (\varphi, 1)/\sqrt{5}$: the ratio f_{k+1}/f_k of consecutive Fibonacci numbers tends to φ. When A is not diagonalizable, (7.4.3) is replaced by terms $k^i\lambda^k$ coming from the Jordan form, and (a) still holds, but the proof needs Section 10.3.

7.4.6 Systems of linear differential equations BEYOND 21-242

Differential equations are not part of 21-242.

The continuous analog of $\mathbf{x}_{k+1} = A\mathbf{x}_k$ is the system $\mathbf{x}' = A\mathbf{x}$, where $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ is differentiable and the derivative is taken coordinate by coordinate. For $n = 1$ this is $x' = \lambda x$, whose solutions are $x(t) = x(0)e^{\lambda t}$: if $x' = \lambda x$, then $\left(e^{-\lambda t}x(t)\right)' = e^{-\lambda t}(x' - \lambda x) = 0$, so $e^{-\lambda t}x(t)$ is constant. A diagonalization decouples a system into n such equations.

Proposition 7.4.34 (Solving $\mathbf{x}' = A\mathbf{x}$). Let $A = SDS^{-1} \in \mathbb{R}^{n\times n}$ with $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ real, and let $\mathbf{s}_1, \ldots, \mathbf{s}_n$ be the columns of S. A differentiable $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ satisfies $\mathbf{x}' = A\mathbf{x}$ if and only if

$$\mathbf{x}(t) = c_1e^{\lambda_1 t}\mathbf{s}_1 + c_2e^{\lambda_2 t}\mathbf{s}_2 + \cdots + c_ne^{\lambda_n t}\mathbf{s}_n$$

for some $c_1, \ldots, c_n \in \mathbb{R}$, namely $(c_1, \ldots, c_n) = S^{-1}\mathbf{x}(0)$.

Proof. STANDARD

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. Change variables: let $\mathbf{y}(t) = S^{-1}\mathbf{x}(t)$. Differentiation commutes with multiplication by a constant matrix.
Hint 2. Then $\mathbf{y}' = D\mathbf{y}$, which is n separate equations $y_{i'} = \lambda_i y_i$.
Sketch. With $\mathbf{y} = S^{-1}\mathbf{x}$, $\mathbf{y}' = S^{-1}\mathbf{x}'$ (each coordinate is a fixed linear combination), so $\mathbf{x}' = A\mathbf{x}$ iff $\mathbf{y}' = D\mathbf{y}$, iff $y_{i'} = \lambda_i y_i$ for all i, iff $y_i(t) = c_ie^{\lambda_i t}$. Then $\mathbf{x} = S\mathbf{y} = \sum c_ie^{\lambda_i t}\mathbf{s}_i$.

Let $\mathbf{y}(t) = S^{-1}\mathbf{x}(t)$. Each coordinate of $\mathbf{y}$ is a fixed linear combination of the coordinates of $\mathbf{x}$, so $\mathbf{y}$ is differentiable with $\mathbf{y}' = S^{-1}\mathbf{x}'$, and similarly $\mathbf{x}' = S\mathbf{y}'$. Hence

$$\mathbf{x}' = A\mathbf{x} \iff S\mathbf{y}' = SDS^{-1}S\mathbf{y} \iff \mathbf{y}' = D\mathbf{y},$$

multiplying by S^{-1} or S. The system $\mathbf{y}' = D\mathbf{y}$ consists of the n separate equations $y_{i'} = \lambda_i y_i$, whose solutions are exactly $y_i(t) = c_ie^{\lambda_i t}$ with $c_i = y_i(0)$, by the case $n = 1$ above. So $\mathbf{x}' = A\mathbf{x}$ if and only if $\mathbf{x} = S\mathbf{y} = \sum_i c_ie^{\lambda_i t}\mathbf{s}_i$, and then $(c_1, \ldots, c_n) = \mathbf{y}(0) = S^{-1}\mathbf{x}(0)$. □

For $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{x}(0) = (1, 0)$, Example 7.4.4 gives $S^{-1}\mathbf{x}(0) = (-1, 1)$, so

$$\mathbf{x}(t) = -e^{2t}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{3t}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2e^{3t} - e^{2t} \\ e^{3t} - e^{2t} \end{pmatrix}.$$

In matrix form, the solution is $\mathbf{x}(t) = S\operatorname{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right)S^{-1}\mathbf{x}(0)$, and the matrix in front of $\mathbf{x}(0)$ is the **matrix exponential** e^{tA}, which Technique 7.4.22 computes from e^{tD}. Complex eigenvalues produce oscillations: for the rotation generator $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, diagonalizable over $\mathbb{C}$ with eigenvalues $\pm i$, the solutions are $\mathbf{x}(t) = (\alpha\cos t - \beta\sin t, \alpha\sin t + \beta\cos t)$, circles traversed at unit angular speed. A non-diagonalizable A produces terms $t^je^{\lambda t}$, which come from the Jordan form (Section 10.3).

Exercises

Exercise 7.4.35. ROUTINE For each matrix over $\mathbb{Q}$, find an invertible S and a diagonal D with $A = SDS^{-1}$, or show that none exist.

(a) $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$.

(b) $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$.

(c) $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

Solution on page 1586

Exercise 7.4.36. STANDARD Let $A = \begin{pmatrix} 7 & -10 \\ 3 & -4 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

(a) Find a formula for A^k valid for every $k \in \mathbb{Z}$.

(b) Find all $B \in \mathbb{R}^{2\times 2}$ with $B^2 = A$.

Solution on page 1586

Exercise 7.4.37. STANDARD For which $a, b, c \in F$ is the matrix

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 2 \end{pmatrix}$$

diagonalizable?

Solution on page 1587

Exercise 7.4.38. STANDARD Let $n \geq 2$ and $A, B \in F^{n\times n}$. Prove or disprove:

(a) If A is diagonalizable, then A is invertible.

(b) If A is invertible, then A is diagonalizable.

(c) If A^2 is diagonalizable, then A is diagonalizable.

(d) If A and B are diagonalizable, then so is $A + B$.

(e) If A is diagonalizable and B is similar to A, then B is diagonalizable.

Solution on page 1587

Exercise 7.4.39. ROUTINE Find the error in the following argument.

"Every $A \in \mathbb{C}^{n\times n}$ is diagonalizable. By the Fundamental Theorem of Algebra, $p_A = (x - \lambda_1)\cdots(x - \lambda_n)$ for some $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Each λ_i is a root of p_A, hence an eigenvalue; choose an eigenvector $\mathbf{v}_i$ for λ_i. Eigenvectors for different eigenvalues are independent, so $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an independent list of length n, hence a basis of $\mathbb{C}^n$, and it is an eigenbasis."

Solution on page 1588

Exercise 7.4.40. STANDARD Let $A \in F^{n\times n}$ have rank 1. Show that A is diagonalizable if and only if $\operatorname{tr} A \neq 0$, and that in this case A is similar to $\operatorname{diag}(\operatorname{tr} A, 0, \ldots, 0)$.

Solution on page 1588

Exercise 7.4.41. STANDARD Let $n \geq 2$, and let $\tau \in \mathrm{L}(F^{n\times n})$ be the transpose map, $\tau(X) = X^{\mathsf{T}}$. Show that τ is diagonalizable if and only if $1 + 1 \neq 0$ in F. In that case, find $\mathrm{gm}(1)$ and $\mathrm{gm}(-1)$.

Solution on page 1589

Exercise 7.4.42. STANDARD

(a) Show that if $T \in \mathrm{L}(V)$ is diagonalizable, then $V = \ker T \oplus \operatorname{im} T$.

(b) Give an example of a non-diagonalizable T with $V = \ker T \oplus \operatorname{im} T$, and one with $\ker T \cap \operatorname{im} T \neq \{\mathbf{0}\}$.

Solution on page 1589

Exercise 7.4.43. STANDARD

(a) Let $A, B \in F^{n\times n}$ be diagonalizable. Show that A and B are similar if and only if $p_A = p_B$.

(b) Show that (a) fails if only one of A, B is assumed diagonalizable.

Solution on page 1589

Exercise 7.4.44. STANDARD Let f_k be the Fibonacci numbers, and work in $\mathbb{F}_{11}$.

(a) Using Example 7.4.20, show that $f_k = 3\left(8^k - 4^k\right)$ in $\mathbb{F}_{11}$ for all $k \geq 0$.

(b) Deduce that $f_{k+10} = f_k$ in $\mathbb{F}_{11}$ for all k, that is, the Fibonacci numbers modulo 11 repeat with period dividing 10.

Solution on page 1590

Exercise 7.4.45. ROUTINE Find a closed formula for the sequence of real numbers with $s_0 = 3, s_1 = 2, s_2 = 6$, and

$$s_{j+3} = 2s_{j+2} + s_{j+1} - 2s_j \quad (j \geq 0).$$

Solution on page 1590

Exercise 7.4.46. STANDARD Two quantities evolve by

$$x_{k+1} = 0.7x_k + 0.4y_k, \quad y_{k+1} = 0.2x_k + 0.9y_k.$$

(a) Find (x_k, y_k) in closed form in terms of (x_0, y_0).

(b) Show that if $x_0, y_0 \geq 0$ are not both 0, then $x_k/y_k \to 1$ and both quantities eventually grow by about 10% per step.

(c) For which $(x_0, y_0) \in \mathbb{R}^2$ does $(x_k, y_k) \to (0, 0)$?

Solution on page 1590

Exercise 7.4.47. STANDARD Let $A \in \mathbb{R}^{n\times n}$ be diagonalizable over $\mathbb{C}$, and suppose all its complex eigenvalues are real. Show that A is diagonalizable over $\mathbb{R}$.

Solution on page 1591

Exercise 7.4.48. STANDARD Find all $A \in \mathbb{R}^{2\times 2}$ with $A^2 = I_2$. Describe them in terms of their entries, and geometrically.

Solution on page 1591

Exercise 7.4.49. HARD Let $q = (x - r_1)(x - r_2)\cdots(x - r_k) \in F[x]$ with $r_1, \ldots, r_k \in F$ distinct, and suppose $q(T) = 0$.

(a) Show that T is diagonalizable. (Use the Lagrange polynomials $L_1, \ldots, L_k$ for the nodes $r_1, \ldots, r_k$: show that $L_1 + \cdots + L_k = 1$ and that $\operatorname{im} L_j(T) \subseteq E_{r_j}$.)

(b) Deduce that every $A \in \mathbb{C}^{n\times n}$ with $A^m = I_n$ for some $m \geq 1$ is diagonalizable over $\mathbb{C}$.

(c) Let p be prime. Deduce that every $A \in \mathbb{F}_p^{n\times n}$ with $A^p = A$ is diagonalizable over $\mathbb{F}_p$.

Solution on page 1592

Exercise 7.4.50. HARD Let $T \in \mathrm{L}(V)$ be diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, and let $U \subseteq V$ be a T-invariant subspace. Show that

$$U = \big(U \cap E_{\lambda_1}\big) \oplus \cdots \oplus \big(U \cap E_{\lambda_m}\big),$$

and deduce that the restriction of T to U is diagonalizable.

Solution on page 1592

Exercise 7.4.51. HARD BEYOND 21-242 Let $S, T \in \mathrm{L}(V)$ be diagonalizable with $ST = TS$. Show that S and T are **simultaneously diagonalizable**: there is a basis of V consisting of vectors that are eigenvectors of both S and T. Show by example that the hypothesis $ST = TS$ cannot be dropped.

Solution on page 1593

Exercise 7.4.52. HARD Let $A \in \mathbb{C}^{n\times n}$ satisfy $A^3 = A$. Show that $\operatorname{rank} A = \operatorname{tr}\big(A^2\big)$.

Solution on page 1593

Exercise 7.4.53. CHALLENGE Let V be a finite-dimensional complex vector space and $T \in \mathrm{L}(V)$. Show that T is diagonalizable if and only if

$$\ker(T - \lambda\,\mathrm{id}) \cap \operatorname{im}(T - \lambda\,\mathrm{id}) = \{\mathbf{0}\} \quad \text{for every } \lambda \in \mathbb{C}.$$

Solution on page 1594

7.5 Upper-triangular form and Schur's theorem

Not every operator can be diagonalized, even over $\mathbb{C}$. The shift $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has the single eigenvalue 0 and $\ker N = \operatorname{span}(\mathbf{e}_1)$, so $\mathbb{C}^2$ has no basis of eigenvectors of N (Section 7.4). But N is already upper triangular, and an upper-triangular matrix shows most of what a diagonal one does: its eigenvalues, its determinant, and its trace can all be read off the diagonal. This section shows that over $\mathbb{C}$ every operator has an upper-triangular matrix (Theorem 7.5.13), and that on an inner product space the basis can be taken orthonormal (Schur's theorem, Theorem 7.5.18). The cost of settling for triangular is the entries above the diagonal; in exchange, the result holds for every operator.

Throughout, V is a finite-dimensional vector space over a field F with $\dim V = n \geq 1$, and $T \in \mathrm{L}(V)$. We abbreviate $[T]_{\mathcal{B}\mathcal{B}}$ to $[T]_{\mathcal{B}}$.

7.5.1 Upper-triangular matrices and invariant flags

Recall (Definition 3.3.31) that $A \in F^{n\times n}$ is upper triangular if $a_{ij} = 0$ whenever $i > j$. For the matrix of an operator this condition has a clean meaning. Column j of $[T]_{\mathcal{B}}$ holds the coordinates of $T\mathbf{b}_j$, so the zeros

below the diagonal in column j say exactly that $T\mathbf{b}_j$ is a combination of $\mathbf{b}_1, \ldots, \mathbf{b}_j$ alone: each basis vector is sent into the span of itself and the vectors before it.

Definition 7.5.1 (Triangularizable). Let $\mathcal{B}$ be a basis of V. The operator T is **upper triangular with respect to $\mathcal{B}$** if $[T]_{\mathcal{B}}$ is upper triangular, and T is **triangularizable** if it is upper triangular with respect to some basis of V. A matrix $A \in F^{n\times n}$ is **triangularizable** if it is similar to an upper-triangular matrix, that is (Theorem 4.5.25), if T_A is triangularizable.

Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces). Let $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V, and let $V_j = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$ for $0 \le j \le n$, so that $V_0 = \{\mathbf{0}\}$. The following are equivalent.

(I) $[T]_{\mathcal{B}}$ is upper triangular.

(II) $T\mathbf{b}_j \in V_j$ for each $j = 1, \ldots, n$.

(III) V_j is T-invariant for each $j = 1, \ldots, n$.

Proof. ROUTINE

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.14 (Coordinates), Proposition 7.2.4 (Invariance is checked on a spanning list), Definition 7.2.1 (Invariant subspace)
Techniques. Proving TFAE (Technique 2.7.1), Check it on a basis (Technique 2.2.7)
Hint 1. Column j of $[T]_{\mathcal{B}}$ is $[T\mathbf{b}_j]_{\mathcal{B}}$. When does a vector lie in V_j, in terms of its coordinates?
Hint 2. For (II) $\Rightarrow$ (III), V_j is spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_j$, and $T\mathbf{b}_i \in V_i \subseteq V_j$ for $i \le j$. Use Proposition 7.2.4.
Sketch. A vector lies in V_j exactly when its coordinates after position j vanish, so (I) $\Leftrightarrow$ (II) column by column. Since $V_i \subseteq V_j$ for $i \le j$, (II) gives $T\mathbf{b}_i \in V_j$ for all $i \le j$, and invariance follows from Proposition 7.2.4. Conversely, $\mathbf{b}_j \in V_j$, so invariance of V_j gives (II).

We prove (I) $\Leftrightarrow$ (II) and (II) $\Leftrightarrow$ (III).

(I) $\Leftrightarrow$ (II). A vector of V lies in V_j if and only if its $\mathcal{B}$-coordinates in positions $j+1, \ldots, n$ are 0, by uniqueness of coordinates (Definition 1.6.14). Column j of $[T]_{\mathcal{B}}$ is $[T\mathbf{b}_j]_{\mathcal{B}}$ (Definition 2.3.1), so $T\mathbf{b}_j \in V_j$ if and only if column j has zeros in rows $j+1, \ldots, n$. These conditions for $j = 1, \ldots, n$ together say that $a_{ij} = 0$ whenever $i > j$.

(II) $\Rightarrow$ (III). Fix j. For $i \le j$ we have $T\mathbf{b}_i \in V_i \subseteq V_j$. So T maps each vector of the spanning list $\mathbf{b}_1, \ldots, \mathbf{b}_j$ of V_j into V_j, and V_j is T-invariant by Proposition 7.2.4.

(III) $\Rightarrow$ (II). $\mathbf{b}_j \in V_j$, and V_j is T-invariant, so $T\mathbf{b}_j \in V_j$. □

The subspaces $V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n$ in the proposition are nested and grow by one dimension at a time. Such chains deserve a name.

Definition 7.5.3 (Flag). A **flag** in V is a list of subspaces $V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n$ with $\dim V_j = j$ for each j (so $V_0 = \{\mathbf{0}\}$ and $V_n = V$). The **flag of a basis** $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is given by $V_j = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$. A flag is **$T$-invariant** if every V_j is T-invariant.

Every flag is the flag of some basis: choose $\mathbf{b}_j \in V_j$ with $\mathbf{b}_j \notin V_{j-1}$ for each j. By induction on j, the list $\mathbf{b}_1, \ldots, \mathbf{b}_j$ is linearly independent (Proposition 1.5.14) and lies in V_j, so it is a basis of V_j (Theorem 4.1.10). In $\mathbb{R}^3$, a flag is a line through $\mathbf{0}$ lying in a plane through $\mathbf{0}$. So Proposition 7.5.2 says that T is triangularizable if and only if V has a T-invariant flag, and that then every basis of that flag triangularizes T. Only the flag matters, not the particular vectors. In particular, the first vector $\mathbf{b}_1$ of a triangularizing basis spans an invariant line, so it is an eigenvector of T, with eigenvalue the $(1,1)$ entry (Proposition 7.2.7). A triangularizable operator has at least one eigenvalue.

Technique 7.5.4 (Build a flag of invariant subspaces). To find a basis in which T is upper triangular, look for nested invariant subspaces of every dimension instead of for matrix entries (Proposition 7.5.2). The flag can be built from the bottom (first an eigenvector, then a vector sent into the span of it and

itself, and so on) or from the top (a large invariant subspace, then invariant subspaces inside it). Any basis adapted to the flag will do, so the vectors can afterwards be changed freely within the flag, for instance orthonormalized.

Reach for it whenever a statement is about triangular matrices of an operator, both to produce them and to use them: "$[T]_{\mathcal{B}}$ is upper triangular" is the statement "T maps $\operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$ into itself for every j", which survives any operation that preserves these spans. The canonical example is the second proof of Theorem 7.5.13. Later uses: Schur's theorem (Theorem 7.5.18), where Gram–Schmidt keeps the flag, and simultaneous triangularization of commuting operators (Section 7.6).

Used in. Theorem 7.5.13, Theorem 7.5.18, Lemma 7.6.21, Theorem 7.6.32, Theorem 8.2.12, Exercise 7.5.30, Exercise 7.5.34.

Example 7.5.5 (Differentiation and shifts).

(a) Let $D \in \mathrm{L}\big(F[x]_{\le 3}\big)$ be differentiation, and use the basis $1, x, x^2, x^3$. Differentiation lowers degrees, so D maps $F[x]_{\le k}$ into $F[x]_{\le k-1} \subseteq F[x]_{\le k}$. The flag $F[x]_{\le 0} \subseteq F[x]_{\le 1} \subseteq F[x]_{\le 2} \subseteq F[x]_{\le 3}$ (plus $\{\mathbf{0}\}$ at the bottom) is D-invariant, and indeed

$$[D]_{1,x,x^2,x^3} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) The shift $N \in \mathrm{L}(F^n)$ with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \ge 2$ (Example 7.2.20) is upper triangular in the standard basis: the flag of coordinate subspaces $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_j)$ is invariant.

(c) Every upper-triangular $A \in F^{n\times n}$ preserves the coordinate subspaces $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_j)$, the vectors whose last $n - j$ coordinates vanish.

Non-example 7.5.6 (A rotation). Let $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, the rotation by a right angle. If T_R were upper triangular with respect to a basis $\mathbf{b}_1, \mathbf{b}_2$ of $\mathbb{R}^2$, then $\mathbf{b}_1$ would be an eigenvector. But R has no real eigenvalues (Example 7.2.18): it moves every line through $\mathbf{0}$. So R is not triangularizable over $\mathbb{R}$. Over $\mathbb{C}$ it is, and even diagonalizable: $R(1, -i) = (i, 1) = i(1, -i)$ and $R(1, i) = -i(1, i)$, and these two vectors form a basis of $\mathbb{C}^2$.

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \quad A\mathbf{e}_1 = 2\mathbf{e}_1, \quad A\mathbf{e}_2 = \mathbf{e}_1 + \mathbf{e}_2$$

The upper-triangular matrix $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. The x-axis is invariant, since $A\mathbf{e}_1 = 2\mathbf{e}_1$, and $\{\mathbf{0}\} \subseteq \operatorname{span}(\mathbf{e}_1) \subseteq \mathbb{R}^2$ is an invariant flag. In the interactive version, drag $A\mathbf{e}_1$ off the axis: the matrix is no longer upper triangular in the standard basis, but it is still triangularizable over $\mathbb{R}$ as long as some line through $\mathbf{0}$ is mapped into itself. A rotation by an angle that is not a multiple of π leaves no line in place. (Interactive in the web version.)

Exercise 7.5.7. ROUTINE Let $\mathcal{B} = \mathbf{b}_1, \dots, \mathbf{b}_n$ be a basis of V and $\mathcal{B}' = \mathbf{b}_n, \dots, \mathbf{b}_1$ the same basis in reverse order. Show that $[T]_{\mathcal{B}}$ is upper triangular if and only if $[T]_{\mathcal{B}'}$ is lower triangular. Write down $[D]_{\mathcal{B}'}$ for the differentiation operator of Example 7.5.5(a) and $\mathcal{B}' = x^3, x^2, x, 1$.

Solution on page 1594

Reading the diagonal

The diagonal of a triangular matrix carries the eigenvalues. We give two proofs: one computes a determinant, the other sees where a zero on the diagonal loses a dimension.

Proposition 7.5.8 (The diagonal of a triangular matrix). Suppose $[T]_{\mathcal{B}}$ is upper triangular with diagonal entries $\lambda_1, \dots, \lambda_n$.

(a) T is invertible if and only if $\lambda_j \neq 0$ for every j.

(b) The eigenvalues of T are exactly $\lambda_1, \dots, \lambda_n$.

Proof (via determinants). ROUTINE

Uses. Definition 5.4.18 (Determinant of an operator), Proposition 5.3.15 (Triangular matrices), Proposition 5.4.20 (Properties of the determinant of an operator), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 2.4.12 (Matrices of sums and scalar multiples)
Hint 1. What is $\det T$, computed from $[T]_{\mathcal{B}}$?
Hint 2. For (b), $[T - \lambda \mathrm{id}_V]_{\mathcal{B}} = [T]_{\mathcal{B}} - \lambda I_n$ is again upper triangular. Apply (a) to $T - \lambda \mathrm{id}_V$.
Sketch. $\det T = \det [T]_{\mathcal{B}} = \lambda_1 \cdots \lambda_n$, which is nonzero exactly when every λ_j is. For (b), $T - \lambda \mathrm{id}_V$ has the upper-triangular matrix $[T]_{\mathcal{B}} - \lambda I_n$ with diagonal $\lambda_j - \lambda$, and λ is an eigenvalue exactly when $T - \lambda \mathrm{id}_V$ is not invertible.

(a) By Definition 5.4.18 and Proposition 5.3.15, $\det T = \det [T]_{\mathcal{B}} = \lambda_1 \lambda_2 \cdots \lambda_n$. By Proposition 5.4.20(c), T is invertible if and only if this product is nonzero, that is, if and only if every λ_j is nonzero.

(b) Let $\lambda \in F$. By Proposition 2.4.12, $[T - \lambda \mathrm{id}_V]_{\mathcal{B}} = [T]_{\mathcal{B}} - \lambda I_n$, which is upper triangular with diagonal entries $\lambda_1 - \lambda, \dots, \lambda_n - \lambda$. By Theorem 7.2.9, λ is an eigenvalue of T if and only if $T - \lambda \mathrm{id}_V$ is not invertible, which by (a) happens if and only if $\lambda_j - \lambda = 0$ for some j. □

Intuition (Where the dimension goes). Suppose $\lambda_j = 0$. Then $T\mathbf{b}_j$ has no $\mathbf{b}_j$-component, so T maps $V_j = \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$ into the smaller space V_{j-1}. A map from a j-dimensional space into a $(j-1)$-dimensional one cannot be injective. Conversely, if every $\lambda_j \neq 0$, we can solve for $\mathbf{b}_j$ in $T\mathbf{b}_j = \lambda_j \mathbf{b}_j + (\cdots)$ and see, one index at a time, that every basis vector lies in the image. This argument uses no determinants, and it is the version that survives where determinants are unavailable.

Proof (without determinants). STANDARD

Uses. Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 4.2.11 (Rank–nullity), Corollary 4.2.18 (Injective if and only if surjective), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 2.4.12 (Matrices of sums and scalar multiples)
Techniques. Dimension counting (Technique 4.2.24), Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Let $V_j = \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$. If $\lambda_j = 0$, where does T send V_j?
Hint 2. If $\lambda_j = 0$, then T maps the j-dimensional V_j into V_{j-1}, so $T|_{V_j}$ has a nonzero kernel by rank–nullity.
Hint 3. If all $\lambda_j \neq 0$, show $\mathbf{b}_j \in \operatorname{im} T$ by strong induction on j, solving $T\mathbf{b}_j = \lambda_j \mathbf{b}_j + \sum_{i<j} a_{ij}\mathbf{b}_i$ for $\mathbf{b}_j$.
Sketch. If $\lambda_j = 0$, then $T\mathbf{b}_i \in V_i \subseteq V_{j-1}$ for $i < j$ and $T\mathbf{b}_j = \sum_{i<j} a_{ij}\mathbf{b}_i \in V_{j-1}$, so $T|_{V_j} : V_j \to V_{j-1}$ is not injective. If all $\lambda_j \neq 0$, then $\mathbf{b}_j = \lambda_j^{-1}\big(T\mathbf{b}_j - \sum_{i<j} a_{ij}\mathbf{b}_i\big) \in \operatorname{im} T$ by induction, so T is surjective, hence invertible. (b) follows from (a) applied to $T - \lambda \mathrm{id}_V$.

Write $[T]_{\mathcal{B}} = (a_{ij})$, so $a_{jj} = \lambda_j$, and $V_j = \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$. By Proposition 7.5.2, $T\mathbf{b}_j = \lambda_j \mathbf{b}_j + \sum_{i<j} a_{ij}\mathbf{b}_i$ for each j.

(a) Suppose $\lambda_j = 0$ for some j. For $i < j$, $T\mathbf{b}_i \in V_i \subseteq V_{j-1}$, and $T\mathbf{b}_j = \sum_{i<j} a_{ij}\mathbf{b}_i \in V_{j-1}$. So T maps the spanning list $\mathbf{b}_1, \dots, \mathbf{b}_j$ of V_j into the subspace V_{j-1}, and $\operatorname{im}\big(T|_{V_j}\big) \subseteq V_{j-1}$. By rank–nullity (Theorem 4.2.11) applied to $T|_{V_j} : V_j \to V$,

$$\dim \ker\big(T|_{V_j}\big) = j - \dim \operatorname{im}\big(T|_{V_j}\big) \geq j - (j-1) = 1.$$

So some nonzero vector of V_j is sent to $\mathbf{0}$, and T is not invertible.

Conversely, suppose every $\lambda_j \neq 0$. We show $\mathbf{b}_j \in \operatorname{im} T$ by strong induction on j. Assume $\mathbf{b}_i \in \operatorname{im} T$ for all $i < j$ (for $j = 1$ there is nothing to assume). Then

$$\mathbf{b}_j = \lambda_j^{-1}\left(T\mathbf{b}_j - \sum_{i<j} a_{ij}\mathbf{b}_i\right)$$

is a linear combination of $T\mathbf{b}_j$ and the $\mathbf{b}_i$ with $i < j$, all of which lie in the subspace $\operatorname{im} T$. So $\mathbf{b}_j \in \operatorname{im} T$. Hence $\operatorname{im} T$ contains a basis of V, and T is surjective, so it is invertible by Corollary 4.2.18.

(b) As in the first proof, $[T-\lambda \mathrm{id}_V]_{\mathcal{B}}$ is upper triangular with diagonal entries $\lambda_j - \lambda$, and λ is an eigenvalue if and only if $T-\lambda\mathrm{id}_V$ is not invertible (Theorem 7.2.9), that is, by (a), if and only if $\lambda = \lambda_j$ for some j. □

The diagonal also records how often each eigenvalue occurs. Write p_T for the characteristic polynomial of T (Definition 7.3.4), so that $p_T = \det(xI_n - A)$ for every matrix A of T. Here the determinant of a matrix with entries in $F[x]$ is defined by the Leibniz formula, and Section 7.3 shows that the results of Chapter 5 proved by expanding that formula, such as Proposition 5.3.15 and Theorem 5.4.26, hold for such matrices.

Proposition 7.5.9 (Multiplicities on the diagonal). Suppose $[T]_{\mathcal{B}}$ is upper triangular with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then

$$p_T = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n).$$

Consequently each eigenvalue λ of T occurs among $\lambda_1, \ldots, \lambda_n$ exactly as many times as its algebraic multiplicity (Definition 7.4.11). In particular, this number is the same for every basis in which T is upper triangular.

Proof. ROUTINE

Uses. Definition 7.3.4 (Characteristic polynomial), Proposition 5.3.15 (Triangular matrices), Proposition 7.1.34 (Roots of a split polynomial), Definition 7.4.11 (Algebraic multiplicity)
Hint 1. $xI_n - [T]_{\mathcal{B}}$ is upper triangular. What is its determinant?
Hint 2. The algebraic multiplicity of λ is its multiplicity as a root of p_T. Use Proposition 7.1.34.
Sketch. $xI_n - [T]_{\mathcal{B}}$ is upper triangular with diagonal entries $x-\lambda_j$, so its determinant is their product. A root of a split polynomial occurs in its list of linear factors exactly as often as its multiplicity.

The matrix $xI_n - [T]_{\mathcal{B}}$ has entries in $F[x]$; it is upper triangular, with diagonal entries $x-\lambda_1, \ldots, x-\lambda_n$. By Proposition 5.3.15 (whose proof only uses the Leibniz formula), $p_T = \det(xI_n - [T]_{\mathcal{B}}) = (x-\lambda_1)\cdots(x-\lambda_n)$. By Proposition 7.1.34, each root of p_T occurs in the list $\lambda_1, \ldots, \lambda_n$ exactly as many times as its multiplicity as a root of p_T, which is its algebraic multiplicity as an eigenvalue (Definition 7.4.11). The polynomial p_T does not depend on the basis, so neither do these counts. □

When p_T splits over F, say $p_T = (x-\lambda_1)\cdots(x-\lambda_n)$ (Definition 7.1.33), we call $\lambda_1, \ldots, \lambda_n$ the eigenvalues of T **counted with multiplicity**. Each eigenvalue occurs in this list as often as its algebraic multiplicity, and the list is determined by T up to order (Proposition 7.1.34). Over $\mathbb{C}$, p_T always splits (Corollary 7.1.35).

Exercise 7.5.10. ROUTINE Let $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ be a basis of a complex vector space V, and let $T \in \mathrm{L}(V)$ satisfy

$$T\mathbf{b}_1 = 3\mathbf{b}_1, \quad T\mathbf{b}_2 = \mathbf{b}_1 - \mathbf{b}_2, \quad T\mathbf{b}_3 = \mathbf{b}_1 + \mathbf{b}_2 + 3\mathbf{b}_3.$$

Find the eigenvalues of T with their algebraic multiplicities, decide whether T is invertible, and find an eigenvector for each eigenvalue. Is T diagonalizable?

Solution on page 1595

7.5.2 Triangularization over $\mathbb{C}$

Intuition (How to find an invariant flag). By Proposition 7.5.2 we need an invariant flag, and its bottom V_1 is an invariant line, spanned by an eigenvector. So we need an eigenvalue, and over $\mathbb{C}$ there always is one (Theorem 7.3.27). Over $\mathbb{R}$ the rotation of Non-example 7.5.6 shows that the theorem can fail at this very first step.

Having found an eigenvector $\mathbf{b}_1$, we would like to finish by induction on $\dim V$. The obvious move is to restrict T to a complement W of $\mathrm{span}(\mathbf{b}_1)$ and use the induction hypothesis there. It fails: T need not map W into W, so $T{\restriction_W}$ is a map $W \to V$, not an operator on W, and the induction hypothesis says nothing about it. (Exercise A.2.13 makes exactly this mistake.) The fix is to throw away the part of $T\mathbf{w}$ that lies in $\mathrm{span}(\mathbf{b}_1)$, by projecting $T\mathbf{w}$ back into W along $\mathrm{span}(\mathbf{b}_1)$. That produces an honest operator on W. The

part thrown away is harmless, because $\operatorname{span}(\mathbf{b}_1)$ sits at the bottom of the flag, and every V_j will contain it anyway.

Lemma 7.5.11 (Compressing an operator to a complement). Let U be a T-invariant subspace of V, and let $V = U \oplus W$. The **compression** of T to W is the operator

$$S = \operatorname{Pr}_{W,U} \circ T\restriction_W \in \mathrm{L}(W),$$

where $\operatorname{Pr}_{W,U}$ is the projection onto W along U (Definition 4.4.24).

(a) $T\mathbf{w} - S\mathbf{w} \in U$ for every $\mathbf{w} \in W$.

(b) Let $\mathcal{C}$ be a basis of U and $\mathcal{D}$ a basis of W, both nonempty. Then $\mathcal{B} = \mathcal{C}, \mathcal{D}$ is a basis of V, and

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T\restriction_U]_{\mathcal{C}} & X \\ 0 & [S]_{\mathcal{D}} \end{pmatrix}$$

for some matrix X, where $T\restriction_U$ is regarded as an operator on U.

Proof. ROUTINE

Uses. Definition 4.4.24 (Projection along a complement), Proposition 4.4.26 (Properties of projections), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Definition 2.3.1 (Matrix of a linear map)
Techniques. Choose a basis adapted to a preserved subspace (Technique 4.5.36)
Hint 1. For (a), $\operatorname{Pr}_{U,W} + \operatorname{Pr}_{W,U} = \mathrm{id}_V$. Apply both sides to $T\mathbf{w}$.
Hint 2. For (b), Proposition 4.5.35 gives the zero block and the top left block. For a column belonging to $\mathbf{d}_j \in \mathcal{D}$, write $T\mathbf{d}_j = (T\mathbf{d}_j - S\mathbf{d}_j) + S\mathbf{d}_j$ and read off the $\mathcal{D}$-coordinates.
Sketch. $T\mathbf{w} = \operatorname{Pr}_{U,W}(T\mathbf{w}) + \operatorname{Pr}_{W,U}(T\mathbf{w})$, and the first term lies in U; this is (a). In (b), $\mathcal{B}$ is a basis because the sum is direct, Proposition 4.5.35 gives the block form with $[T\restriction_U]_{\mathcal{C}}$ in the corner, and by (a) the $\mathcal{D}$-coordinates of $T\mathbf{d}_j$ are those of $S\mathbf{d}_j$, which form column j of $[S]_{\mathcal{D}}$.

(a) By Proposition 4.4.26, $\operatorname{Pr}_{U,W} + \operatorname{Pr}_{W,U} = \mathrm{id}_V$ and $\operatorname{im} \operatorname{Pr}_{U,W} = U$. So for $\mathbf{w} \in W$,

$$T\mathbf{w} - S\mathbf{w} = T\mathbf{w} - \operatorname{Pr}_{W,U}(T\mathbf{w}) = \operatorname{Pr}_{U,W}(T\mathbf{w}) \in U.$$

(b) Since $V = U \oplus W$, the list $\mathcal{B} = \mathcal{C}, \mathcal{D}$ is a basis of V (Theorem 4.4.13). Let k be the length of $\mathcal{C}$. Since $U = \operatorname{span}(\mathcal{C})$ is T-invariant, Proposition 4.5.35 shows that $[T]_{\mathcal{B}}$ has a zero block below the first k columns and that its top left block is $[T\restriction_U]_{\mathcal{C}}$. Let $\mathcal{D} = \mathbf{d}_1, \ldots, \mathbf{d}_m$ and $[S]_{\mathcal{D}} = (s_{ij})$, so $S\mathbf{d}_j = \sum_i s_{ij}\mathbf{d}_i$ (Definition 2.3.1). By (a), $T\mathbf{d}_j = \mathbf{u} + \sum_i s_{ij}\mathbf{d}_i$ for some $\mathbf{u} \in U = \operatorname{span}(\mathcal{C})$. So the last m coordinates of $T\mathbf{d}_j$ in $\mathcal{B}$ are $s_{1j}, \ldots, s_{mj}$, and the bottom right block of $[T]_{\mathcal{B}}$ is $[S]_{\mathcal{D}}$. □

The compression depends on the choice of W, but its matrix is always the bottom right block of $[T]_{\mathcal{B}}$. This is the move the triangularization theorem needs, and it is an instance of a general technique.

Technique 7.5.12 (Induction on dimension via an invariant subspace). To prove that every operator on every finite-dimensional space (of some kind) has a property, induct on $\dim V$. In the inductive step, find a T-invariant subspace U with $\{\mathbf{0}\} \neq U \neq V$, apply the induction hypothesis to an operator on a smaller space built from U, and glue the pieces with the block triangular form of Proposition 4.5.35. The smaller operator can be

- the restriction $T\restriction_U$, an operator on U because U is invariant;
- the compression of T to a complement W of U (Lemma 7.5.11), or to $U^\perp$ when there is an inner product;
- the quotient operator T/U on V/U (Proposition 4.6.32), if quotients are available.

The hypothesis applies only to operators, so whatever is passed down must map the smaller space into itself; a restriction to a subspace that is not invariant does not qualify (Exercise A.2.13). The invariant subspace usually comes from an eigenvector, which is why such inductions tend to need $\mathbb{C}$, or from a kernel or image of a polynomial in T (Technique 7.2.31).

Reach for it for existence statements about bases adapted to an operator: triangular, block triangular, orthonormal and triangular, diagonal. The canonical example is Theorem 7.5.13. Later uses: Schur's theorem (Theorem 7.5.18), the spectral theorems (Section 8.2), and the structure of nilpotent operators (Section 10.3).

Used in. Theorem 7.5.13, Theorem 7.5.17, Theorem 7.5.18, Theorem 7.5.27, Exercise 7.5.38, Exercise 7.5.42, Theorem 7.6.32, Exercise 7.7.13, Problem 7.14, Theorem 8.2.5, Theorem 8.2.12, Exercise 8.2.35, Exercise 8.2.48, Theorem 8.4.15, Theorem 8.4.25, Exercise 8.4.44.

Theorem 7.5.13 (Upper-triangular form). Let V be a finite-dimensional complex vector space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. Then there is a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is upper triangular. Its diagonal entries are the eigenvalues of T, each occurring as many times as its algebraic multiplicity.

Proof (compress to a complement). STANDARD

Uses. Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Proposition 4.4.16 (Complements exist), Lemma 7.5.11 (Compressing an operator to a complement), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 7.5.9 (Multiplicities on the diagonal)

Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Build a flag of invariant subspaces (Technique 7.5.4)

Hint 1. Induct on $n = \dim V$. Over $\mathbb{C}$, T has an eigenvector $\mathbf{b}_1$, and $\operatorname{span}(\mathbf{b}_1)$ is invariant. It should be the bottom of the flag.

Hint 2. Choose a complement W of $\operatorname{span}(\mathbf{b}_1)$. The restriction of T to W is not an operator on W; what is?

Hint 3. Apply the induction hypothesis to the compression S of T to W, getting $\mathbf{b}_2, \dots, \mathbf{b}_n$. Then use $T\mathbf{w} - S\mathbf{w} \in \operatorname{span}(\mathbf{b}_1)$ to check that $T\mathbf{b}_j \in \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$.

Sketch. Induct on n; $n = 1$ is trivial. Take an eigenvector $\mathbf{b}_1$ (Theorem 7.3.27), a complement W of $U = \operatorname{span}(\mathbf{b}_1)$, and the compression $S \in \mathrm{L}(W)$. By induction, W has a basis $\mathbf{b}_2, \dots, \mathbf{b}_n$ with $S\mathbf{b}_j \in \operatorname{span}(\mathbf{b}_2, \dots, \mathbf{b}_j)$. Then $T\mathbf{b}_1 = \lambda_1 \mathbf{b}_1$ and $T\mathbf{b}_j = (T\mathbf{b}_j - S\mathbf{b}_j) + S\mathbf{b}_j \in \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$, so $\mathbf{b}_1, \dots, \mathbf{b}_n$ triangularizes T. The diagonal is described by Proposition 7.5.9.

The statement about the diagonal is Proposition 7.5.9, so we prove existence, by induction on $n = \dim V$. If $n = 1$, every 1×1 matrix is upper triangular, and any basis works.

Let $n \geq 2$, and assume the theorem for complex vector spaces of dimension $n - 1$. By Theorem 7.3.27, T has an eigenvalue λ_1; let $\mathbf{b}_1$ be an eigenvector for it. Then $U = \operatorname{span}(\mathbf{b}_1)$ is T-invariant (Proposition 7.2.7). By Proposition 4.4.16, U has a complement W, and $\dim W = n - 1$. Let $S \in \mathrm{L}(W)$ be the compression of T to W (Lemma 7.5.11). By the induction hypothesis and Proposition 7.5.2, W has a basis $\mathbf{b}_2, \dots, \mathbf{b}_n$ with

$$S\mathbf{b}_j \in \operatorname{span}(\mathbf{b}_2, \dots, \mathbf{b}_j) \quad \text{for } j = 2, \dots, n.$$

The list $\mathcal{B} = \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ is a basis of $V = U \oplus W$ (Theorem 4.4.13). We check condition (II) of Proposition 7.5.2. First, $T\mathbf{b}_1 = \lambda_1 \mathbf{b}_1 \in \operatorname{span}(\mathbf{b}_1)$. For $j \geq 2$, by Lemma 7.5.11(a),

$$T\mathbf{b}_j = \underbrace{(T\mathbf{b}_j - S\mathbf{b}_j)}_{\in \operatorname{span}(\mathbf{b}_1)} + \underbrace{S\mathbf{b}_j}_{\in \operatorname{span}(\mathbf{b}_2, \dots, \mathbf{b}_j)} \in \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j).$$

So $[T]_{\mathcal{B}}$ is upper triangular. In block form (Lemma 7.5.11(b)), $[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & X \\ 0 & [S]_{\mathbf{b}_2, \dots, \mathbf{b}_n} \end{pmatrix}$. □

Intuition (A second route, from the top). The first proof builds the flag from the bottom: an invariant line first, then whatever the smaller problem returns. The second proof (LADR's) works from the top. Fix an eigenvalue λ, and consider $U = \operatorname{im}(T - \lambda \mathrm{id}_V)$. Since $T\mathbf{v} = \lambda\mathbf{v} + (T - \lambda \mathrm{id}_V)\mathbf{v}$, every subspace that contains U is T-invariant: T moves each vector only within its own span plus U. And $U \neq V$, because $T - \lambda \mathrm{id}_V$ is not injective. So above U the flag can be chosen arbitrarily, and the only real work happens inside U, where the restriction of T is an honest operator on a smaller space. This route needs no complement and no projection, only restriction.

Proof (via the invariant subspace $\operatorname{im}(T - \lambda \mathrm{id}_V)$). STANDARD

Uses. Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Theorem 7.2.9 (Eigenvalues and invertibility), Corollary 4.2.18 (Injective if and only if surjective), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Theorem 4.1.16 (Extending to a basis), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Build a flag of invariant subspaces (Technique 7.5.4), Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Let λ be an eigenvalue and $U = \operatorname{im}(T - \lambda \mathrm{id}_V)$. Show that U is T-invariant and $\dim U < n$.
Hint 2. Apply the induction hypothesis to $T|_U$ (strong induction, since $\dim U$ can be anything below n), and extend the resulting basis of U to a basis of V.
Hint 3. For an added vector $\mathbf{v}$, write $T\mathbf{v} = \lambda\mathbf{v} + (T - \lambda \mathrm{id}_V)\mathbf{v}$.
Sketch. By strong induction on n. Let λ be an eigenvalue and $U = \operatorname{im}(T - \lambda \mathrm{id}_V)$; it is T-invariant because T commutes with $T - \lambda \mathrm{id}_V$, and $U \neq V$ because $T - \lambda \mathrm{id}_V$ is not injective, hence not surjective. Triangularize $T|_U$ by $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and extend by $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Then $T\mathbf{v}_j = \lambda\mathbf{v}_j + (T - \lambda \mathrm{id}_V)\mathbf{v}_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_j)$.

We use strong induction on $n = \dim V$. For $n = 1$ there is nothing to prove. Let $n \geq 2$, and assume the theorem for all complex spaces of dimension less than n.

By Theorem 7.3.27, T has an eigenvalue λ. Let $U = \operatorname{im}(T - \lambda \mathrm{id}_V)$. Since T commutes with $T - \lambda \mathrm{id}_V$, the image U is T-invariant (Proposition 7.2.30). Since λ is an eigenvalue, $T - \lambda \mathrm{id}_V$ is not injective (Theorem 7.2.9), hence not surjective (Corollary 4.2.18), so $m = \dim U < n$.

If $m = 0$, then $T = \lambda \mathrm{id}_V$, and every basis triangularizes T. Otherwise $T|_U$ is an operator on U, and by the induction hypothesis and Proposition 7.5.2, U has a basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ with $T\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$ for each j. Extend it to a basis $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_k$ of V (Theorem 4.1.16). For each $j \leq k$,

$$T\mathbf{v}_j = \lambda\mathbf{v}_j + \underbrace{(T - \lambda \mathrm{id}_V)\mathbf{v}_j}_{\in U} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_j).$$

So every vector of the basis is sent into the span of itself and the vectors before it, and T is upper triangular with respect to this basis by Proposition 7.5.2. □

The two proofs pass different objects down the induction. The first passes a compression to a space of dimension exactly $n - 1$, and so needs a complement and a projection; the second passes a restriction to a space of unknown smaller dimension, and so needs strong induction. The second proof also says something about the diagonal: the last k basis vectors have diagonal entry λ, and $k = n - \dim \operatorname{im}(T - \lambda \mathrm{id}_V) = \dim \ker(T - \lambda \mathrm{id}_V)$ by rank–nullity. With Proposition 7.5.9, the algebraic multiplicity of λ is at least $\dim E_\lambda$, its geometric multiplicity. So this argument recovers Theorem 7.4.15 for complex operators.

Corollary 7.5.14 (Triangularizing a matrix). Every $A \in \mathbb{C}^{n \times n}$ is similar to an upper-triangular matrix: $A = SRS^{-1}$ with $S \in \mathbb{C}^{n \times n}$ invertible and R upper triangular. The columns of S form a basis of $\mathbb{C}^n$ whose flag is T_A-invariant, and the diagonal of R lists the eigenvalues of A counted with multiplicity.

Proof. ROUTINE

Uses. Theorem 7.5.13 (Upper-triangular form), Theorem 4.5.17 (Change of basis)
Hint. Apply Theorem 7.5.13 to T_A and translate the change of basis into matrices.
Sketch. If $\mathcal{B}$ triangularizes T_A and S has the vectors of $\mathcal{B}$ as columns, then $S = [\mathrm{id}]_{\mathcal{E}\mathcal{B}}$ and $[T_A]_{\mathcal{B}} = S^{-1}AS$ by Theorem 4.5.17.

By Theorem 7.5.13 there is a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{C}^n$ with $R = [T_A]_{\mathcal{B}}$ upper triangular. Let S be the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_n$, that is, $S = [\mathrm{id}]_{\mathcal{E}\mathcal{B}}$ for the standard basis $\mathcal{E}$. It is invertible, and Theorem 4.5.17 gives $R = S^{-1}[T_A]_{\mathcal{E}}S = S^{-1}AS$, so $A = SRS^{-1}$. The rest is Theorem 7.5.13. □

Example 7.5.15 (Triangularizing a 3×3 matrix). Let

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -2 \\ -1 & 1 & 0 \end{pmatrix} \in \mathbb{C}^{3 \times 3}.$$

We follow the first proof of Theorem 7.5.13. Its characteristic polynomial is $(x-2)(x-1)^2$, and $\mathbf{b}_1 = (1,1,0)$ is an eigenvector: $A\mathbf{b}_1 = (2,2,0) = 2\mathbf{b}_1$. A convenient complement of $\operatorname{span}(\mathbf{b}_1)$ is $W = \operatorname{span}(\mathbf{e}_2, \mathbf{e}_3)$, since $\mathbf{b}_1$ has a nonzero first coordinate. Every vector decomposes as

$$(x_1, x_2, x_3) = x_1\mathbf{b}_1 + (0, x_2 - x_1, x_3),$$

so the projection onto W along $\operatorname{span}(\mathbf{b}_1)$ is $(x_1, x_2, x_3) \mapsto (0, x_2 - x_1, x_3)$. The compression S sends

$$\mathbf{e}_2 \mapsto \Pr(A\mathbf{e}_2) = \Pr(0,2,1) = 2\mathbf{e}_2 + \mathbf{e}_3, \quad \mathbf{e}_3 \mapsto \Pr(A\mathbf{e}_3) = \Pr(-1,-2,0) = -\mathbf{e}_2,$$

so $[S]_{\mathbf{e}_2,\mathbf{e}_3} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$. This 2×2 problem has the eigenvalue 1 with eigenvector $(1,1)$ in these coordinates, that is, $\mathbf{b}_2 = \mathbf{e}_2 + \mathbf{e}_3$, and we complete it with $\mathbf{b}_3 = \mathbf{e}_3$. Now

$$A\mathbf{b}_1 = 2\mathbf{b}_1, \quad A\mathbf{b}_2 = (-1,0,1) = -\mathbf{b}_1 + \mathbf{b}_2, \quad A\mathbf{b}_3 = (-1,-2,0) = -\mathbf{b}_1 - \mathbf{b}_2 + \mathbf{b}_3,$$

so with $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, whose columns are $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$,

$$P^{-1}AP = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The diagonal $2, 1, 1$ matches $p_A = (x-2)(x-1)^2$. The matrix A is not diagonalizable: $A - I_3$ has rank 2, so E_1 is a line, while 1 has algebraic multiplicity 2. The entries above the diagonal cannot all be removed.

Remark 7.5.16 (The quotient version). BEYOND 21-242 With quotient spaces (Section 4.6), the first proof needs no choice of complement. Let $U = \operatorname{span}(\mathbf{b}_1)$. The operator T/U on V/U (Proposition 4.6.32) is triangularized by some $\mathbf{b}_2 + U, \ldots, \mathbf{b}_n + U$, by induction, since $\dim V/U = n-1$. Then $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis of V (Proposition 4.6.28), and Proposition 4.6.32(b) gives $[T] = \begin{pmatrix} \lambda_1 & X \\ 0 & [T/U] \end{pmatrix}$, which is upper triangular. The compression is T/U in disguise: the isomorphism $W \to V/U$, $\mathbf{w} \mapsto \mathbf{w} + U$ (Proposition 4.6.29), carries S to T/U, because $S\mathbf{w} + U = T\mathbf{w} + U$ by Lemma 7.5.11(a). So different complements give different compressions, but they are all similar to one operator.

Only one step of the first proof used $\mathbb{C}$: the existence of an eigenvalue, for T and then for each compression. Over a general field we can ask for exactly that, and the characteristic polynomial tells us when it holds.

Theorem 7.5.17 (Triangularizable means the characteristic polynomial splits). BEYOND 21-242 Let V be a finite-dimensional vector space over any field F, with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. Then T is triangularizable if and only if p_T splits over F.

Proof. HARD

Uses. Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Lemma 7.5.11 (Compressing an operator to a complement), Theorem 5.4.26 (Block-triangular matrices), Proposition 7.1.34 (Roots of a split polynomial), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 4.4.16 (Complements exist)

Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)

Hint 1. One direction is Proposition 7.5.9. For the other, run the first proof of Theorem 7.5.13. What has to be checked at each step?

Hint 2. The eigenvalue for the first step is a root of p_T. For the induction, show that $p_T = (x - \lambda_1)p_S$ for the compression S, using Lemma 7.5.11(b) and block triangular determinants.

Hint 3. If $(x-\lambda_1)q = (x-\mu_1)\cdots(x-\mu_n)$, then λ_1 is one of the μ_i, and cancelling shows that q splits.

Sketch. If T is upper triangular, $p_T = \prod(x - \lambda_j)$ splits. Conversely, by induction on n: a root λ_1 of p_T is an eigenvalue; take an eigenvector, a complement W and the compression S. The block form of Lemma 7.5.11(b) gives $p_T = (x - \lambda_1)p_S$, so p_S splits after cancelling the factor $x - \lambda_1$. Triangularize S by induction and finish as in the first proof of Theorem 7.5.13.

If $[T]_{\mathcal{B}}$ is upper triangular, then p_T splits by Proposition 7.5.9.

Conversely, suppose p_T splits. We induct on $n = \dim V$; for $n = 1$ every basis works. Let $n \geq 2$ and $p_T = (x-\mu_1)\cdots(x-\mu_n)$. Then $\lambda_1 = \mu_1$ is a root of p_T, hence an eigenvalue of T (Theorem 7.3.6); let $\mathbf{b}_1$ be an eigenvector, $U = \operatorname{span}(\mathbf{b}_1)$, W a complement of U (Proposition 4.4.16), and S the compression of T to W. For any basis $\mathcal{D}$ of W and $\mathcal{B} = \mathbf{b}_1, \mathcal{D}$, Lemma 7.5.11(b) gives

$$xI_n - [T]_{\mathcal{B}} = \begin{pmatrix} x-\lambda_1 & -X \\ 0 & xI_{n-1} - [S]_{\mathcal{D}} \end{pmatrix},$$

so by Theorem 5.4.26 (over $F[x]$), $p_T = (x-\lambda_1)p_S$. Hence $(x-\mu_1)p_S = (x-\mu_1)(x-\mu_2)\cdots(x-\mu_n)$, and cancelling the nonzero factor $x-\mu_1$ (Corollary 7.1.7) gives $p_S = (x-\mu_2)\cdots(x-\mu_n)$, which splits. By the induction hypothesis, W has a basis $\mathbf{b}_2, \ldots, \mathbf{b}_n$ in which S is upper triangular, and the computation in the first proof of Theorem 7.5.13, which used only Lemma 7.5.11(a), shows that T is upper triangular with respect to $\mathbf{b}_1, \ldots, \mathbf{b}_n$. □

For example, a real matrix is triangularizable over $\mathbb{R}$ exactly when all its complex eigenvalues are real, and $\left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{Q}^{2\times 2}$ is triangularizable over $\mathbb{R}$ but not over $\mathbb{Q}$, since $x^2 - 2$ splits over $\mathbb{R}$ but not over $\mathbb{Q}$.

7.5.3 Schur's theorem

Now let V be a finite-dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (Definition 6.1.2). Orthonormal bases are the best bases to compute in: coordinates are inner products (Proposition 6.3.18), the (i,j) entry of $[T]_{\mathcal{O}}$ is $\langle T\mathbf{u}_j, \mathbf{u}_i\rangle$, and the change-of-basis matrix to an orthonormal basis of $\mathbb{F}^n$ is unitary, so its inverse is its conjugate transpose (Definition 6.3.33). Can a triangularizing basis be chosen orthonormal?

By Proposition 7.5.2, being upper triangular depends only on the flag of the basis. And Gram–Schmidt replaces a basis by an orthonormal one with the same flag: $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j) = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$ for every j (Theorem 6.4.15(d)). So the answer is yes, at no extra cost.

Theorem 7.5.18 (Schur's theorem). Let V be a finite-dimensional complex inner product space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$.

(a) There is an orthonormal basis $\mathcal{O}$ of V such that $[T]_{\mathcal{O}}$ is upper triangular.

(b) Every $A \in \mathbb{C}^{n\times n}$ can be written as $A = URU^*$ with U unitary and R upper triangular. The diagonal entries of R are the eigenvalues of A, counted with multiplicity.

Proof (Gram–Schmidt on a triangularizing basis). ROUTINE

Uses. Theorem 7.5.13 (Upper-triangular form), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 6.4.15 (Gram–Schmidt is correct), Theorem 4.5.17 (Change of basis), Definition 6.3.33 (Unitary and orthogonal matrices), Theorem 6.3.29 (Orthonormal columns)
Techniques. Build a flag of invariant subspaces (Technique 7.5.4)
Hint 1. Start from a basis $\mathcal{B}$ with $[T]_{\mathcal{B}}$ upper triangular. What does Proposition 7.5.2 say about the subspaces $\operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$?
Hint 2. Gram–Schmidt does not change these subspaces (Theorem 6.4.15(d)).
Sketch. Triangularize T in some basis $\mathcal{B}$ (Theorem 7.5.13), so the flag of $\mathcal{B}$ is invariant. Gram–Schmidt turns $\mathcal{B}$ into an orthonormal basis $\mathcal{O}$ with the same flag, so T is upper triangular with respect to $\mathcal{O}$. For (b), apply (a) to T_A on $\mathbb{C}^n$ with the dot product; the matrix U whose columns are the vectors of $\mathcal{O}$ is unitary.

(a) By Theorem 7.5.13, there is a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of V with $[T]_{\mathcal{B}}$ upper triangular, and by Proposition 7.5.2 every $V_j = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$ is T-invariant. Apply Gram–Schmidt to $\mathcal{B}$. Since $\mathcal{B}$ is linearly independent, Theorem 6.4.15(c),(d) gives an orthonormal basis $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ of V with $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j) = V_j$ for every j. These subspaces are T-invariant, so $[T]_{\mathcal{O}}$ is upper triangular by Proposition 7.5.2.

(b) Apply (a) to T_A on $\mathbb{C}^n$ with the dot product, and let U be the matrix whose columns are the vectors of $\mathcal{O}$. Its columns are orthonormal, so U is unitary (Theorem 6.3.29, Definition 6.3.33) and $U^{-1} = U^*$. By

Theorem 4.5.17, $R = [T_A]_{\mathcal{O}} = U^{-1}AU = U^*AU$ is upper triangular, and $A = URU^*$. The diagonal of R is described by Theorem 7.5.13. □

Intuition (The lecture's proof). The second proof does the algebra and the geometry at once, without quoting Theorem 7.5.13. It is the first proof of Theorem 7.5.13 with the complement chosen orthogonal: take a unit eigenvector $\mathbf{u}_1$, use $W = \operatorname{span}(\mathbf{u}_1)^\perp$ as the complement, and compress T to W with the orthogonal projection. The compression is the operator on a smaller space to which the induction hypothesis applies (Technique 7.5.12). This is the template for the spectral theorem in Section 8.2: for a self-adjoint T, the complement W turns out to be invariant itself, the compression is just the restriction, and the entries above the diagonal vanish too.

Proof (induction with an orthogonal complement). STANDARD

Uses. Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Theorem 6.4.24 (Orthogonal decomposition), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 4.4.13 (Bases and dimension of a sum), Definition 6.3.15 (Orthonormal list, orthonormal basis), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Definition 6.1.2 (Inner product)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Induct on $n = \dim V$. Take a unit eigenvector $\mathbf{u}_1$ and $W = \operatorname{span}(\mathbf{u}_1)^\perp$. Which operator on W should the induction hypothesis be applied to?
Hint 2. Define $S\mathbf{w} = T\mathbf{w} - \langle T\mathbf{w}, \mathbf{u}_1\rangle \mathbf{u}_1$. Check that $S\mathbf{w} \in W$.
Hint 3. With an orthonormal basis $\mathbf{u}_2, \dots, \mathbf{u}_n$ of W triangularizing S, write $T\mathbf{u}_j = \langle T\mathbf{u}_j, \mathbf{u}_1\rangle \mathbf{u}_1 + S\mathbf{u}_j$.
Sketch. Induct on n. Take a unit eigenvector $\mathbf{u}_1$ and $W = \operatorname{span}(\mathbf{u}_1)^\perp$, of dimension $n-1$. The map $S\mathbf{w} = T\mathbf{w} - \langle T\mathbf{w}, \mathbf{u}_1\rangle \mathbf{u}_1$ is an operator on W (the compression of T). By induction, W has an orthonormal basis $\mathbf{u}_2, \dots, \mathbf{u}_n$ triangularizing S. Then $\mathbf{u}_1, \dots, \mathbf{u}_n$ is orthonormal, and $T\mathbf{u}_j = \langle T\mathbf{u}_j, \mathbf{u}_1\rangle \mathbf{u}_1 + S\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_j)$.

We prove (a) by induction on $n = \dim V$; part (b) follows from (a) as in the first proof. If $n = 1$, any unit vector forms an orthonormal basis, and 1×1 matrices are upper triangular.

Let $n \geq 2$, and assume (a) for complex inner product spaces of dimension $n-1$. By Theorem 7.3.27, T has an eigenvector; dividing it by its norm gives an eigenvector $\mathbf{u}_1$ with $\|\mathbf{u}_1\| = 1$ and $T\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$. Let $W = \operatorname{span}(\mathbf{u}_1)^\perp$, the set of vectors orthogonal to $\mathbf{u}_1$. By Theorem 6.4.24 and Corollary 6.4.25, $V = \operatorname{span}(\mathbf{u}_1) \oplus W$ and $\dim W = n - 1$; with the inner product of V restricted to it, W is a complex inner product space. Define

$$S : W \to W, \quad S\mathbf{w} = T\mathbf{w} - \langle T\mathbf{w}, \mathbf{u}_1\rangle \mathbf{u}_1.$$

This lands in W, since $\langle S\mathbf{w}, \mathbf{u}_1\rangle = \langle T\mathbf{w}, \mathbf{u}_1\rangle - \langle T\mathbf{w}, \mathbf{u}_1\rangle\langle \mathbf{u}_1, \mathbf{u}_1\rangle = 0$, and it is linear because T is linear and the inner product is linear in its first slot (Definition 6.1.2). ($S\mathbf{w}$ is the orthogonal projection of $T\mathbf{w}$ onto W, so S is the compression of T to W.)

By the induction hypothesis and Proposition 7.5.2, W has an orthonormal basis $\mathbf{u}_2, \dots, \mathbf{u}_n$ with $S\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_2, \dots, \mathbf{u}_j)$ for $2 \leq j \leq n$. Each $\mathbf{u}_j$ with $j \geq 2$ lies in W, so it is orthogonal to $\mathbf{u}_1$, and $\mathcal{O} = \mathbf{u}_1, \dots, \mathbf{u}_n$ is orthonormal (Definition 6.3.15); it is a basis of $V = \operatorname{span}(\mathbf{u}_1) \oplus W$ by Theorem 4.4.13. Finally, $T\mathbf{u}_1 = \lambda_1 \mathbf{u}_1$, and for $j \geq 2$,

$$T\mathbf{u}_j = \langle T\mathbf{u}_j, \mathbf{u}_1\rangle \mathbf{u}_1 + S\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j).$$

By Proposition 7.5.2, $[T]_{\mathcal{O}}$ is upper triangular. □

The first proof separates the two ingredients: the algebra (an invariant flag exists over $\mathbb{C}$) and the geometry (Gram–Schmidt keeps flags). It shows that any triangularizing basis can be orthonormalized, so for instance the basis of Example 7.5.15 turns into a Schur basis of that matrix. The second proof is self-contained and computes the entries as it goes: $[T]_{\mathcal{O}}$ has (i, j) entry $\langle T\mathbf{u}_j, \mathbf{u}_i\rangle = \langle S\mathbf{u}_j, \mathbf{u}_i\rangle$ for $i, j \geq 2$.

Example 7.5.19 (A Schur form). Let $A = \begin{pmatrix} 1 & 2 \\ -2 & 5 \end{pmatrix} \in \mathbb{C}^{2\times 2}$. Then $\det(A - \lambda I_2) = (1-\lambda)(5-\lambda) + 4 = (\lambda - 3)^2$, so 3 is the only eigenvalue, and $A - 3I_2 = \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix}$ gives the unit eigenvector $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1)$. Complete it to the orthonormal basis $\mathbf{u}_1, \mathbf{u}_2$ with $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-1,1)$. Since $A\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1,7)$,

$$\langle A\mathbf{u}_2, \mathbf{u}_1 \rangle = \frac{1}{2}(1+7) = 4, \quad \langle A\mathbf{u}_2, \mathbf{u}_2 \rangle = \frac{1}{2}(-1+7) = 3,$$

and so

$$A = URU^*, \quad U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 3 & 4 \\ 0 & 3 \end{pmatrix}.$$

Here E_3 is a line, so A is not diagonalizable, and no unitary change of basis can remove the entry 4. In fact no change of basis at all can make it 0, although a non-orthonormal one can make it 1.

Remark 7.5.20 (What is unique). A Schur form is far from unique. The eigenvalues can be made to appear on the diagonal in any order (Exercise 7.5.38), and even for a fixed order, replacing U by UD with $D = \operatorname{diag}\left(e^{i\theta_1}, \ldots, e^{i\theta_n}\right)$ replaces R by D^*RD, which multiplies r_{ij} by $e^{i(\theta_j - \theta_i)}$. What does not change is the list of diagonal entries, up to order (Proposition 7.5.9), and the total size $\sum_{i,j} |r_{ij}|^2 = \operatorname{tr}(A^*A)$ (Exercise 7.5.41).

For a real matrix with nonreal eigenvalues, U and R cannot both be real, since a real upper-triangular R has real eigenvalues. If all eigenvalues of $A \in \mathbb{R}^{n\times n}$ are real, then $A = QRQ^{\mathsf{T}}$ with Q orthogonal and R real (Exercise 7.5.42). For general real matrices, see Section 7.5.5.

Exercise 7.5.21. ROUTINE Find a unitary U and an upper-triangular R with $\begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} = URU^*$.

Solution on page 1595

7.5.4 Trace and determinant from eigenvalues

Every operator on a complex space has a triangular matrix whose diagonal lists its eigenvalues. Trace and determinant are both read off the diagonal of a triangular matrix, and both are independent of the basis. Putting these facts together gives formulas for them that mention no basis at all.

Theorem 7.5.22 (Trace and determinant from eigenvalues). Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, let $T \in \mathrm{L}(V)$, and suppose that p_T splits over F, as it always does when $F = \mathbb{C}$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T counted with multiplicity. Then

$$\operatorname{tr} T = \lambda_1 + \lambda_2 + \cdots + \lambda_n \quad \text{and} \quad \det T = \lambda_1\lambda_2\cdots\lambda_n.$$

In particular, for $A \in \mathbb{C}^{n\times n}$, and so for every real matrix regarded as a complex one, $\operatorname{tr} A$ is the sum and $\det A$ the product of the complex eigenvalues of A, counted with multiplicity.

Proof (triangularize). ROUTINE

Uses. Theorem 7.5.13 (Upper-triangular form), Theorem 7.5.17 (Triangularizable means the characteristic polynomial splits), Definition 4.5.30 (Trace of an operator), Definition 5.4.18 (Determinant of an operator), Definition 2.4.26 (Trace), Proposition 5.3.15 (Triangular matrices), Proposition 7.5.9 (Multiplicities on the diagonal)
Techniques. Prove it for triangular matrices first (Technique 7.5.23), Choose a convenient basis (Technique 4.5.33)
Hint 1. Both sides do not depend on the basis. Which basis makes the left sides easy to compute?
Hint 2. In a triangularizing basis $\mathcal{B}$, what are $\operatorname{tr}[T]_{\mathcal{B}}$ and $\det[T]_{\mathcal{B}}$, and what is the diagonal?
Sketch. Choose $\mathcal{B}$ with $R = [T]_{\mathcal{B}}$ upper triangular (Theorem 7.5.13 over $\mathbb{C}$, Theorem 7.5.17 in general). Then $\operatorname{tr} T = \operatorname{tr} R$ and $\det T = \det R$ are the sum and the product of the diagonal entries, and the diagonal entries are the eigenvalues counted with multiplicity (Proposition 7.5.9).

By Theorem 7.5.13 when $F = \mathbb{C}$, and by Theorem 7.5.17 in general, there is a basis $\mathcal{B}$ with $R = [T]_{\mathcal{B}}$ upper triangular. By Definition 4.5.30 and Definition 5.4.18, $\operatorname{tr} T = \operatorname{tr} R$ and $\det T = \det R$. By Definition 2.4.26 and Proposition 5.3.15,

$$\operatorname{tr} R = r_{11} + \cdots + r_{nn}, \quad \det R = r_{11} r_{22} \cdots r_{nn}.$$

By Proposition 7.5.9, $p_T = (x - r_{11}) \cdots (x - r_{nn})$, so $r_{11}, \ldots, r_{nn}$ are the eigenvalues of T counted with multiplicity, in some order. Sums and products do not depend on the order. □

Technique 7.5.23 (Prove it for triangular matrices first). If a statement about an operator T (or a square matrix A) does not change under change of basis (or similarity), it is enough to prove it for an upper-triangular matrix of T, which exists over $\mathbb{C}$ (Theorem 7.5.13), or over any field once the characteristic polynomial splits. Triangular matrices are easy to compute with: products, powers, inverses, and polynomials of upper-triangular matrices are upper triangular, and their diagonals are computed entry by entry (Proposition 2.4.36, Exercise 7.5.34), while the diagonal itself is the list of eigenvalues.

Reach for it for identities that involve eigenvalues together with similarity invariants: trace, determinant, the characteristic polynomial, $\operatorname{tr}\left(A^k\right)$, invertibility, nilpotence. It is a special case of Technique 4.5.33. Over $\mathbb{R}$, either check that the characteristic polynomial splits, or regard the real matrix as a complex one first. The canonical example is Theorem 7.5.22. Later it proves the Cayley–Hamilton theorem (Section 7.6).

Used in. Theorem 7.5.22, Exercise 7.5.34, Exercise 7.5.36, Exercise 7.5.45, Lemma 7.6.21, Theorem 7.6.28, Problem 7.5, Problem 7.10, Theorem 8.2.5, Theorem 8.2.12, Theorem 8.4.25, Exercise 7.6.51.

The second proof avoids triangularization. It compares two descriptions of the coefficients of the characteristic polynomial.

Proof (read off the characteristic polynomial). STANDARD

Uses. Definition 7.3.4 (Characteristic polynomial), Proposition 7.1.36 (Sum and product of the roots)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12), Polynomial identity arguments (Technique 5.5.24)
Hint 1. Section 7.3 computes two coefficients of $p_A = \det(xI_n - A)$ directly from the entries of A. Which ones?
Hint 2. Compare them with the coefficients of $(x - \lambda_1) \cdots (x - \lambda_n)$ from Proposition 7.1.36.
Sketch. For any matrix A of T, Section 7.3 shows $p_T = x^n - (\operatorname{tr} A)x^{n-1} + \cdots + (-1)^n \det A$: the constant term is $\det(-A)$, and in the Leibniz expansion only the identity permutation reaches degree $n - 1$. On the other hand $p_T = \prod(x - \lambda_i)$ has these coefficients equal to $-\sum \lambda_i$ and $(-1)^n \prod \lambda_i$.

Let $A = [T]_{\mathcal{B}}$ for some basis $\mathcal{B}$, so $p_T = \det(xI_n - A)$, $\operatorname{tr} T = \operatorname{tr} A$, and $\det T = \det A$. By Section 7.3, the coefficient of x^{n-1} in p_T is $-\operatorname{tr} A$ and its constant term is $(-1)^n \det A$. (The constant term is $\det(-A)$, by setting $x = 0$; for the other coefficient, a permutation other than the identity moves at least two indices, so its term in the Leibniz expansion has at most $n - 2$ factors containing x, and only $(x - a_{11}) \cdots (x - a_{nn})$ contributes.) On the other hand $p_T = (x - \lambda_1) \cdots (x - \lambda_n)$, whose coefficient of x^{n-1} is $-(\lambda_1 + \cdots + \lambda_n)$ and whose constant term is $(-1)^n \lambda_1 \cdots \lambda_n$ (Proposition 7.1.36). Comparing coefficients, $\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n$ and $\det A = \lambda_1 \cdots \lambda_n$. □

The first proof is the more flexible one: the same reduction gives the eigenvalues of A^k and of A^{-1} (Exercise 7.5.34). The second proof needs no triangular form, only two coefficients of the characteristic polynomial, which are $-\operatorname{tr} T$ and $(-1)^n \det T$ over every field, whether or not p_T splits.

Example 7.5.24 (Rotations and a 3×3 check). The rotation $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathbb{R}^{2\times 2}$ has the complex eigenvalues $\cos\theta \pm i \sin\theta$ (Exercise 7.2.42). Their sum is $2\cos\theta = \operatorname{tr} R_\theta$ and their product is $\cos^2\theta + \sin^2\theta = 1 = \det R_\theta$. For a real matrix the nonreal eigenvalues come in conjugate pairs (Proposition 7.1.38), and each pair $\lambda, \overline{\lambda}$ contributes $2 \operatorname{Re} \lambda$ to the trace and $|\lambda|^2 > 0$ to the determinant.

For the matrix A of Example 7.5.15, with eigenvalues $2, 1, 1$, the theorem predicts $\operatorname{tr} A = 4$ and $\det A = 2$. Directly: $\operatorname{tr} A = 2 + 2 + 0 = 4$, and expanding along the first row, $\det A = 2(0 + 2) - 0 + (-1)(0 + 2) = 2$.

The trace of an operator

The trace of an operator was defined through a matrix (Definition 4.5.30). Over $\mathbb{C}$, Theorem 7.5.22 gives it a meaning without any basis: it is the sum of the eigenvalues. The matrix identities of Section 2.4.7 also have versions for maps, and for those, maps between different spaces are allowed.

Proposition 7.5.25 (Cyclicity of the trace for linear maps). Let V and W be finite-dimensional, both of dimension at least 1, and let $S \in \mathrm{L}(V, W)$ and $T \in \mathrm{L}(W, V)$. Then

$$\operatorname{tr}(TS) = \operatorname{tr}(ST),$$

where $TS \in \mathrm{L}(V)$ and $ST \in \mathrm{L}(W)$. In particular $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ for all $S, T \in \mathrm{L}(V)$.

Proof. ROUTINE

Uses. Definition 4.5.30 (Trace of an operator), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.28 (Cyclicity of the trace)
Techniques. Translate between maps and matrices (Technique 2.3.27), Cyclicity of trace (Technique 2.4.29)
Hint 1. Choose bases $\mathcal{B}$ of V and $\mathcal{C}$ of W, and write the matrices of TS and ST as products.
Hint 2. Proposition 2.4.28 holds for rectangular matrices.
Sketch. With $X = [S]_{\mathcal{C}\mathcal{B}}$ and $Y = [T]_{\mathcal{B}\mathcal{C}}$, we have $[TS]_{\mathcal{B}} = YX$ and $[ST]_{\mathcal{C}} = XY$, and $\operatorname{tr}(YX) = \operatorname{tr}(XY)$.

Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V and W, with $\dim V = n$ and $\dim W = m$, and let $X = [S]_{\mathcal{C}\mathcal{B}} \in F^{m \times n}$ and $Y = [T]_{\mathcal{B}\mathcal{C}} \in F^{n \times m}$. By Theorem 2.4.2, $[TS]_{\mathcal{B}} = YX$ and $[ST]_{\mathcal{C}} = XY$. So by Definition 4.5.30 and Proposition 2.4.28,

$$\operatorname{tr}(TS) = \operatorname{tr}(YX) = \operatorname{tr}(XY) = \operatorname{tr}(ST). \qquad \square$$

Corollary 7.5.26 ($ST - TS \neq \mathrm{id}_V$). Let V be finite-dimensional with $n = \dim V \geq 1$, and suppose that $n * 1 \neq 0$ in F (for instance, $F = \mathbb{R}$ or $\mathbb{C}$). Then $ST - TS \neq \mathrm{id}_V$ for all $S, T \in \mathrm{L}(V)$.

Proof. ROUTINE

Uses. Proposition 7.5.25 (Cyclicity of the trace for linear maps), Definition 4.5.30 (Trace of an operator), Proposition 2.4.12 (Matrices of sums and scalar multiples), Proposition 2.4.27 (The trace is linear)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint. Take the trace of both sides.
Sketch. $\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0$, while $\operatorname{tr} \mathrm{id}_V = \operatorname{tr} I_n = n * 1 \neq 0$.

The trace of operators is linear, by Definition 4.5.30, Proposition 2.4.12, and Proposition 2.4.27. So by Proposition 7.5.25, $\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0$. But $\operatorname{tr} \mathrm{id}_V = \operatorname{tr} I_n = n * 1 \neq 0$. Hence $ST - TS \neq \mathrm{id}_V$. $\square$

This is Proposition 2.4.30 restated for maps, and the hypotheses matter in the same ways. On the infinite-dimensional space $F[x]$, differentiation and multiplication by x satisfy $DM - MD = \mathrm{id}$ (Exercise 2.4.51), and over $\mathbb{F}_p$ there are $p \times p$ matrices with $AB - BA = I_p$ (Exercise 2.4.52). The trace argument shows more than the corollary states: $ST - TS$ always has trace 0. Over $\mathbb{C}$ the converse also holds, a theorem of Shoda: every matrix of trace 0 is of the form $AB - BA$. The practice bank has a proof.

7.5.5 Real operators: block upper-triangular form BEYOND 21-242

The real block form and the real Schur form are not covered in lecture.

Over $\mathbb{R}$ an operator need not have any eigenvalue, so it need not be triangularizable. But it always comes close: the bottom of an invariant flag may have to be a plane instead of a line, and the matrix becomes upper triangular with some 2×2 blocks on the diagonal. The key fact, proved in Section 7.3, is that small invariant subspaces always exist: every operator on a nonzero finite-dimensional real vector space has an invariant subspace of dimension 1 or 2. One proof complexifies a matrix A of T: a complex eigenvector

$\mathbf{x} + i\mathbf{y}$ ($\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$) for an eigenvalue $a + bi$ gives $A\mathbf{x} = a\mathbf{x} - b\mathbf{y}$ and $A\mathbf{y} = b\mathbf{x} + a\mathbf{y}$, by comparing real and imaginary parts, so the vectors with coordinates $\mathbf{x}$ and $\mathbf{y}$ span an invariant subspace. The induction of Theorem 7.5.13 then goes through with this subspace in place of an eigenvector.

Theorem 7.5.27 (Real block upper-triangular form). Let V be a real vector space with $1 \le \dim V < \infty$, and let $T \in \mathrm{L}(V)$.

(a) There is a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is block upper triangular, with every diagonal block of size 1×1 or 2×2, and with no 2×2 diagonal block having a real eigenvalue.

(b) If V is a real inner product space, $\mathcal{B}$ can be chosen orthonormal. In particular, every $A \in \mathbb{R}^{n \times n}$ can be written as $A = QRQ^{\mathsf{T}}$ with Q orthogonal and R block upper triangular as in (a) (the **real Schur form**).

Proof. HARD

Uses. an invariant subspace of dimension 1 or 2 exists (Section 7.3), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Proposition 4.4.16 (Complements exist), Lemma 7.5.11 (Compressing an operator to a complement), Theorem 6.4.24 (Orthogonal decomposition), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 4.5.17 (Change of basis)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Strong induction on $\dim V$. Choose an invariant subspace U of dimension 1 if T has a real eigenvalue, and of dimension 2 otherwise.
Hint 2. Compress T to a complement W of U, apply the induction hypothesis to the compression, and read off the block form from Lemma 7.5.11(b).
Hint 3. For (b), use the orthogonal complement and orthonormal bases throughout.
Sketch. If T has a real eigenvalue, let U be an eigenvector's span; otherwise an invariant subspace of dimension 1 or 2 (Section 7.3) must be a plane U, and $T|_U$ has no real eigenvalue. If $U = V$ we are done. Otherwise compress T to a complement W (for (b), to $U^\perp$), use the induction hypothesis on the compression, and assemble with Lemma 7.5.11(b): the first diagonal block is $[T|_U]$, the rest are the blocks of the compression.

(a) We use strong induction on $n = \dim V$. Choose a T-invariant subspace U as follows. If T has a real eigenvalue, let U be the span of an eigenvector, of dimension 1 (Proposition 7.2.7). Otherwise, by Section 7.3 there is an invariant subspace U of dimension 1 or 2, and dimension 1 is impossible since it would be spanned by an eigenvector; so $\dim U = 2$. In that case $T|_U$ has no real eigenvalue, since an eigenvector of $T|_U$ would be an eigenvector of T.

If $U = V$, any basis of V works: $[T]_{\mathcal{B}}$ is a single block of size 1, or of size 2 with no real eigenvalue. Otherwise let W be a complement of U (Proposition 4.4.16), so $1 \le \dim W < n$, and let $S \in \mathrm{L}(W)$ be the compression of T to W. By the induction hypothesis, W has a basis $\mathcal{D}$ with $[S]_{\mathcal{D}}$ block upper triangular as in (a). Let $\mathcal{C}$ be any basis of U, and $\mathcal{B} = \mathcal{C}, \mathcal{D}$. By Lemma 7.5.11(b),

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_U]_{\mathcal{C}} & X \\ 0 & [S]_{\mathcal{D}} \end{pmatrix},$$

which is block upper triangular, with first diagonal block $[T|_U]_{\mathcal{C}}$ and the remaining diagonal blocks those of $[S]_{\mathcal{D}}$. All of them have the required form.

(b) Run the same induction with $W = U^\perp$, which is a complement of U by Theorem 6.4.24, and with an orthonormal basis $\mathcal{C}$ of U and, by the induction hypothesis applied to the real inner product space W, an orthonormal basis $\mathcal{D}$ of W. Then $\mathcal{C}, \mathcal{D}$ is orthonormal, since vectors of U and of $U^\perp$ are orthogonal. For $A \in \mathbb{R}^{n \times n}$, apply this to T_A on $\mathbb{R}^n$ with the dot product, and let Q have the resulting orthonormal basis as columns. Then Q is orthogonal, and $R = Q^{\mathsf{T}}AQ = [T_A]_{\mathcal{B}}$ by Theorem 4.5.17. □

Example 7.5.28 (An invariant plane). Let

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -2 & 2 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Its characteristic polynomial is $(x-1)\big(x^2+1\big)$, so 1 is its only real eigenvalue, with eigenvector $\mathbf{v} = (1,1,0)$. Over $\mathbb{C}$, $\mathbf{z} = (1,0,-1+i)$ is an eigenvector for i, as one checks: $A\mathbf{z} = (i,0,-1-i) = i\mathbf{z}$. As above, split $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x} = (1,0,-1)$ and $\mathbf{y} = (0,0,1)$. Then

$$A\mathbf{x} = (0,0,-1) = -\mathbf{y}, \quad A\mathbf{y} = (1,0,-1) = \mathbf{x},$$

so $\operatorname{span}(\mathbf{x},\mathbf{y})$ is an invariant plane on which A acts by $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, a rotation by a right angle, with eigenvalues $\pm i$. In the basis $\mathbf{x}, \mathbf{y}, \mathbf{v}$,

$$[T_A] = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here the plane happens to have an invariant complement, $\operatorname{span}(\mathbf{v})$, so the matrix is even block diagonal. As a check, $\operatorname{tr} A = 1 = i + (-i) + 1$ and $\det A = 1 = i \cdot (-i) \cdot 1$.

Exercises

Exercise 7.5.29. ROUTINE Find an invertible S and an upper-triangular R with diagonal $1, 1, 2$ such that $A = SRS^{-1}$, where

$$A = \begin{pmatrix} 2 & -2 & 0 \\ 2 & -1 & -2 \\ -1 & 0 & 3 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

Solution on page 1596

Exercise 7.5.30. STANDARD

(a) Let V be a complex vector space with $\dim V = n \geq 1$, and $T \in \mathrm{L}(V)$. Show that for each $k = 0, 1, \ldots, n$, there is a T-invariant subspace of V of dimension k.

(b) Let $T \in \mathrm{L}\big(\mathbb{R}^4\big)$ be $T(x_1,x_2,x_3,x_4) = (-x_2, x_1, -x_4, x_3)$. Show that T has no invariant subspace of dimension 1 or 3.

Solution on page 1596

Exercise 7.5.31. STANDARD How many of the 16 matrices in $\mathbb{F}_2^{2\times 2}$ are triangularizable over $\mathbb{F}_2$? List the ones that are not.

Solution on page 1597

Exercise 7.5.32. ROUTINE

(a) Find $A \in \mathbb{Q}^{2\times 2}$ that is triangularizable over $\mathbb{R}$ but not over $\mathbb{Q}$.

(b) Find $A \in \mathbb{R}^{3\times 3}$ that has a real eigenvalue but is not triangularizable over $\mathbb{R}$.

Solution on page 1597

Exercise 7.5.33. STANDARD Prove or disprove.

(a) If $A \in F^{n\times n}$ is triangularizable, then so is A^{T}.

(b) If $A, B \in \mathbb{R}^{2\times 2}$ are triangularizable over $\mathbb{R}$, then so is $A + B$.

(c) If T is upper triangular with respect to both $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mathbf{c}_1, \ldots, \mathbf{c}_n$, then $\mathbf{c}_1$ is a multiple of $\mathbf{b}_1$.

(d) Two upper-triangular matrices with the same diagonal are similar.

Solution on page 1598

Exercise 7.5.34. STANDARD Let $A \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, counted with multiplicity.

(a) Show that for every $k \geq 1$, the eigenvalues of A^k counted with multiplicity are $\lambda_1^k, \ldots, \lambda_n^k$. Deduce that $\operatorname{tr}\left(A^k\right) = \lambda_1^k + \cdots + \lambda_n^k$.

(b) Suppose A is invertible. Show that the eigenvalues of A^{-1} counted with multiplicity are $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$, and that $\operatorname{tr}\left(A^{-1}\right) = \lambda_1^{-1} + \cdots + \lambda_n^{-1}$.

Solution on page 1598

Exercise 7.5.35. STANDARD A matrix $A \in \mathbb{C}^{3\times 3}$ satisfies $\operatorname{tr} A = 2$, $\operatorname{tr}\left(A^2\right) = 6$, and $\det A = -2$. Find the eigenvalues of A, and decide whether A is diagonalizable.

Solution on page 1598

Exercise 7.5.36. STANDARD

(a) Let $A \in \mathbb{C}^{n\times n}$. Show that 0 is the only eigenvalue of A if and only if $A^n = 0$.

(b) Find $A \in \mathbb{R}^{3\times 3}$ whose only real eigenvalue is 0 but with $A^k \neq 0$ for every $k \geq 1$.

Solution on page 1599

Exercise 7.5.37. STANDARD Find the error in the following argument.

Claim. If $A, B \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$ (counted with multiplicity), then the eigenvalues of AB are $\lambda_1\mu_1, \ldots, \lambda_n\mu_n$, in some order.

Proof. By Schur's theorem, $A = URU^*$ and $B = UR'U^*$ with U unitary and R, R' upper triangular, with diagonals $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$. Then $AB = URR'U^*$, and RR' is upper triangular with diagonal $\lambda_1\mu_1, \ldots, \lambda_n\mu_n$.

Solution on page 1599

Exercise 7.5.38. HARD Let V be a complex vector space with $\dim V = n \geq 1$, and let $T \in \mathrm{L}(V)$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, counted with multiplicity and listed in any order chosen in advance. Show that there is a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is upper triangular with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ in this order. If V is an inner product space, show that $\mathcal{B}$ can be taken orthonormal.

Solution on page 1600

Exercise 7.5.39. STANDARD BEYOND 21-242 Let $A \in \mathbb{C}^{n\times n}$, and write $A = SRS^{-1}$ as in Corollary 7.5.14. Let $S = QR'$ be the QR factorization of S (Theorem 6.4.38). Show that $A = QR''Q^*$ with $R'' = R'RR'^{-1}$ upper triangular with the same diagonal as R. Explain why this is the first proof of Theorem 7.5.18 in matrix form.

Solution on page 1600

Exercise 7.5.40. STANDARD

(a) Show that an upper-triangular unitary matrix $R \in \mathbb{C}^{n\times n}$ is diagonal, with $|r_{jj}| = 1$ for every j.

(b) Deduce that every unitary $A \in \mathbb{C}^{n\times n}$ can be written as $A = UDU^*$ with U unitary and D diagonal.

Solution on page 1601

Exercise 7.5.41. HARD Let $A \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, counted with multiplicity. Prove **Schur's inequality**

$$\sum_{j=1}^{n} |\lambda_j|^2 \le \sum_{i,j=1}^{n} |a_{ij}|^2,$$

with equality if and only if $A = UDU^*$ for some unitary U and diagonal D.

Solution on page 1601

Exercise 7.5.42. HARD Let $A \in \mathbb{R}^{n\times n}$, and suppose that p_A splits over $\mathbb{R}$, that is, all eigenvalues of A are real. Show that $A = QRQ^{\mathsf{T}}$ with $Q \in \mathbb{R}^{n\times n}$ orthogonal and $R \in \mathbb{R}^{n\times n}$ upper triangular.

Solution on page 1601

Exercise 7.5.43. HARD Let $A \in \mathbb{C}^{n\times n}$ with $\operatorname{tr}\left(A^k\right) = 0$ for $k = 1, 2, \ldots, n$. Show that $A^n = 0$. Then show that the statement fails over $\mathbb{F}_p$.

Solution on page 1602

Exercise 7.5.44. HARD Let $A, B \in \mathbb{C}^{n\times n}$ satisfy $AB - BA = A$. Show that $A^n = 0$.

Solution on page 1603

Solution on page 1603

Exercise 7.5.45. HARD Let $A \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, counted with multiplicity, and let $L_A \in \mathrm{L}(\mathbb{C}^{n\times n})$ be $L_A(X) = AX - XA$. Use the matrix units $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Lemma 2.4.24).

(a) Suppose A is upper triangular. List the matrix units row by row, and within each row from right to left: $U_{1n}, U_{1,n-1}, \ldots, U_{11}, U_{2n}, \ldots, U_{21}, \ldots, U_{n1}$. Show that L_A is upper triangular with respect to this basis, with diagonal entry $a_{ii} - a_{jj}$ at U_{ij}.

(b) Deduce that for every A, the eigenvalues of L_A counted with multiplicity are the n^2 numbers $\lambda_i - \lambda_j$ $(1 \le i, j \le n)$.

(c) Deduce again that $\operatorname{tr} L_A = 0$, and that L_A is not invertible.

Solution on page 1603

Exercise 7.5.46. STANDARD BEYOND 21-242 Let $A = \begin{pmatrix} 0 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(a) Check that $(1, 0, -i)$ is an eigenvector of A over $\mathbb{C}$, and use it to find a two-dimensional T_A-invariant subspace of $\mathbb{R}^3$.

(b) Find a basis of $\mathbb{R}^3$ in which T_A has the form of Theorem 7.5.27(a).

(c) Find an orthogonal Q such that $Q^{\mathsf{T}}AQ$ is a real Schur form.

Solution on page 1604

7.6 Minimal polynomials, Cayley–Hamilton, and commuting operators BEYOND 21-242

By Proposition 4.1.31, every operator on an n-dimensional space is a root of some nonzero polynomial of degree at most n^2. Polynomial identities of this kind have already done a lot of work. A polynomial that kills T confines the eigenvalues of T to its roots (Proposition 7.2.28), and a diagonalizable operator is killed by the product of the factors $x-\lambda$ over its eigenvalues (Proposition 7.4.26). This section studies the best such polynomial, the one of least degree. Its roots are exactly the eigenvalues, and whether it has a repeated root decides whether T is diagonalizable, a test that needs no eigenvectors at all. The Cayley–Hamilton theorem says that the characteristic polynomial is always among the polynomials that kill T, which brings the degree bound down from n^2 to n; we prove it four ways. The last two topics also run on polynomials in an operator: commuting operators, which over $\mathbb{C}$ can be triangularized by a single basis (and, if each is diagonalizable, diagonalized by one over any field), and Gershgorin's theorem, which locates eigenvalues without computing them.

Throughout, V is a finite-dimensional vector space over a field F with $\dim V = n \geq 1$, and $T \in \mathrm{L}(V)$.

7.6.1 Annihilating polynomials and the minimal polynomial

Definition 7.6.1 (Annihilating polynomial). A polynomial $p \in F[x]$ **annihilates** T if $p(T) = 0$. Similarly, p annihilates $A \in F^{n\times n}$ if $p(A) = 0$.

The zero polynomial annihilates every operator, so the question is which nonzero polynomials do. We have met several. A projection ($P^2 = P$) is annihilated by $x^2 - x$, an operator with $N^k = 0$ by x^k, the cyclic permutation Z of Example 7.2.36 by $x^3 - 1$, and a diagonalizable operator with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ by $(x-\lambda_1)\cdots(x-\lambda_m)$. Annihilating polynomials never come alone: if $p(T) = 0$, then $(qp)(T) = q(T)p(T) = 0$ for every $q \in F[x]$ (Proposition 7.2.27(a)). So the natural candidate for "the" polynomial of T is one of least degree.

Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial). There is exactly one monic polynomial $m \in F[x]$ such that $m(T) = 0$ and $\deg m \leq \deg p$ for every nonzero $p \in F[x]$ with $p(T) = 0$. Moreover, for every $p \in F[x]$,

$$p(T) = 0 \quad \Longleftrightarrow \quad m \mid p.$$

Proof. STANDARD

Uses. Proposition 4.1.31 (Every square matrix satisfies a polynomial), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 7.1.14 (Division with remainder), Definition 7.1.5 (Monic polynomial), Definition 7.1.16 (Divisibility)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Uniqueness via two candidates (Technique 1.1.8), Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. Among the nonzero annihilating polynomials, take one of smallest degree. Why is there one, and how do you make it monic?
Hint 2. For the divisibility statement, divide p by m with remainder and evaluate the identity at T. The remainder annihilates T too. What does the choice of m say about it?
Hint 3. For uniqueness, if m and m' both qualify, they have the same degree. Look at $m - m'$.
Sketch. Nonzero annihilating polynomials exist by Proposition 4.1.31(b); take one of least degree d and divide it by its leading coefficient. If $p(T) = 0$, write $p = qm + r$ with $r = \mathbf{0}$ or $\deg r < d$; then $r(T) = p(T) - q(T)m(T) = 0$, so $r = \mathbf{0}$ by minimality. Multiples of m annihilate T. A second candidate m' also has degree d, and $m - m'$ annihilates T with degree less than d (the leading terms cancel), so $m = m'$.

Existence. By Proposition 4.1.31(b), some nonzero polynomial annihilates T. So the degrees of the nonzero annihilating polynomials form a nonempty set of nonnegative integers, which has a least element d. Let p_0 be a nonzero annihilating polynomial of degree d, with leading coefficient $a \neq 0$, and put $m = a^{-1}p_0$.

Then m is monic of degree d, and $m(T) = a^{-1}p_0(T) = 0$ by Proposition 7.2.27(a). By the choice of d, $\deg m \leq \deg p$ for every nonzero p with $p(T) = 0$.

Divisibility. If $m \mid p$, say $p = qm$, then $p(T) = q(T)m(T) = 0$ by Proposition 7.2.27(a). Conversely, suppose $p(T) = 0$. By Theorem 7.1.14 there are $q, r \in F[x]$ with $p = qm + r$, where $r = \mathbf{0}$ or $\deg r < d$. Then

$$r(T) = p(T) - q(T)m(T) = 0 - 0 = 0.$$

If r were nonzero, it would be a nonzero annihilating polynomial of degree less than d, contradicting the choice of d. So $r = \mathbf{0}$, and $m \mid p$.

Uniqueness. Suppose m' is also monic, $m'(T) = 0$, and $\deg m' \leq \deg p$ for every nonzero annihilating p. Applying this to $p = m$, and the same property of m to $p = m'$, gives $\deg m' = \deg m = d$. Both are monic of degree d, so the coefficients of x^d cancel in $m - m'$, which is therefore $\mathbf{0}$ or of degree less than d. It annihilates T, so $m - m' = \mathbf{0}$ by the minimality of d. □

Definition 7.6.3 (Minimal polynomial). The polynomial m of Theorem 7.6.2 is the **minimal polynomial** of T, written m_T. For $A \in F^{n\times n}$, the **minimal polynomial** m_A of A is the minimal polynomial of $T_A \in \mathrm{L}(F^n)$. Since $p(T_A) = T_{p(A)}$ (Proposition 7.2.27(c)), m_A is the monic polynomial of least degree with $m_A(A) = 0$.

So m_T is minimal in two senses: no nonzero annihilating polynomial has smaller degree, and every annihilating polynomial is a multiple of it. Since $V \neq \{\mathbf{0}\}$, a nonzero constant c gives $c\mathrm{id}_V \neq 0$, so $\deg m_T \geq 1$. Moreover $\deg m_T = 1$ exactly when $T - \lambda\mathrm{id}_V = 0$ for some $\lambda \in F$, that is, when $T = \lambda\mathrm{id}_V$ is a scalar multiple of the identity; then $m_T = x - \lambda$.

ASIDE: *Ideals*

The set $I_T = \{p \in F[x] \mid p(T) = 0\}$ is closed under addition and under multiplication by arbitrary polynomials. Such a subset of a commutative ring is called an **ideal**. The proof of Theorem 7.6.2 used nothing about I_T except these two closure properties and $I_T \neq \{\mathbf{0}\}$, so it shows that every nonzero ideal of $F[x]$ consists of the multiples of one monic polynomial. This makes $F[x]$ a *principal ideal domain*, the setting of Section 14.4. The greatest common divisor (Theorem 7.1.45) is the same phenomenon for the ideal $\{up + vs \mid u, v \in F[x]\}$.

The least degree can be found by looking for the first linear dependence among the powers of T, exactly as the Krylov polynomial of a vector was found among $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ in Proposition 7.2.35.

Proposition 7.6.4 (The minimal polynomial is the first dependence among the powers). Let $d = \deg m_T$. Then $\mathrm{id}_V, T, \ldots, T^{d-1}$ are linearly independent in $\mathrm{L}(V)$, and d is the smallest integer $k \geq 1$ with $T^k \in \mathrm{span}\big(\mathrm{id}_V, T, \ldots, T^{k-1}\big)$. If

$$T^d = -\big(c_0\mathrm{id}_V + c_1T + \cdots + c_{d-1}T^{d-1}\big),$$

then $m_T = x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0$. Moreover, $\mathrm{id}_V, T, \ldots, T^{d-1}$ is a basis of the subspace $F[T] = \{p(T) \mid p \in F[x]\}$ of $\mathrm{L}(V)$, so $\dim F[T] = \deg m_T$.

Proof. ROUTINE

Uses. Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 7.1.14 (Division with remainder), Theorem 1.5.5 (Independence is decided at zero), Definition 7.2.26 (Polynomials in an operator)
Techniques. Too many vectors force a dependence (Technique 4.1.30)
Hint 1. A linear relation $b_0\mathrm{id}_V + b_1T + \cdots + b_{k-1}T^{k-1} = 0$ says that the polynomial $b_0 + b_1x + \cdots + b_{k-1}x^{k-1}$ annihilates T.
Hint 2. For the basis of $F[T]$, divide an arbitrary p by m_T with remainder.
Sketch. A nontrivial relation among $\mathrm{id}_V, \ldots, T^{d-1}$ would be a nonzero annihilating polynomial of degree less than d. The relation $m_T(T) = 0$ expresses T^d through lower powers, and since those are independent, the coefficients are unique,

so they are those of m_T. If T^k were a combination of lower powers for some $k < d$, there would be a monic annihilating polynomial of degree $k < d$. Finally, $p = qm_T + r$ gives $p(T) = r(T) \in \operatorname{span}\left(\mathrm{id}_V, \ldots, T^{d-1}\right)$.

For scalars $b_0, \ldots, b_{k-1}$, the relation $\sum_{i<k} b_i T^i = 0$ says exactly that $q = \sum_{i<k} b_i x^i$ annihilates T (Definition 7.2.26). If $\mathrm{id}_V, T, \ldots, T^{d-1}$ satisfied a nontrivial relation, q would be a nonzero annihilating polynomial of degree less than d, which Theorem 7.6.2 rules out. So the list is independent.

Writing $m_T = x^d + a_{d-1}x^{d-1} + \cdots + a_0$, the equation $m_T(T) = 0$ says $T^d = -\left(a_0\mathrm{id}_V + \cdots + a_{d-1}T^{d-1}\right)$. So T^d lies in the span of the lower powers, and since these are independent, its coefficients are unique (Theorem 1.5.5). Hence $c_i = a_i$ for every i. If $T^k = \sum_{i<k} b_i T^i$ for some $1 \le k < d$, then $x^k - \sum_{i<k} b_i x^i$ would be a monic annihilating polynomial of degree $k < d$, which is impossible. So d is the smallest such k.

Finally, let $p \in F[x]$. By Theorem 7.1.14, $p = qm_T + r$ with $r = \mathbf{0}$ or $\deg r < d$, so $p(T) = q(T)m_T(T) + r(T) = r(T)$, a combination of $\mathrm{id}_V, \ldots, T^{d-1}$. Thus this independent list spans $F[T]$, and it is a basis. □

In the language of Proposition 7.2.35: the Krylov sequence of the "vector" $\mathrm{id}_V \in \mathrm{L}(V)$ under the operator $S \mapsto TS$ on $\mathrm{L}(V)$ is $\mathrm{id}_V, T, T^2, \ldots$, and its Krylov polynomial is m_T.

Example 7.6.5 (First minimal polynomials).

(a) *Projections.* Let $P \in \mathrm{L}(V)$ with $P^2 = P$ and $P \neq 0, \mathrm{id}_V$. Then $x^2 - x$ annihilates P. If $P = c\mathrm{id}_V$, then $c^2 = c$, so $c \in \{0, 1\}$; this is excluded, so no polynomial of degree 1 annihilates P, and $m_P = x^2 - x = x(x-1)$.

(b) *The shift.* Let $N \in \mathrm{L}(F^n)$ with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \ge 2$ (Example 7.2.20). Then $N^n = 0$. For scalars $b_0, \ldots, b_{n-1}$,

$$\left(b_0\,\mathrm{id} + b_1N + \cdots + b_{n-1}N^{n-1}\right)\mathbf{e}_n = b_0\mathbf{e}_n + b_1\mathbf{e}_{n-1} + \cdots + b_{n-1}\mathbf{e}_1,$$

which is $\mathbf{0}$ only if every $b_i = 0$. So $\mathrm{id}, N, \ldots, N^{n-1}$ are independent, and Proposition 7.6.4 gives $\deg m_N \ge n$. Since x^n annihilates N, $m_N \mid x^n$, so $\deg m_N = n$, and a monic divisor of x^n of degree n is x^n itself. Hence $m_N = x^n$.

(c) *A rotation.* Let $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then $R^2 = -I_2$, so $x^2 + 1$ annihilates R, and R is not a scalar matrix, so $m_R = x^2 + 1$, over $\mathbb{R}$ and over $\mathbb{C}$ alike. Over $\mathbb{R}$ this has no roots, and R has no eigenvalues; over $\mathbb{C}$ its roots are $\pm i$, which are the eigenvalues of R (Non-example 7.5.6).

(d) *The cyclic permutation.* Let $Z(x, y, z) = (z, x, y)$ on F^3, so $Z^3 = \mathrm{id}$. The matrices of id, Z, Z^2 have their 1's in disjoint positions,

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad [Z] = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad [Z]^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

so they are independent, and $m_Z = x^3 - 1$ over every field.

Exercise 7.6.6. ROUTINE Let $A \in F^{n\times n}$ and $c \in F$.

(a) Show that $m_{A^{\mathsf{T}}} = m_A$.

(b) Show that $m_{A+cI_n}(x) = m_A(x - c)$. Deduce the minimal polynomial of $\lambda I_n + N$, where N is the shift of Example 7.6.5(b).

Solution on page 1605

7.6.2 Roots of the minimal polynomial

In each part of Example 7.6.5, the roots of m_T in F were exactly the eigenvalues of T: 0 and 1 for a projection, 0 for the shift, none for the real rotation, $\pm i$ for the complex one, and the cube roots of 1 in F for Z. This always happens.

Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues). Let $\lambda \in F$. Then λ is an eigenvalue of T if and only if $m_T(\lambda) = 0$.

One direction is Proposition 7.2.28. For the other, peel the factor $x - \lambda$ off m_T. What is left has smaller degree, so it does not annihilate T; it produces some nonzero vector, and $T - \lambda\,\mathrm{id}$ kills that vector.

Proof. ROUTINE

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.1.20 (Factor theorem), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 7.2.6 (Eigenvalue, eigenvector)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. The direction "eigenvalue ⇒ root" was proved in Section 7.2. Where?
Hint 2. If $m_T(\lambda) = 0$, write $m_T = (x - \lambda)q$. Why is $q(T) \neq 0$?
Hint 3. Choose $\mathbf{v}$ with $q(T)\mathbf{v} \neq \mathbf{0}$, and apply $T - \lambda\mathrm{id}_V$ to $q(T)\mathbf{v}$.
Sketch. If λ is an eigenvalue, then $m_T(\lambda) = 0$ by Proposition 7.2.28, as $m_T(T) = 0$. Conversely, if $m_T(\lambda) = 0$, the factor theorem gives $m_T = (x - \lambda)q$ with q monic of degree $\deg m_T - 1$, so $q(T) \neq 0$ by minimality. For $\mathbf{v}$ with $\mathbf{w} = q(T)\mathbf{v} \neq \mathbf{0}$, we get $(T - \lambda\mathrm{id}_V)\mathbf{w} = m_T(T)\mathbf{v} = \mathbf{0}$, so $\mathbf{w}$ is an eigenvector for λ.

If λ is an eigenvalue of T, then $m_T(\lambda) = 0$ by Proposition 7.2.28, because $m_T(T) = 0$.

Conversely, suppose $m_T(\lambda) = 0$. By the factor theorem (Theorem 7.1.20), $m_T = (x - \lambda)q$ for some $q \in F[x]$ of degree $\deg m_T - 1$, and q is monic because m_T is. Since $q \neq \mathbf{0}$ has smaller degree than m_T, Theorem 7.6.2 gives $q(T) \neq 0$. Choose $\mathbf{v} \in V$ with $\mathbf{w} = q(T)\mathbf{v} \neq \mathbf{0}$. By Proposition 7.2.27(a),

$$(T - \lambda\mathrm{id}_V)\mathbf{w} = (T - \lambda\mathrm{id}_V)q(T)\mathbf{v} = m_T(T)\mathbf{v} = \mathbf{0}.$$

So $T\mathbf{w} = \lambda\mathbf{w}$ with $\mathbf{w} \neq \mathbf{0}$, and λ is an eigenvalue of T (Definition 7.2.6). □

So m_T and the characteristic polynomial p_T have the same roots in F, namely the eigenvalues (Corollary 7.3.13). They can differ in the multiplicities of those roots: $m_{\mathrm{id}_V} = x - 1$, while $p_{\mathrm{id}_V} = (x - 1)^n$.

Corollary 7.6.8 (The shape of the minimal polynomial over $\mathbb{C}$). Suppose $F = \mathbb{C}$, and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then

$$m_T = (x - \lambda_1)^{k_1}(x - \lambda_2)^{k_2}\cdots(x - \lambda_m)^{k_m}$$

for some integers $k_1, \ldots, k_m \geq 1$.

Proof. ROUTINE

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues)
Hint. Factor the monic polynomial m_T completely over $\mathbb{C}$. Which numbers can appear in the factors, and which must?
Sketch. By Corollary 7.1.35, m_T is a product of factors $x - \mu$ over its roots μ. By Theorem 7.6.7, the roots are exactly $\lambda_1, \ldots, \lambda_m$, so each occurs at least once; group equal factors.

Since m_T is monic of degree at least 1, Corollary 7.1.35 writes it as $(x - \mu_1)\cdots(x - \mu_d)$, where $\mu_1, \ldots, \mu_d$ are its roots listed with multiplicity. By Theorem 7.6.7, the set of roots is $\{\lambda_1, \ldots, \lambda_m\}$. Grouping equal factors gives the stated form, with $k_i \geq 1$ the number of j with $\mu_j = \lambda_i$. □

Two quick consequences show what the minimal polynomial is good for.

Corollary 7.6.9 (Inverses are polynomials). T is invertible if and only if $m_T(0) \neq 0$. In that case, write $m_T = xs + c_0$ with $c_0 = m_T(0)$ and $s \in F[x]$. Then

$$T^{-1} = -c_0^{-1}s(T),$$

a polynomial in T of degree $\deg m_T - 1$.

Proof. ROUTINE

Uses. Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues), Theorem 7.1.20 (Factor theorem), Proposition 7.2.27 (Evaluating polynomials at an operator)
Hint 1. T is invertible if and only if 0 is not an eigenvalue. Use Theorem 7.6.7.
Hint 2. Evaluate $m_T = xs + c_0$ at T and solve for id_V.
Sketch. Since V is finite-dimensional, T is invertible exactly when 0 is not an eigenvalue, that is, when $m_T(0) \neq 0$. Then $0 = m_T(T) = Ts(T) + c_0\mathrm{id}_V$ gives $T\big(-c_0^{-1}s(T)\big) = \mathrm{id}_V$, and $s(T)$ commutes with T.

By Theorem 7.2.9 (with $\lambda = 0$, using $\dim V < \infty$), T is invertible if and only if 0 is not an eigenvalue of T, which by Theorem 7.6.7 happens if and only if $m_T(0) \neq 0$.

Suppose $c_0 = m_T(0) \neq 0$. By Theorem 7.1.20 with $c = 0$, $m_T = xs + c_0$ with $\deg s = \deg m_T - 1$. Evaluating at T (Proposition 7.2.27(a)),

$$0 = m_T(T) = Ts(T) + c_0\mathrm{id}_V, \quad \text{so} \quad T\big(-c_0^{-1}s(T)\big) = \mathrm{id}_V.$$

By Proposition 7.2.27(b), $s(T)$ commutes with T, so also $\big(-c_0^{-1}s(T)\big)T = \mathrm{id}_V$, and $-c_0^{-1}s(T)$ is the inverse of T. □

For example, $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ has $A^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} = 3A - I_2$, and A is not a scalar matrix, so $m_A = x^2 - 3x + 1$. Here $m_A(0) = 1$ and $m_A = x(x-3) + 1$, so $A^{-1} = -(A - 3I_2) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, as one checks by multiplying.

Proposition 7.6.10 (Invariant subspaces and similarity).

(a) If $U \neq \{\mathbf{0}\}$ is a T-invariant subspace of V, then $m_{T|_U}$ divides m_T, where $T|_U$ is regarded as an operator on U.

(b) If $V = U \oplus W$ with U and W both T-invariant and nonzero, then for every $p \in F[x]$,

$$p(T) = 0 \iff p(T|_U) = 0 \text{ and } p(T|_W) = 0.$$

(c) For every basis $\mathcal{B}$ of V, $m_{[T]_{\mathcal{B}\mathcal{B}}} = m_T$. In particular, similar matrices have the same minimal polynomial.

Proof. ROUTINE

Uses. Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 4.2.1 (Restriction), Proposition 2.4.12 (Matrices of sums and scalar multiples), Theorem 4.5.25 (Similar matrices are one operator in two bases)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. First show that $p(T)$ restricted to an invariant U is $p(T|_U)$. Start with powers.
Hint 2. For (c), use $[p(T)]_{\mathcal{B}\mathcal{B}} = p([T]_{\mathcal{B}\mathcal{B}})$ and the injectivity of $S \mapsto [S]_{\mathcal{B}\mathcal{B}}$.
Sketch. For invariant U, $T^k|_U = (T|_U)^k$ by induction, so $p(T)|_U = p(T|_U)$. Then (a): $m_T(T|_U) = m_T(T)|_U = 0$, so $m_{T|_U} \mid m_T$. (b): write $\mathbf{v} = \mathbf{u} + \mathbf{w}$. (c): $p(T) = 0$ if and only if $p([T]_{\mathcal{B}\mathcal{B}}) = [p(T)]_{\mathcal{B}\mathcal{B}} = 0$, so T and $[T]_{\mathcal{B}\mathcal{B}}$ have the same annihilating polynomials; similar matrices are matrices of one operator.

Let U be T-invariant. For $\mathbf{u} \in U$, induction on k gives $T^k\mathbf{u} = (T|_U)^k\mathbf{u}$: both sides are $\mathbf{u}$ for $k = 0$, and $T^{k+1}\mathbf{u} = T(T|_U)^k\mathbf{u} = (T|_U)^{k+1}\mathbf{u}$ because $(T|_U)^k\mathbf{u} \in U$. Taking linear combinations, $p(T)\mathbf{u} = p(T|_U)\mathbf{u}$ for every $p \in F[x]$ and $\mathbf{u} \in U$.

(a) By the above, $m_T(T|_U)\mathbf{u} = m_T(T)\mathbf{u} = \mathbf{0}$ for every $\mathbf{u} \in U$. So m_T annihilates $T|_U$, and $m_{T|_U} \mid m_T$ by Theorem 7.6.2.

(b) If $p(T) = 0$, then $p(T|_U)\mathbf{u} = p(T)\mathbf{u} = \mathbf{0}$ for all $\mathbf{u} \in U$, and likewise on W. Conversely, if both restrictions are annihilated, write $\mathbf{v} \in V$ as $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$, $\mathbf{w} \in W$. Then $p(T)\mathbf{v} = p(T|_U)\mathbf{u} + p(T|_W)\mathbf{w} = \mathbf{0}$.

(c) By Proposition 7.2.27(c), $[p(T)]_{\mathcal{B}\mathcal{B}} = p([T]_{\mathcal{B}\mathcal{B}})$, and $S \mapsto [S]_{\mathcal{B}\mathcal{B}}$ is injective (Proposition 2.4.12). So $p(T) = 0$ if and only if $p([T]_{\mathcal{B}\mathcal{B}}) = 0$. Thus T and $[T]_{\mathcal{B}\mathcal{B}}$ have the same annihilating polynomials, hence the same minimal polynomial by the uniqueness in Theorem 7.6.2. If $A \sim B$, then A and B are matrices of a single operator T (Theorem 4.5.25), so $m_A = m_T = m_B$. □

Remark 7.6.11 (Neither polynomial determines the other). The matrices I_2 and $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ have the same characteristic polynomial $(x-1)^2$. But $m_{I_2} = x - 1$, while $J - I_2 \neq 0$ and $(J - I_2)^2 = 0$, so $m_J = (x-1)^2$. In the other direction, $\operatorname{diag}(2,2,3)$ and $\operatorname{diag}(2,3,3)$ have the same minimal polynomial $(x-2)(x-3)$ but different characteristic polynomials. Even the two together do not determine a matrix up to similarity (Exercise 7.6.43); the Jordan form of Section 10.3 does.

7.6.3 Diagonalizability

Proposition 7.4.26 shows that a diagonalizable operator is killed by a product of distinct linear factors. The converse is the most useful property of the minimal polynomial: to show that T is diagonalizable, it is enough to find one polynomial identity of the right shape, without looking for a single eigenvector.

Intuition (Why distinct factors should split the space). If $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$, each summand is killed by its own factor $T - \lambda_i \operatorname{id}$, so the product of all the factors kills V. Running this backwards: if $(T - c_1 \operatorname{id}) \cdots (T - c_k \operatorname{id}) = 0$ with distinct c_i, we want to write each $\mathbf{v}$ as a sum of vectors $\mathbf{v}_i$ with $(T - c_i \operatorname{id})\mathbf{v}_i = \mathbf{0}$. Polynomials in T are the only operators we have, so we look for polynomials L_i with $L_1 + \cdots + L_k = 1$ and $(x - c_i)L_i$ divisible by the product $q = (x - c_1) \cdots (x - c_k)$. Then $\mathbf{v} = \sum L_i(T)\mathbf{v}$, and $L_i(T)\mathbf{v}$ is killed by $T - c_i \operatorname{id}$. The Lagrange polynomials of the nodes $c_1, \ldots, c_k$ (Section 7.1.4) are exactly such polynomials. If a factor is repeated, as in $(x - c)^2$, no such decomposition can work, and the shift shows it does not.

Proposition 7.6.12 (Distinct linear factors force a diagonalization). Let $c_1, \ldots, c_k \in F$ be distinct, and suppose $q = (x - c_1)(x - c_2) \cdots (x - c_k)$ annihilates T. Then

$$V = E_{c_1} \oplus E_{c_2} \oplus \cdots \oplus E_{c_k},$$

where $E_c = \ker(T - c\mathrm{id}_V)$ (some of these may be $\{\mathbf{0}\}$). In particular, T is diagonalizable, and its eigenvalues are among $c_1, \ldots, c_k$.

Proof (Lagrange interpolation in T). STANDARD

Uses. Proposition 7.1.29 (Lagrange polynomials in closed form), Corollary 7.1.23 (Polynomials and their values), Proposition 7.2.27 (Evaluating polynomials at an operator), Corollary 7.2.23 (Eigenspaces are independent), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Techniques. Lagrange interpolation in T (Technique 7.6.13), Use an annihilating polynomial (Technique 7.6.17)
Hint 1. Let $L_1, \ldots, L_k$ be the Lagrange polynomials for the nodes $c_1, \ldots, c_k$. What is $L_1 + \cdots + L_k$?
Hint 2. $(x - c_i)L_i$ is a constant multiple of q. What does that say about $\operatorname{im} L_i(T)$?
Hint 3. Write $\mathbf{v} = \sum_i L_i(T)\mathbf{v}$. Why is the resulting sum of eigenspaces direct?
Sketch. Let $L_i = \prod_{j \neq i} (x - c_j)/(c_i - c_j)$. Then $\sum_i L_i = 1$, since both sides have degree less than k and agree at the k nodes. Also $(x - c_i)L_i = a_i q$ for a constant a_i, so $(T - c_i \operatorname{id})L_i(T) = 0$ and $\operatorname{im} L_i(T) \subseteq E_{c_i}$. So $\mathbf{v} = \sum L_i(T)\mathbf{v} \in E_{c_1} + \cdots + E_{c_k}$, and the sum is direct by Corollary 7.2.23. Concatenating bases of the summands gives an eigenbasis.

If $k = 1$, then $T - c_1 \mathrm{id}_V = 0$, so $V = E_{c_1}$ and every basis is an eigenbasis. Let $k \geq 2$, and let

$$L_i = \prod_{j \neq i} \frac{x - c_j}{c_i - c_j} \quad (i = 1, \ldots, k)$$

be the Lagrange polynomials of the nodes $c_1, \ldots, c_k$ (Proposition 7.1.29), so $L_i(c_j) = 1$ if $i = j$ and 0 otherwise.

First, $L_1 + \cdots + L_k = 1$. Both sides have degree at most $k - 1$, and at each node c_j the left side takes the value $\sum_i L_i(c_j) = 1$. By Corollary 7.1.23(a), they are equal.

Second, $(x - c_i)L_i = a_i q$ with $a_i = \prod_{j \neq i} (c_i - c_j)^{-1}$, since multiplying L_i by $x - c_i$ supplies the one missing factor of q.

Let $P_i = L_i(T)$. Evaluating the two identities at T (Proposition 7.2.27(a)) gives

$$P_1 + \cdots + P_k = \mathrm{id}_V, \quad (T - c_i \mathrm{id}_V)P_i = a_i q(T) = 0.$$

The second says $\operatorname{im} P_i \subseteq E_{c_i}$. So every $\mathbf{v} \in V$ satisfies

$$\mathbf{v} = P_1\mathbf{v} + \cdots + P_k\mathbf{v} \in E_{c_1} + \cdots + E_{c_k},$$

and $V = E_{c_1} + \cdots + E_{c_k}$. The c_i are distinct, so this sum is direct (Corollary 7.2.23). Concatenating bases of the nonzero E_{c_i} gives a basis of V (Theorem 4.4.13) consisting of eigenvectors of T, so T is diagonalizable (Theorem 7.4.2). Every eigenvalue is a root of q by Proposition 7.2.28, so it is one of the c_i. □

The operators $P_i = L_i(T)$ in this proof are worth keeping. P_i maps V into E_{c_i}, and on each E_{c_j} it acts as the scalar $L_i(c_j)$ (Proposition 7.2.28), which is 1 for $j = i$ and 0 otherwise. So P_i is the projection of V onto E_{c_i} along the sum of the other eigenspaces, and it is a polynomial in T. The same argument, without the Lagrange formula, is Exercise 7.4.49.

> **Technique 7.6.13** (Lagrange interpolation in T). If $q(T) = 0$ for $q = (x - c_1)\cdots(x - c_k)$ with distinct c_i, evaluate the Lagrange polynomials L_i of the nodes c_i at T. The operators $P_i = L_i(T)$ satisfy $\sum P_i = \mathrm{id}$, $P_iP_j = 0$ for $i \neq j$ (as $q \mid L_iL_j$), and $P_i^2 = P_i$, and P_i is the projection onto E_{c_i} along the other eigenspaces. More generally, for any values $d_1, \ldots, d_k$, the interpolating polynomial $p = \sum d_iL_i$ gives an operator $p(T)$ that acts as d_i on E_{c_i}.
>
> Reach for it to write spectral projections as polynomials in T, to build operators with a prescribed action on each eigenspace (square roots, inverses, operators commuting with T), or to decompose a vector along the eigenspaces without finding an eigenbasis. The canonical example is the first proof of Proposition 7.6.12. Later: the spectral decomposition (Section 8.2).
>
> *Used in.* Proposition 7.6.12, Exercise 7.6.45.

The second proof avoids interpolation. It peels off one eigenspace at a time, like the second proof of Theorem 7.5.13, and the only polynomial fact it uses is that $q(c_k) \neq 0$ after the factor $x - c_k$ is removed.

Proof (induction on the number of factors). STANDARD

> *Uses.* Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 4.2.11 (Rank–nullity), Theorem 4.4.7 (Criteria for direct sums), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
> *Techniques.* Dimension counting (Technique 4.2.24), Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Restrict to a subspace (Technique 4.2.5)
> *Hint 1.* Induct on k. Let $N = T - c_k\mathrm{id}_V$ and $U = \operatorname{im} N$. Which polynomial annihilates $T|_U$?
> *Hint 2.* Show $E_{c_k} \cap U = \{\mathbf{0}\}$: for $\mathbf{v} = N\mathbf{w}$ in E_{c_k}, compute $q'(T)\mathbf{v}$ in two ways, where $q' = (x - c_1)\cdots(x - c_{k-1})$.
> *Hint 3.* Rank–nullity gives $\dim E_{c_k} + \dim U = n$, so $V = E_{c_k} \oplus U$.
> *Sketch.* By induction on k; $k = 1$ means $T = c_1\mathrm{id}_V$. Let $N = T - c_k\mathrm{id}_V$, $U = \operatorname{im} N$ (which is T-invariant), and $q' = q/(x - c_k)$. Then $q'(T|_U) = 0$, since $q'(T)N\mathbf{w} = q(T)\mathbf{w}$, so by induction U is the direct sum of the eigenspaces of $T|_U$. If $\mathbf{v} = N\mathbf{w} \in E_{c_k}$, then $q'(c_k)\mathbf{v} = q'(T)\mathbf{v} = q(T)\mathbf{w} = \mathbf{0}$ with $q'(c_k) \neq 0$, so $E_{c_k} \cap U = \{\mathbf{0}\}$; with rank–nullity, $V = E_{c_k} \oplus U$.

We prove the statement for all finite-dimensional spaces, by induction on k. If $k = 1$, then $T = c_1\mathrm{id}_V$ and $V = E_{c_1}$.

Let $k \geq 2$, and assume the statement for $k - 1$ distinct factors. Let $N = T - c_k\mathrm{id}_V$, $U = \operatorname{im} N$, and $q' = (x - c_1)\cdots(x - c_{k-1})$, so $q = q'(x - c_k)$. The subspace U is T-invariant by Proposition 7.2.30(a). For $\mathbf{u} = N\mathbf{w} \in U$,

$$q'(T)\mathbf{u} = q'(T)(T - c_k\mathrm{id}_V)\mathbf{w} = q(T)\mathbf{w} = \mathbf{0},$$

so q' annihilates $T|_U$. If $U \neq \{\mathbf{0}\}$, the induction hypothesis applied to $T|_U \in \mathrm{L}(U)$ shows that U is the direct sum of the eigenspaces of $T|_U$, so U has a basis of eigenvectors of $T|_U$, which are eigenvectors of T. If $U = \{\mathbf{0}\}$, then $T = c_k\mathrm{id}_V$ and there is nothing to prove.

Next, $E_{c_k} \cap U = \{\mathbf{0}\}$. Let $\mathbf{v} \in E_{c_k} \cap U$, say $\mathbf{v} = N\mathbf{w}$. As above, $q'(T)\mathbf{v} = q(T)\mathbf{w} = \mathbf{0}$. On the other hand, if $\mathbf{v} \neq \mathbf{0}$, it is an eigenvector for c_k, and Proposition 7.2.28 gives $q'(T)\mathbf{v} = q'(c_k)\mathbf{v}$, where $q'(c_k) = \prod_{j<k}(c_k - c_j) \neq 0$. Then $q'(c_k)\mathbf{v} = \mathbf{0}$ forces $\mathbf{v} = \mathbf{0}$, a contradiction.

By rank–nullity (Theorem 4.2.11) applied to N, $\dim E_{c_k} + \dim U = \dim \ker N + \dim \operatorname{im} N = n$. So the sum $E_{c_k} + U$ is direct (Theorem 4.4.7) of dimension n (Theorem 4.4.13(c)), hence equal to V (Proposition 4.1.21), and $V = E_{c_k} \oplus U$. Concatenating a basis of E_{c_k} with an eigenbasis of U gives an eigenbasis of V, so T is diagonalizable (Theorem 7.4.2). The eigenspace decomposition follows: the eigenvalues of T are roots of q (Proposition 7.2.28), so the eigenbasis lies in $E_{c_1} \cup \cdots \cup E_{c_k}$, which gives $V = E_{c_1} + \cdots + E_{c_k}$, and this sum is direct by Corollary 7.2.23. □

Remark 7.6.14 (Coprime factors). The identity $\sum L_i = 1$ in the first proof is a special case of Bézout's identity: if p and s are coprime, then $up + vs = 1$ for some $u, v \in F[x]$ (Corollary 7.1.48). The same argument then shows that if $(ps)(T) = 0$, then $V = \ker p(T) \oplus \ker s(T)$, with both summands T-invariant (Exercise 7.2.55). Applied to the factorization $m_T = (x - \lambda_1)^{k_1} \cdots (x - \lambda_m)^{k_m}$ over $\mathbb{C}$, whose factors are pairwise coprime, it splits V into the subspaces $\ker(T - \lambda_i \operatorname{id})^{k_i}$ even when T is not diagonalizable. That is the generalized eigenspace decomposition of Section 10.2.

Theorem 7.6.15 (Diagonalizability and the minimal polynomial). The following are equivalent.

(I) T is diagonalizable.

(II) $m_T = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_m)$ for some distinct $\lambda_1, \ldots, \lambda_m \in F$.

(III) Some product $(x - c_1) \cdots (x - c_k)$ of distinct linear factors, with $c_1, \ldots, c_k \in F$, annihilates T.

In that case, the λ_i in (II) are exactly the distinct eigenvalues of T.

Proof. STANDARD

Uses. Proposition 7.4.26 (Polynomials that kill a diagonalizable operator), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues), Proposition 7.1.22 (Counting roots), Proposition 7.6.12 (Distinct linear factors force a diagonalization)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. For (I) ⇒ (II), Proposition 7.4.26 gives an annihilating polynomial. How does it compare with m_T?
Hint 2. Both m_T and that polynomial have every eigenvalue as a root. Show that each divides the other.
Sketch. (I) ⇒ (II): $q = \prod(x - \lambda_i)$ over the distinct eigenvalues annihilates T, so $m_T \mid q$. Each λ_i is a root of m_T, so $q \mid m_T$ by Proposition 7.1.22. Both are monic, so $m_T = q$. (II) ⇒ (III) is immediate, and (III) ⇒ (I) is Proposition 7.6.12.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (I).

(I) ⇒ (II). Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T and $q = (x - \lambda_1) \cdots (x - \lambda_m)$. By Proposition 7.4.26, $q(T) = 0$, so $m_T \mid q$ (Theorem 7.6.2). By Theorem 7.6.7, each λ_i is a root of m_T, so $q \mid m_T$ by Proposition 7.1.22. Two monic polynomials that divide each other are equal (their quotient is a constant, which is 1 by comparing leading coefficients). So $m_T = q$.

(II) ⇒ (III). Take $c_i = \lambda_i$; m_T annihilates T.

(III) ⇒ (I). This is Proposition 7.6.12.

Finally, by Theorem 7.6.7 the roots of m_T are the eigenvalues, so in (II) the λ_i are exactly the distinct eigenvalues. □

Corollary 7.6.16 (Restrictions of diagonalizable operators). If T is diagonalizable and $U \neq \{\mathbf{0}\}$ is T-invariant, then $T{\restriction_U} \in \mathrm{L}(U)$ is diagonalizable.

Proof. ROUTINE

Uses. Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Proposition 7.6.10 (Invariant subspaces and similarity)
Hint. Which polynomial annihilates $T{\restriction_U}$?
Sketch. m_T is a product of distinct linear factors and annihilates $T{\restriction_U}$ (Proposition 7.6.10(a)), so (III) of Theorem 7.6.15 holds for $T{\restriction_U}$.

By Theorem 7.6.15, m_T is a product of distinct linear factors. By the proof of Proposition 7.6.10(a), $m_T(T\restriction_U) = 0$. So $T\restriction_U$ satisfies condition (III) of Theorem 7.6.15, and it is diagonalizable. (Compare Exercise 7.4.50, whose first solution argues directly with Lagrange polynomials.) □

Technique 7.6.17 (Use an annihilating polynomial). When an operator satisfies a polynomial identity $q(T) = 0$, factor q and read off:

- every eigenvalue of T is a root of q (Proposition 7.2.28), and the minimal polynomial divides q;
- if q is a product of distinct linear factors, T is diagonalizable (Theorem 7.6.15);
- if $q(0) \neq 0$, then T is invertible and T^{-1} is a polynomial in T (Corollary 7.6.9);
- powers T^k reduce to polynomials of degree less than $\deg q$, by dividing x^k by q.

Reach for it whenever the hypothesis is an identity such as $A^2 = A$, $A^3 = A$, $A^k = I$, or $T^2 = 3T - 2\,\mathrm{id}$, and whenever the characteristic polynomial is available, since it always annihilates T (Theorem 7.6.20). Watch the field: $x^2 - 1 = (x-1)(x+1)$ has distinct roots only if $1 + 1 \neq 0$, and $x^2 + 1$ splits over $\mathbb{C}$ and over $\mathbb{F}_5$ but not over $\mathbb{R}$ or $\mathbb{F}_3$. The canonical example is Theorem 7.6.15.

Used in. Proposition 7.6.12, Exercise 7.6.19, Exercise 7.6.27, Exercise 7.6.42, Exercise 7.6.50, Exercise 7.7.4, Exercise 7.7.7, Exercise 7.7.8, Exercise 7.7.10, Problem 7.14, Problem 7.16, Problem 7.17, Exercise 7.6.41, Exercise 7.6.43, Exercise 7.6.45, Exercise 7.6.46.

Example 7.6.18 (Diagonalizability from an identity).

(a) *Projections.* If $P^2 = P$, then $x(x-1)$ annihilates P, and $0 \neq 1$ in every field. So every projection is diagonalizable, with eigenvalues in $\{0, 1\}$, over every field.

(b) *Involutions.* If $T^2 = \mathrm{id}_V$, then $(x-1)(x+1)$ annihilates T. When $1 + 1 \neq 0$, the roots 1 and -1 are distinct, and T is diagonalizable with $V = E_1 \oplus E_{-1}$. Over $\mathbb{F}_2$ this fails: $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$ satisfies $J^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$, but $J \neq I_2$, so $m_J = x^2 - 1 = (x-1)^2$ in $\mathbb{F}_2[x]$, and J is not diagonalizable.

(c) *Nilpotent and diagonalizable.* If $T^k = 0$ and T is diagonalizable, then every eigenvalue is a root of x^k, so 0 is the only eigenvalue, and (II) of Theorem 7.6.15 gives $m_T = x$. So $T = 0$.

(d) *Roots of unity.* The cyclic permutation Z of Example 7.6.5(d) has $m_Z = x^3 - 1$. Over $\mathbb{C}$ the three cube roots of 1 are distinct (Theorem B.1.26), so Z is diagonalizable over $\mathbb{C}$. Over $\mathbb{F}_7$, $x^3 - 1 = (x-1)(x-2)(x-4)$ (as $2^3 = 8$ and $4^3 = 64$ are 1 in $\mathbb{F}_7$), so Z is diagonalizable over $\mathbb{F}_7$. Over $\mathbb{F}_3$, $(x-1)^3 = x^3 - 3x^2 + 3x - 1 = x^3 - 1$, so $m_Z = (x-1)^3$ and Z is not diagonalizable. Over $\mathbb{R}$, $x^3 - 1 = (x-1)\big(x^2 + x + 1\big)$ does not split, and Z is not diagonalizable.

Exercise 7.6.19. ROUTINE For each matrix in $\mathbb{Q}^{n\times n}$, find a polynomial of degree at most 2 that annihilates it, find its minimal polynomial, and decide whether it is diagonalizable.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & -2 \\ 4 & -3 \end{pmatrix}.$$

Solution on page 1605

7.6.4 The Cayley–Hamilton theorem

In every example so far, m_T divided the characteristic polynomial p_T: for the shift both are x^n, for a projection $x(x-1)$ divides $x^r(x-1)^{n-r}$, and for $\mathrm{diag}(2, 2, 3)$, $(x-2)(x-3)$ divides $(x-2)^2(x-3)$. By Theorem 7.6.2, this is the statement that p_T annihilates T.

Theorem 7.6.20 (Cayley–Hamilton). $p_T(T) = 0$. Equivalently, m_T divides p_T. In particular, $p_A(A) = 0$ for every $A \in F^{n\times n}$, and $\deg m_T \leq n$.

Intuition (Four ways in). For a diagonal matrix $D = \operatorname{diag}(d_1, \dots, d_n)$ the theorem is clear: $p_D = (x - d_1)\cdots(x - d_n)$, and $p_D(D) = \operatorname{diag}(p_D(d_1), \dots, p_D(d_n)) = 0$ because each d_i is a root. The four proofs below get from this case to the general one in different ways.

- *Triangular form.* Over $\mathbb{C}$, T is upper triangular in a basis whose flag is $V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n$, and $p_T = (x - \lambda_1)\cdots(x - \lambda_n)$ with the λ_j on the diagonal. Each factor $T - \lambda_j\,\mathrm{id}$ pushes V_j down into V_{j-1}, so applying all n factors pushes V down to $\{\mathbf{0}\}$.
- *One Krylov subspace at a time.* To show that $p_T(T)$ kills $\mathbf{v}$, only the Krylov subspace of $\mathbf{v}$ matters. There T acts by a companion matrix, which visibly satisfies its own characteristic polynomial, and that polynomial divides p_T.
- *The adjugate.* The tempting one-liner "$p_A(A) = \det(AI - A) = \det 0 = 0$" is wrong (Exercise 7.3.47). Its idea can be rescued: substitute A not into a determinant, but into the matrix identity $(xI - A)\operatorname{adj}(xI - A) = p_A I$.
- *Perturbation.* Diagonalizable matrices are "most" matrices. Nudge A to a diagonalizable matrix, where the theorem is easy, and let the nudge go to zero. Because everything is polynomial, "go to zero" can be done with algebra instead of limits.

The first proof rests on a lemma that is worth having on its own.

Lemma 7.6.21 (A triangular operator is killed by its diagonal). Let $\mathcal{B} = \mathbf{b}_1, \dots, \mathbf{b}_n$ be a basis with $[T]_{\mathcal{B}}$ upper triangular, with diagonal entries $\lambda_1, \dots, \lambda_n$, and let $V_j = \operatorname{span}(\mathbf{b}_1, \dots, \mathbf{b}_j)$. Then $(T - \lambda_j \mathrm{id}_V)(V_j) \subseteq V_{j-1}$ for each j, and

$$(T - \lambda_1 \mathrm{id}_V)(T - \lambda_2 \mathrm{id}_V)\cdots(T - \lambda_n \mathrm{id}_V) = 0.$$

Proof. ROUTINE

Uses. Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Definition 2.3.1 (Matrix of a linear map), Proposition 7.2.4 (Invariance is checked on a spanning list)
Techniques. Build a flag of invariant subspaces (Technique 7.5.4)
Hint 1. Check $(T - \lambda_j \mathrm{id}_V)\mathbf{b}_i \in V_{j-1}$ for $i < j$ and for $i = j$ separately.
Hint 2. For the product, show by downward induction on j that $(T - \lambda_j \mathrm{id}_V)\cdots(T - \lambda_n \mathrm{id}_V)$ maps V into V_{j-1}.
Sketch. For $i < j$, $(T - \lambda_j\,\mathrm{id})\mathbf{b}_i \in V_i \subseteq V_{j-1}$ since V_i is invariant; and $(T - \lambda_j\,\mathrm{id})\mathbf{b}_j = \sum_{i<j} a_{ij}\mathbf{b}_i \in V_{j-1}$. So $T - \lambda_j\,\mathrm{id}$ maps V_j into V_{j-1}. Applying $T - \lambda_n\,\mathrm{id}, T - \lambda_{n-1}\,\mathrm{id}, \dots, T - \lambda_1\,\mathrm{id}$ in turn moves $V = V_n$ into V_{n-1}, then V_{n-2}, and finally into $V_0 = \{\mathbf{0}\}$.

Write $[T]_{\mathcal{B}} = (a_{ij})$, so $a_{jj} = \lambda_j$ and, by Proposition 7.5.2, $T\mathbf{b}_j = \lambda_j \mathbf{b}_j + \sum_{i<j} a_{ij}\mathbf{b}_i$. Fix j. For $i < j$, $\mathbf{b}_i \in V_i$ and V_i is T-invariant (Proposition 7.5.2), so $(T - \lambda_j \mathrm{id}_V)\mathbf{b}_i \in V_i \subseteq V_{j-1}$. For $i = j$,

$$(T - \lambda_j \mathrm{id}_V)\mathbf{b}_j = \sum_{i<j} a_{ij}\mathbf{b}_i \in V_{j-1}.$$

So $T - \lambda_j \mathrm{id}_V$ maps the spanning list $\mathbf{b}_1, \dots, \mathbf{b}_j$ of V_j into the subspace V_{j-1}, hence maps V_j into V_{j-1}.

Now let $Q_j = (T - \lambda_j \mathrm{id}_V)\cdots(T - \lambda_n \mathrm{id}_V)$. We show $Q_j(V) \subseteq V_{j-1}$ by downward induction on j. For $j = n$, $Q_n(V) = (T - \lambda_n \mathrm{id}_V)(V_n) \subseteq V_{n-1}$. If $Q_{j+1}(V) \subseteq V_j$, then $Q_j(V) = (T - \lambda_j \mathrm{id}_V)\big(Q_{j+1}(V)\big) \subseteq (T - \lambda_j \mathrm{id}_V)(V_j) \subseteq V_{j-1}$. For $j = 1$ this says $Q_1(V) \subseteq V_0 = \{\mathbf{0}\}$, so $Q_1 = 0$. □

Proof (via triangular form). STANDARD

Uses. Theorem 7.5.13 (Upper-triangular form), Theorem 7.5.17 (Triangularizable means the characteristic polynomial splits), Proposition 7.5.9 (Multiplicities on the diagonal), Lemma 7.6.21 (A triangular operator is killed by its diagonal), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 7.3.1 (Determinant over a commutative ring)
Techniques. Prove it for triangular matrices first (Technique 7.5.23), Build a flag of invariant subspaces (Technique 7.5.4)
Hint 1. Over $\mathbb{C}$, choose a basis in which T is upper triangular. What is p_T in terms of the diagonal?
Hint 2. Write $p_T(T)$ as a product of the factors $T - \lambda_j \mathrm{id}_V$ in the order of Lemma 7.6.21.
Hint 3. For a real matrix A, regard it as a complex matrix. Do p_A and $p_A(A)$ change?

Sketch. If p_T splits (always over $\mathbb{C}$), T is upper triangular in some basis with diagonal $\lambda_1, \dots, \lambda_n$, and $p_T = (x-\lambda_1)\cdots(x-\lambda_n)$. Then $p_T(T) = (T-\lambda_1\,\mathrm{id})\cdots(T-\lambda_n\,\mathrm{id}) = 0$ by Lemma 7.6.21. A real matrix has the same characteristic polynomial and the same $p_A(A)$ when regarded as a complex matrix, so the complex case covers it.

Suppose first that p_T splits over F, as it always does when $F = \mathbb{C}$. By Theorem 7.5.17 (or Theorem 7.5.13 when $F = \mathbb{C}$), V has a basis $\mathcal{B}$ in which $[T]_{\mathcal{B}}$ is upper triangular, with diagonal entries $\lambda_1, \dots, \lambda_n$, and then $p_T = (x-\lambda_1)\cdots(x-\lambda_n)$ by Proposition 7.5.9. By Proposition 7.2.27(a) and Lemma 7.6.21,

$$p_T(T) = (T-\lambda_1 \mathrm{id}_V)(T-\lambda_2 \mathrm{id}_V)\cdots(T-\lambda_n \mathrm{id}_V) = 0.$$

Now let $A \in \mathbb{R}^{n\times n}$, and regard it as an element of $\mathbb{C}^{n\times n}$. Its characteristic polynomial is the same polynomial in both views, because $\det(xI_n - A)$ is computed by the Leibniz formula from the entries of A alone (Definition 7.3.1). The matrix $p_A(A)$ is also the same, since it is computed from the entries of A and the coefficients of p_A by sums and products. By the complex case, $p_A(A) = 0$. For a real operator T, apply this to $A = [T]_{\mathcal{B}\mathcal{B}}$ and use $[p_T(T)]_{\mathcal{B}\mathcal{B}} = p_T(A)$ (Proposition 7.2.27(c), Definition 7.3.12).

The same argument works over any subfield of $\mathbb{C}$, such as $\mathbb{Q}$. Over a general field it needs a larger field in which p_T splits; such a field always exists (Section 14.2), but the next two proofs avoid the question. □

The second proof works over every field. All of its work was done in Section 7.3, where Proposition 7.3.26 identified the action of T on a Krylov subspace with a companion matrix; Exercise 7.3.51 already carried it out.

Proof (one Krylov subspace at a time). ROUTINE

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Proposition 7.3.26 (Krylov subspaces and the characteristic polynomial), Proposition 7.2.27 (Evaluating polynomials at an operator)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. It suffices to show $p_T(T)\mathbf{v} = \mathbf{0}$ for every $\mathbf{v} \neq \mathbf{0}$. Which polynomial kills $\mathbf{v}$?
Hint 2. The Krylov polynomial p of $\mathbf{v}$ kills $\mathbf{v}$, and by Proposition 7.3.26 it divides p_T.
Sketch. For $\mathbf{v} \neq \mathbf{0}$, let p be the Krylov polynomial of $\mathbf{v}$ (Proposition 7.2.35), so $p(T)\mathbf{v} = \mathbf{0}$. By Proposition 7.3.26, $p_T = gp$ for some g, so $p_T(T)\mathbf{v} = g(T)p(T)\mathbf{v} = \mathbf{0}$.

Let $\mathbf{v} \in V$. If $\mathbf{v} = \mathbf{0}$, then $p_T(T)\mathbf{v} = \mathbf{0}$. Otherwise let p be the monic polynomial of Proposition 7.2.35 for $\mathbf{v}$, so $p(T)\mathbf{v} = \mathbf{0}$. By Proposition 7.3.26(c), $p_T = gp$ for some $g \in F[x]$. By Proposition 7.2.27(a),

$$p_T(T)\mathbf{v} = g(T)(p(T)\mathbf{v}) = g(T)\mathbf{0} = \mathbf{0}.$$

So $p_T(T)$ sends every vector to $\mathbf{0}$, and $p_T(T) = 0$. □

The third proof is pure algebra: it uses only that $F[x]$ is a commutative ring, so it works verbatim for matrices with entries in any commutative ring, such as $\mathbb{Z}$. It turns the naive substitution into a correct argument.

Proof (via the adjugate). HARD

Uses. Definition 7.3.4 (Characteristic polynomial), Proposition 7.3.2 (Determinants over a commutative ring), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Proposition 7.1.4 (Degree of a product), Remark 7.1.11 (Polynomials in a matrix)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. In $F[x]^{n\times n}$, $(xI_n - A)\,\mathrm{adj}(xI_n - A) = p_A I_n$ by Proposition 7.3.2(f). The entries of $\mathrm{adj}(xI_n - A)$ have degree at most $n-1$; collect powers of x to write $\mathrm{adj}(xI_n - A) = \sum_{k=0}^{n-1} x^k B_k$ with $B_k \in F^{n\times n}$.
Hint 2. Expand the left side and compare the coefficients of each x^k with those of $p_A I_n = \sum_k c_k x^k I_n$. You get $n+1$ matrix equations.
Hint 3. Multiply the equation for x^k on the left by A^k and add them all up. The left sides telescope.
Sketch. Write $\mathrm{adj}(xI_n - A) = \sum_{k=0}^{n-1} x^k B_k$ and $p_A = \sum_{k=0}^{n} c_k x^k$. Comparing coefficients in $(xI_n - A)\sum_k x^k B_k = p_A I_n$ gives $B_{n-1} = I_n$, $B_{k-1} - AB_k = c_k I_n$ for $1 \le k \le n-1$, and $-AB_0 = c_0 I_n$. Multiplying the x^k equation by A^k and summing, the left sides telescope to 0 and the right sides add up to $p_A(A)$.

It suffices to treat a matrix $A \in F^{n\times n}$; for an operator, take $A = [T]_{\mathcal{B}\mathcal{B}}$ as in the first proof. Write $p_A = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_0$, with $c_n = 1$, and let $M = xI_n - A \in F[x]^{n\times n}$. By Proposition 7.3.2(f) (for $n \geq 2$; for $n = 1$, $\operatorname{adj} M = (1)$ and the identity is clear),

$$M \operatorname{adj}(M) = (\det M) I_n = p_A I_n. \tag{7.6.1}$$

Each entry of $\operatorname{adj}(M)$ is, up to sign, the determinant of an $(n-1)\times(n-1)$ submatrix of M (Definition 5.5.6). Each Leibniz term of such a determinant is a product of $n-1$ entries of degree at most 1, so every entry of $\operatorname{adj}(M)$ is $\mathbf{0}$ or has degree at most $n-1$ (Proposition 7.1.4). Collecting, in each entry, the coefficient of x^k, we get matrices $B_0, \ldots, B_{n-1} \in F^{n\times n}$ with

$$\operatorname{adj}(M) = B_0 + xB_1 + \cdots + x^{n-1}B_{n-1},$$

an identity in $F[x]^{n\times n}$, in which $x^k B_k$ means the matrix with entries $x^k (B_k)_{ij}$. Since x commutes with everything in $F[x]$, the left side of (7.6.1) is

$$(xI_n - A)\sum_{k=0}^{n-1} x^k B_k = \sum_{k=0}^{n-1} x^{k+1} B_k - \sum_{k=0}^{n-1} x^k A B_k.$$

Two matrices with entries in $F[x]$ are equal exactly when, for every k, the matrices of coefficients of x^k are equal. Comparing with the right side $\sum_{k=0}^{n} c_k x^k I_n$:

$$\begin{aligned} x^n: &\quad B_{n-1} = I_n, \\ x^k \text{ for } 1 \leq k \leq n-1: &\quad B_{k-1} - AB_k = c_k I_n, \\ x^0: &\quad -AB_0 = c_0 I_n. \end{aligned}$$

Multiply the equation for x^k on the left by A^k, for $k = 0, 1, \ldots, n$, and add. The right sides add up to $\sum_k c_k A^k = p_A(A)$. The left sides add up to

$$\sum_{k=1}^{n} A^k B_{k-1} - \sum_{k=0}^{n-1} A^{k+1} B_k = \sum_{k=1}^{n} A^k B_{k-1} - \sum_{j=1}^{n} A^j B_{j-1} = 0,$$

where the first sum collects the terms $A^k B_{k-1}$ (from x^k with $1 \leq k \leq n$) and the second the terms $-A^{k+1}B_k$ (from x^k with $0 \leq k \leq n-1$), reindexed by $j = k+1$. So $p_A(A) = 0$. □

Remark 7.6.22 (Where the naive substitution goes wrong). The bogus argument of Exercise 7.3.47 substitutes A for x inside a determinant, which turns scalars a_{ij} into nothing in particular. The adjugate proof substitutes only into identities between polynomials with matrix coefficients, and only after multiplying each coefficient by A^k on the left. That step is safe here because the coefficients I_n and $-A$ of the left factor $xI_n - A$ commute with A. For a factor whose coefficients did not commute with A, "evaluate the product at A" and "multiply the evaluations" could give different answers.

The fourth proof introduces a technique that recurs whenever a statement is easy for "generic" matrices.

> **Technique 7.6.23** (Perturbation and density). To prove an identity for every matrix, first prove it for matrices in general position (diagonalizable, invertible, with distinct eigenvalues), where it is easy. Then move an arbitrary matrix A into general position along a path A_t that depends polynomially on a parameter t, with $A_0 = A$ and A_t in general position for all but finitely many t. If both sides of the identity are polynomials in the entries, the difference is a polynomial in t that vanishes at infinitely many points, hence identically, and in particular at $t = 0$. With analysis available, the same idea reads: the general-position matrices are dense, and both sides are continuous.
>
> Reach for it when a statement is clear for diagonalizable or invertible matrices and both sides are polynomial in the entries, over $\mathbb{C}$ or another infinite field. It fails over finite fields, where there is no

room to perturb. The canonical example is the fourth proof of Theorem 7.6.20. Compare Technique 5.3.18, which is the same move with $A + tI_n$ to reach invertible matrices.
Used in. Lemma 7.6.21, Exercise 7.7.14.

Proof (by perturbation, over $\mathbb{C}$). HARD

Uses. Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix), Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Corollary 7.5.14 (Triangularizing a matrix), Proposition 7.5.8 (The diagonal of a triangular matrix), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization), Theorem 5.3.3 (The Leibniz formula), Corollary 7.1.23 (Polynomials and their values)
Techniques. Perturbation and density (Technique 7.6.23), det is a polynomial in the entries (Technique 5.3.18)
Hint 1. Prove $p_A(A) = 0$ first for diagonalizable $A = SDS^{-1}$.
Hint 2. Triangularize $A = SRS^{-1}$ and add $t\,\mathrm{diag}(1, 2, \ldots, n)$ to R. For which t are the new diagonal entries distinct?
Hint 3. Each entry of $p_{A_t}(A_t)$ is a polynomial in t. How many roots does it have?
Sketch. For $A = SDS^{-1}$ with D diagonal, $p_A = p_D$ and $p_A(A) = S\,\mathrm{diag}(p_D(d_i))S^{-1} = 0$. In general write $A = SRS^{-1}$ with R triangular and let $A_t = S(R + t\,\mathrm{diag}(1, \ldots, n))S^{-1}$. Except for finitely many t, A_t has n distinct eigenvalues, so $p_{A_t}(A_t) = 0$. Each entry of $p_{A_t}(A_t)$ is a polynomial in t with infinitely many roots, hence $\mathbf{0}$; evaluate at $t = 0$.

Diagonalizable matrices. Let $A = SDS^{-1}$ with $D = \mathrm{diag}(d_1, \ldots, d_n)$. By Proposition 7.3.11, $p_A = p_D = (x - d_1)\cdots(x - d_n)$, and by Proposition 7.4.23,

$$p_A(A) = S\,\mathrm{diag}(p_A(d_1), \ldots, p_A(d_n))S^{-1} = S0S^{-1} = 0.$$

Perturbing. Let $A \in \mathbb{C}^{n\times n}$ be arbitrary. By Corollary 7.5.14, $A = SRS^{-1}$ with S invertible and R upper triangular with diagonal entries $r_{11}, \ldots, r_{nn}$. For $t \in \mathbb{C}$ let $E = S\,\mathrm{diag}(1, 2, \ldots, n)S^{-1}$ and

$$A_t = S(R + t\,\mathrm{diag}(1, 2, \ldots, n))S^{-1} = A + tE.$$

The matrix $R + t\,\mathrm{diag}(1, \ldots, n)$ is upper triangular with diagonal entries $r_{jj} + jt$. For $i < j$, $r_{ii} + it = r_{jj} + jt$ holds for exactly one t, namely $t = (r_{ii} - r_{jj})/(j - i)$. So outside a finite set of values of t, these diagonal entries are distinct, A_t has n distinct eigenvalues (Proposition 7.5.8(b), as A_t is similar to a triangular matrix), A_t is diagonalizable (Corollary 7.4.5), and $p_{A_t}(A_t) = 0$ by the first step.

Passing to $t = 0$. Each entry of $A_t = A + tE$ is a polynomial in t of degree at most 1. By the Leibniz formula (Theorem 5.3.3), each coefficient of $p_{A_t} = \det(xI_n - A_t)$ is a sum of products of entries of A_t and constants, hence a polynomial in t; so is each entry of every power A_t^k. Therefore each entry of the matrix $p_{A_t}(A_t)$ is a polynomial $f_{ij}(t) \in \mathbb{C}[t]$. We showed that f_{ij} vanishes at all but finitely many $t \in \mathbb{C}$, that is, at infinitely many points, so $f_{ij} = \mathbf{0}$ (Corollary 7.1.23(b)). In particular $f_{ij}(0) = 0$ for all i, j, and $A_0 = A$, so $p_A(A) = 0$. □

Each proof shows something the others do not. The triangular proof is the geometric one: it shows each factor $T - \lambda_j\,\mathrm{id}$ doing its share of the work, but it needs the eigenvalues to lie in F. The Krylov proof reduces everything to companion matrices and works over every field. The adjugate proof never mentions eigenvalues or vectors, and holds over every commutative ring. The perturbation proof shows that the diagonalizable case, which is nearly trivial, already implies the general one.

Corollary 7.6.24 (The minimal polynomial divides the characteristic polynomial).

(a) $m_T \mid p_T$. Hence $\deg m_T \le n$, the operators $\mathrm{id}_V, T, \ldots, T^n$ are linearly dependent, and $\dim F[T] \le n$.

(b) m_T and p_T have the same roots in F, namely the eigenvalues of T.

(c) If $p_T = (x - \lambda_1)^{a_1}\cdots(x - \lambda_m)^{a_m}$ with distinct $\lambda_i \in F$ (as always happens over $\mathbb{C}$), then

$$m_T = (x - \lambda_1)^{k_1}\cdots(x - \lambda_m)^{k_m} \quad \text{with } 1 \le k_i \le a_i.$$

Proof. STANDARD

Uses. Theorem 7.6.20 (Cayley–Hamilton), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Proposition 7.6.4 (The minimal polynomial is the first dependence among the powers), Theorem 7.6.7 (The roots of

the minimal polynomial are the eigenvalues), Corollary 7.3.13 (Eigenvalues of an operator), Theorem 7.1.20 (Factor theorem), Corollary 7.1.7 (No zero divisors, cancellation, and units)
Hint 1. (a) and (b) are direct. For (c), show that a monic divisor of a product of linear factors is a product of some of those factors.
Hint 2. Induct on the number of linear factors. If $q = st$ and μ is a root of q, then μ is a root of s or of t; cancel $x - \mu$.
Sketch. (a) By Cayley–Hamilton and Theorem 7.6.2, $m_T \mid p_T$; then Proposition 7.6.4 gives the rest. (b) Both root sets are the set of eigenvalues. (c) A monic divisor s of a product q of N linear factors is a subproduct: some root μ of q is a root of s or of the cofactor; cancel $x - \mu$ and induct on N. So $m_T = \prod (x - \lambda_i)^{k_i}$ with $k_i \le a_i$, and $k_i \ge 1$ by (b).

(a) By Theorem 7.6.20, p_T annihilates T, so $m_T \mid p_T$ by Theorem 7.6.2, and $\deg m_T \le \deg p_T = n$. By Proposition 7.6.4, $T^{\deg m_T}$ is a combination of lower powers, so $\mathrm{id}_V, \dots, T^n$ are dependent, and $\dim F[T] = \deg m_T \le n$.

(b) By Theorem 7.6.7 and Corollary 7.3.13, the roots of m_T and of p_T in F are both exactly the eigenvalues of T.

(c) We first show: if $q = (x - \mu_1)\cdots(x - \mu_N)$ with $\mu_j \in F$ (repetitions allowed) and s is a monic divisor of q, then s is the product of some of the factors $x - \mu_j$. We induct on N. If $N = 0$, then $q = 1 = st$, so s is a nonzero constant, and $s = 1$ since it is monic. Let $N \ge 1$ and write $q = st$. Since $q(\mu_N) = 0$, we get $s(\mu_N) = 0$ or $t(\mu_N) = 0$, because F has no zero divisors. In the first case, $s = (x - \mu_N)s'$ by Theorem 7.1.20, and cancelling $x - \mu_N$ (Corollary 7.1.7) in $(x - \mu_N)(x - \mu_1)\cdots(x - \mu_{N-1}) = (x - \mu_N)s't$ shows that the monic polynomial s' divides $(x - \mu_1)\cdots(x - \mu_{N-1})$; by induction s' is a product of some of these factors, and so is s. In the second case, $t = (x - \mu_N)t'$, and cancelling shows that s divides $(x - \mu_1)\cdots(x - \mu_{N-1})$; again the induction hypothesis applies.

Apply this with $q = p_T$ and $s = m_T$, which divides p_T by (a). So $m_T = \prod_i (x - \lambda_i)^{k_i}$ with $0 \le k_i \le a_i$, and $k_i \ge 1$ because each λ_i is a root of m_T by (b). □

Example 7.6.25 (Finding a minimal polynomial from the characteristic polynomial). Let

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Both have $p = (x - 2)^2(x - 3)$. By Corollary 7.6.24(c), each minimal polynomial is $(x - 2)(x - 3)$ or $(x - 2)^2(x - 3)$, so it suffices to test the smaller candidate. For D, $(D - 2I)(D - 3I) = \mathrm{diag}(0, 0, 1)\,\mathrm{diag}(-1, -1, 0) = 0$, so $m_D = (x - 2)(x - 3)$. For J,

$$(J - 2I)(J - 3I) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0,$$

so $m_J = (x - 2)^2(x - 3)$, and J is not diagonalizable (Theorem 7.6.15).

Example 7.6.26 (Powers without a diagonalization). Let $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$. Then $\mathrm{tr}\, A = 4$ and $\det A = 4$, so $p_A = x^2 - 4x + 4 = (x - 2)^2$, and $A \neq 2I_2$, so $m_A = (x - 2)^2$ and A is not diagonalizable. The method of Section 7.4.3 is not available, but Cayley–Hamilton is. By the binomial theorem in the commutative ring $F[x]$,

$$x^k = (2 + (x - 2))^k = 2^k + k2^{k-1}(x - 2) + (x - 2)^2 g_k$$

for some $g_k \in F[x]$. Evaluating at A and using $(A - 2I_2)^2 = 0$,

$$A^k = 2^k I_2 + k2^{k-1}(A - 2I_2) = 2^{k-1}\begin{pmatrix} 2 + k & k \\ -k & 2 - k \end{pmatrix}.$$

In general, $A^k = r(A)$, where r is the remainder of x^k on division by p_A (or by m_A), so A^k is always a combination of $I, A, \dots, A^{n-1}$.

Exercise 7.6.27. ROUTINE Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Use the Cayley–Hamilton theorem to write A^{-1} and A^3 in the form $aI_2 + bA$, and compute both.

Solution on page 1605

7.6.5 Spectral mapping

On an eigenvector, $p(T)$ acts as the scalar $p(\lambda)$ (Proposition 7.2.28). So every $p(\lambda)$, for λ an eigenvalue of T, is an eigenvalue of $p(T)$. Over $\mathbb{C}$ there are no others, and even the multiplicities match.

Theorem 7.6.28 (Spectral mapping theorem). Let V be a finite-dimensional complex vector space with $\dim V = n \geq 1$, let $T \in \mathrm{L}(V)$ and $p \in \mathbb{C}[x]$, and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T counted with multiplicity.

(a) The eigenvalues of $p(T)$ are exactly the numbers $p(\lambda)$ with λ an eigenvalue of T.

(b) $p_{p(T)} = (x - p(\lambda_1)) \cdots (x - p(\lambda_n))$. That is, the eigenvalues of $p(T)$ counted with multiplicity are $p(\lambda_1), \dots, p(\lambda_n)$.

Proof (via triangular form). STANDARD

Uses. Theorem 7.5.13 (Upper-triangular form), Proposition 7.5.9 (Multiplicities on the diagonal), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 2.4.1 (Matrix product), Corollary 7.3.13 (Eigenvalues of an operator)
Techniques. Prove it for triangular matrices first (Technique 7.5.23), Compute entry by entry (Technique 2.4.15)
Hint 1. Triangularize T. What does p do to an upper-triangular matrix?
Hint 2. Show that a product of upper-triangular matrices is upper triangular, with diagonal entries the products of the diagonal entries.
Sketch. In a basis with $R = [T]_{\mathcal{B}}$ upper triangular with diagonal $\lambda_1, \dots, \lambda_n$, products and sums of upper-triangular matrices are upper triangular with the diagonal computed entrywise, so $[p(T)]_{\mathcal{B}} = p(R)$ is upper triangular with diagonal $p(\lambda_1), \dots, p(\lambda_n)$. Then Proposition 7.5.9 gives (b), and (a) follows by taking roots.

By Theorem 7.5.13, there is a basis $\mathcal{B}$ with $R = [T]_{\mathcal{B}}$ upper triangular, and its diagonal entries are $\lambda_1, \dots, \lambda_n$ (Proposition 7.5.9).

If $X, Y \in \mathbb{C}^{n \times n}$ are upper triangular, then $(XY)_{ij} = \sum_l x_{il} y_{lj}$, and $x_{il} y_{lj} \neq 0$ requires $i \leq l \leq j$ (Definition 2.4.1). For $i > j$ there is no such l, and for $i = j$ only $l = i$, so XY is upper triangular with $(XY)_{ii} = x_{ii} y_{ii}$. Sums and scalar multiples of upper-triangular matrices are upper triangular, with diagonals added and scaled. By induction on k, R^k is upper triangular with diagonal entries λ_j^k, and so $p(R)$ is upper triangular with diagonal entries $p(\lambda_1), \dots, p(\lambda_n)$.

Since $[p(T)]_{\mathcal{B}} = p(R)$ (Proposition 7.2.27(c)), Proposition 7.5.9 applied to $p(T)$ gives (b). By Corollary 7.3.13, the eigenvalues of $p(T)$ are the roots of $p_{p(T)}$, which by (b) are $p(\lambda_1), \dots, p(\lambda_n)$; this is (a). □

The second proof gives only (a), but it uses no triangularization, only the factorization of one polynomial and the fact that polynomials in T commute.

Proof (factor $p - \mu$). STANDARD

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Fix $\mu \in \mathbb{C}$. When is $p(T) - \mu \mathrm{id}_V$ not injective? Factor $p - \mu$ over $\mathbb{C}$.
Hint 2. A product of commuting operators is injective if and only if each factor is. Why does commuting matter?
Sketch. If p is constant the claim is clear, since T has an eigenvalue. Otherwise $p - \mu = a(x - r_1) \cdots (x - r_d)$, so $p(T) - \mu \,\mathrm{id} = a(T - r_1 \,\mathrm{id}) \cdots (T - r_d \,\mathrm{id})$. These factors commute, so the product fails to be injective exactly when some factor does, that is, when some root r_i of $p - \mu$ is an eigenvalue of T, that is, when $\mu = p(\lambda)$ for some eigenvalue λ.

We prove (a). Let $\mu \in \mathbb{C}$. If $p = c$ is constant, then $p(T) = c\mathrm{id}_V$, whose only eigenvalue is c, and $p(\lambda) = c$ for every eigenvalue λ of T, of which there is at least one (Theorem 7.3.27). So assume $\deg p = d \geq 1$. Then $p - \mu$ has degree d and some leading coefficient $a \neq 0$, and by Corollary 7.1.35,

$$p - \mu = a(x - r_1)(x - r_2)\cdots(x - r_d),$$

where $r_1, \ldots, r_d$ are the roots of $p - \mu$, that is, the solutions of $p(r) = \mu$. By Proposition 7.2.27(a), $p(T) - \mu\mathrm{id}_V = aS_1S_2\cdots S_d$ with $S_i = T - r_i\mathrm{id}_V$, and these factors commute with each other (Proposition 7.2.27(b)).

We claim that $S_1\cdots S_d$ is injective if and only if every S_i is. If every S_i is injective, so is their composite. Conversely, suppose some S_i is not injective, say $S_i\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$. Since the factors commute, $S_1\cdots S_d = \left(\prod_{j\neq i} S_j\right)S_i$, which sends $\mathbf{v}$ to $\mathbf{0}$, so the composite is not injective.

Hence, by Theorem 7.2.9, μ is an eigenvalue of $p(T)$ if and only if $p(T) - \mu\mathrm{id}_V$ is not injective, if and only if some $T - r_i\mathrm{id}_V$ is not injective, if and only if some r_i is an eigenvalue of T. Since the r_i are exactly the solutions of $p(r) = \mu$, this says that $\mu = p(\lambda)$ for some eigenvalue λ of T. □

The second proof never used $\dim V < \infty$ except to know that T has an eigenvalue when p is constant, so part (a) holds for operators on any complex vector space, with nonconstant p.

Non-example 7.6.29 (Spectral mapping over $\mathbb{R}$). Over any field, every $p(\lambda)$ is an eigenvalue of $p(T)$, by Proposition 7.2.28. The converse needs $\mathbb{C}$. The rotation $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ has no real eigenvalues, but $R^2 = -I_2$ has the eigenvalue -1. In the second proof, the polynomial $x^2 - (-1) = x^2 + 1$ cannot be factored over $\mathbb{R}$.

Corollary 7.6.30 (Trace, determinant, and invertibility of $p(T)$). In the setting of Theorem 7.6.28,

$$\operatorname{tr} p(T) = p(\lambda_1) + \cdots + p(\lambda_n), \quad \det p(T) = p(\lambda_1)\cdots p(\lambda_n),$$

and $p(T)$ is invertible if and only if $p(\lambda) \neq 0$ for every eigenvalue λ of T.

Proof. ROUTINE

Uses. Theorem 7.6.28 (Spectral mapping theorem), Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem 7.2.9 (Eigenvalues and invertibility)
Hint. Apply Theorem 7.5.22 to the operator $p(T)$.
Sketch. By Theorem 7.6.28(b), the eigenvalues of $p(T)$ with multiplicity are the $p(\lambda_j)$, so Theorem 7.5.22 gives the trace and determinant. $p(T)$ is invertible exactly when 0 is not among its eigenvalues.

By Theorem 7.6.28(b), $p(\lambda_1), \ldots, p(\lambda_n)$ are the eigenvalues of $p(T)$ counted with multiplicity, and Theorem 7.5.22 applied to $p(T)$ gives the two formulas. By Theorem 7.2.9, $p(T)$ is invertible if and only if 0 is not an eigenvalue of $p(T)$, which by Theorem 7.6.28(a) means $p(\lambda) \neq 0$ for every eigenvalue λ of T. □

For example, $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$ has $p_A = x^2 - 3x + 2 = (x - 1)(x - 2)$, so its eigenvalues are 1 and 2. For $p = x^2 + x + 1$, the matrix $p(A)$ has eigenvalues $p(1) = 3$ and $p(2) = 7$, so $\operatorname{tr} p(A) = 10$ and $\det p(A) = 21$. Directly, $p(A) = A^2 + A + I_2 = \begin{pmatrix} -1 & 4 \\ -8 & 11 \end{pmatrix}$, with trace 10 and determinant $-11 + 32 = 21$.

7.6.6 Commuting operators

If $ST = TS$, then each eigenspace of T is mapped into itself by S (Proposition 7.2.30(b)). Inside an eigenspace of T, every nonzero vector is an eigenvector of T, so an eigenvector of S found there serves both operators. Over $\mathbb{C}$, this is enough to find common eigenvectors, and then common triangular and diagonal forms.

Theorem 7.6.31 (Commuting operators have a common eigenvector). Let V be a finite-dimensional complex vector space with $V \neq \{\mathbf{0}\}$, and let $T_1, \ldots, T_k \in \mathrm{L}(V)$ commute pairwise, that is, $T_iT_j = T_jT_i$ for all i, j. Then some nonzero $\mathbf{v} \in V$ is an eigenvector of every T_i. More generally, every nonzero subspace W that is invariant under all of the T_i contains such a $\mathbf{v}$.

Proof (one operator at a time). STANDARD

> *Uses.* Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Definition 7.2.1 (Invariant subspace), Definition 7.2.13 (Eigenspace)
> *Techniques.* Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Restrict to a subspace (Technique 4.2.5)
> *Hint 1.* Prove the "more generally" statement by induction on k. For $k = 1$, restrict T_1 to W.
> *Hint 2.* Let λ be an eigenvalue of $T_k|_W$. Which subspace of W consists of eigenvectors of T_k (and $\mathbf{0}$), and why is it invariant under $T_1, \dots, T_{k-1}$?
> *Sketch.* Induct on k. For $k = 1$, $T_1|_W$ is an operator on a nonzero complex space and has an eigenvector. For $k \geq 2$, let λ be an eigenvalue of $T_k|_W$ and $W' = W \cap E_\lambda(T_k) \neq \{\mathbf{0}\}$. Each T_i preserves W, and preserves $E_\lambda(T_k)$ because it commutes with T_k, so it preserves W'. By induction, W' contains a common eigenvector of $T_1, \dots, T_{k-1}$, and every nonzero vector of W' is an eigenvector of T_k.

We prove the second statement by induction on k; the first is the case $W = V$. Let $W \neq \{\mathbf{0}\}$ be invariant under every T_i.

If $k = 1$, then $T_1|_W$ is an operator on the nonzero finite-dimensional complex space W, so it has an eigenvector $\mathbf{v} \in W$ (Theorem 7.3.27), and $T_1\mathbf{v} = T_1|_W\mathbf{v}$ is a multiple of $\mathbf{v}$.

Let $k \geq 2$, and assume the statement for $k - 1$ operators. Let λ be an eigenvalue of $T_k|_W$ and

$$W' = W \cap E_\lambda(T_k) = \{\mathbf{w} \in W \mid T_k\mathbf{w} = \lambda\mathbf{w}\},$$

which contains an eigenvector of $T_k|_W$, so $W' \neq \{\mathbf{0}\}$. For $i < k$, T_i commutes with T_k, so $E_\lambda(T_k)$ is T_i-invariant (Proposition 7.2.30(b)); W is T_i-invariant by hypothesis; so W' is T_i-invariant. By the induction hypothesis applied to $T_1, \dots, T_{k-1}$ and W', some nonzero $\mathbf{v} \in W'$ is an eigenvector of $T_1, \dots, T_{k-1}$. Since $\mathbf{v} \in E_\lambda(T_k)$, it is also an eigenvector of T_k. □

> *Intuition (A smallest invariant subspace).* The first proof shrinks W once for each operator. The second goes straight to the end: take a nonzero subspace that is invariant under all the operators and as small as possible. Each operator must act on it as a scalar, since otherwise one of its eigenspaces inside would be a smaller invariant subspace. This extremal argument never counts the operators, so it works for any set of commuting operators, finite or not.

Proof (a minimal common invariant subspace). STANDARD

> *Uses.* Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 4.1.21 (Dimensions of subspaces)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9), Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
> *Hint 1.* Among the nonzero subspaces of W invariant under every T_i, choose one, W_0, of smallest dimension.
> *Hint 2.* Let λ be an eigenvalue of $T_i|_{W_0}$. Show that $W_0 \cap E_\lambda(T_i)$ is again invariant under every T_j. What does minimality force?
> *Sketch.* Pick a nonzero subspace $W_0 \subseteq W$, invariant under all T_i, of least dimension. For each i, take an eigenvalue λ of $T_i|_{W_0}$; then $W_0 \cap E_\lambda(T_i)$ is nonzero, invariant under all T_j, and contained in W_0, so it equals W_0. Thus every T_i acts on W_0 as a scalar, and every nonzero vector of W_0 is a common eigenvector.

Consider the nonzero subspaces of W that are invariant under every T_i. There is at least one, W itself, so we may choose one, W_0, of smallest dimension. Fix i. By Theorem 7.3.27, $T_i|_{W_0}$ has an eigenvalue λ, so $W_1 = W_0 \cap E_\lambda(T_i)$ is nonzero. For each j, $E_\lambda(T_i)$ is T_j-invariant (Proposition 7.2.30(b)), so W_1 is T_j-invariant. By the choice of W_0, $\dim W_1 \geq \dim W_0$, and as $W_1 \subseteq W_0$, we get $W_1 = W_0$ (Proposition 4.1.21). That is, $T_i\mathbf{w} = \lambda\mathbf{w}$ for every $\mathbf{w} \in W_0$.

So each T_i acts on W_0 as a scalar, and every nonzero $\mathbf{v} \in W_0$ is an eigenvector of every T_i. The argument used neither the number of operators nor their order, so it applies to any set of pairwise commuting operators on V. □

Both hypotheses are needed. Over $\mathbb{R}$ even a single operator, a rotation, may have no eigenvector. And for non-commuting operators the conclusion fails: the eigenvectors of $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ are the nonzero multiples of $\mathbf{e}_1$, and those of N^{T} the nonzero multiples of $\mathbf{e}_2$.

Theorem 7.6.32 (Simultaneous triangularization). Let V be a finite-dimensional complex vector space with $\dim V \geq 1$, and let $T_1, \dots, T_k \in \mathrm{L}(V)$ commute pairwise. Then there is a basis $\mathcal{B}$ of V such that $[T_i]_{\mathcal{B}}$ is upper triangular for every i.

Proof. STANDARD

Uses. Theorem 7.6.31 (Commuting operators have a common eigenvector), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Proposition 4.6.32 (Operators on a quotient), Theorem 4.6.27 (Dimension of a quotient), Proposition 4.6.28 (Lifting a basis of a quotient), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Build a flag of invariant subspaces (Technique 7.5.4), Pass to a quotient (Technique 4.6.31)
Hint 1. Induct on $n = \dim V$, as in the proof of Theorem 7.5.13. The bottom of the flag should be a common eigenvector.
Hint 2. Pass to V/U with $U = \operatorname{span}(\mathbf{b}_1)$. Do the quotient operators T_i/U still commute?
Hint 3. Lift a basis that triangularizes all T_i/U, and put $\mathbf{b}_1$ in front.
Sketch. Induct on n. Take a common eigenvector $\mathbf{b}_1$ (Theorem 7.6.31); $U = \operatorname{span}(\mathbf{b}_1)$ is invariant under every T_i. The quotient operators commute, since $(T_i/U)\big(T_j/U\big)(\mathbf{v}+U) = T_iT_j\mathbf{v}+U$. By induction, V/U has a basis $\mathbf{b}_2+U, \dots, \mathbf{b}_n+U$ triangularizing all of them. Then $\mathbf{b}_1, \dots, \mathbf{b}_n$ is a basis of V, and $T_i\mathbf{b}_j \in \operatorname{span}\big(\mathbf{b}_1, \dots, \mathbf{b}_j\big)$ because this holds modulo $U = \operatorname{span}(\mathbf{b}_1)$.

We induct on $n = \dim V$. For $n = 1$ every basis works.

Let $n \geq 2$ and assume the theorem in dimension $n-1$. By Theorem 7.6.31 there is a common eigenvector $\mathbf{b}_1$ of $T_1, \dots, T_k$, and $U = \operatorname{span}(\mathbf{b}_1)$ is invariant under every T_i (Proposition 7.2.7). Let $\overline{T}_i = T_i/U \in \mathrm{L}(V/U)$ be the quotient operators (Proposition 4.6.32(a)). They commute: for $\mathbf{v} \in V$,

$$\overline{T}_i\overline{T}_j(\mathbf{v}+U) = \overline{T}_i\big(T_j\mathbf{v}+U\big) = T_iT_j\mathbf{v}+U = T_jT_i\mathbf{v}+U = \overline{T}_j\overline{T}_i(\mathbf{v}+U).$$

Since $\dim V/U = n-1$ (Theorem 4.6.27), the induction hypothesis and Proposition 7.5.2 give $\mathbf{b}_2, \dots, \mathbf{b}_n \in V$ such that $\mathbf{b}_2+U, \dots, \mathbf{b}_n+U$ is a basis of V/U and, for all i and all $j \geq 2$,

$$\overline{T}_i\big(\mathbf{b}_j+U\big) \in \operatorname{span}\big(\mathbf{b}_2+U, \dots, \mathbf{b}_j+U\big).$$

By Proposition 4.6.28, $\mathcal{B} = \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ is a basis of V. Fix i and $j \geq 2$, and write $\overline{T}_i\big(\mathbf{b}_j+U\big) = \sum_{l=2}^{j} c_l(\mathbf{b}_l + U) = \Big(\sum_{l=2}^{j} c_l\mathbf{b}_l\Big) + U$. Since also $\overline{T}_i\big(\mathbf{b}_j+U\big) = T_i\mathbf{b}_j+U$, the difference $T_i\mathbf{b}_j - \sum_{l=2}^{j} c_l\mathbf{b}_l$ lies in $U = \operatorname{span}(\mathbf{b}_1)$, so $T_i\mathbf{b}_j \in \operatorname{span}\big(\mathbf{b}_1, \dots, \mathbf{b}_j\big)$. Also $T_i\mathbf{b}_1 \in \operatorname{span}(\mathbf{b}_1)$. By Proposition 7.5.2, $[T_i]_{\mathcal{B}}$ is upper triangular for every i. □

With an inner product, the basis can be taken orthonormal: all the T_i preserve the flag of $\mathcal{B}$, and Gram–Schmidt produces an orthonormal basis with the same flag, exactly as in the proof of Theorem 7.5.18 (Technique 7.5.4). So commuting complex matrices $A_1, \dots, A_k$ can be written $A_i = UR_iU^*$ with one unitary U and upper-triangular R_i. The converse fails: upper-triangular matrices need not commute (Exercise 7.6.51).

Corollary 7.6.33 (Eigenvalues of sums and products of commuting operators). Let V be a finite-dimensional complex vector space with $\dim V = n \geq 1$, and let $S, T \in \mathrm{L}(V)$ commute. Then the eigenvalues of S and of T, counted with multiplicity, can be listed as $\lambda_1, \dots, \lambda_n$ and $\mu_1, \dots, \mu_n$ so that the eigenvalues of $S+T$ and of ST, counted with multiplicity, are $\lambda_1+\mu_1, \dots, \lambda_n+\mu_n$ and $\lambda_1\mu_1, \dots, \lambda_n\mu_n$. In particular, every eigenvalue of $S+T$ is $\lambda+\mu$ for some eigenvalue λ of S and some eigenvalue μ of T.

Proof. ROUTINE

Uses. Theorem 7.6.32 (Simultaneous triangularization), Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.12 (Matrices of sums and scalar multiples)
Hint. Triangularize S and T in one basis. What are the diagonals of $[S+T]$ and $[ST]$?

Sketch. In a common triangularizing basis, $[S+T] = [S] + [T]$ and $[ST] = [S][T]$ are upper triangular, with diagonal entries $\lambda_j + \mu_j$ and $\lambda_j \mu_j$ (as in the proof of Theorem 7.6.28). Read the eigenvalues off the diagonals with Proposition 7.5.9.

By Theorem 7.6.32, there is a basis $\mathcal{B}$ in which $[S]_{\mathcal{B}}$ and $[T]_{\mathcal{B}}$ are both upper triangular; let their diagonal entries be $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$. By Proposition 7.5.9, these are the eigenvalues of S and of T counted with multiplicity. Now $[S+T]_{\mathcal{B}} = [S]_{\mathcal{B}} + [T]_{\mathcal{B}}$ (Proposition 2.4.12) and $[ST]_{\mathcal{B}} = [S]_{\mathcal{B}}[T]_{\mathcal{B}}$ (Theorem 2.4.2) are upper triangular, with diagonal entries $\lambda_j + \mu_j$ and $\lambda_j \mu_j$ respectively (see the proof of Theorem 7.6.28). Applying Proposition 7.5.9 to $S+T$ and to ST gives the claim. □

Without commuting, nothing of the kind holds: $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and N^{T} both have 0 as their only eigenvalue, while $N + N^{\mathsf{T}} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has eigenvalues ± 1 and $NN^{\mathsf{T}} = \operatorname{diag}(1, 0)$ has the eigenvalue 1.

For diagonalization, commuting is also necessary, since diagonal matrices commute.

Theorem 7.6.34 (Simultaneous diagonalization). Let V be a finite-dimensional vector space over any field F, with $\dim V \geq 1$, and let $T_1, \ldots, T_k \in \mathrm{L}(V)$ be diagonalizable. Then V has a basis consisting of vectors that are eigenvectors of every T_i (equivalently, a basis $\mathcal{B}$ in which every $[T_i]_{\mathcal{B}}$ is diagonal) if and only if $T_i T_j = T_j T_i$ for all i, j.

Proof. STANDARD

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.12 (Matrices of sums and scalar multiples), Corollary 7.6.16 (Restrictions of diagonalizable operators), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Corollary 7.2.23 (Eigenspaces are independent), Theorem 4.4.13 (Bases and dimension of a sum)

Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Glue bases of the summands (Technique 4.4.14)

Hint 1. For "only if", compute $\left[T_i T_j\right]_{\mathcal{B}}$ and $\left[T_j T_i\right]_{\mathcal{B}}$.

Hint 2. For "if", induct on k, proving: every nonzero subspace W invariant under all T_i has a basis of common eigenvectors.

Hint 3. Split W into the eigenspaces of $T_k|_W$. Why is each piece invariant under $T_1, \ldots, T_{k-1}$, and why is $T_k|_W$ diagonalizable?

Sketch. If every $[T_i]_{\mathcal{B}}$ is diagonal, the matrices commute, hence so do the operators. Conversely, induct on k with the stronger statement for invariant subspaces W. $T_k|_W$ is diagonalizable (Corollary 7.6.16), so W is the direct sum of the pieces $W \cap E_\mu(T_k)$. Each piece is invariant under $T_1, \ldots, T_{k-1}$ because they commute with T_k, so by induction it has a basis of common eigenvectors of $T_1, \ldots, T_{k-1}$, which are also eigenvectors of T_k. Concatenate.

Only if. Let $\mathcal{B}$ be a basis with every $[T_i]_{\mathcal{B}}$ diagonal (Theorem 7.4.2). Diagonal matrices commute, so by Theorem 2.4.2, $\left[T_i T_j\right]_{\mathcal{B}} = [T_i]_{\mathcal{B}}\left[T_j\right]_{\mathcal{B}} = \left[T_j\right]_{\mathcal{B}}[T_i]_{\mathcal{B}} = \left[T_j T_i\right]_{\mathcal{B}}$. Since $S \mapsto [S]_{\mathcal{B}}$ is injective (Proposition 2.4.12), $T_i T_j = T_j T_i$.

If. We show by induction on k: every nonzero subspace W that is invariant under all of $T_1, \ldots, T_k$ has a basis of common eigenvectors of $T_1, \ldots, T_k$. The theorem is the case $W = V$.

For $k = 1$, $T_1|_W$ is diagonalizable by Corollary 7.6.16, and its eigenvectors are eigenvectors of T_1.

Let $k \geq 2$. By Corollary 7.6.16, $T_k|_W$ is diagonalizable. Let $\mu_1, \ldots, \mu_r$ be its distinct eigenvalues and $W_l = W \cap E_{\mu_l}(T_k)$, the eigenspaces of $T_k|_W$. Grouping an eigenbasis of $T_k|_W$ by eigenvalue shows $W = W_1 + \cdots + W_r$, and the sum is direct by Corollary 7.2.23. For $i < k$, T_i commutes with T_k, so $E_{\mu_l}(T_k)$ is T_i-invariant (Proposition 7.2.30(b)), and so is W; hence each W_l is invariant under $T_1, \ldots, T_{k-1}$. Each W_l is nonzero, so by the induction hypothesis it has a basis of common eigenvectors of $T_1, \ldots, T_{k-1}$, and these lie in $E_{\mu_l}(T_k)$, so they are eigenvectors of T_k as well. Concatenating these bases gives a basis of W (Theorem 4.4.13) of common eigenvectors of $T_1, \ldots, T_k$. □

This is the statement of Exercise 7.4.51; the minimal polynomial supplies the one step that needed work there, that a diagonalizable operator stays diagonalizable on an invariant subspace. The theorem extends to an arbitrary set Ω of commuting diagonalizable operators: $\operatorname{span}(\Omega)$ is a subspace of the finite-dimensional space $\mathrm{L}(V)$, so it is spanned by finitely many elements $T_1, \ldots, T_k$ of Ω (keep adding elements of Ω outside

the span of those chosen; the list stays independent, so this stops). A basis that diagonalizes $T_1, \dots, T_k$ diagonalizes every linear combination of them, hence all of Ω.

Example 7.6.35 (A common eigenbasis). Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{in } \mathbb{Q}^{3\times 3}.$$

Both are block diagonal with blocks of sizes 2 and 1, and A is a scalar on each block, so $AB = BA$. Both are diagonalizable: A is diagonal, and B has the distinct eigenvalues $1, -1, 5$. The eigenspaces of A are $E_1(A) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ and $E_2(A) = \operatorname{span}(\mathbf{e}_3)$, and B preserves both. On $E_1(A)$, B swaps $\mathbf{e}_1$ and $\mathbf{e}_2$, with eigenvectors $\mathbf{e}_1 + \mathbf{e}_2$ (eigenvalue 1) and $\mathbf{e}_1 - \mathbf{e}_2$ (eigenvalue -1). So in the basis $(1,1,0), (1,-1,0), (0,0,1)$,

$$A \mapsto \operatorname{diag}(1,1,2), \quad B \mapsto \operatorname{diag}(1,-1,5).$$

Not every eigenbasis of A works: $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is one, but $\mathbf{e}_1$ is not an eigenvector of B. The common basis has to be chosen inside the eigenspaces of A, using B, which is how the proof proceeds.

7.6.7 Gershgorin's disk theorem

The eigenvalues of $A \in \mathbb{C}^{n\times n}$ are the roots of a polynomial of degree n, and for $n \geq 5$ there is in general no formula for them. Often a location is enough: is 0 an eigenvalue, that is, is A singular? Do all eigenvalues have negative real part, so that the solutions of $\mathbf{x}' = A\mathbf{x}$ decay (Proposition 7.4.34)? If the off-diagonal entries of A were 0, the eigenvalues would be the diagonal entries. Gershgorin's theorem says that off-diagonal entries can move the eigenvalues away from the diagonal entries, but only as far as the size of the off-diagonal entries in a row allows.

Theorem 7.6.36 (Gershgorin's disk theorem). Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$. For $i = 1, \dots, n$, let $R_i = \sum_{j\neq i} |a_{ij}|$ and

$$D_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i\},$$

the closed disk with center a_{ii} and radius R_i. Then every eigenvalue of A lies in $D_1 \cup D_2 \cup \cdots \cup D_n$. More precisely, if $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and i is an index with $|v_i| = \max_j |v_j|$, then $\lambda \in D_i$.

Technique 7.6.37 (Look at the largest coordinate). To bound an eigenvalue λ with $A\mathbf{v} = \lambda\mathbf{v}$, look at the row of this equation belonging to a coordinate v_i of largest absolute value. After dividing by v_i, every other coordinate enters only through the ratio $|v_j/v_i| \leq 1$, so the row becomes an inequality involving the entries of A alone.

Reach for it when you need a bound on eigenvalues in terms of entries, or to show that a matrix with a dominant diagonal, or with nonnegative rows summing to 1, is invertible or has its eigenvalues in a given region. The canonical example is Theorem 7.6.36. Later: $|\lambda| \leq 1$ for stochastic matrices (Section 11.1) and the Perron–Frobenius theorem (Section 11.2).

Used in. Theorem 7.6.36, Exercise 7.6.54, Exercise 7.6.55.

Proof. STANDARD

Uses. Definition 2.3.8 (Matrix–vector product), Theorem B.1.11 (Triangle inequality in $\mathbb{C}$), Proposition B.1.8 (Properties of the modulus)
Techniques. Look at the largest coordinate (Technique 7.6.37)
Hint 1. Write out row i of $A\mathbf{v} = \lambda\mathbf{v}$ and move the diagonal term to the right.
Hint 2. Take absolute values, use the triangle inequality, and bound each $|v_j|$ by $|v_i|$.
Sketch. Row i of $A\mathbf{v} = \lambda\mathbf{v}$ gives $(\lambda - a_{ii})v_i = \sum_{j\neq i} a_{ij}v_j$. Taking absolute values and using $|v_j| \leq |v_i|$ gives $|\lambda - a_{ii}||v_i| \leq R_i|v_i|$, and $|v_i| > 0$ since $\mathbf{v} \neq \mathbf{0}$.

Let $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and choose i with $|v_i| = \max_j |v_j|$; then $|v_i| > 0$. Row i of $A\mathbf{v} = \lambda\mathbf{v}$ reads $\sum_j a_{ij}v_j = \lambda v_i$ (Definition 2.3.8), that is,

$$(\lambda - a_{ii})v_i = \sum_{j \neq i} a_{ij}v_j.$$

By Proposition B.1.8(d) and the triangle inequality (Theorem B.1.11),

$$|\lambda - a_{ii}||v_i| \leq \sum_{j \neq i} |a_{ij}||v_j| \leq \sum_{j \neq i} |a_{ij}||v_i| = R_i|v_i|.$$

Dividing by $|v_i| > 0$ gives $|\lambda - a_{ii}| \leq R_i$, that is, $\lambda \in D_i$. □

Example 7.6.38 (Three disks). Let

$$A = \begin{pmatrix} 3 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & -4 \end{pmatrix}.$$

The disks are D_1 with center 3 and radius 2, D_2 with center 0 and radius 2, and D_3 with center -4 and radius 1. Expanding $\det(xI_3 - A)$ along the third column gives $p_A = (x+4)\big(x^2 - 3x + 4\big)$, so the eigenvalues are -4 and $\big(3 \pm i\sqrt{7}\big)/2$. The complex pair satisfies $|\lambda|^2 = (9+7)/4 = 4$ and $|\lambda - 3|^2 = 9/4 + 7/4 = 4$: each lies on the boundary of both D_1 and D_2, at the two points where the circles cross. So the disks cannot be shrunk in general. Here Gershgorin cannot decide whether A is invertible, because $0 \in D_2$; in fact $\det A = -16$.

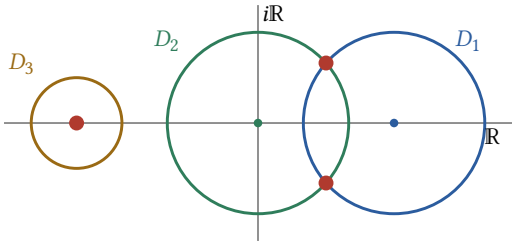


Figure 7.6.1. The Gershgorin disks of the matrix in Example 7.6.38 and its three eigenvalues (red). The eigenvalue -4 is the center of D_3; the other two sit where the circles bounding D_1 and D_2 meet.

Corollary 7.6.39 (Diagonally dominant matrices are invertible). Let $A \in \mathbb{C}^{n\times n}$ be **strictly diagonally dominant**: $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ for every i. Then A is invertible. The same holds for real or rational matrices.

Proof. ROUTINE

Uses. Theorem 7.6.36 (Gershgorin's disk theorem), Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 5.4.7 (Invertibility and the determinant)
Hint. Is 0 in any Gershgorin disk?
Sketch. $|0 - a_{ii}| > R_i$, so $0 \notin D_i$ for every i, and 0 is not an eigenvalue. A real matrix that is invertible over $\mathbb{C}$ has nonzero determinant, so it is invertible over $\mathbb{R}$.

For every i, $|0 - a_{ii}| = |a_{ii}| > R_i$, so $0 \notin D_i$. By Theorem 7.6.36, 0 is not an eigenvalue of A, so A is invertible (Theorem 7.2.9). If A has real (or rational) entries, then $\det A \neq 0$ by Theorem 5.4.7, and the determinant does not depend on whether A is regarded as real or complex, so A is invertible over $\mathbb{R}$ (or $\mathbb{Q}$) as well. □

Since A^{T} has the same characteristic polynomial as A (Exercise 7.3.17), the theorem also holds with columns in place of rows, and every eigenvalue lies in the intersection of the union of the row disks with the union of the column disks.

ASIDE: *Counting eigenvalues in the disks*

A sharper form of the theorem says: if the union of some k of the disks is disjoint from the union of the other $n-k$, then it contains exactly k eigenvalues, counted with multiplicity. The proof follows the eigenvalues of $A_t = \operatorname{diag}(a_{11},\dots,a_{nn}) + t(A - \operatorname{diag}(a_{11},\dots,a_{nn}))$ as t runs from 0 to 1. The disks of A_t have radii tR_i, so they only grow, and the eigenvalues, being roots of a polynomial whose coefficients depend continuously on t, move continuously and cannot jump from one group of disks to the other. Making "the roots move continuously" precise takes some analysis, which we do not do here. For Example 7.6.38 it says that D_3 contains exactly one eigenvalue. For a real matrix, an isolated disk containing exactly one eigenvalue λ also contains $\overline{\lambda}$ (the disk is symmetric about $\mathbb{R}$, and $\overline{\lambda}$ is an eigenvalue by Proposition 7.1.38), so λ is real.

Exercises

Exercise 7.6.40. ROUTINE Find the characteristic and minimal polynomials of each matrix in $\mathbb{Q}^{3\times 3}$, and decide whether it is diagonalizable.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Solution on page 1606

Exercise 7.6.41. STANDARD

(a) Let $n \geq 2$ and let $J \in F^{n\times n}$ be the matrix with every entry 1. Show that $J^2 = nJ$, where n means $1 + \cdots + 1$ (n terms) in F. Find m_J, and decide when J is diagonalizable. (Distinguish whether $n = 0$ in F.)

(b) Let $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Find m_R over $\mathbb{F}_2$, $\mathbb{F}_3$, and $\mathbb{F}_5$, and decide in each case whether R is diagonalizable over that field. When it is, find an eigenbasis.

Solution on page 1606

Exercise 7.6.42. STANDARD Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$.

(a) Find p_A and m_A.

(b) Show that $A^k = (k-1)A^2 + (2-k)A$ for every $k \geq 1$.

(c) Let $B = A + I_3$. Use (a) and Exercise 7.6.6 to find m_B, and write B^{-1} as a polynomial in B. Compute B^{-1}.

Solution on page 1607

Exercise 7.6.43. STANDARD Let $A, B \in F^{n\times n}$. Prove or disprove each statement.

(a) If $p_A = p_B$, then $m_A = m_B$.

(b) If $p_A = p_B$ and $m_A = m_B$, then A and B are similar.

(c) If A^2 is diagonalizable, then A is diagonalizable.

(d) If $F = \mathbb{C}$, A is invertible, and A^2 is diagonalizable, then A is diagonalizable.

Solution on page 1607

Exercise 7.6.44. STANDARD Find the error in the following argument.

Claim. For every $A \in \mathbb{C}^{n\times n}$, $p_A(A) = 0$.

"Proof." Let λ be an eigenvalue of A and $\mathbf{v}$ an eigenvector for λ. By Proposition 7.2.28, $p_A(A)\mathbf{v} = p_A(\lambda)\mathbf{v} = \mathbf{0}$, since λ is a root of p_A. So $p_A(A)$ sends every eigenvector of A to $\mathbf{0}$. A linear map that sends every vector of a spanning set to $\mathbf{0}$ is zero, so $p_A(A) = 0$.

Which matrices does the argument handle correctly? Explain why the fourth proof of Theorem 7.6.20 can be read as a repair of it.

Solution on page 1608

Exercise 7.6.45. STANDARD Let $T \in \mathrm{L}(V)$ satisfy $T^2 = T + 2\mathrm{id}_V$.

(a) Suppose $1 + 1 + 1 \neq 0$ in F. Show that T is diagonalizable with eigenvalues in $\{2, -1\}$, so that $V = E_2 \oplus E_{-1}$, and write the projection of V onto E_2 along E_{-1} as a polynomial in T.

(b) Find $A \in \mathbb{F}_3^{2\times 2}$ with $A^2 = A + 2I_2$ that is not diagonalizable.

Solution on page 1608

Exercise 7.6.46. STANDARD

(a) Find all $A \in \mathbb{F}_2^{2\times 2}$ with $A^2 + A + I_2 = 0$.

(b) For one such A, show that the four matrices $0, I_2, A, A + I_2$ are closed under addition and multiplication, and that each nonzero one has an inverse among them.

Solution on page 1609

Exercise 7.6.47. HARD A vector $\mathbf{v}$ is **cyclic** for T if $K_T(\mathbf{v}) = V$ (Definition 7.2.33).

(a) Show that if T has a cyclic vector, then $m_T = p_T$.

(b) Suppose T is diagonalizable. Show that T has a cyclic vector if and only if T has n distinct eigenvalues.

Solution on page 1609

Exercise 7.6.48. STANDARD Let $\mathbf{v}$ be a cyclic vector for T (as in Exercise 7.6.47). Show that $S \in \mathrm{L}(V)$ commutes with T if and only if $S = q(T)$ for some $q \in F[x]$. Deduce that the operators commuting with T form a subspace of $\mathrm{L}(V)$ of dimension n.

Solution on page 1609

Exercise 7.6.49. HARD Let V be a finite-dimensional complex vector space and $T \in \mathrm{L}(V)$. Show that T is diagonalizable if and only if every T-invariant subspace U of V has a T-invariant complement, that is, a T-invariant W with $V = U \oplus W$. Does "$\Longleftarrow$" hold over $\mathbb{R}$?

Solution on page 1610

Exercise 7.6.50. HARD Let $A \in \mathbb{R}^{n\times n}$ satisfy $A^3 = A + I_n$. Show that $\det A > 0$.

Solution on page 1610

Exercise 7.6.51. STANDARD Let $A, B \in \mathbb{C}^{n\times n}$.

(a) Suppose $AB = BA$ and $A^k = B^k = 0$ for some k. Show that $A + B$ and AB are nilpotent.

(b) Show by example that (a) fails without $AB = BA$.

(c) Suppose A and B are both upper triangular. Show that $AB - BA$ is nilpotent. Deduce that $AB - BA$ is nilpotent whenever A and B can be triangularized by the same basis, and give two upper-triangular matrices that do not commute.

Solution on page 1611

Exercise 7.6.52. STANDARD Let $A, B \in \mathbb{R}^{n\times n}$ with $AB = BA$. Show that $\det(A^2 + B^2) \geq 0$. Find $A, B \in \mathbb{R}^{2\times 2}$ with $\det(A^2 + B^2) < 0$.

Solution on page 1611

Exercise 7.6.53. STANDARD Let

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Check that $AB = BA$ and that A and B are diagonalizable over $\mathbb{R}$.

(b) Find a basis of $\mathbb{R}^3$ consisting of common eigenvectors of A and B.

(c) Is there such a basis of $\mathbb{Q}^3$?

Solution on page 1612

Exercise 7.6.54. STANDARD Let $G = \begin{pmatrix} 6 & 1 & -1 \\ 1 & 5 & 2 \\ 0 & -1 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(a) Show that every complex eigenvalue λ of G satisfies $\operatorname{Re}\lambda \geq 2$. Deduce that G is invertible.

(b) Show that $\det G > 0$ without computing it.

(c) Show that $|\lambda| \leq 7$ for every eigenvalue, and explain why the row disks alone give only $|\lambda| \leq 8$.

Solution on page 1612

Exercise 7.6.55. HARD Let $A \in \mathbb{R}^{n\times n}$ have nonnegative entries, with every row summing to 1.

(a) Show that every complex eigenvalue λ of A satisfies $|\lambda| \leq 1$, and that 1 is an eigenvalue.

(b) Suppose moreover that $a_{ii} > 0$ for every i. Show that 1 is the only eigenvalue with $|\lambda| = 1$.

(c) Show that (b) fails without the assumption $a_{ii} > 0$.

Solution on page 1613

Exercise 7.6.56. CHALLENGE Let $A, B \in \mathbb{C}^{n\times n}$, and suppose $C = AB - BA$ commutes with A. Show that C is nilpotent. Show that the statement fails over $\mathbb{F}_p$.

Solution on page 1613

7.7 Proof patterns

The chapter asked one question about an operator T on a finite-dimensional space: in which basis is its matrix as simple as possible, and how simple can it be made? Diagonal when there is an eigenbasis, triangular when the characteristic polynomial splits, and in every case the answer is controlled by polyno-

mials in T. Most of the chapter's proofs use a few moves to turn a hypothesis about T into information about vectors: apply a polynomial in T to kill terms, follow a vector along its chain $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$, look for invariant subspaces among kernels of operators that commute with T, test an identity on eigenvectors, or triangularize and read off the diagonal. This section sorts the chapter's arguments by the question they answer, solves three problems in several ways to compare the moves, and lists the errors that the chapter's exercises warn against. The exercises at the end practice one move at a time, and the problem set (Problems for Chapter 7) mixes them with earlier chapters.

7.7.1 Polynomials in T kill terms

On an eigenvector, every polynomial in T acts as a scalar: $p(T)\mathbf{v} = p(\lambda)\mathbf{v}$ (Proposition 7.2.28). So a polynomial that vanishes at some eigenvalues and not at others removes exactly the terms it vanishes on, and this is the most common move of the chapter.

- *Kill all terms but one* (Technique 4.1.25). To show that eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ for distinct eigenvalues are independent, apply $\prod_{j \neq i}(T - \lambda_j \operatorname{id})$ to a relation; only the ith term survives, multiplied by the nonzero scalar $\prod_{j \neq i}(\lambda_i - \lambda_j)$ (Theorem 7.2.21). With the normalized product, the Lagrange polynomial L_i, the surviving term comes out with coefficient 1, and $L_i(T)$ is the projection onto E_{λ_i} (Technique 7.6.13).
- *Read eigenvalues off an annihilating polynomial* (Technique 7.6.17). If $p(T) = 0$, every eigenvalue is a root of p (Proposition 7.2.28). If p is a product of distinct linear factors, T is diagonalizable (Proposition 7.6.12). If $p(0) \neq 0$, then T is invertible and T^{-1} is a polynomial in T, read off from p (Corollary 7.6.9). Note the direction: a root of p need not be an eigenvalue, since $T = 0$ on $V \neq \{\mathbf{0}\}$ is killed by $x(x-1)$ and 1 is not an eigenvalue of it.
- *Bézout in T.* If $up + vq = 1$ in $F[x]$, then $u(T)p(T) + v(T)q(T) = \mathrm{id}_V$, so a vector killed by both $p(T)$ and $q(T)$ is $\mathbf{0}$, and when $p(T)q(T) = 0$ the space splits as $\ker p(T) \oplus \ker q(T)$ (Corollary 7.1.48, Exercise 7.2.55). Lagrange interpolation is the case of linear factors.

Behind all three is the fact that polynomials in T commute with each other and with T (Proposition 7.2.27), so they can be applied to a relation in any order.

7.7.2 Chains

The chain move of Section 4.8.1 returns with T replaced by $N = T - \lambda \operatorname{id}$. If $N^k\mathbf{v} \neq \mathbf{0}$ and $N^{k+1}\mathbf{v} = \mathbf{0}$, then $\mathbf{v}, N\mathbf{v}, \ldots, N^k\mathbf{v}$ is independent (Proposition 7.2.37), its span is the Krylov subspace $K_T(\mathbf{v})$, the smallest T-invariant subspace containing $\mathbf{v}$, and the last nonzero vector $N^k\mathbf{v}$ is an eigenvector with eigenvalue λ (Corollary 7.2.38). Three things make this the most useful single fact of the chapter.

- *It turns "killed by a power" into "an eigenvector exists."* Whenever some $(T - \lambda \operatorname{id})^k$ kills a nonzero $\mathbf{v}$, walk along the chain until the next step gives $\mathbf{0}$; the last vector is an eigenvector. This is how Technique 7.3.28 produces eigenvectors: the first dependence among $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots$ is a polynomial p with $p(T)\mathbf{v} = \mathbf{0}$ (Proposition 7.2.35), and peeling off one factor $x - \lambda$ of p leaves a nonzero vector killed by $T - \lambda \operatorname{id}$. Over $\mathbb{C}$, p always has such a factor, which is the determinant-free proof that eigenvalues exist (Theorem 7.3.27).
- *It counts.* A chain is independent, so it has at most $\dim V$ vectors, and a chain of length $\dim V$ is a basis in which T is λ on the diagonal and 1 just above it (Corollary 7.2.38(c)). Combined with the kernel count $\dim \ker N^{j+1} \leq \dim \ker N^j + \dim \ker N$, it decides when one chain fills the space (Problem 7.3).
- *It works one vector at a time*, with no hypothesis on T beyond the two conditions on $\mathbf{v}$, and in any dimension and over any field. That is why it decides problems where nothing is known about T globally, such as whether a matrix has a square root (Exercise 7.7.9).

Example 7.7.1 (Two eigenvalues cannot share a generalized eigenvector, three ways). Let $T \in \mathrm{L}(V)$, let $\lambda \neq \mu$ in F, let $k, m \geq 1$, and let $\mathbf{v} \in V$ satisfy

$$(T-\lambda\,\mathrm{id})^k\mathbf{v}=\mathbf{0}\quad\text{and}\quad(T-\mu\,\mathrm{id})^m\mathbf{v}=\mathbf{0}.$$

We show $\mathbf{v}=\mathbf{0}$. Here V is any vector space over any field. Write $A=T-\lambda\,\mathrm{id}$ and $B=T-\mu\,\mathrm{id}$; these commute, and $B-A=(\lambda-\mu)\,\mathrm{id}$.

Kill every term. By the binomial theorem for commuting operators (Exercise A.2.8), with $r=k+m-1$,

$$(\lambda-\mu)^r\mathbf{v}=(B-A)^r\mathbf{v}=\sum_{j=0}^{r}\binom{r}{j}(-1)^{r-j}A^{r-j}B^j\mathbf{v}.$$

In each term, either $j\geq m$, and then $B^j\mathbf{v}=\mathbf{0}$, or $j\leq m-1$, and then $r-j\geq k$ and $A^{r-j}B^j\mathbf{v}=B^jA^{r-j}\mathbf{v}=\mathbf{0}$. So $(\lambda-\mu)^r\mathbf{v}=\mathbf{0}$, and since $\lambda-\mu\neq 0$, $\mathbf{v}=\mathbf{0}$.

A chain. Suppose $\mathbf{v}\neq\mathbf{0}$. Since $A^k\mathbf{v}=\mathbf{0}$, there is a largest j with $0\leq j<k$ and $A^j\mathbf{v}\neq\mathbf{0}$. Then $\mathbf{w}=A^j\mathbf{v}$ satisfies $A\mathbf{w}=\mathbf{0}$, so it is an eigenvector of T with eigenvalue λ (Corollary 7.2.38(a)). By Proposition 7.2.28, $B^m\mathbf{w}=(\lambda-\mu)^m\mathbf{w}\neq\mathbf{0}$. But $B^m\mathbf{w}=A^jB^m\mathbf{v}=\mathbf{0}$, a contradiction.

Restrict, then invert. Let $U=\ker A^k\cap\ker B^m$. Both kernels are T-invariant (Proposition 7.2.30(a)), so U is, and on U the restriction $N=A|_U$ satisfies $N^k=0$. Now $B|_U=N+(\lambda-\mu)\mathrm{id}_U=(\lambda-\mu)(\mathrm{id}_U+cN)$ with $c=(\lambda-\mu)^{-1}$, and id_U+cN is invertible, with inverse $\mathrm{id}_U-cN+c^2N^2-\cdots+(-c)^{k-1}N^{k-1}$ (the computation of Exercise 4.1.44). So $B|_U$ is invertible, and so is its mth power. But ${B|_U}^m=0$. An invertible zero operator lives only on $\{\mathbf{0}\}$, so $U=\{\mathbf{0}\}$ and $\mathbf{v}=\mathbf{0}$.

The first proof is Bézout's identity for $(x-\lambda)^k$ and $(x-\mu)^m$, written out by hand, and it is the one that generalizes: for any coprime p,q, a vector killed by $p(T)$ and $q(T)$ is $\mathbf{0}$. The chain proof reduces the question to eigenvectors, where polynomials act as scalars, and it is the shortest once the chain is in your hands. The third proof sees the hypothesis as a statement about the operator on U: there T is "λ plus nilpotent," and a nilpotent perturbation of the nonzero scalar $\lambda-\mu$ cannot be 0. All three are the key step in splitting V into generalized eigenspaces (Section 10.2).

Practice: Exercise 7.7.9, Problem 7.3, Problem 7.12, and Problem 7.9.

7.7.3 Invariant subspaces from commuting operators

To find subspaces that T maps into themselves, look at kernels, images, and eigenspaces of operators that commute with T (Technique 7.2.31, Proposition 7.2.30). Polynomials in T always qualify, so $\ker(T-\lambda\,\mathrm{id})^k$ and $\mathrm{im}\,p(T)$ are always available, and Problem 7.4 finds every invariant subspace of a 3×3 matrix this way. When two given operators S,T commute, each eigenspace of S is a smaller space on which T still acts, and the problem can be continued there. This is the whole proof that commuting operators have a common eigenvector (Theorem 7.6.31): T has an eigenvalue on $E_\lambda(S)$ when the field is $\mathbb{C}$. Two cautions. The restriction of T to $E_\lambda(S)$ inherits diagonalizability (Corollary 7.6.16), but nothing forces it to be diagonalizable when T is not, and the eigenspaces of S need not fill V; so commuting alone gives a common eigenvector and a common triangular form (Theorem 7.6.32), and a common eigenbasis only for diagonalizable operators (Theorem 7.6.34, Exercise 7.7.6). The kernel move does not need eigenvalues at all, which makes it work over any field: two commuting nilpotent operators have a common kernel vector, because $\ker S$ is T-invariant and T is still nilpotent there (Exercise 7.7.13).

Practice: Exercise 7.7.6, Exercise 7.7.13, Exercise 7.7.12, and Problem 7.4.

7.7.4 Proving and refuting diagonalizability

There are four standard ways to show that T is diagonalizable, in roughly increasing order of generality:

- n distinct eigenvalues (Corollary 7.4.5);
- an annihilating polynomial that is a product of distinct linear factors (Theorem 7.6.15), which is the route for any identity such as $T^2=T$, $T^3=T$, or $T^k=\mathrm{id}$ over $\mathbb{C}$;

- a direct sum decomposition of V into eigenspaces, often produced by idempotents (Technique 4.4.28) or by Lagrange polynomials in T;
- counting: p_T splits and $\operatorname{gm}(\lambda) = \operatorname{am}(\lambda)$ for each eigenvalue (Theorem 7.4.16), with $\operatorname{gm}(\lambda) = n - \operatorname{rank}(A - \lambda I_n)$ computed as a rank (Lemma 7.4.13).

To refute it, find one of the obstructions of Obstructions to diagonalizability: p_T does not split over F (the rotation, Example 7.4.9), or some eigenvalue has gm < am (the shear, Example 7.4.8). Two shortcuts catch most examples. A diagonalizable operator with a single eigenvalue λ is λ id (Proposition 7.4.7), so a nonzero nilpotent operator, or $I + N$ with $N \neq 0$ nilpotent, is never diagonalizable. And over a field where the roots of the relevant polynomial collide, the polynomial route fails: $x^2 - 1 = (x-1)^2$ over $\mathbb{F}_2$, and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ satisfies $A^2 = I_2$ there without being diagonalizable.

Example 7.7.2 (An idempotent is diagonalizable, three ways). Let $\dim V = n$ and let $P \in \mathrm{L}(V)$ satisfy $P^2 = P$. We show that P is diagonalizable, with P similar to $\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$, where the number of ones is $\operatorname{rank} P$.

An annihilating polynomial. The polynomial $x^2 - x = x(x-1)$ kills P, and its factors are distinct because $0 \neq 1$ in every field. By Proposition 7.6.12, $V = E_0 \oplus E_1$ and P is diagonalizable.

Image and kernel. By Theorem 4.4.27, $V = \operatorname{im} P \oplus \ker P$ and $\operatorname{im} P = \{\mathbf{v} \mid P\mathbf{v} = \mathbf{v}\}$. So $\operatorname{im} P = E_1$ and $\ker P = E_0$, and a basis of $\operatorname{im} P$ followed by a basis of $\ker P$ is a basis of V (Theorem 4.4.13) consisting of eigenvectors. There are $\operatorname{rank} P$ vectors with eigenvalue 1.

Eigenvalues and rank. Every eigenvalue is a root of $x^2 - x$ (Proposition 7.2.28), so it is 0 or 1. By Lemma 7.4.13, $\operatorname{gm}(0) = n - \operatorname{rank} P$ and $\operatorname{gm}(1) = n - \operatorname{rank}(P - \mathrm{id})$. Since $P + (\mathrm{id} - P) = \mathrm{id}$, subadditivity gives $\operatorname{rank} P + \operatorname{rank}(\mathrm{id} - P) \geq n$ (Proposition 4.3.20); since $P(\mathrm{id} - P) = 0$, Sylvester's inequality gives $0 \geq \operatorname{rank} P + \operatorname{rank}(\mathrm{id} - P) - n$ (Exercise 4.3.30). So $\operatorname{gm}(0) + \operatorname{gm}(1) = 2n - (\operatorname{rank} P + \operatorname{rank}(\mathrm{id} - P)) = n$, and Theorem 7.4.16(III) applies.

The polynomial proof is the shortest and generalizes at once: if $(T - a\,\mathrm{id})(T - b\,\mathrm{id}) = 0$ with $a \neq b$, then T is diagonalizable. It is also the one that breaks first when the field changes, as the involution over $\mathbb{F}_2$ above shows. The image-and-kernel proof produces the eigenspaces themselves, and it explains the polynomial proof: for $(T - a\,\mathrm{id})(T - b\,\mathrm{id}) = 0$, the Lagrange polynomial $(T - b\,\mathrm{id})/(a - b)$ is an idempotent with image E_a. The rank proof is the computational one, and run backwards it shows that $\operatorname{rank} P + \operatorname{rank}(\mathrm{id} - P) = n$ characterizes idempotents (Exercise 4.3.42).

Practice: Exercise 7.7.4, Exercise 7.7.11, Problem 7.2, and Problem 7.6.

7.7.5 Proving an identity for every operator

To prove an identity such as $\operatorname{tr} T = \sum \lambda_i$, $\det p(T) = \prod p(\lambda_i)$, or $p_T(T) = 0$, there are four routes, and they apply to fewer operators the more concrete they are.

- *Test on eigenvectors* (Technique 7.4.25). For diagonalizable T, two polynomials in T agree as soon as they agree on an eigenbasis, which reduces the identity to scalars. For other operators the eigenvectors do not span, and the test proves nothing (Exercise 7.7.7).
- *Diagonalize, solve, conjugate back* (Technique 7.4.22). For computations with a diagonalizable matrix: powers, square roots, recurrences, matrix equations. The equation $AX = XB$ becomes $DY = YE$ with diagonal D, E, which is solved entry by entry (Exercise 7.7.11).
- *Triangularize, then read the diagonal* (Technique 7.5.23). Over $\mathbb{C}$ every operator is triangular in some basis (Theorem 7.5.13), and products and polynomials of triangular matrices have the expected diagonals. This proves anything that depends only on the diagonal: characteristic polynomials, traces, determinants, eigenvalues with multiplicity (Theorem 7.5.22, Theorem 7.6.28).

- *A polynomial identity, or perturbation* (Technique 5.5.24, Technique 7.6.23). If both sides are polynomials in the entries, it is enough to prove the identity for $A - tI_n$ (or another path through A) at infinitely many values of t, for instance at those t for which it is invertible or has distinct eigenvalues; both sides are then polynomials in t that agree infinitely often. This route works over every infinite field and needs no eigenvalues of the original matrix.

Example 7.7.3 (A unipotent matrix has determinant 1, four ways). Let $N \in F^{n\times n}$ satisfy $N^k = 0$ for some $k \geq 1$. We show $\det(I_n + N) = 1$.

Triangularize (over $\mathbb{C}$). The only eigenvalue of N is 0, since every eigenvalue is a root of x^k (Proposition 7.2.28). By Corollary 7.5.14, $N = SRS^{-1}$ with R upper triangular, and the diagonal entries of R are eigenvalues of N (Proposition 7.5.8), so they are 0. Then $I_n + N = S(I_n + R)S^{-1}$, and $I_n + R$ is upper triangular with ones on the diagonal, so $\det(I_n + N) = \det(I_n + R) = 1$ (Proposition 5.4.17).

The characteristic polynomial (any field). By Exercise 7.3.48, $p_N = x^n$. Evaluating $p_N = \det(xI_n - N)$ at $x = -1$ is allowed, since evaluation respects sums and products (Proposition 7.1.10, Proposition 7.3.2(g)), and gives $\det(-I_n - N) = (-1)^n$. So $\det(I_n + N) = (-1)^n \det(-I_n - N) = 1$.

A flag of kernels (any field). The kernels $\ker N \subseteq \ker N^2 \subseteq \cdots \subseteq \ker N^k = F^n$ are nested. Choose a basis of $\ker N$, extend it to a basis of $\ker N^2$, and so on up to F^n. If a basis vector $\mathbf{b}$ was added at stage $j+1$, then $N\mathbf{b} \in \ker N^j$, which is spanned by basis vectors added strictly before $\mathbf{b}$. So in this basis N is strictly upper triangular, and $I_n + N$ is similar to an upper triangular matrix with ones on the diagonal (Technique 7.5.4).

A polynomial identity in t (any field). In $F[t]^{n\times n}$, the matrices $I_n + tN$ and $G = \sum_{j=0}^{k-1} (-t)^j N^j$ satisfy $(I_n + tN)G = I_n - (-tN)^k = I_n$, by telescoping. Let $f(t) = \det(I_n + tN) \in F[t]$. By multiplicativity over $F[t]$ (Proposition 7.3.2(c)), $f(t) \det G = 1$, so f is a unit of $F[t]$, hence a nonzero constant (Corollary 7.1.7(c)). Its value at $t = 0$ is $\det I_n = 1$, so $f = 1$, and at $t = 1$ this is $\det(I_n + N) = 1$.

The triangular proof is the quickest, but it needs the eigenvalues of N in the field, which is why it is stated over $\mathbb{C}$; the characteristic polynomial proof removes that need but uses Exercise 7.3.48. The flag proof is the most informative: it shows that over any field a nilpotent matrix is similar to a strictly upper triangular one, so $I_n + N$ has trace n as well. The last proof is the strangest and the most portable. At $t = 1$ alone, $(I_n + N)G = I_n$ shows only that $\det(I_n + N) \neq 0$; the parameter t is what pins the value down, because a polynomial with a polynomial inverse cannot vary. That is the shape of every perturbation argument in Technique 7.6.23.

Practice: Exercise 7.7.7, Exercise 7.7.14, Problem 7.10, and Problem 7.13.

7.7.6 Induction on dimension

The existence theorems of Section 7.5 and Section 7.6.6 all have the same shape (Technique 7.5.12): find an invariant subspace U with $\{\mathbf{0}\} \neq U \neq V$, apply the induction hypothesis to a smaller operator, and glue with a block triangular matrix. The smaller operator can be the restriction to U, the compression to a complement W (Lemma 7.5.11), or the quotient operator on V/U (Section 4.6). The point to check in every such proof is that the smaller operator satisfies the hypotheses again. A restriction of a complex operator is a complex operator, so triangularization goes through; the compressions of two commuting operators to a complement of a common invariant subspace commute, so simultaneous triangularization goes through (Exercise 7.7.13). A compression is not a restriction, and treating it as one is a real error (Problem 7.14). For Schur's theorem one compresses to $U^\perp$, which keeps orthonormality (Theorem 7.5.18), and the same move proves the spectral theorems in Section 8.2.

7.7.7 Real matrices BEYOND 21-242

Complexification is not covered in lecture.

A real matrix need not have a real eigenvalue, and the arguments of the chapter that start from an eigenvector then have nothing to start from. Three moves get around this. In odd dimension there is always a real eigenvalue, by the intermediate value theorem applied to p_A (Proposition 7.3.30). Otherwise, regard A as a complex matrix, use complex eigenvalues and eigenvectors, and come back by taking real and imaginary parts (Technique 7.3.35): a complex eigenvector $\mathbf{u} + i\mathbf{w}$ gives the real invariant plane $\operatorname{span}(\mathbf{u}, \mathbf{w})$ (Proposition 7.3.34), and the nonreal eigenvalues come in conjugate pairs with equal multiplicities (Proposition 7.1.38), so $\det A$ is a product of real eigenvalues and factors $|\lambda|^2 > 0$. That last fact settles sign questions such as $\det\left(A^2 + B^2\right) \geq 0$ for commuting A, B (Exercise 7.6.52). Finally, a real identity can sometimes be factored over $\mathbb{C}$ as $Z\overline{Z}$, whose determinant is $|\det Z|^2 \geq 0$ (Problem 7.18).

Practice: Exercise 7.7.8, Exercise 7.7.12, Problem 7.11, and Problem 7.17.

7.7.8 Reading the chapter's proofs

Which field? Each statement of the chapter depends on the field in one of three ways. Some hold over every field: eigenvectors for distinct eigenvalues are independent, chains are independent, $\operatorname{gm} \leq \operatorname{am}$, and the characteristic polynomial has tr and det as coefficients. Some hold whenever the relevant polynomial splits: triangularization (Theorem 7.5.17), and tr and det as sum and product of eigenvalues. Over $\mathbb{C}$ that is automatic by the Fundamental Theorem of Algebra, and over a finite field it can fail for every size $n \geq 2$ (Exercise 7.3.44). And some depend on the characteristic: "an involution is diagonalizable" needs $1 + 1 \neq 0$, and "$\operatorname{tr} A^k = 0$ for all k forces A nilpotent" needs characteristic 0. When a proof divides by an integer, or uses that roots such as ± 1 are distinct, that is where the field enters.

Where finite dimension enters. Eigenvectors for distinct eigenvalues are independent in any vector space, and so are chains. Existence of eigenvalues is where finite dimension is needed: multiplication by x on $F[x]$ has none (Remark 7.2.10), because the Krylov sequence $\mathbf{v}, T\mathbf{v}, \ldots$ never becomes dependent.

Two polynomials. The characteristic polynomial records multiplicities: am, trace, determinant. The minimal polynomial records structure: the eigenvalues again, and whether T is diagonalizable. Neither determines the other (Remark 7.6.11), and by Cayley–Hamilton the minimal polynomial divides the characteristic one (Theorem 7.6.20). For a vector $\mathbf{v}$, the first dependence in its Krylov sequence is a third polynomial, and the minimal polynomial is assembled from these one basis vector at a time (Problem 7.9).

Similarity invariants. Characteristic polynomial, minimal polynomial, geometric multiplicities, and the ranks of $(A - \lambda I_n)^k$ are all invariant under similarity, and each of the first three fails to determine the similarity class on its own (Exercise 7.3.38, Exercise 7.6.43). When p_A splits, as over $\mathbb{C}$, the ranks of the powers do determine it; that is the Jordan form (Section 10.3).

7.7.9 Common errors

Most of these appear in the chapter as find-the-error exercises.

- *Assuming the eigenvectors span.* An argument that checks an identity on eigenvectors and concludes it everywhere has assumed that T is diagonalizable (Exercise 7.7.7, Exercise 7.6.44). The same gap hides in "choose an eigenbasis" and in "the eigenspaces of A add up to V" (Exercise 7.7.6).
- *Algebraic and geometric multiplicity confused.* $\dim \ker A = \operatorname{gm}(0)$, while "the eigenvalue 0 counted with multiplicity" means $\operatorname{am}(0)$, which can be larger (Exercise 7.7.5). The Fundamental Theorem of Algebra gives am's that add up to n, not gm's (Exercise 7.4.39).
- *Assuming eigenvalues exist.* A real matrix need not have a real eigenvalue, and a matrix over a finite field need not have any (Example 7.4.9, Exercise 7.3.44). "The only real eigenvalue is 1" does not make 1 the only eigenvalue over $\mathbb{C}$ (Exercise 7.7.8).
- *A one-way implication used both ways.* "$\det A = 0$ if and only if 0 is an eigenvalue" is correct (Theorem 7.2.9). But $p(T) = 0$ makes every eigenvalue a root of p, not every root an eigenvalue;

A diagonalizable makes A^2 diagonalizable, not conversely (Exercise 7.4.38); and equal characteristic polynomials follow from similarity without implying it (Exercise 7.3.38).

- *Eigenvalues of sums and products.* For noncommuting A, B the eigenvalues of $A+B$ and AB have nothing to do with sums and products of eigenvalues of A and B (Exercise 7.2.43, Exercise 7.5.37); commuting is what makes them pair up (Corollary 7.6.33).
- *Substituting a matrix for x inside a determinant*, as in the false proof of Cayley–Hamilton (Exercise 7.3.47). The entries of $xI_n - A$ are polynomials; $AI_n - A$ is a matrix with matrix entries, and its "determinant" is not $p_A(A)$.
- *Eigenvectors are not closed under addition*, and an eigenvector of T^2 need not be one of T (Exercise 7.2.25, Exercise 7.2.53).
- *Restriction and compression confused.* Only an invariant subspace carries a restriction of T; a complement usually is not invariant (Problem 7.14).

Exercises

Exercise 7.7.4. ROUTINE For each statement, decide whether it is true, and name the fact that decides it.

(a) Every $A \in \mathbb{Q}^{3\times 3}$ with $A^3 = A$ is diagonalizable over $\mathbb{Q}$. Does the same hold over $\mathbb{F}_2$?

(b) Every $A \in \mathbb{C}^{4\times 4}$ with $p_A = x^2(x-1)^2$, $\operatorname{rank} A = 2$, and $\operatorname{rank}(A - I_4) = 3$ is diagonalizable.

(c) Every $A \in \mathbb{R}^{5\times 5}$ has a real eigenvalue.

(d) If $A \in \mathbb{C}^{3\times 3}$ has eigenvalues $1, 2, 3$, then $\operatorname{rank}\left(A^2 - 3A + 2I_3\right) = 1$.

(e) Every $A \in \mathbb{R}^{2\times 2}$ with $\operatorname{tr} A = 0$ and $\det A < 0$ is diagonalizable over $\mathbb{R}$.

Solution on page 1614

Exercise 7.7.5. ROUTINE Find the error in the following argument, decide whether the claim is true, and prove a correct version.

Claim. If $A \in \mathbb{C}^{n\times n}$ has rank r, then exactly r of the eigenvalues of A, counted with multiplicity, are nonzero.

"Proof." By rank–nullity, $\dim\ker A = n - r$. Since $\ker A = E_0$, the eigenvalue 0 has multiplicity $n - r$. There are n eigenvalues counted with multiplicity, so the other r are nonzero.

Solution on page 1615

Exercise 7.7.6. ROUTINE Find the errors in the following argument, and give a pair of matrices showing that the claim is false.

Claim. If $A, B \in \mathbb{C}^{n\times n}$ commute, then $\mathbb{C}^n$ has a basis of common eigenvectors of A and B.

"Proof." Let λ be an eigenvalue of A. For $\mathbf{v} \in E_\lambda(A)$, $A(B\mathbf{v}) = BA\mathbf{v} = \lambda B\mathbf{v}$, so $E_\lambda(A)$ is B-invariant. Choose a basis of $E_\lambda(A)$ consisting of eigenvectors of B. Doing this for each eigenvalue of A and putting the bases together gives a basis of $\mathbb{C}^n$ consisting of common eigenvectors.

Solution on page 1615

Exercise 7.7.7. STANDARD Find the error in the following argument, show that the claim is false, and then prove the two correct statements below.

Claim. If $A \in \mathbb{C}^{n\times n}$ satisfies $A^3 = A^2$, then $A^2 = A$.

"Proof." Let $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Then $\lambda^3\mathbf{v} = A^3\mathbf{v} = A^2\mathbf{v} = \lambda^2\mathbf{v}$, so $\lambda^2(\lambda - 1) = 0$ and $\lambda \in \{0, 1\}$. Hence $\lambda^2 = \lambda$, and $A^2\mathbf{v} = \lambda^2\mathbf{v} = \lambda\mathbf{v} = A\mathbf{v}$. So A^2 and A agree on every eigenvector of A, and therefore $A^2 = A$.

(a) If $A^3 = A^2$, then A^2 is idempotent.

(b) If $A^3 = A^2$, then $A^2 = A$ if and only if A is diagonalizable.

Solution on page 1615

Exercise 7.7.8. STANDARD Find the error in the following argument, give a counterexample with integer entries, and prove the correct statements below.

Claim. If $A \in \mathbb{R}^{3\times 3}$ satisfies $A^3 = I_3$, then $A = I_3$.

"Proof." If λ is an eigenvalue of A, then $\lambda^3 = 1$ by Proposition 7.2.28, and the only real cube root of 1 is 1. So 1 is the only eigenvalue of A. The minimal polynomial divides $x^3 - 1$ and its roots are the eigenvalues of A, so $m_A = x - 1$, that is, $A = I_3$.

(a) 1 is an eigenvalue of A.

(b) $A = I_3$ if and only if A is diagonalizable over $\mathbb{R}$.

Solution on page 1616

Exercise 7.7.9. STANDARD Let $n \geq 2$, and let $N \in F^{n\times n}$ satisfy $N^n = 0$ and $N^{n-1} \neq 0$. Show that there is no $B \in F^{n\times n}$ with $B^k = N$ for some $k \geq 2$. Give two proofs: one using powers of B, and one using kernels.

Solution on page 1616

Exercise 7.7.10. STANDARD BEYOND 21-242 Let $A \in F^{m\times n}$ and $B \in F^{n\times m}$ with $m, n \geq 1$.

(a) Show that $Bf(AB)A = (xf)(BA)$ for every $f \in F[x]$, and deduce that m_{BA} divides xm_{AB}.

(b) Deduce that AB and BA have the same nonzero eigenvalues, and give an example with $m_{AB} \neq m_{BA}$.

(c) Show that $\operatorname{gm}_{AB}(\lambda) = \operatorname{gm}_{BA}(\lambda)$ for every $\lambda \neq 0$.

(d) Suppose AB is diagonalizable and invertible. Show that BA is diagonalizable, and show that "invertible" cannot be dropped.

Solution on page 1617

Exercise 7.7.11. STANDARD Let $A \in F^{m\times m}$ and $B \in F^{n\times n}$ be diagonalizable, and let $W = \{X \in F^{m\times n} \mid AX = XB\}$.

(a) Show that $\dim W = \sum_\lambda \operatorname{gm}_A(\lambda)\operatorname{gm}_B(\lambda)$, the sum over the common eigenvalues of A and B.

(b) Deduce that $AX = XB$ has only the solution $X = 0$ if and only if A and B have no common eigenvalue.

(c) Deduce that the matrices commuting with A form a subspace of dimension at least m, with equality if and only if A has m distinct eigenvalues.

Solution on page 1617

Exercise 7.7.12. STANDARD BEYOND 21-242 Let $A, B \in \mathbb{R}^{n\times n}$ with $n \geq 1$ and $AB = BA$.

(a) Show that there is a subspace $U \subseteq \mathbb{R}^n$ of dimension 1 or 2 with $AU \subseteq U$ and $BU \subseteq U$.

(b) Show by an example that dimension 1 cannot always be achieved.

Solution on page 1618

Exercise 7.7.13. HARD Let $A, B \in F^{n\times n}$, $n \geq 1$, be nilpotent with $AB = BA$, over any field F.

(a) Show that there is a nonzero $\mathbf{v}$ with $A\mathbf{v} = B\mathbf{v} = \mathbf{0}$.

(b) Show that there is an invertible S such that $S^{-1}AS$ and $S^{-1}BS$ are both strictly upper triangular.

(c) Deduce that $A + B$ and AB are nilpotent.

Solution on page 1619

Exercise 7.7.14. HARD BEYOND 21-242 Let $A \in F^{n\times n}$ with $n \geq 2$, and write $p_A = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$. Show that

$$\operatorname{adj}(A) = (-1)^{n+1}\bigl(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_2 A + c_1 I_n\bigr),$$

first for invertible A, and then for all A when F is infinite. Check the formula for $n = 2$.

Solution on page 1619

Problems for Chapter 7

These problems combine the sections of Chapter 7 with each other and with earlier chapters: rank inequalities, direct sums, determinants, transposes, and inner products. The first three are review prompts: state a main result, prove it without looking, and then answer a question that the text did not ask. The rest are ordered roughly by difficulty. Problems that use the minimal polynomial or complexification go beyond the course and are marked, and problems of kind "competition" are harder than anything the course expects. Section 7.7 compares the techniques these problems call for.

Review: state and prove

Problem 7.1. STANDARD

(a) State the theorem on eigenvectors for distinct eigenvalues.

(b) Prove it twice: by applying products of the operators $T - \lambda_j \operatorname{id}$ to a relation, and by a relation of minimal length.

(c) Let V be the real vector space of infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$, and $D^2 \in \mathrm{L}(V)$ the second derivative. Show that $1, \cos x, \cos 2x, \ldots, \cos nx$ are linearly independent. Then show that the $2n + 1$ functions $1, \cos x, \sin x, \ldots, \cos nx, \sin nx$ are linearly independent, although $\cos kx$ and $\sin kx$ are eigenvectors of D^2 for the same eigenvalue.

Solution on page 1620

Problem 7.2. STANDARD

(a) State the criteria for diagonalizability in terms of eigenspaces and of multiplicities.

(b) Prove that T is diagonalizable if and only if $\operatorname{gm}(\lambda_1) + \cdots + \operatorname{gm}(\lambda_m) = \dim V$, and that this holds if and only if p_T splits and $\operatorname{gm}(\lambda_i) = \operatorname{am}(\lambda_i)$ for every i.

(c) Let $B \in F^{k\times k}$ and $D \in F^{l\times l}$ be diagonalizable with no common eigenvalue, and let $C \in F^{k\times l}$. Show that $M = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$ is diagonalizable. Show that the hypothesis on eigenvalues cannot be dropped.

Solution on page 1621

Problem 7.3. STANDARD

(a) State the proposition on chains $\mathbf{v}, N\mathbf{v}, \ldots, N^k\mathbf{v}$ with $N = T - \lambda\,\mathrm{id}$, and what it says about the last vector and the span of the chain.

(b) Prove that the chain is independent and that $N^k\mathbf{v}$ is an eigenvector.

(c) Let V be a complex vector space with $\dim V = n \geq 1$, and let $T \in \mathrm{L}(V)$ satisfy $p_T = (x-\lambda)^n$ and $\mathrm{gm}_T(\lambda) = 1$. Show that $(T - \lambda\,\mathrm{id})^{n-1} \neq 0$, and that V has a basis in which the matrix of T has every diagonal entry λ, every entry just above the diagonal 1, and all other entries 0.

Solution on page 1622

Synthesis problems

Problem 7.4. STANDARD Let $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$. Find all T_A-invariant subspaces of $\mathbb{Q}^3$. Begin by showing that every invariant subspace U satisfies $U = (U \cap G) \oplus (U \cap E)$, where $G = \ker(A - 2I_3)^2$ and $E = \ker(A - 3I_3)$.

Solution on page 1622

Problem 7.5. STANDARD Let $A, B, N \in \mathbb{C}^{n\times n}$ with $n \geq 1$. Prove or disprove:

(a) If A^2 is similar to A, then every eigenvalue of A is 0 or a root of unity.

(b) If A^2 is similar to B^2, then A is similar to B.

(c) If A is invertible and $A + A^{-1}$ is diagonalizable, then A is diagonalizable.

(d) If N is nilpotent and $I_n + N$ is diagonalizable, then $N = 0$.

Solution on page 1623

Problem 7.6. STANDARD Let $A \in F^{n\times n}$ and $M = \begin{pmatrix} A & A \\ A & A \end{pmatrix} \in F^{2n\times 2n}$.

(a) Show that $p_M = x^n p_{2A}$ over every field.

(b) Suppose $1 + 1 \neq 0$ in F. Show that M is similar to $\begin{pmatrix} 2A & 0 \\ 0 & 0 \end{pmatrix}$, and that M is diagonalizable if and only if A is.

(c) Over $\mathbb{F}_2$, show that $M^2 = 0$, and that M is diagonalizable if and only if $A = 0$.

Solution on page 1623

Problem 7.7. STANDARD Let $A \in F^{n\times n}$.

(a) Suppose $A\mathbf{v} = \lambda\mathbf{v}$ and $A^{\mathsf{T}}\mathbf{w} = \mu\mathbf{w}$ with $\lambda \neq \mu$. Show that $\mathbf{w}^{\mathsf{T}}\mathbf{v} = 0$. (Over $\mathbb{R}$, this says $\mathbf{v} \perp \mathbf{w}$ for the dot product.)

(b) Suppose $A = SDS^{-1}$ with $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be the columns of S, and let $\mathbf{w}_1^{\mathsf{T}}, \ldots, \mathbf{w}_n^{\mathsf{T}}$ be the rows of S^{-1}. Show that $A^{\mathsf{T}}\mathbf{w}_i = \lambda_i\mathbf{w}_i$, that $\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j$ is 1 if $i = j$ and 0 otherwise, and that

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{w}_i^{\mathsf{T}}.$$

(c) With $P_i = \mathbf{v}_i\mathbf{w}_i^{\mathsf{T}}$, show that $P_iP_j = 0$ for $i \neq j$, $P_i^2 = P_i$, $P_1 + \cdots + P_n = I_n$, and $A^k = \sum_i \lambda_i^k P_i$ for $k \geq 0$.

Solution on page 1624

Problem 7.8. STANDARD For $n \geq 1$, let $A_n \in \mathbb{R}^{n\times n}$ have entries 1 just above and just below the diagonal and 0 elsewhere.

(a) For $\theta \in \mathbb{R}$, let $\mathbf{v} = (\sin\theta, \sin 2\theta, \ldots, \sin n\theta)$. Show that $A_n\mathbf{v} = 2\cos\theta \cdot \mathbf{v} - \sin((n+1)\theta)\mathbf{e}_n$.

(b) Deduce that the eigenvalues of A_n are $2\cos(k\pi/(n+1))$ for $k = 1, \ldots, n$, and that A_n is diagonalizable.

(c) Compute $\det A_n$ from the tridiagonal recurrence, and deduce the value of $\prod_{k=1}^{n} 2\cos(k\pi/(n+1))$.

Solution on page 1625

Problem 7.9. STANDARD BEYOND 21-242 Let V be finite-dimensional and $T \in \mathrm{L}(V)$. For $\mathbf{v} \in V$, let $p_\mathbf{v}$ be the monic polynomial of least degree with $p_\mathbf{v}(T)\mathbf{v} = \mathbf{0}$ (for $\mathbf{v} \neq \mathbf{0}$ it is the polynomial of Proposition 7.2.35, and $p_\mathbf{0} = 1$).

(a) Show that $f(T)\mathbf{v} = \mathbf{0}$ if and only if $p_\mathbf{v}$ divides f.

(b) Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V. Show that $f(T) = 0$ if and only if every $p_{\mathbf{b}_i}$ divides f. Deduce that m_T is the monic polynomial of least degree divisible by $p_{\mathbf{b}_1}, \ldots, p_{\mathbf{b}_n}$.

(c) Use (b) with the standard basis to find m_A for $A = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$, and decide whether A is diagonalizable.

Solution on page 1625

Problem 7.10. STANDARD

(a) Show that every $A \in \mathbb{C}^{n\times n}$ is the sum of two diagonalizable matrices.

(b) Write $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ explicitly as such a sum.

(c) Show that every $A \in \mathbb{C}^{n\times n}$ is the sum of two invertible matrices.

Solution on page 1626

Problem 7.11. HARD Let $A \in \mathbb{R}^{n\times n}$ satisfy $A^2 = -I_n$.

(a) Show that n is even, say $n = 2m$, and that $\mathbb{R}^n$ has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_m, A\mathbf{v}_1, \ldots, A\mathbf{v}_m$.

(b) Deduce that A is similar to $J = \begin{pmatrix} 0 & -I_m \\ I_m & 0 \end{pmatrix}$, so that any two real matrices of the same size with square $-I$ are similar.

Solution on page 1626

Problem 7.12. HARD Let $n \geq 1$, $V = F[x]_{\leq n}$, and $T \in \mathrm{L}(V)$, $T(f) = f(x+1)$ (Definition 7.1.9). Let $\Delta = T - \mathrm{id}_V$.

(a) Let $F = \mathbb{Q}$. Show that $\deg \Delta f = \deg f - 1$ for every nonconstant f, that $x^n, \Delta x^n, \ldots, \Delta^n x^n$ is a basis of V, and that T has 1 as its only eigenvalue, with $\mathrm{gm}(1) = 1$. Is T diagonalizable?

(b) Let $F = \mathbb{F}_p$ with p prime and $n \geq p$. Show that $T^p = \mathrm{id}_V$ and $\Delta^p = 0$, and that $\mathrm{gm}(1) \geq 2$.

Solution on page 1627

Problem 7.13. HARD For $A, B \in F^{n\times n}$, let $L \in \mathrm{L}(F^{n\times n})$ be $L(X) = AXB$.

(a) Show that if $A\mathbf{v} = \lambda\mathbf{v}$ and $B^{\mathsf{T}}\mathbf{w} = \mu\mathbf{w}$ with $\mathbf{v}, \mathbf{w} \neq \mathbf{0}$, then $\mathbf{v}\mathbf{w}^{\mathsf{T}}$ is an eigenvector of L with eigenvalue $\lambda\mu$.

(b) Suppose A and B are diagonalizable, with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$ counted with multiplicity. Show that L is diagonalizable with eigenvalues $\lambda_i\mu_j$ $(1 \leq i, j \leq n)$, and deduce $\operatorname{tr} L = \operatorname{tr} A \operatorname{tr} B$ and $\det L = (\det A)^n (\det B)^n$.

(c) Show that $\det L = (\det A)^n (\det B)^n$ for all A, B, over every field.

Solution on page 1628

Problem 7.14. HARD BEYOND 21-242 Find the error in the following argument, and then prove the claim correctly.

Claim. Let V be finite-dimensional, let $T \in \mathrm{L}(V)$ be diagonalizable, let U be a T-invariant subspace, and let W be any complement of U. Then the compression $S = \mathrm{Pr}_{W,U} \circ T\restriction_W \in \mathrm{L}(W)$ (Lemma 7.5.11) is diagonalizable, and its eigenvalues are eigenvalues of T.

"Proof." S is the restriction of T to W, and the restriction of a diagonalizable operator is diagonalizable (Corollary 7.6.16). Every eigenvector of S is an eigenvector of T with the same eigenvalue.

Solution on page 1628

Problem 7.15. HARD BEYOND 21-242 Let $V \neq \{\mathbf{0}\}$ be a finite-dimensional complex vector space and $T \in \mathrm{L}(V)$, with $m_T = (x - \lambda_1)^{k_1} \cdots (x - \lambda_r)^{k_r}$ for distinct λ_i. With $p_{\mathbf{v}}$ as in Problem 7.9, show that there is $\mathbf{v} \in V$ with $p_{\mathbf{v}} = m_T$. Deduce that $\dim K_T(\mathbf{v}) = \deg m_T$, and that V has a cyclic vector (a $\mathbf{v}$ with $K_T(\mathbf{v}) = V$) if and only if $m_T = p_T$.

Solution on page 1629

Problem 7.16. HARD Does there exist $A \in \mathbb{R}^{3\times 3}$ with $\operatorname{tr} A = 0$ and $A^2 + A^{\mathsf{T}} = I_3$?

Solution on page 1630

Problem 7.17. HARD BEYOND 21-242 Let $n \geq 2$, and let $A, B \in \mathbb{R}^{n\times n}$ be invertible with $A + B$ invertible and

$$A^{-1} + B^{-1} = (A + B)^{-1}.$$

Show that $\det A = \det B$. Does the same conclusion hold for complex matrices?

Solution on page 1630

Problem 7.18. CHALLENGE Let $A, B \in \mathbb{R}^{n\times n}$ satisfy $A^2 + B^2 = AB$, and suppose $AB - BA$ is invertible. Show that n is divisible by 3.

Solution on page 1631

Chapter 8

The spectral theorems

Self-adjoint and normal operators, the real and complex spectral theorems, positive operators, and isometries.

8.1 Self-adjoint and normal operators

An operator with an orthonormal basis of eigenvectors is the best kind of operator to have on an inner product space. In that basis it is a diagonal matrix, coordinates are inner products (Proposition 6.3.18), and lengths are computed by Pythagoras. This chapter determines which operators have such a basis. The answer, proved in Section 8.2, depends on the field: over $\mathbb{R}$ they are the self-adjoint operators, and over $\mathbb{C}$ they are the *normal* operators, the ones that commute with their adjoint. This section studies the two classes before that theorem is available. Self-adjoint operators have real eigenvalues and mutually orthogonal eigenspaces, and on a nonzero finite-dimensional real space they always have an eigenvalue. Normal operators are exactly the operators T for which T and T^* stretch every vector by the same amount, and this one fact gives them orthogonal eigenspaces as well.

Throughout, V is an inner product space over $\mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, and we write $T\mathbf{v}$ for $T(\mathbf{v})$. Whenever T^* appears, T is assumed to have an adjoint; by Theorem 6.6.15 this is automatic when V is finite-dimensional. Results that need finite dimension say so.

8.1.1 Self-adjoint operators and their eigenvalues

Recall that $T \in \mathrm{L}(V)$ is self-adjoint if $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$ (Definition 6.6.26), and that in an orthonormal basis this says that the matrix of T is Hermitian (Corollary 6.6.28). Section 6.6.5 collected examples: orthogonal projections, the operators T^*T, and the maps $\mathbf{x} \mapsto A\mathbf{x}$ for real symmetric and complex Hermitian matrices A. Over $\mathbb{C}$, the self-adjoint operators are the ones whose values $\langle T\mathbf{v}, \mathbf{v}\rangle$ are all real (Proposition 6.6.30). Their eigenvalues are real as well, and this is the first thing the defining identity gives when we feed it eigenvectors.

Intuition. If $T\mathbf{v} = \lambda\mathbf{v}$, then T acts on $\mathbf{v}$ as the scalar λ, and the defining identity $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle$ moves this scalar from the first slot to the second. Pulled out of the first slot it is λ; pulled out of the second it is $\overline{\lambda}$ (Proposition 6.1.4). So $\lambda = \overline{\lambda}$. With eigenvectors $\mathbf{v}$ and $\mathbf{w}$ for two different eigenvalues, the same identity, applied to the pair, gives $\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle$, and that forces $\langle \mathbf{v}, \mathbf{w}\rangle = 0$.

Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators). Let $T \in \mathrm{L}(V)$ be self-adjoint.

(a) Every eigenvalue of T is real.

(b) If $T\mathbf{v} = \lambda\mathbf{v}$ and $T\mathbf{w} = \mu\mathbf{w}$ with $\lambda \neq \mu$, then $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. That is, eigenspaces of T for distinct eigenvalues are orthogonal: every vector of one is orthogonal to every vector of the other.

Proof (by moving T across the inner product). ROUTINE

Uses. Definition 6.6.26 (Self-adjoint operator), Definition 7.2.6 (Eigenvalue, eigenvector), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm), Notation 6.1.1 (Real and complex scalars)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Let $\mathbf{v}$ be an eigenvector for λ. Compute $\langle T\mathbf{v}, \mathbf{v}\rangle$ and $\langle \mathbf{v}, T\mathbf{v}\rangle$ separately.
Hint 2. The first is $\lambda\|\mathbf{v}\|^2$ and the second is $\overline{\lambda}\|\mathbf{v}\|^2$, and $\|\mathbf{v}\| \neq 0$.
Hint 3. For (b), compute both sides of $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$, and use (a) to drop the conjugate on μ.
Sketch. (a) $\lambda\|\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\lambda}\|\mathbf{v}\|^2$, and $\mathbf{v} \neq \mathbf{0}$, so $\lambda = \overline{\lambda}$. (b) $\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle = \overline{\mu}\langle \mathbf{v}, \mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle$, and $\lambda \neq \mu$.

(a) Let λ be an eigenvalue of T and $\mathbf{v}$ an eigenvector for it. Then

$$\lambda\|\mathbf{v}\|^2 = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \overline{\lambda}\|\mathbf{v}\|^2,$$

where the middle equality is self-adjointness and the last uses Proposition 6.1.4. Since $\mathbf{v} \neq \mathbf{0}$, $\|\mathbf{v}\|^2 \neq 0$ (Proposition 6.1.27), so $\lambda = \overline{\lambda}$, which says that $\lambda \in \mathbb{R}$ (Notation 6.1.1). Over $\mathbb{R}$ there was nothing to prove.

(b) By (a), μ is real, so

$$\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle = \overline{\mu}\langle \mathbf{v}, \mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle.$$

Hence $(\lambda - \mu)\langle \mathbf{v}, \mathbf{w}\rangle = 0$, and $\lambda \neq \mu$ gives $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. If $\mathbf{v} \in E_\lambda$ and $\mathbf{w} \in E_\mu$ are arbitrary, then either one of them is $\mathbf{0}$, and $\langle \mathbf{v}, \mathbf{w}\rangle = 0$ trivially, or both are eigenvectors and the computation applies. So E_λ and E_μ are orthogonal. □

Proof (for (a), through the values $\langle T\mathbf{v}, \mathbf{v}\rangle$). ROUTINE

Uses. Definition 6.6.26 (Self-adjoint operator), Proposition 6.6.30 (Self-adjoint means real values), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. For self-adjoint T, every number $\langle T\mathbf{v}, \mathbf{v}\rangle$ is real. Express an eigenvalue through such a number.
Hint 2. If $T\mathbf{v} = \lambda\mathbf{v}$, then $\langle T\mathbf{v}, \mathbf{v}\rangle = \lambda\|\mathbf{v}\|^2$.
Sketch. $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v}, \mathbf{v}\rangle}$ is real for every $\mathbf{v}$. For an eigenvector, $\lambda = \langle T\mathbf{v}, \mathbf{v}\rangle/\|\mathbf{v}\|^2$ is a real number divided by a positive one.

For every $\mathbf{v} \in V$, self-adjointness and conjugate symmetry give $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v}, \mathbf{v}\rangle}$, so $\langle T\mathbf{v}, \mathbf{v}\rangle \in \mathbb{R}$. (This is the easy direction of Proposition 6.6.30; it needs neither $\mathbb{C}$ nor finite dimension.) If $T\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\langle T\mathbf{v}, \mathbf{v}\rangle = \lambda\|\mathbf{v}\|^2$, so

$$\lambda = \frac{\langle T\mathbf{v}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2}$$

is a real number divided by a positive real number (Proposition 6.1.27), hence real. □

The second proof shows where eigenvalues live. The function $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{v}\rangle/\|\mathbf{v}\|^2$ on nonzero vectors, the *Rayleigh quotient* of T, takes only real values, and at an eigenvector it equals the eigenvalue. So every eigenvalue of T lies between the infimum and the supremum of the Rayleigh quotient. Section 12.3 shows that in finite dimensions the smallest and largest eigenvalues are exactly the minimum and maximum of the Rayleigh quotient.

Example 8.1.2 (A Hermitian 2×2 matrix). Let $A = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}$, which is Hermitian (Example 6.6.27). Its characteristic polynomial is

$$p_A = (x-2)(x-3) - (1-i)(1+i) = x^2 - 5x + 4 = (x-1)(x-4),$$

so its eigenvalues are 1 and 4, real although the entries of A are not. For $\lambda = 1$, the first row of $A - I_2 = \begin{pmatrix} 1 & 1-i \\ 1+i & 2 \end{pmatrix}$ gives $x_1 + (1-i)x_2 = 0$, and the second row is $1+i$ times the first, so $E_1 = \operatorname{span}((-1+i, 1))$. For $\lambda = 4$, the first row of $A - 4I_2 = \begin{pmatrix} -2 & 1-i \\ 1+i & -1 \end{pmatrix}$ gives $E_4 = \operatorname{span}((1-i, 2))$. As predicted,

$$(-1+i, 1) \cdot (1-i, 2) = (-1+i)\overline{(1-i)} + 2 = (-1+i)(1+i) + 2 = -2 + 2 = 0.$$

The squared norms are 3 and 6, so $\mathbf{u}_1 = \frac{1}{\sqrt{3}}(-1+i, 1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{6}}(1-i, 2)$ is an orthonormal basis of $\mathbb{C}^2$ consisting of eigenvectors, and with the unitary matrix U whose columns are $\mathbf{u}_1$ and $\mathbf{u}_2$, we get $U^* A U = \operatorname{diag}(1, 4)$.

Part (b) says nothing about two eigenvectors for the *same* eigenvalue: for the identity, every nonzero vector is an eigenvector. An orthonormal list of eigenvectors is obtained by choosing an orthonormal basis inside each eigenspace, with Gram–Schmidt (Algorithm 6.4.14), and Exercise 8.1.23 does this in an example.

Real self-adjoint operators have eigenvalues

Over $\mathbb{R}$, part (a) of Proposition 8.1.1 has no content, since the eigenvalues of an operator on a real space are real by definition. The real question is whether there are any; the rotation by $\pi/2$ has none (Example 7.2.18). Self-adjoint operators always do. The idea is to let a real symmetric matrix act on $\mathbb{C}^n$, where eigenvalues exist, and to apply part (a) there.

Theorem 8.1.3 (Self-adjoint operators have eigenvalues). Let V be a finite-dimensional inner product space with $V \neq \{\mathbf{0}\}$, and let $T \in \mathrm{L}(V)$ be self-adjoint. Then T has an eigenvalue, and it is real. In particular, every real symmetric matrix $A \in \mathbb{R}^{n \times n}$ with $n \geq 1$ has a real eigenvalue.

Proof (by letting the matrix act on $\mathbb{C}^n$). STANDARD

Uses. Theorem 6.4.20 (Orthonormal bases exist), Corollary 6.6.28 (Self-adjointness in an orthonormal basis), Example 6.6.10 (The adjoint of a matrix), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Definition 7.3.12 (Characteristic polynomial of an operator), Corollary 7.3.13 (Eigenvalues of an operator)
Techniques. Pass to orthonormal coordinates (Technique 6.7.2)
Hint 1. Over $\mathbb{C}$ this is Theorem 7.3.27 together with Proposition 8.1.1(a). For $\mathbb{R}$, pass to the matrix A of T in an orthonormal basis.
Hint 2. A is real and symmetric. Regard it as a complex matrix acting on $\mathbb{C}^n$ with the dot product. What kind of operator is $\mathbf{x} \mapsto A\mathbf{x}$ there?
Hint 3. That operator is self-adjoint on $\mathbb{C}^n$, so it has an eigenvalue λ, which is real. Now λ is a real root of $p_A = p_T$.
Sketch. Choose an orthonormal basis $\mathcal{O}$; then $A = [T]_{\mathcal{O}}$ is Hermitian. On $\mathbb{C}^n$ with the dot product, T_A is self-adjoint since $A^* = A$, so it has an eigenvalue λ (Theorem 7.3.27), which is real by Proposition 8.1.1(a). Hence λ is a root of p_A, which is the same polynomial whether A is read as real or complex, and $p_A = p_T$. So $\lambda \in \mathbb{F}$ is a root of p_T, hence an eigenvalue of T.

Let $n = \dim V \geq 1$. By Theorem 6.4.20, V has an orthonormal basis $\mathcal{O}$, and $A = [T]_{\mathcal{O}} \in \mathbb{F}^{n \times n}$ satisfies $A^* = A$ by Corollary 6.6.28. Regard A as an element of $\mathbb{C}^{n \times n}$; when $\mathbb{F} = \mathbb{R}$ its entries are real and still $A^* = A^{\mathsf{T}} = A$. By Example 6.6.10, the operator $T_A : \mathbf{x} \mapsto A\mathbf{x}$ on $\mathbb{C}^n$ with the dot product has adjoint $T_{A^*} = T_A$, so it is self-adjoint. By Theorem 7.3.27, T_A has an eigenvalue $\lambda \in \mathbb{C}$, and $\lambda \in \mathbb{R}$ by Proposition 8.1.1(a). By Theorem 7.3.6, $p_A(\lambda) = 0$.

The polynomial $p_A = \det(xI_n - A)$ is computed from the entries of A alone, so it is the same polynomial whether A is regarded as a real or a complex matrix, and it equals p_T (Definition 7.3.12). So λ is a root of p_T lying in $\mathbb{R} \subseteq \mathbb{F}$, and by Corollary 7.3.13 it is an eigenvalue of T. It is real by Proposition 8.1.1(a). The last sentence of the theorem is the case $V = \mathbb{R}^n$ with the dot product and $T = T_A$, which is self-adjoint by Example 6.6.10. □

The proof shows more. Every complex root of p_A is an eigenvalue of the self-adjoint operator T_A on $\mathbb{C}^n$, hence real, so by Corollary 7.1.35

$$p_A = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n) \quad \text{with all } \lambda_j \in \mathbb{R}.$$

The characteristic polynomial of a Hermitian matrix, and in particular of a real symmetric matrix, splits over $\mathbb{R}$. This is what the real spectral theorem in Section 8.2 needs.

The first proof leans on determinants and lets a real matrix act on $\mathbb{C}^n$. The second avoids determinants and never complexifies the operator, although the factorization over $\mathbb{R}$ that it uses still rests on the Fundamental Theorem of Algebra. Following the determinant-free proof of Theorem 7.3.27, it factors a polynomial that kills a vector. Over $\mathbb{R}$ some factors may be quadratics without real roots, and these are exactly where a real operator can fail to have eigenvalues; for a self-adjoint operator they turn out to be injective, hence harmless.

Proof (without determinants). HARD BEYOND 21-242

> *Uses.* Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.1.39 (Factorization over $\mathbb{R}$), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 7.2.9 (Eigenvalues and invertibility)
> *Techniques.* Krylov sequences must become dependent (Technique 7.3.28), A sum of nonnegative terms (Technique 6.1.9)
> *Hint 1.* Over $\mathbb{C}$, combine Theorem 7.3.27 with (a) of Proposition 8.1.1. Over $\mathbb{R}$, imitate the second proof of Theorem 7.3.27. What goes wrong over $\mathbb{R}$, and how could self-adjointness repair it?
> *Hint 2.* Over $\mathbb{R}$, the polynomial p with $p(T)\mathbf{v} = \mathbf{0}$ factors into linear factors and quadratics $x^2 + bx + c$ with $b^2 < 4c$ (Theorem 7.1.39). Show that each such quadratic gives an injective operator $T^2 + bT + c\,\mathrm{id}$.
> *Hint 3.* Compute $\langle (T^2 + bT + c\,\mathrm{id})\mathbf{v}, \mathbf{v}\rangle = \|T\mathbf{v}\|^2 + b\langle T\mathbf{v}, \mathbf{v}\rangle + c\|\mathbf{v}\|^2$, bound the middle term with Cauchy–Schwarz, and complete the square.
> *Sketch.* For $b^2 < 4c$ and $\mathbf{v} \neq \mathbf{0}$, $\langle (T^2 + bT + c\,\mathrm{id})\mathbf{v}, \mathbf{v}\rangle \geq (\|T\mathbf{v}\| - |b|\|\mathbf{v}\|/2)^2 + (c - b^2/4)\|\mathbf{v}\|^2 > 0$, so the quadratic factors are injective at T. Take $\mathbf{v} \neq \mathbf{0}$ and a monic p of positive degree with $p(T)\mathbf{v} = \mathbf{0}$ (Proposition 7.2.35). Factor p over $\mathbb{R}$. Then $p(T)$ is the composite of its factors evaluated at T, and it is not injective, so some linear factor $T - \lambda\,\mathrm{id}$ is not injective.

If $\mathbb{F} = \mathbb{C}$, then T has an eigenvalue by Theorem 7.3.27, and it is real by Proposition 8.1.1(a). So let $\mathbb{F} = \mathbb{R}$.

Quadratics without real roots are injective at T. Let $b, c \in \mathbb{R}$ with $b^2 < 4c$, and let $Q = T^2 + bT + c\,\mathrm{id}$. For $\mathbf{v} \in V$, self-adjointness gives $\langle T^2\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$, and Cauchy–Schwarz (Theorem 6.2.1) gives $b\langle T\mathbf{v}, \mathbf{v}\rangle \geq -|b|\|T\mathbf{v}\|\|\mathbf{v}\|$. Hence

$$\begin{aligned}\langle Q\mathbf{v}, \mathbf{v}\rangle &\geq \|T\mathbf{v}\|^2 - |b|\|T\mathbf{v}\|\|\mathbf{v}\| + c\|\mathbf{v}\|^2 \\ &= \left(\|T\mathbf{v}\| - \frac{|b|}{2}\|\mathbf{v}\|\right)^2 + \left(c - \frac{b^2}{4}\right)\|\mathbf{v}\|^2.\end{aligned}$$

If $\mathbf{v} \neq \mathbf{0}$, the last term is positive, so $\langle Q\mathbf{v}, \mathbf{v}\rangle > 0$ and $Q\mathbf{v} \neq \mathbf{0}$. So Q is injective.

Some linear factor is not injective. Fix $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$. By Proposition 7.2.35 there is a monic $p \in \mathbb{R}[x]$ of degree $d \geq 1$ with $p(T)\mathbf{v} = \mathbf{0}$. By Theorem 7.1.39, and since p is monic,

$$p = (x-\lambda_1)\cdots(x-\lambda_k)\left(x^2 + b_1x + c_1\right)\cdots\left(x^2 + b_lx + c_l\right)$$

with $b_j^2 < 4c_j$ for each j. By Proposition 7.2.27(a),

$$p(T) = (T - \lambda_1\,\mathrm{id})\cdots(T - \lambda_k\,\mathrm{id})Q_1\cdots Q_l, \quad Q_j = T^2 + b_jT + c_j\,\mathrm{id}.$$

Each Q_j is injective by the first part. If every $T - \lambda_i\,\mathrm{id}$ were injective too (this includes the case $k = 0$), then $p(T)$ would be a composite of injective maps, hence injective, contradicting $p(T)\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$. So some $T - \lambda_i\,\mathrm{id}$ is not injective, and λ_i is an eigenvalue of T by Theorem 7.2.9. □

For the rotation $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ the Krylov polynomial of $\mathbf{e}_1$ is $x^2 + 1$, and $J^2 + \mathrm{id} = 0$ is as far from injective as possible. The positivity estimate is exactly what a rotation violates: $\langle J^2\mathbf{v}, \mathbf{v}\rangle = -\|\mathbf{v}\|^2$, whereas for self-adjoint T the term $\langle T^2\mathbf{v}, \mathbf{v}\rangle = \|T\mathbf{v}\|^2$ is never negative.

8.1.2 Testing self-adjoint operators on the diagonal

By Proposition 6.1.41, an operator T on a complex inner product space with $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$ is 0, and over $\mathbb{R}$ this fails for the rotation J (Example 6.1.39). The rotation is far from self-adjoint, and for self-adjoint operators the real case behaves like the complex one. We will need this to characterize normal operators over both fields at once.

> *Intuition.* Put $\mathbf{v} = \mathbf{x} + \mathbf{y}$ into $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$. Over either field the result is $\langle T\mathbf{x}, \mathbf{y}\rangle + \langle T\mathbf{y}, \mathbf{x}\rangle = 0$ (Proposition 6.1.41, first proof), and for J the two terms cancel without being 0. If T is self-adjoint, then $\langle T\mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{y}, T\mathbf{x}\rangle$ is the conjugate of $\langle T\mathbf{x}, \mathbf{y}\rangle$, so the sum is $2\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle$. Then $\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle = 0$ for all $\mathbf{y}$, and the choice $\mathbf{y} = T\mathbf{x}$ gives $\|T\mathbf{x}\|^2 = 0$.

Proposition 8.1.4 (Self-adjoint operators are determined by their diagonal values). Let $T \in \mathrm{L}(V)$ be self-adjoint, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. If $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v} \in V$, then $T = 0$. Consequently, if $S, T \in \mathrm{L}(V)$ are self-adjoint and $\langle S\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, \mathbf{v}\rangle$ for every $\mathbf{v} \in V$, then $S = T$.

Proof (by polarization). STANDARD

> *Uses.* Definition 6.6.26 (Self-adjoint operator), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.6.18 (Properties of the adjoint)
> *Techniques.* Polarization (Technique 6.1.33), Show a vector is zero via its norm (Technique 6.1.36)
> *Hint 1.* Expand $\langle T(\mathbf{x}+\mathbf{y}), \mathbf{x}+\mathbf{y}\rangle - \langle T(\mathbf{x}-\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle$.
> *Hint 2.* With self-adjointness, the result is $4\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle$. So $\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle = 0$ for all $\mathbf{x}, \mathbf{y}$.
> *Hint 3.* Choose $\mathbf{y} = T\mathbf{x}$.
> *Sketch.* Expanding, and using $\langle T\mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{y}, T\mathbf{x}\rangle = \overline{\langle T\mathbf{x}, \mathbf{y}\rangle}$, we get $0 = \langle T(\mathbf{x}+\mathbf{y}), \mathbf{x}+\mathbf{y}\rangle - \langle T(\mathbf{x}-\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle = 4\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle$. With $\mathbf{y} = T\mathbf{x}$ this is $4\|T\mathbf{x}\|^2$. For the last claim, apply this to $S - T$, which is self-adjoint.

Let $\mathbf{x}, \mathbf{y} \in V$. Expanding by linearity of T and sesquilinearity, the terms $\langle T\mathbf{x}, \mathbf{x}\rangle$ and $\langle T\mathbf{y}, \mathbf{y}\rangle$ cancel, and

$$\langle T(\mathbf{x}+\mathbf{y}), \mathbf{x}+\mathbf{y}\rangle - \langle T(\mathbf{x}-\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle = 2\langle T\mathbf{x}, \mathbf{y}\rangle + 2\langle T\mathbf{y}, \mathbf{x}\rangle.$$

Since T is self-adjoint, $\langle T\mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{y}, T\mathbf{x}\rangle = \overline{z}$, where $z = \langle T\mathbf{x}, \mathbf{y}\rangle$. So the right side is $2(z + \overline{z}) = 4\operatorname{Re} z$, and the left side is 0 by hypothesis. Hence $\operatorname{Re}\langle T\mathbf{x}, \mathbf{y}\rangle = 0$ for all $\mathbf{x}, \mathbf{y} \in V$. Taking $\mathbf{y} = T\mathbf{x}$ gives $\|T\mathbf{x}\|^2 = \operatorname{Re}\langle T\mathbf{x}, T\mathbf{x}\rangle = 0$, so $T\mathbf{x} = \mathbf{0}$ (Proposition 6.1.27). As $\mathbf{x}$ was arbitrary, $T = 0$.

For the last statement, $S - T$ is self-adjoint, since $(S-T)^* = S^* - T^* = S - T$ by (a) and (b) of Proposition 6.6.18, and $\langle (S-T)\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, \mathbf{v}\rangle - \langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$. So $S - T = 0$. □

Proof (with the test vector $\mathbf{v} + T\mathbf{v}$). STANDARD

> *Uses.* Definition 6.6.26 (Self-adjoint operator), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
> *Techniques.* Plug in clever test vectors (Technique 6.1.40), Show a vector is zero via its norm (Technique 6.1.36)
> *Hint 1.* We want $\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle T^2\mathbf{v}, \mathbf{v}\rangle$ to vanish. Feed the hypothesis a vector built from $\mathbf{v}$ and $T\mathbf{v}$.
> *Hint 2.* Try $\mathbf{v} + T\mathbf{v}$. Two of the four terms in the expansion vanish by the hypothesis, applied at $\mathbf{v}$ and at $T\mathbf{v}$.
> *Sketch.* $\langle T(\mathbf{v}+T\mathbf{v}), \mathbf{v}+T\mathbf{v}\rangle = \langle T\mathbf{v}, \mathbf{v}\rangle + \langle T\mathbf{v}, T\mathbf{v}\rangle + \langle T^2\mathbf{v}, \mathbf{v}\rangle + \langle T(T\mathbf{v}), T\mathbf{v}\rangle$. The first and last terms vanish by hypothesis, and $\langle T^2\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle$ by self-adjointness, so $0 = 2\|T\mathbf{v}\|^2$.

Let $\mathbf{v} \in V$ and put $\mathbf{w} = T\mathbf{v}$. Applying the hypothesis to $\mathbf{v} + \mathbf{w}$ and expanding,

$$0 = \langle T\mathbf{v}, \mathbf{v}\rangle + \langle T\mathbf{v}, \mathbf{w}\rangle + \langle T\mathbf{w}, \mathbf{v}\rangle + \langle T\mathbf{w}, \mathbf{w}\rangle.$$

The first and last terms are 0 by hypothesis, applied at $\mathbf{v}$ and at $\mathbf{w}$. The second is $\langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$, and the third is $\langle T(T\mathbf{v}), \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$ because T is self-adjoint. So $2\|T\mathbf{v}\|^2 = 0$, and $T\mathbf{v} = \mathbf{0}$ (Proposition 6.1.27). The last statement follows as in the first proof. □

The first proof is the polarization argument of Proposition 6.1.41, with self-adjointness supplying the second equation that the scalar i supplied there. The second proof chooses its test vector by looking at the goal. We want $\|T\mathbf{v}\|^2$, which by self-adjointness is $\langle T^2\mathbf{v}, \mathbf{v}\rangle$, and $\mathbf{v} + T\mathbf{v}$ is the simplest vector whose expansion contains that term.

Without self-adjointness, the real case is described by Exercise 6.1.52(b): an operator T on a real inner product space satisfies $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$ exactly when $\langle T\mathbf{v}, \mathbf{w}\rangle = -\langle \mathbf{v}, T\mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w}$, that is, when $T^* = -T$. The rotation J is of this kind, and such operators have a name.

Definition 8.1.5 (Skew-adjoint operators, skew-Hermitian matrices). An operator $T \in \mathrm{L}(V)$ is **skew-adjoint** if it has an adjoint and $T^* = -T$. A matrix $A \in \mathbb{F}^{n\times n}$ is **skew-Hermitian** if $A^* = -A$. A real skew-Hermitian matrix, one with $A^{\mathsf{T}} = -A$, is **skew-symmetric**.

The matrices $\begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}$ with $b \in \mathbb{R}$, among them J, are skew-symmetric, and every skew-symmetric matrix has zeros on its diagonal. A skew-Hermitian matrix has purely imaginary diagonal entries, as in $\begin{pmatrix} i & 1+i \\ -1+i & 0 \end{pmatrix}$. Differentiation on trigonometric polynomials is skew-adjoint (Exercise 6.6.39). Over $\mathbb{C}$, T is skew-adjoint if and only if iT is self-adjoint, because $(iT)^* = -iT^*$ (Proposition 6.6.18(b)). For every T, the operator $T - T^*$ is skew-adjoint.

Over $\mathbb{R}$, every operator T splits as $T = S + K$ with $S = (T + T^*)/2$ self-adjoint and $K = (T - T^*)/2$ skew-adjoint, and $\langle K\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$. So $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, \mathbf{v}\rangle$, and by Proposition 8.1.4 these values determine S and nothing more. The values $\langle T\mathbf{v}, \mathbf{v}\rangle$ see exactly the self-adjoint part of a real operator.

Exercise 8.1.6. ROUTINE Let $K \in \mathrm{L}(V)$ be skew-adjoint and $c \in \mathbb{R}$. Show that

$$\|(K - c\,\mathrm{id})\mathbf{v}\|^2 = \|K\mathbf{v}\|^2 + c^2\|\mathbf{v}\|^2 \quad \text{for every } \mathbf{v} \in V.$$

Deduce that K has no nonzero real eigenvalue, and that $\mathrm{id} + K$ is injective.

Solution on page 1631

8.1.3 Normal operators

Suppose V is finite-dimensional and has an orthonormal basis $\mathcal{O}$ of eigenvectors of T. What does this force on T? In the basis $\mathcal{O}$, T is a diagonal matrix D, and T^* is D^*, which is diagonal too. Diagonal matrices commute, so T and T^* commute. Over $\mathbb{R}$, moreover, $D^* = D$.

Proposition 8.1.7 (What an orthonormal eigenbasis forces). Let V be finite-dimensional and $T \in \mathrm{L}(V)$. Suppose that V has an orthonormal basis consisting of eigenvectors of T. Then $T^*T = TT^*$. If moreover every eigenvalue of T is real, which is automatic when $\mathbb{F} = \mathbb{R}$, then T is self-adjoint.

Proof. ROUTINE

Uses. Theorem 6.6.15 (Existence of adjoints), Definition 6.3.26 (Conjugate transpose), Theorem 2.4.2 (The matrix of a composite), Theorem 2.3.20 (Linear maps correspond to matrices)
Techniques. Pass to orthonormal coordinates (Technique 6.7.2), Translate between maps and matrices (Technique 2.3.27)
Hint 1. Write down the matrix D of T in the given basis, and use Theorem 6.6.15(b) for the matrix of T^*.
Hint 2. D^* is diagonal with entries $\overline{\lambda_j}$, and diagonal matrices commute.
Sketch. $[T]_{\mathcal{O}} = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, so $[T^*]_{\mathcal{O}} = D^* = \mathrm{diag}\big(\overline{\lambda_1}, \ldots, \overline{\lambda_n}\big)$. Then $[T^*T]_{\mathcal{O}} = D^*D = DD^* = [TT^*]_{\mathcal{O}}$, so $T^*T = TT^*$. If all λ_j are real, $D^* = D$ and $T^* = T$.

Let $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ be the orthonormal basis, with $T\mathbf{u}_j = \lambda_j\mathbf{u}_j$. Then $[T]_{\mathcal{O}} = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, and by Theorem 6.6.15(b) and Definition 6.3.26, $[T^*]_{\mathcal{O}} = D^* = \mathrm{diag}\big(\overline{\lambda_1}, \ldots, \overline{\lambda_n}\big)$. By Theorem 2.4.2, $[T^*T]_{\mathcal{O}} = D^*D$ and $[TT^*]_{\mathcal{O}} = DD^*$, and both equal $\mathrm{diag}\big(|\lambda_1|^2, \ldots, |\lambda_n|^2\big)$. So $T^*T = TT^*$ by Theorem 2.3.20. If every λ_j is real, then $D^* = D$, so $[T^*]_{\mathcal{O}} = [T]_{\mathcal{O}}$ and $T^* = T$. □

So orthonormal eigenbases occur only for operators that commute with their adjoints, and over $\mathbb{R}$ only for self-adjoint ones. The spectral theorems of Section 8.2 prove the converses in finite dimensions: over $\mathbb{C}$ the first condition suffices, and over $\mathbb{R}$ the second. Over $\mathbb{C}$, self-adjointness would be too much to ask, since a diagonal matrix such as $\mathrm{diag}(1, i)$ has an orthonormal eigenbasis without being Hermitian. The weaker condition gets a name.

Definition 8.1.8 (Normal operator, normal matrix). An operator $T \in \mathrm{L}(V)$ is **normal** if it has an adjoint and

$$T^*T = TT^*.$$

A matrix $A \in \mathbb{F}^{n\times n}$ is **normal** if $A^*A = AA^*$.

For $A \in \mathbb{F}^{n\times n}$, the operator T_A on $\mathbb{F}^n$ with the dot product is normal exactly when A is, because $(T_A)^* = T_{A^*}$ (Example 6.6.10) and the matrix of a composite is the product of the matrices.

Example 8.1.9 (Normal operators).

(a) Self-adjoint and skew-adjoint operators are normal: if $T^* = T$, then $T^*T = T^2 = TT^*$, and if $T^* = -T$, then $T^*T = -T^2 = TT^*$.

(b) Unitary matrices are normal, since $U^*U = I_n = UU^*$ by Proposition 6.3.34; in particular orthogonal matrices are normal. More generally, every operator with $T^*T = TT^* = \mathrm{id}$ is normal. These are the unitary operators (Definition 8.4.9) of Section 8.4.

(c) For $a, b \in \mathbb{R}$, the matrix $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = aI_2 + bJ$ satisfies

$$A^{\mathsf{T}}A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = (a^2+b^2)I_2 = AA^{\mathsf{T}}.$$

When $a \neq 0$, $b \neq 0$, and $a^2 + b^2 \neq 1$, it is neither symmetric, skew-symmetric, nor orthogonal. Identifying $\mathbb{R}^2$ with $\mathbb{C}$ by $(x, y) \leftrightarrow x + iy$, the map $\mathbf{x} \mapsto A\mathbf{x}$ is multiplication by $a + bi$: it rotates by the argument of $a + bi$ and scales by $|a + bi|$.

(d) Every diagonal matrix D is normal, as $D^*D = DD^*$ is diagonal with entries $|d_{jj}|^2$. For instance $\mathrm{diag}(1, 2i)$ is normal but not Hermitian, not skew-Hermitian, and not unitary.

Non-example 8.1.10 (Operators that are not normal).

- The nilpotent $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$: here $N^*N = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ while $NN^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.
- The shear $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (Example 7.4.8), which is invertible: $S^{\mathsf{T}}S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ while $SS^{\mathsf{T}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.
- $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, which is diagonalizable, with eigenvectors $(1, 0)$ for 1 and $(1, 1)$ for 2: $B^{\mathsf{T}}B = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}$ while $BB^{\mathsf{T}} = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$.
- The complex symmetric matrix $C = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$ of Exercise 6.6.34: $C^*C = \begin{pmatrix} 2 & 2i \\ -2i & 2 \end{pmatrix}$ while $CC^* = \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix}$.

None of these has an orthonormal eigenbasis, by Proposition 8.1.7. For N, S, and C there is no eigenbasis at all: each has a single eigenvalue ($C^2 = 0$, so 0 is the only eigenvalue of C) and is not a multiple of the identity. The matrix B does have an eigenbasis, but its eigenspaces are two lines that are not perpendicular, so no choice of eigenvectors is orthogonal.

Sums of normal operators need not be normal: $N = (N + N^*)/2 + (N - N^*)/2$ is a sum of a self-adjoint and a skew-adjoint operator (Exercise 8.1.24). So the normal operators do not form a subspace of $\mathrm{L}(V)$, unlike the self-adjoint operators over $\mathbb{R}$.

Corollary 8.1.11 (Normality in an orthonormal basis). Let V be finite-dimensional, let $\mathcal{O}$ be an orthonormal basis of V, let $T \in \mathrm{L}(V)$, and let $A = [T]_{\mathcal{O}}$. Then T is normal, self-adjoint, skew-adjoint, or satisfies $T^*T = TT^* = \mathrm{id}$ if and only if A is normal, Hermitian, skew-Hermitian, or unitary, respectively.

Proof. ROUTINE

Uses. Theorem 6.6.15 (Existence of adjoints), Theorem 2.4.2 (The matrix of a composite), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 6.3.34 (Characterizations of unitary matrices), Example 2.3.6 (The zero map and the identity map)

Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Translate each operator identity into a matrix identity with $[T^*]_{\mathcal{O}} = A^*$ (Theorem 6.6.15(b)).
Hint 2. Two operators are equal if and only if their matrices in $\mathcal{O}$ are equal.
Sketch. $[T^*]_{\mathcal{O}} = A^*$, $[T^*T]_{\mathcal{O}} = A^*A$, $[TT^*]_{\mathcal{O}} = AA^*$, $[-T]_{\mathcal{O}} = -A$, and $[\mathrm{id}_V]_{\mathcal{O}} = I_n$. Since $T \mapsto [T]_{\mathcal{O}}$ is injective, each identity among operators holds iff the same identity holds among the matrices.

By Theorem 6.6.15(b), $[T^*]_{\mathcal{O}} = A^*$. By Theorem 2.4.2, $[T^*T]_{\mathcal{O}} = A^*A$ and $[TT^*]_{\mathcal{O}} = AA^*$; also $[-T]_{\mathcal{O}} = -A$, since every entry of the matrix of $-T$ is the negative of the corresponding entry of A, and $[\mathrm{id}_V]_{\mathcal{O}} = I_n$ (Example 2.3.6). By Theorem 2.3.20, two operators are equal if and only if their matrices in $\mathcal{O}$ are equal. So $T^*T = TT^*$ iff $A^*A = AA^*$, $T^* = T$ iff $A^* = A$, $T^* = -T$ iff $A^* = -A$, and $T^*T = TT^* = \mathrm{id}_V$ iff $A^*A = AA^* = I_n$, which by Proposition 6.3.34 says that A is unitary. □

Orthonormality matters here, as it did for Corollary 6.6.28. The matrix B of Non-example 8.1.10 is diagonal in the basis $(1,0), (1,1)$, so its matrix in that basis is normal, but T_B is not normal for the dot product.

Condition	Operator	Matrix	Eigenvalues
$T^* = T$	self-adjoint	Hermitian (symmetric)	real
$T^* = -T$	skew-adjoint	skew-Hermitian (skew-symmetric)	in $i\mathbb{R}$
$T^*T = TT^* = \mathrm{id}$	unitary	unitary (orthogonal)	$\lvert\lambda\rvert = 1$
$T^*T = TT^*$	normal	normal	any

Table 8.1.1. Four classes of normal operators. The third column names the matrix of the operator in an orthonormal basis (Corollary 8.1.11), with the name used for real matrices in parentheses. The last column is Corollary 8.1.20. Under the analogy between $T \mapsto T^*$ and $z \mapsto \overline{z}$ (Section 6.6.5), the rows correspond to the real numbers ($\overline{z} = z$), the imaginary numbers ($\overline{z} = -z$), the unit circle ($\overline{z}z = 1$), and all of $\mathbb{C}$, where $\overline{z}z = z\overline{z}$ always holds.

Exercise 8.1.12. ROUTINE Let $T \in \mathrm{L}(V)$ be normal and $c \in \mathbb{F}$. Show that T^*, cT, and T^k for $k \geq 1$ are normal. Show also that if T is invertible and T^{-1} has an adjoint, then T^{-1} is normal.

Solution on page 1632

8.1.4 Normal operators and their adjoints

The definition compares the operators T^*T and TT^*. Feeding both the same vector and taking an inner product with that vector turns this into a comparison of lengths.

Proposition 8.1.13 (Normal means $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$). Let $T \in \mathrm{L}(V)$ have an adjoint. Then T is normal if and only if

$$\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \quad \text{for every } \mathbf{v} \in V.$$

Proof (through the diagonal values of $T^*T - TT^*$). STANDARD

Uses. Definition 6.6.8 (Adjoint), Proposition 6.6.18 (Properties of the adjoint), Proposition 8.1.4 (Self-adjoint operators are determined by their diagonal values), Definition 8.1.8 (Normal operator, normal matrix)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Write $\|T\mathbf{v}\|^2$ and $\|T^*\mathbf{v}\|^2$ in the form $\langle X\mathbf{v}, \mathbf{v}\rangle$ for suitable operators X.
Hint 2. $\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2 = \langle (T^*T - TT^*)\mathbf{v}, \mathbf{v}\rangle$. Which proposition turns "this vanishes for every $\mathbf{v}$" into "$T^*T - TT^* = 0$"? Check its hypothesis.
Sketch. Moving operators across, $\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2 = \langle S\mathbf{v}, \mathbf{v}\rangle$ with $S = T^*T - TT^*$. The operator S is self-adjoint, so by Proposition 8.1.4, $S = 0$ if and only if $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$.

For $\mathbf{v} \in V$, the defining identities of T^* and of $(T^*)^* = T$ (Proposition 6.6.18(c)) give

$$\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle T^*T\mathbf{v}, \mathbf{v}\rangle, \quad \|T^*\mathbf{v}\|^2 = \langle T^*\mathbf{v}, T^*\mathbf{v}\rangle = \langle TT^*\mathbf{v}, \mathbf{v}\rangle.$$

So $\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2 = \langle S\mathbf{v}, \mathbf{v}\rangle$, where $S = T^*T - TT^*$. If T is normal, then $S = 0$ and the norms agree. Conversely, S is self-adjoint: by (a)–(d) of Proposition 6.6.18,

$$S^* = (T^*T)^* - (TT^*)^* = T^*(T^*)^* - (T^*)^*T^* = T^*T - TT^* = S.$$

If $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for every $\mathbf{v}$, then $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$, and Proposition 8.1.4 gives $S = 0$. □

Proof (by polarization). STANDARD

Uses. Proposition 6.1.32 (Polarization identities), Corollary 6.1.38 (Linear maps are determined by inner products), Definition 6.6.8 (Adjoint), Proposition 6.6.18 (Properties of the adjoint)
Techniques. Polarization (Technique 6.1.33)
Hint 1. Equal norms for all vectors should give equal inner products. Which identity recovers $\langle \mathbf{x}, \mathbf{y}\rangle$ from norms?
Hint 2. Apply Proposition 6.1.32 to $\langle T\mathbf{v}, T\mathbf{w}\rangle$ and to $\langle T^*\mathbf{v}, T^*\mathbf{w}\rangle$, using $T(\mathbf{v} + c\mathbf{w}) = T\mathbf{v} + cT\mathbf{w}$.
Sketch. By polarization, $\langle T\mathbf{v}, T\mathbf{w}\rangle$ is a fixed combination of the numbers $\|T(\mathbf{v} + c\mathbf{w})\|^2$ with $c \in \{\pm 1\}$ (over $\mathbb{R}$) or $c \in \{\pm 1, \pm i\}$ (over $\mathbb{C}$), and the same combination of the numbers $\|T^*(\mathbf{v} + c\mathbf{w})\|^2$ gives $\langle T^*\mathbf{v}, T^*\mathbf{w}\rangle$. So $\langle T^*T\mathbf{v}, \mathbf{w}\rangle = \langle TT^*\mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w}$, and $T^*T = TT^*$.

If T is normal, then $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \|T^*\mathbf{v}\|^2$ for every $\mathbf{v}$.

Conversely, suppose $\|T\mathbf{x}\| = \|T^*\mathbf{x}\|$ for every $\mathbf{x} \in V$, and let $\mathbf{v}, \mathbf{w} \in V$. Over $\mathbb{C}$, Proposition 6.1.32(b) and linearity of T and T^* give

$$\langle T\mathbf{v}, T\mathbf{w}\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \left\|T\left(\mathbf{v} + i^k\mathbf{w}\right)\right\|^2 = \frac{1}{4}\sum_{k=0}^{3} i^k \left\|T^*\left(\mathbf{v} + i^k\mathbf{w}\right)\right\|^2 = \langle T^*\mathbf{v}, T^*\mathbf{w}\rangle,$$

and over $\mathbb{R}$ the same holds with Proposition 6.1.32(a), which uses only $\mathbf{v} \pm \mathbf{w}$. Moving the operators across, the two ends of this chain are $\langle T^*T\mathbf{v}, \mathbf{w}\rangle$ and $\langle TT^*\mathbf{v}, \mathbf{w}\rangle$. They agree for all $\mathbf{v}, \mathbf{w} \in V$, so $T^*T = TT^*$ by Corollary 6.1.38. □

The first proof locates normality in a single self-adjoint operator, $T^*T - TT^*$, whose diagonal values are the differences $\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2$. The second proof produces something visibly stronger along the way: T and T^* give the same inner products, $\langle T\mathbf{v}, T\mathbf{w}\rangle = \langle T^*\mathbf{v}, T^*\mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w}$.

Intuition. T and T^* need not be equal, but for a normal operator they stretch every vector by the same factor; they may send it in different directions. For $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ this fails as badly as possible: $N\mathbf{e}_2 = \mathbf{e}_1$ has length 1, while $N^*\mathbf{e}_2 = \mathbf{0}$. Everything else in this section follows from the equality of lengths. The first consequence is about kernels: a vector killed by T is killed by T^*.

Proposition 8.1.14 (Kernels and images of normal operators). Let $T \in \mathrm{L}(V)$ be normal.

(a) $\ker T = \ker T^*$.

(b) For every $\lambda \in \mathbb{F}$, the operator $T - \lambda\,\mathrm{id}$ is normal, and $(T - \lambda\,\mathrm{id})^* = T^* - \overline{\lambda}\,\mathrm{id}$.

(c) If V is finite-dimensional, then $\operatorname{im} T = \operatorname{im} T^* = (\ker T)^\perp$. So $V = \ker T \oplus \operatorname{im} T$, and $\ker T$ and $\operatorname{im} T$ are orthogonal.

Proof. ROUTINE

Uses. Proposition 8.1.13 (Normal means $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$), Proposition 6.1.27 (Basic properties of the norm), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.6.21 (Kernels and images of adjoints), Theorem 6.4.24 (Orthogonal decomposition)
Techniques. Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. For (a), $T\mathbf{v} = \mathbf{0}$ if and only if $\|T\mathbf{v}\| = 0$.
Hint 2. For (b), compute both $(T - \lambda\,\mathrm{id})^*(T - \lambda\,\mathrm{id})$ and $(T - \lambda\,\mathrm{id})(T - \lambda\,\mathrm{id})^*$, using $(T - \lambda\,\mathrm{id})^* = T^* - \overline{\lambda}\,\mathrm{id}$.
Hint 3. For (c), combine (a) with (c) and (d) of Theorem 6.6.21.

Sketch. (a) $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$, so one is 0 iff the other is. (b) Both products equal $T^*T - \lambda T^* - \overline{\lambda}T + |\lambda|^2 \operatorname{id}$. (c) $\operatorname{im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{im} T^*$, and $V = \ker T \oplus (\ker T)^\perp$.

(a) By Proposition 8.1.13, $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for every $\mathbf{v}$, and a vector is $\mathbf{0}$ if and only if its norm is 0 (Proposition 6.1.27). So $T\mathbf{v} = \mathbf{0}$ if and only if $T^*\mathbf{v} = \mathbf{0}$.

(b) By (a), (b), and (e) of Proposition 6.6.18, $T - \lambda \operatorname{id}$ has the adjoint $T^* - \overline{\lambda} \operatorname{id}$. Expanding,

$$\big(T^* - \overline{\lambda} \operatorname{id}\big)(T - \lambda \operatorname{id}) = T^*T - \lambda T^* - \overline{\lambda}T + |\lambda|^2 \operatorname{id},$$
$$(T - \lambda \operatorname{id})\big(T^* - \overline{\lambda} \operatorname{id}\big) = TT^* - \overline{\lambda}T - \lambda T^* + |\lambda|^2 \operatorname{id},$$

and these are equal because $T^*T = TT^*$.

(c) By (d) and (c) of Theorem 6.6.21, together with (a),

$$\operatorname{im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{im} T^*.$$

By Theorem 6.4.24, applied to the finite-dimensional subspace $\ker T$, we get $V = \ker T \oplus (\ker T)^\perp = \ker T \oplus \operatorname{im} T$, and the two summands are orthogonal by the definition of $(\ker T)^\perp$. □

Part (a) has a second proof that never mentions norms: by Proposition 6.6.23(a), applied to T and to T^* (whose adjoint is T), $\ker T = \ker(T^*T) = \ker(TT^*) = \ker T^*$.

Part (c) says that a normal operator splits a finite-dimensional V into two perpendicular pieces. It kills $\ker T$, and it maps $\operatorname{im} T$ into itself injectively, since $\operatorname{im} T \cap \ker T = \{\mathbf{0}\}$; so on $\operatorname{im} T$ it is a bijection. The nilpotent N shows what normality rules out: $\ker N = \operatorname{im} N = \operatorname{span}(\mathbf{e}_1)$, so the image of N falls into its kernel. For a normal T with $T^2 = 0$ this would put $\operatorname{im} T$ inside $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$, so $T = 0$, which is one of the three proofs in Example 6.7.3.

Part (b) is what lets us pass from kernels to eigenvectors. The move is common enough to name.

Technique 8.1.15 (Shift by a scalar). A statement about the kernel of an operator becomes a statement about eigenvectors once it is applied to $T - \lambda \operatorname{id}$, because $E_\lambda(T) = \ker(T - \lambda \operatorname{id})$ (Definition 7.2.13). So prove the statement for kernels, check that the hypotheses survive the shift $T \mapsto T - \lambda \operatorname{id}$, and apply it to the shifted operator. Hypotheses that survive include being normal, being self-adjoint when λ is real, commuting with a fixed operator, and being upper triangular in a fixed basis.

Reach for it when you want a fact about every eigenvalue and the case $\lambda = 0$ is easier to see. The canonical example is Theorem 8.1.16, which is $\ker T = \ker T^*$ applied to $T - \lambda \operatorname{id}$.

Used in. Theorem 8.1.16, Theorem 8.1.18, Exercise 8.1.28.

Theorem 8.1.16 (Eigenvectors of a normal operator and of its adjoint). Let $T \in \mathrm{L}(V)$ be normal, $\mathbf{v} \in V$, and $\lambda \in \mathbb{F}$. Then

$$T\mathbf{v} = \lambda\mathbf{v} \quad \Longleftrightarrow \quad T^*\mathbf{v} = \overline{\lambda}\mathbf{v}.$$

In particular $E_\lambda(T) = E_{\overline{\lambda}}(T^*)$, and λ is an eigenvalue of T if and only if $\overline{\lambda}$ is an eigenvalue of T^*, with the same eigenvectors.

Proof. ROUTINE

Uses. Proposition 8.1.14 (Kernels and images of normal operators), Definition 7.2.13 (Eigenspace)
Techniques. Shift by a scalar (Technique 8.1.15)
Hint 1. Rewrite $T\mathbf{v} = \lambda\mathbf{v}$ as a statement about a kernel.
Hint 2. Apply Proposition 8.1.14(a) to the normal operator $T - \lambda \operatorname{id}$, whose adjoint is $T^* - \overline{\lambda} \operatorname{id}$.
Sketch. $T\mathbf{v} = \lambda\mathbf{v}$ iff $\mathbf{v} \in \ker(T - \lambda \operatorname{id}) = \ker\big((T - \lambda \operatorname{id})^*\big) = \ker\big(T^* - \overline{\lambda} \operatorname{id}\big)$ iff $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$, using (b) and then (a) of Proposition 8.1.14.

By Proposition 8.1.14(b), $T - \lambda \operatorname{id}$ is normal with adjoint $T^* - \overline{\lambda} \operatorname{id}$. Applying Proposition 8.1.14(a) to it,

$$\ker(T - \lambda\,\mathrm{id}) = \ker\left(T^* - \overline{\lambda}\,\mathrm{id}\right).$$

The left side is $E_\lambda(T)$ and the right side is $E_{\overline{\lambda}}(T^*)$ (Definition 7.2.13), which is the claim. □

Without normality, T and T^* usually have different eigenvectors. For $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $N\mathbf{e}_1 = \mathbf{0}$, but $N^*\mathbf{e}_1 = \mathbf{e}_2$, and $E_0(N^*) = \operatorname{span}(\mathbf{e}_2)$. What survives in general is an orthogonality between the eigenvectors of T and those of T^*.

Lemma 8.1.17 (Eigenvectors of T and of T^*). Let $T \in \mathrm{L}(V)$ have an adjoint. If $T\mathbf{v} = \lambda\mathbf{v}$ and $T^*\mathbf{w} = \nu\mathbf{w}$ with $\lambda \neq \overline{\nu}$, then $\langle \mathbf{v}, \mathbf{w}\rangle = 0$.

Proof. ROUTINE

Uses. Definition 6.6.8 (Adjoint), Proposition 6.1.4 (Sesquilinearity)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint. Compute $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle$ in two ways.
Sketch. $\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle = \overline{\nu}\langle \mathbf{v}, \mathbf{w}\rangle$, and $\lambda \neq \overline{\nu}$.

By the defining identity of T^* and Proposition 6.1.4,

$$\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T^*\mathbf{w}\rangle = \langle \mathbf{v}, \nu\mathbf{w}\rangle = \overline{\nu}\langle \mathbf{v}, \mathbf{w}\rangle.$$

So $(\lambda - \overline{\nu})\langle \mathbf{v}, \mathbf{w}\rangle = 0$, and since $\lambda \neq \overline{\nu}$, $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. □

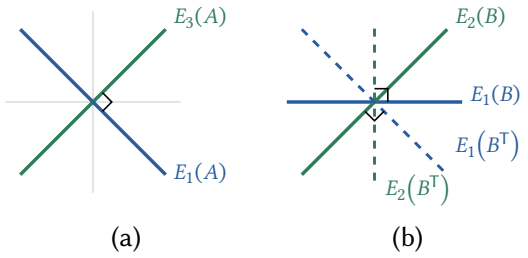


Figure 8.1.1. (a) The symmetric matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has the eigenspaces $E_1(A) = \operatorname{span}((1,-1))$ and $E_3(A) = \operatorname{span}((1,1))$, which are perpendicular. (b) $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ has the eigenspaces $E_1(B) = \operatorname{span}(\mathbf{e}_1)$ and $E_2(B) = \operatorname{span}((1,1))$ (solid), at an angle of $\pi/4$. The eigenspaces of B^{T} (dashed) are $E_1\left(B^{\mathsf{T}}\right) = \operatorname{span}((1,-1))$ and $E_2\left(B^{\mathsf{T}}\right) = \operatorname{span}(\mathbf{e}_2)$. Each eigenspace of B is perpendicular to the eigenspace of B^{T} for the other eigenvalue (Lemma 8.1.17). For a normal matrix the solid and dashed lines coincide (Theorem 8.1.16), and the eigenspaces themselves become perpendicular.

Theorem 8.1.18 (Eigenspaces of a normal operator are orthogonal). Let $T \in \mathrm{L}(V)$ be normal. If $T\mathbf{v} = \lambda\mathbf{v}$ and $T\mathbf{w} = \mu\mathbf{w}$ with $\lambda \neq \mu$, then $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. So eigenspaces of T for distinct eigenvalues are orthogonal.

Proof (through the eigenvectors of T^*). ROUTINE

Uses. Theorem 8.1.16 (Eigenvectors of a normal operator and of its adjoint), Lemma 8.1.17 (Eigenvectors of T and of T^*)
Hint 1. What does T^* do to $\mathbf{w}$?
Hint 2. $T^*\mathbf{w} = \overline{\mu}\mathbf{w}$. Apply Lemma 8.1.17 with $\nu = \overline{\mu}$.
Sketch. By Theorem 8.1.16, $T^*\mathbf{w} = \overline{\mu}\mathbf{w}$, and $\lambda \neq \mu = \overline{\overline{\mu}}$, so Lemma 8.1.17 applies.

By Theorem 8.1.16, $T^*\mathbf{w} = \overline{\mu}\mathbf{w}$. Since $\lambda \neq \mu = \overline{\overline{\mu}}$, Lemma 8.1.17 with $\nu = \overline{\mu}$ gives $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. For arbitrary $\mathbf{v} \in E_\lambda(T)$ and $\mathbf{w} \in E_\mu(T)$, either one of them is $\mathbf{0}$ or both are eigenvectors, so the eigenspaces are orthogonal. □

Proof (through kernels and images). STANDARD

Uses. Proposition 8.1.14 (Kernels and images of normal operators), Theorem 6.6.21 (Kernels and images of adjoints), Definition 6.4.22 (Orthogonal complement)
Techniques. Shift by a scalar (Technique 8.1.15)
Hint 1. Let $R = T - \mu\,\mathrm{id}$. Then $\mathbf{w} \in \ker R$. Show that $\mathbf{v} \in \operatorname{im} R$.

Hint 2. $R\mathbf{v} = (\lambda - \mu)\mathbf{v}$ with $\lambda - \mu \neq 0$. Now relate $\ker R$ to $\operatorname{im} R$ using that R is normal.
Sketch. With $R = T - \mu\,\mathrm{id}$, $\mathbf{v} = R(\mathbf{v}/(\lambda - \mu)) \in \operatorname{im} R$ and $\mathbf{w} \in \ker R = \ker R^* = (\operatorname{im} R)^\perp$, by Proposition 8.1.14 and Theorem 6.6.21(a).

Let $R = T - \mu\,\mathrm{id}$, which is normal by Proposition 8.1.14(b). Since $R\mathbf{v} = (\lambda - \mu)\mathbf{v}$ and $\lambda - \mu \neq 0$, we have $\mathbf{v} = R\big((\lambda - \mu)^{-1}\mathbf{v}\big) \in \operatorname{im} R$. On the other hand $\mathbf{w} \in \ker R$, and by Proposition 8.1.14(a) and Theorem 6.6.21(a),

$$\ker R = \ker R^* = (\operatorname{im} R)^\perp.$$

So $\mathbf{w}$ is orthogonal to every vector of $\operatorname{im} R$, in particular to $\mathbf{v}$. The statement about eigenspaces follows as in the first proof. □

The first proof reduces the theorem to a fact about single eigenvectors: T^* has the same eigenvectors as T, so the orthogonality of Lemma 8.1.17 applies to eigenvectors of T alone. The second proof is about subspaces and proves more. Every eigenvector $\mathbf{w}$ for μ is orthogonal to all of $\operatorname{im}(T - \mu\,\mathrm{id})$, which contains the eigenvectors for every other eigenvalue but in general much more. Part (b) of Proposition 8.1.1 is the special case of self-adjoint T.

Exercise 8.1.19. ROUTINE Let V be a complex inner product space, let $T \in \mathrm{L}(V)$ be normal, and let $\mathbf{v}, \mathbf{w} \in V$ satisfy $T\mathbf{v} = 2\mathbf{v}$, $T\mathbf{w} = i\mathbf{w}$, $\|\mathbf{v}\| = 1$, and $\|\mathbf{w}\| = 3$. Compute $\|T(\mathbf{v} + \mathbf{w})\|$, $\|T^*(\mathbf{v} - \mathbf{w})\|$, and $\langle T(\mathbf{v} + \mathbf{w}), \mathbf{v} - \mathbf{w}\rangle$.

Solution on page 1632

Applied to the special classes of Example 8.1.9, Theorem 8.1.16 says what their eigenvalues can be.

Corollary 8.1.20 (Eigenvalues of self-adjoint, skew-adjoint, and unitary operators). Let $T \in \mathrm{L}(V)$ have an adjoint, and let λ be an eigenvalue of T.

(a) If $T^* = T$, then $\lambda \in \mathbb{R}$.

(b) If $T^* = -T$, then $\overline{\lambda} = -\lambda$, that is, $\lambda \in i\mathbb{R}$.

(c) If $T^*T = TT^* = \mathrm{id}$, then $|\lambda| = 1$.

Proof. ROUTINE
Uses. Example 8.1.9 (Normal operators), Theorem 8.1.16 (Eigenvectors of a normal operator and of its adjoint), Proposition 1.2.35 (Zero products)
Hint 1. Each such T is normal. Let $\mathbf{v}$ be an eigenvector for λ. What is $T^*\mathbf{v}$?
Hint 2. $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$. Compare with what the hypothesis says about $T^*\mathbf{v}$, or for (c) compute $T^*T\mathbf{v}$.
Sketch. Each T is normal, so $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$. (a) $\overline{\lambda}\mathbf{v} = T\mathbf{v} = \lambda\mathbf{v}$. (b) $\overline{\lambda}\mathbf{v} = -\lambda\mathbf{v}$. (c) $\mathbf{v} = T^*T\mathbf{v} = \lambda\overline{\lambda}\mathbf{v}$.

In each case T is normal (Example 8.1.9). Let $\mathbf{v}$ be an eigenvector for λ; by Theorem 8.1.16, $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$.

(a) $\overline{\lambda}\mathbf{v} = T^*\mathbf{v} = T\mathbf{v} = \lambda\mathbf{v}$, so $\big(\overline{\lambda} - \lambda\big)\mathbf{v} = \mathbf{0}$, and $\overline{\lambda} = \lambda$ since $\mathbf{v} \neq \mathbf{0}$ (Proposition 1.2.35).

(b) $\overline{\lambda}\mathbf{v} = T^*\mathbf{v} = -T\mathbf{v} = -\lambda\mathbf{v}$, so $\overline{\lambda} = -\lambda$ in the same way.

(c) $\mathbf{v} = T^*T\mathbf{v} = T^*(\lambda\mathbf{v}) = \lambda\overline{\lambda}\mathbf{v} = |\lambda|^2\mathbf{v}$, so $|\lambda|^2 = 1$. □

Part (a) is a third proof of Proposition 8.1.1(a), and part (c) returns in Section 8.4. The converses fail without normality: $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ has real eigenvalues and is not symmetric, and the shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has only the eigenvalue 1 and is not orthogonal. For normal operators on finite-dimensional complex spaces the converses hold, because by Section 8.2 such an operator is determined by its eigenvalues on an orthonormal eigenbasis.

8.1.5 The Cartesian decomposition BEYOND 21-242

The Cartesian decomposition is in LADR 7A but not in the lecture.

Every complex number is $z = x + iy$ with x, y real, and then $|z|^2 = x^2 + y^2$. Operators split the same way: $T = S + K$ with $S = (T + T^*)/2$ self-adjoint and $K = (T - T^*)/2$ skew-adjoint, and this is the only such

splitting (over $\mathbb{C}$ this is Exercise 6.6.37, with $K = iB$). For numbers the two parts commute automatically. For operators they need not, and their commuting is exactly normality.

Proposition 8.1.21 (Normal means commuting parts). Let $T \in \mathrm{L}(V)$ have an adjoint, and let $S = (T + T^*)/2$ and $K = (T - T^*)/2$. Then S is self-adjoint, K is skew-adjoint, $T = S + K$, and

$$T^*T - TT^* = 2(SK - KS).$$

Hence T is normal if and only if $SK = KS$. When $\mathbb{F} = \mathbb{C}$, writing $K = iB$ with $B = (T - T^*)/(2i)$ self-adjoint, T is normal if and only if $SB = BS$.

Proof. ROUTINE

Uses. Proposition 6.6.18 (Properties of the adjoint), Definition 8.1.5 (Skew-adjoint operators, skew-Hermitian matrices)
Hint 1. $T^* = S^* + K^* = S - K$.
Hint 2. Expand $(S - K)(S + K) - (S + K)(S - K)$.
Sketch. $S^* = S$ and $K^* = -K$ by Proposition 6.6.18, so $T^* = S - K$. Then $T^*T - TT^* = (S - K)(S + K) - (S + K)(S - K) = 2(SK - KS)$. Over $\mathbb{C}$, $SK - KS = i(SB - BS)$.

By (a)–(c) of Proposition 6.6.18, $S^* = (T^* + T)/2 = S$ and $K^* = (T^* - T)/2 = -K$, and clearly $S + K = T$. So $T^* = S - K$, and

$$\begin{aligned} T^*T - TT^* &= (S - K)(S + K) - (S + K)(S - K) \\ &= \left(S^2 + SK - KS - K^2\right) - \left(S^2 - SK + KS - K^2\right) = 2(SK - KS). \end{aligned}$$

So $T^*T = TT^*$ if and only if $SK = KS$. When $\mathbb{F} = \mathbb{C}$, $B = -iK$ satisfies $B^* = \overline{-i}K^* = i(-K) = B$, so B is self-adjoint, $K = iB$, and $SK - KS = i(SB - BS)$, which vanishes if and only if $SB = BS$. □

For $B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, the parts are $S = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 2 \end{pmatrix}$ and $K = \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix}$, and

$$2(SK - KS) = 2\left(\begin{pmatrix} -1/4 & 1/2 \\ -1 & 1/4 \end{pmatrix} - \begin{pmatrix} 1/4 & 1 \\ -1/2 & -1/4 \end{pmatrix}\right) = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = B^{\mathsf{T}}B - BB^{\mathsf{T}},$$

as the proposition predicts, so the parts of B do not commute.

Intuition. For a normal operator, $T^*T = (S - K)(S + K) = S^2 - K^2$, and for every $\mathbf{v}$,

$$\|T\mathbf{v}\|^2 = \langle S^2\mathbf{v}, \mathbf{v}\rangle - \langle K^2\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2 + \|K\mathbf{v}\|^2,$$

since $-\langle K^2\mathbf{v}, \mathbf{v}\rangle = -\langle K\mathbf{v}, K^*\mathbf{v}\rangle = \|K\mathbf{v}\|^2$. This is $|z|^2 = x^2 + y^2$ for operators, and Exercise 8.1.29 shows that it characterizes normal operators.

Exercises

Exercise 8.1.22. ROUTINE For each matrix, decide whether it is Hermitian, skew-Hermitian, unitary, or normal (any number of these may hold).

(a) $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

(b) $\begin{pmatrix} i & 1 \\ -1 & 2i \end{pmatrix}$.

(c) $\begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}$.

(d) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(e) $\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$.

Solution on page 1632

Exercise 8.1.23. ROUTINE Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

(a) Find the eigenvalues and eigenspaces of A. (Write $A = I_3 + \mathbf{1}\mathbf{1}^{\mathsf{T}}$ with $\mathbf{1} = (1,1,1)$.)

(b) Check that the eigenspaces are orthogonal, and find two eigenvectors for the same eigenvalue that are not orthogonal.

(c) Find an orthonormal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.

Solution on page 1633

Exercise 8.1.24. STANDARD Let V be a finite-dimensional inner product space and $S, T \in \mathrm{L}(V)$. Prove or disprove each statement.

(a) If S and T are normal, then $S + T$ is normal.

(b) If T^2 is normal, then T is normal.

(c) If $\ker T = \ker T^*$, then T is normal.

(d) If T is normal, then $T + cT^*$ is normal for every $c \in \mathbb{F}$.

(e) If $\mathbb{F} = \mathbb{R}$, T is normal, and $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v} \in V$, then $T = 0$.

Solution on page 1633

Exercise 8.1.25. STANDARD Show that $A \in \mathbb{R}^{2\times 2}$ is normal if and only if either A is symmetric or $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ for some $a, b \in \mathbb{R}$. Describe the second family geometrically, and show that for $b \neq 0$ it has no real eigenvalues.

Solution on page 1634

Exercise 8.1.26. STANDARD Let $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ be nonzero, and let $A = \mathbf{x}\mathbf{y}^* \in \mathbb{C}^{n\times n}$. Show that A is normal if and only if $\mathbf{y} = c\mathbf{x}$ for some $c \in \mathbb{C}$, and that in this case A is Hermitian if and only if $c \in \mathbb{R}$.

Solution on page 1634

Exercise 8.1.27. STANDARD Find the error in the following argument, and give an operator for which its conclusion fails. Then name a hypothesis on T under which the conclusion is true.

Claim. Let $T \in \mathrm{L}(V)$ have an adjoint. If $T\mathbf{v} = \lambda\mathbf{v}$, then $T^*\mathbf{v} = \overline{\lambda}\mathbf{v}$.

Proof. We have $\langle T^*\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle = \langle \overline{\lambda}\mathbf{v}, \mathbf{v}\rangle$. Since $T^*\mathbf{v}$ and $\overline{\lambda}\mathbf{v}$ have the same inner product with $\mathbf{v}$, they are equal.

Solution on page 1635

Exercise 8.1.28. STANDARD Let V be a complex inner product space, $T \in \mathrm{L}(V)$ self-adjoint, and $\lambda = a + bi$ with $a, b \in \mathbb{R}$.

(a) Show that $\|(T - \lambda\,\mathrm{id})\mathbf{v}\|^2 = \|(T - a\,\mathrm{id})\mathbf{v}\|^2 + b^2\|\mathbf{v}\|^2$ for every $\mathbf{v} \in V$.

(b) Deduce once more that the eigenvalues of T are real.

(c) Suppose V is finite-dimensional and $b \neq 0$. Show that $T - \lambda\,\mathrm{id}$ is invertible and that $\left\|(T - \lambda\,\mathrm{id})^{-1}\mathbf{w}\right\| \leq \|\mathbf{w}\|/|b|$ for every $\mathbf{w} \in V$.

Solution on page 1635

Exercise 8.1.29. STANDARD BEYOND 21-242 Let $T \in \mathrm{L}(V)$ have an adjoint, and let $S = (T + T^*)/2$ and $K = (T - T^*)/2$ be its self-adjoint and skew-adjoint parts.

(a) Show that $\|T\mathbf{v}\|^2 = \|S\mathbf{v}\|^2 + \|K\mathbf{v}\|^2 + \langle (SK - KS)\mathbf{v}, \mathbf{v}\rangle$ for every $\mathbf{v} \in V$.

(b) Deduce that T is normal if and only if $\|T\mathbf{v}\|^2 = \|S\mathbf{v}\|^2 + \|K\mathbf{v}\|^2$ for every $\mathbf{v} \in V$.

Solution on page 1636

Exercise 8.1.30. STANDARD Let $U \in \mathbb{F}^{n\times n}$ be both Hermitian and unitary.

(a) Show that $U^2 = I_n$.

(b) Show that $P = (I_n + U)/2$ satisfies $P^2 = P = P^*$, so that $\mathbf{x} \mapsto P\mathbf{x}$ is an orthogonal projection of $\mathbb{F}^n$, and that $U = 2P - I_n$. Describe $\mathbf{x} \mapsto U\mathbf{x}$ geometrically.

(c) Conversely, show that $2P - I_n$ is Hermitian and unitary whenever $P^2 = P = P^*$.

(d) Find U when P is the matrix of the orthogonal projection of $\mathbb{R}^2$ onto $\operatorname{span}((1, 1))$.

Solution on page 1636

Exercise 8.1.31. STANDARD Theorem 8.1.3 says that every real symmetric matrix has a real eigenvalue. Prove or disprove the analogues over other fields.

(a) Every symmetric $A \in \mathbb{Q}^{2\times 2}$ has an eigenvalue in $\mathbb{Q}$.

(b) Every symmetric $A \in \mathbb{F}_2^{2\times 2}$ has an eigenvalue in $\mathbb{F}_2$.

(c) Which steps of the first proof of Theorem 8.1.3 fail over $\mathbb{Q}$ and over $\mathbb{F}_2$?

Solution on page 1637

Exercise 8.1.32. HARD Let V be finite-dimensional and let $T \in \mathrm{L}(V)$ be normal with $T^3 = T^2$. Show that T is an orthogonal projection, that is, $T = \mathrm{Pr}_W$ for some subspace W of V.

Solution on page 1637

Exercise 8.1.33. HARD Let $A \in \mathbb{F}^{n\times n}$ satisfy $\|A^*\mathbf{x}\| \le \|A\mathbf{x}\|$ for every $\mathbf{x} \in \mathbb{F}^n$ (dot product). Show that A is normal.

Solution on page 1637

Exercise 8.1.34. HARD Let $A \in \mathbb{R}^{n\times n}$ satisfy $A^{\mathsf{T}} = A^2$.

(a) Show that $A^4 = A$ and that A is normal.

(b) Show that A^3 is the matrix of an orthogonal projection.

(c) Find all such A when $n = 2$.

Solution on page 1638

Exercise 8.1.35. HARD Let $K \in \mathbb{F}^{n\times n}$ be skew-Hermitian.

(a) Show that $I_n + K$ and $I_n - K$ are invertible.

(b) Show that $Q = (I_n - K)(I_n + K)^{-1}$ is unitary and that -1 is not an eigenvalue of Q.

(c) Compute Q for $K = \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix}$ with $t \in \mathbb{R}$, and identify it as a rotation.

Solution on page 1639

Exercise 8.1.36. STANDARD Let V be finite-dimensional and let $T \in \mathrm{L}(V)$ be normal and diagonalizable. Show that V has an orthonormal basis consisting of eigenvectors of T. Deduce that a diagonalizable operator has an orthonormal eigenbasis if and only if it is normal.

Solution on page 1639

Exercise 8.1.37. HARD Let $A \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, listed with algebraic multiplicity.

(a) Prove *Schur's inequality*

$$\sum_{j=1}^{n} |\lambda_j|^2 \leq \operatorname{tr}(A^*A) = \sum_{i,j} |a_{ij}|^2.$$

(b) Show that if equality holds, then A is normal.

Solution on page 1640

Exercise 8.1.38. CHALLENGE Let $A \in \mathbb{C}^{n\times n}$ be normal and let $W \subseteq \mathbb{C}^n$ be a subspace with $AW \subseteq W$. Show that $A^*W \subseteq W$, and deduce that $AW^\perp \subseteq W^\perp$. (Let P be the matrix of Pr_W, so that $AW \subseteq W$ says $PAP = AP$, and consider $X = PA(I_n - P)$.)

Solution on page 1640

8.2 The spectral theorems

An operator is easiest to understand when it is diagonal in some basis. On an inner product space we want more: a basis that is diagonal for T *and* orthonormal, so that T stretches V by scalar factors along mutually perpendicular axes, and coordinates are computed by inner products instead of linear systems. This section answers the question of which operators have such a basis. Over $\mathbb{C}$ the answer is the normal operators (Definition 8.1.8), and over $\mathbb{R}$ it is the self-adjoint ones. Both answers come from the same short computation, which shows that these conditions are necessary; the work is in showing that they suffice.

Throughout, V is a finite-dimensional inner product space over $\mathbb{F}$, with $\dim V \geq 1$, and $T \in \mathrm{L}(V)$. As in Section 6.6, inner products are linear in the first slot. For an orthonormal basis $\mathcal{O}$ we write $[T]_{\mathcal{O}}$ for $[T]_{\mathcal{O}\mathcal{O}}$, and we use freely that $[T^*]_{\mathcal{O}} = \big([T]_{\mathcal{O}}\big)^*$ (Theorem 6.6.15) and $[ST]_{\mathcal{O}} = [S]_{\mathcal{O}}[T]_{\mathcal{O}}$ (Theorem 2.4.2).

8.2.1 Orthonormal eigenbases

An **orthonormal eigenbasis** for T is an orthonormal basis of V consisting of eigenvectors of T. By Theorem 7.4.2, an orthonormal basis $\mathcal{O}$ is an eigenbasis exactly when $[T]_{\mathcal{O}}$ is diagonal. For matrices, the change-of-basis matrix is then unitary.

Definition 8.2.1 (Unitarily and orthogonally diagonalizable). A matrix $A \in \mathbb{F}^{n\times n}$ is **unitarily diagonalizable** if

$$A = UDU^*$$

for some unitary $U \in \mathbb{F}^{n\times n}$ and some diagonal $D \in \mathbb{F}^{n\times n}$. When $\mathbb{F} = \mathbb{R}$, U is then an orthogonal matrix Q and $A = QDQ^{\mathsf{T}}$; we say A is **orthogonally diagonalizable**.

Since $U^* = U^{-1}$ for unitary U (Definition 6.3.33), $A = UDU^*$ is a diagonalization $A = SDS^{-1}$ with $S = U$. By the last part of Theorem 7.4.2, it says that the columns of U are eigenvectors of A with eigenvalues the diagonal entries of D; and by Theorem 6.3.29, U is unitary exactly when its columns are orthonormal. So A is unitarily diagonalizable if and only if $\mathbb{F}^n$, with the dot product, has an orthonormal eigenbasis for A.

We start with the easy direction of both spectral theorems, because its proof shows where the conditions "normal" and "self-adjoint" come from.

Proposition 8.2.2 (Necessary conditions). Suppose V has an orthonormal eigenbasis for T.

(a) T is normal: $T^*T = TT^*$.

(b) If every eigenvalue of T is real (in particular, if $\mathbb{F} = \mathbb{R}$), then T is self-adjoint.

Proof. ROUTINE

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 6.6.15 (Existence of adjoints), Theorem 2.4.2 (The matrix of a composite), Theorem 2.3.20 (Linear maps correspond to matrices), Definition 8.1.8 (Normal operator, normal matrix), Definition 6.6.26 (Self-adjoint operator)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint 1. Write down $[T]_{\mathcal{O}}$ for the orthonormal eigenbasis $\mathcal{O}$. What is $[T^*]_{\mathcal{O}}$?
Hint 2. Diagonal matrices commute with each other, and $D^* = D$ for a real diagonal D.
Sketch. In the eigenbasis $\mathcal{O}$, $[T]_{\mathcal{O}} = D$ is diagonal, so $[T^*]_{\mathcal{O}} = D^*$ is diagonal too. Diagonal matrices commute, so $[T^*T]_{\mathcal{O}} = D^*D = DD^* = [TT^*]_{\mathcal{O}}$. If the eigenvalues are real, $D^* = D$, so $T^* = T$.

Let $\mathcal{O}$ be an orthonormal eigenbasis, with $T\mathbf{u}_j = \lambda_j \mathbf{u}_j$. By Theorem 7.4.2, $[T]_{\mathcal{O}} = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and by Theorem 6.6.15, $[T^*]_{\mathcal{O}} = D^* = \operatorname{diag}\left(\overline{\lambda_1}, \ldots, \overline{\lambda_n}\right)$.

(a) Products of diagonal matrices are computed entry by entry, so $D^*D = \operatorname{diag}\left(|\lambda_1|^2, \ldots, |\lambda_n|^2\right) = DD^*$. By Theorem 2.4.2, $[T^*T]_{\mathcal{O}} = D^*D = DD^* = [TT^*]_{\mathcal{O}}$, and maps with the same matrix are equal (Theorem 2.3.20). So $T^*T = TT^*$.

(b) If every λ_j is real, then $D^* = D$, so $[T^*]_{\mathcal{O}} = [T]_{\mathcal{O}}$ and $T^* = T$ by Theorem 2.3.20. □

So normality is what remains of "diagonal in an orthonormal basis" once the basis is forgotten: a diagonal matrix commutes with its conjugate transpose. The spectral theorems say that nothing else is needed. Before proving them, here are the examples that show which hypothesis goes with which field.

Example 8.2.3 (Which matrices have orthonormal eigenbases).

(a) $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ has the distinct eigenvalues 1 and 2, so it is diagonalizable (Corollary 7.4.5), with eigenvectors $(1, 0)$ and $(1, 1)$. These are not orthogonal, and each eigenspace is a line, so no choice of eigenvectors is orthogonal. Correspondingly A is not normal: $(A^*A)_{11} = 1$ while $(AA^*)_{11} = 2$. Diagonalizable and unitarily diagonalizable are different conditions.

(b) The rotation $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ satisfies $J^{\mathsf{T}}J = I_2 = JJ^{\mathsf{T}}$, so T_J is normal on $\mathbb{R}^2$ and on $\mathbb{C}^2$. Over $\mathbb{R}$ it has no eigenvector at all (Example 7.2.18). Over $\mathbb{C}$,

$$J\begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = i\begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad J\begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix} = -i\begin{pmatrix} 1 \\ i \end{pmatrix},$$

and $\frac{1}{\sqrt{2}}(1, -i)$, $\frac{1}{\sqrt{2}}(1, i)$ is an orthonormal eigenbasis of $\mathbb{C}^2$, since $(1, -i) \cdot (1, i) = 1 + (-i)\bar{i} = 1 - 1 = 0$. So over $\mathbb{C}$ normality is the right condition, while over $\mathbb{R}$ we will need the stronger condition $T^* = T$.

Intuition (Two ways to build the basis). Suppose we want an orthonormal eigenbasis for a normal T on a complex space. There are two natural plans, and both work.

The first plan starts from something we already have. Schur's theorem (Theorem 7.5.18) gives an orthonormal basis in which T is upper triangular, so it remains to show that normality kills everything above the diagonal. For an upper-triangular R, the first column has only one nonzero entry, r_{11}, while the first row is $r_{11}, r_{12}, \ldots, r_{1n}$. The diagonal entries of R^*R and RR^* are the squared norms of the columns and of the rows, so $R^*R = RR^*$ forces the first row to have no more mass than the first column. Then the same comparison works for the second row, and so on down the diagonal.

The second plan builds the basis one vector at a time (Technique 7.5.12). Take a unit eigenvector $\mathbf{u}$ and try to recurse on $W = \operatorname{span}(\mathbf{u})^{\perp}$. In the second proof of Schur's theorem this required compressing T to W, because W need not be T-invariant, and the compression produced the entries above the diagonal. For a normal T, $\mathbf{u}$ is also an eigenvector of T^*, and this is exactly what makes W invariant (Exercise 6.6.38). The compression is then a restriction, and nothing appears above the diagonal.

8.2.2 The complex spectral theorem

We start with the matrix fact behind the first plan.

Lemma 8.2.4 (Triangular and normal means diagonal). Let $R \in \mathbb{F}^{n\times n}$ be upper triangular with $R^*R = RR^*$. Then R is diagonal.

Proof. STANDARD

Uses. Definition 2.4.1 (Matrix product), Definition 6.3.26 (Conjugate transpose)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Compute the diagonal entries $(R^*R)_{kk}$ and $(RR^*)_{kk}$ as sums of $|r_{ij}|^2$. Which entries of R appear in each?
Hint 2. $(R^*R)_{kk}$ is the squared norm of column k and $(RR^*)_{kk}$ that of row k. Start with $k = 1$, where the column has a single possibly nonzero entry.
Hint 3. Prove by strong induction on k that row k has no nonzero entries to the right of the diagonal. When you reach row k, the entries of column k above the diagonal lie in earlier rows.
Sketch. $(R^*R)_{kk} = \sum_i |r_{ik}|^2$ is the squared norm of column k, and $(RR^*)_{kk} = \sum_j |r_{kj}|^2$ that of row k. By strong induction on k: if rows $1, \dots, k-1$ vanish off the diagonal, then column k has only the entry r_{kk}, so equating the two norms gives $\sum_{j>k} |r_{kj}|^2 = 0$, and row k vanishes off the diagonal too.

For each k, the definitions of the product and of R^* give

$$(R^*R)_{kk} = \sum_{i=1}^{n} \overline{r_{ik}} r_{ik} = \sum_{i=1}^{n} |r_{ik}|^2, \quad (RR^*)_{kk} = \sum_{j=1}^{n} r_{kj}\overline{r_{kj}} = \sum_{j=1}^{n} |r_{kj}|^2,$$

the squared norms of column k and of row k. We prove by strong induction on k that $r_{kj} = 0$ for all $j > k$. Let $1 \le k \le n$ and suppose this holds for every row $i < k$. Then $r_{ik} = 0$ for $i < k$ by the hypothesis (as $k > i$), and $r_{ik} = 0$ for $i > k$ because R is upper triangular. So column k has the single entry r_{kk}, and $(R^*R)_{kk} = (RR^*)_{kk}$ reads

$$|r_{kk}|^2 = |r_{kk}|^2 + \sum_{j>k} |r_{kj}|^2.$$

Hence $\sum_{j>k} |r_{kj}|^2 = 0$. A sum of nonnegative reals vanishes only if every term does, so $r_{kj} = 0$ for all $j > k$. (For $k = 1$ the hypothesis is empty and the same argument applies.) So every entry above the diagonal is 0, and R is diagonal. □

Theorem 8.2.5 (Complex spectral theorem). Let V be a finite-dimensional complex inner product space with $\dim V \ge 1$, and let $T \in \mathrm{L}(V)$. The following are equivalent:

(I) T is normal;

(II) V has an orthonormal basis consisting of eigenvectors of T;

(III) $[T]_{\mathcal{O}}$ is diagonal for some orthonormal basis $\mathcal{O}$ of V.

In matrix form: $A \in \mathbb{C}^{n\times n}$ satisfies $A^*A = AA^*$ if and only if A is unitarily diagonalizable.

(II) and (III) say the same thing by Theorem 7.4.2, and (II) ⇒ (I) is Proposition 8.2.2(a). The content is (I) ⇒ (III). We give three proofs. The first two carry out the two plans of Intuition (Two ways to build the basis); the third avoids induction altogether.

Proof (Schur, then normality). ROUTINE

Uses. Theorem 7.5.18 (Schur's theorem), Lemma 8.2.4 (Triangular and normal means diagonal), Theorem 6.6.15 (Existence of adjoints), Theorem 2.4.2 (The matrix of a composite), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Proposition 8.2.2 (Necessary conditions), Theorem 4.5.17 (Change of basis), Definition 6.3.33 (Unitary and orthogonal matrices)
Techniques. Proving TFAE (Technique 2.7.1), Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. Start from a Schur basis: an orthonormal $\mathcal{O}$ with $R = [T]_{\mathcal{O}}$ upper triangular. What does normality of T say about R?
Hint 2. Since $\mathcal{O}$ is orthonormal, $[T^*]_{\mathcal{O}} = R^*$. Apply Lemma 8.2.4.
Sketch. (II) $\Leftrightarrow$ (III) is Theorem 7.4.2, and (II) $\Rightarrow$ (I) is Proposition 8.2.2. For (I) $\Rightarrow$ (III), Schur gives an orthonormal $\mathcal{O}$ with $R = [T]_{\mathcal{O}}$ upper triangular, and $R^*R = [T^*T]_{\mathcal{O}} = [TT^*]_{\mathcal{O}} = RR^*$. By Lemma 8.2.4, R is diagonal. For matrices, apply this to T_A on $\mathbb{C}^n$ and use that the columns of an orthonormal basis form a unitary matrix.

(II) and (III) are equivalent by Theorem 7.4.2, and (II) $\Rightarrow$ (I) is Proposition 8.2.2(a). Suppose T is normal. By Theorem 7.5.18 there is an orthonormal basis $\mathcal{O}$ with $R = [T]_{\mathcal{O}}$ upper triangular. By Theorem 6.6.15 and Theorem 2.4.2,

$$R^*R = [T^*]_{\mathcal{O}}[T]_{\mathcal{O}} = [T^*T]_{\mathcal{O}} = [TT^*]_{\mathcal{O}} = RR^*.$$

By Lemma 8.2.4, R is diagonal, which is (III).

For the matrix form, if $A = UDU^*$ with U unitary and D diagonal, then $A^*A = UD^*DU^* = UDD^*U^* = AA^*$. Conversely, if $A^*A = AA^*$, then T_A is normal on $\mathbb{C}^n$ with the dot product (Example 6.6.10), so by (II) there is an orthonormal eigenbasis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of $\mathbb{C}^n$. The matrix U with these columns is unitary (Definition 6.3.33), and by Theorem 4.5.17, $U^*AU = U^{-1}AU = [T_A]_{\mathcal{O}} = D$ is diagonal. Hence $A = UDU^*$. □

Intuition (What the second proof needs from normality). The induction runs on $W = \operatorname{span}(\mathbf{u})^{\perp}$, and two facts about normal operators make it work. Both have one-line proofs, given inside the proof below; Section 8.1 treats them properly.

- *T and T^* stretch every vector equally:* $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for every $\mathbf{v}$. Applied to the normal operator $T - \lambda\,\mathrm{id}$, this says that $T\mathbf{u} = \lambda\mathbf{u}$ implies $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$.
- *Restrictions stay normal:* if W is invariant under both T and T^*, then the adjoint of $T|_W$ is $T^*|_W$.

Proof (induction with invariant complements). STANDARD

Uses. Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Definition 8.1.8 (Normal operator, normal matrix), Definition 6.6.8 (Adjoint), Proposition 6.6.18 (Properties of the adjoint), Proposition 6.6.9 (Adjoints are unique and linear), Exercise 6.6.38, Theorem 6.4.24 (Orthogonal decomposition), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 4.4.13 (Bases and dimension of a sum), Definition 6.3.15 (Orthonormal list, orthonormal basis)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Induct on $n = \dim V$. Take a unit eigenvector $\mathbf{u}$ with $T\mathbf{u} = \lambda\mathbf{u}$, and let $W = \operatorname{span}(\mathbf{u})^{\perp}$. You want W to be T-invariant. By Exercise 6.6.38, what would suffice?
Hint 2. It suffices that $\operatorname{span}(\mathbf{u})$ is T^*-invariant, that is, that $\mathbf{u}$ is an eigenvector of T^*. Show $\|T\mathbf{v}\| = \|T^*\mathbf{v}\|$ for normal T, and apply it to $T - \lambda\,\mathrm{id}$.
Hint 3. To apply the induction hypothesis to $S = T|_W$, check that W is also T^*-invariant and that $S^* = T^*|_W$. Then S is normal.
Sketch. For normal T, $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \|T^*\mathbf{v}\|^2$; since $T - \lambda\,\mathrm{id}$ is normal, $T\mathbf{u} = \lambda\mathbf{u}$ gives $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$. Take a unit eigenvector $\mathbf{u}$ (it exists over $\mathbb{C}$). Then $\operatorname{span}(\mathbf{u})$ is invariant under T and T^*, so $W = \operatorname{span}(\mathbf{u})^{\perp}$ is too (Exercise 6.6.38). On W, $T|_W$ has adjoint $T^*|_W$, so it is normal, and by induction W has an orthonormal eigenbasis. Add $\mathbf{u}$.

We prove (I) $\Rightarrow$ (II) by induction on $n = \dim V$; the rest is as in the first proof. First, two facts about a normal operator N on an inner product space.

- For every $\mathbf{v}$, $\|N\mathbf{v}\| = \|N^*\mathbf{v}\|$, because

$$\|N\mathbf{v}\|^2 = \langle N\mathbf{v}, N\mathbf{v}\rangle = \langle N^*N\mathbf{v}, \mathbf{v}\rangle = \langle NN^*\mathbf{v}, \mathbf{v}\rangle = \langle N^*\mathbf{v}, N^*\mathbf{v}\rangle = \|N^*\mathbf{v}\|^2.$$

- For every $\lambda \in \mathbb{C}$, $N - \lambda\,\mathrm{id}$ is normal. Indeed $(N - \lambda\,\mathrm{id})^* = N^* - \overline{\lambda}\,\mathrm{id}$ by Proposition 6.6.18, and expanding,

$$(N - \lambda\,\mathrm{id})\big(N^* - \overline{\lambda}\,\mathrm{id}\big) = NN^* - \overline{\lambda}N - \lambda N^* + |\lambda|^2\,\mathrm{id} = \big(N^* - \overline{\lambda}\,\mathrm{id}\big)(N - \lambda\,\mathrm{id}),$$

using $NN^* = N^*N$. So if $N\mathbf{u} = \lambda\mathbf{u}$, the first fact applied to $N - \lambda\,\mathrm{id}$ gives $\big\|\big(N^* - \overline{\lambda}\,\mathrm{id}\big)\mathbf{u}\big\| = \|(N - \lambda\,\mathrm{id})\mathbf{u}\| = 0$, that is, $N^*\mathbf{u} = \overline{\lambda}\mathbf{u}$.

Now let T be normal. If $n = 1$, any unit vector $\mathbf{u}$ spans V, so $T\mathbf{u} \in \operatorname{span}(\mathbf{u})$, and $\mathbf{u}$ alone is an orthonormal eigenbasis. Let $n \geq 2$, and assume (I) $\Rightarrow$ (II) for complex inner product spaces of dimension $n - 1$. By Theorem 7.3.27, T has an eigenvector; dividing by its norm, we get $\mathbf{u}$ with $\|\mathbf{u}\| = 1$ and $T\mathbf{u} = \lambda\mathbf{u}$. By the facts above, $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$. Let $U = \operatorname{span}(\mathbf{u})$ and $W = U^\perp$. Then U is invariant under T and under T^*. Applying Exercise 6.6.38 to T shows that T^* maps W into W, and applying it to T^* (whose adjoint is T, by Proposition 6.6.18(c)) shows that T maps W into W.

Let $S = T{\restriction}_W \in \mathrm{L}(W)$, where W carries the inner product of V. For $\mathbf{w}, \mathbf{w}' \in W$ we have $T^*{\restriction}_W\mathbf{w}' = T^*\mathbf{w}' \in W$ and

$$\langle S\mathbf{w}, \mathbf{w}'\rangle = \langle T\mathbf{w}, \mathbf{w}'\rangle = \langle \mathbf{w}, T^*\mathbf{w}'\rangle,$$

so $T^*{\restriction}_W$ is the adjoint of S (Definition 6.6.8, Proposition 6.6.9). Hence $S^*S = (T^*T){\restriction}_W = (TT^*){\restriction}_W = SS^*$, and S is normal. By Theorem 6.4.24 and Corollary 6.4.25, $V = U \oplus W$ and $\dim W = n - 1$. By the induction hypothesis, W has an orthonormal basis $\mathbf{u}_2, \ldots, \mathbf{u}_n$ of eigenvectors of S, which are eigenvectors of T. Each $\mathbf{u}_j$ lies in $W = U^\perp$, so $\mathbf{u}, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is orthonormal (Definition 6.3.15), and it is a basis of $V = U \oplus W$ by Theorem 4.4.13. It consists of eigenvectors of T. □

The third proof looks at all eigenvectors at once. Let E be the sum of all the eigenspaces of T. If $E \neq V$, we look for an eigenvector of T in $E^\perp$, which is impossible since every eigenvector lies in E. Finding one needs $E^\perp$ to be T-invariant, and once again that comes from T^* preserving eigenspaces.

Proof (the sum of the eigenspaces is everything). STANDARD

> *Uses.* Definition 7.2.13 (Eigenspace), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Exercise 6.6.38, Corollary 6.4.27 (Testing whether a subspace is everything), Theorem 6.4.20 (Orthonormal bases exist), Definition 6.3.15 (Orthonormal list, orthonormal basis), Proposition 6.3.13 (Orthogonal lists are independent)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9), Move an operator across the inner product (Technique 6.6.20)
> *Hint 1.* Let $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$ be the sum of the eigenspaces. Use, as in the second proof, that $T\mathbf{u} = \lambda\mathbf{u}$ implies $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$. Which subspaces does T^* preserve?
> *Hint 2.* T^* maps each E_λ into itself, so E is T^*-invariant, and $E^\perp$ is T-invariant. What would an eigenvector of $T{\restriction}_{E^\perp}$ be?
> *Hint 3.* For orthogonality, compute $\langle T\mathbf{u}, \mathbf{v}\rangle$ in two ways for $\mathbf{u} \in E_\lambda$ and $\mathbf{v} \in E_\mu$.
> *Sketch.* Every eigenvector of T is one of T^* (second proof), so T^* preserves each eigenspace and hence $E = \sum_j E_{\lambda_j}$; by Exercise 6.6.38, T preserves $E^\perp$. If $E^\perp \neq \{\mathbf{0}\}$, then $T{\restriction}_{E^\perp}$ has an eigenvector, which lies in $E \cap E^\perp = \{\mathbf{0}\}$, a contradiction. So $E = V$. Eigenspaces for $\lambda \neq \mu$ are orthogonal since $\lambda\langle\mathbf{u}, \mathbf{v}\rangle = \langle T\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, T^*\mathbf{v}\rangle = \mu\langle\mathbf{u}, \mathbf{v}\rangle$. Concatenate orthonormal bases of the eigenspaces.

Let T be normal, let $\lambda_1, \ldots, \lambda_k$ be its distinct eigenvalues (there is at least one, by Theorem 7.3.27), and let $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$. As shown in the second proof, $T\mathbf{u} = \lambda\mathbf{u}$ implies $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$. So T^* maps each E_{λ_j} into itself, and therefore maps E into E. By Exercise 6.6.38 applied to T^*, the operator $T = (T^*)^*$ maps $E^\perp$ into $E^\perp$.

Suppose $E^\perp \neq \{\mathbf{0}\}$. Then $T{\restriction}_{E^\perp}$ is an operator on a nonzero finite-dimensional complex space, so it has an eigenvector $\mathbf{w}$ (Theorem 7.3.27). This $\mathbf{w}$ is an eigenvector of T, so $\mathbf{w} \in E$. Then $\mathbf{w} \in E \cap E^\perp$, so $\langle \mathbf{w}, \mathbf{w}\rangle = 0$ and $\mathbf{w} = \mathbf{0}$, a contradiction. Hence $E^\perp = \{\mathbf{0}\}$, and $E = V$ by Corollary 6.4.27.

Next, eigenspaces for different eigenvalues are orthogonal. Let $\mathbf{u} \in E_\lambda$ and $\mathbf{v} \in E_\mu$ with $\lambda \neq \mu$. Then $T^*\mathbf{v} = \overline{\mu}\mathbf{v}$, and

$$\lambda\langle\mathbf{u}, \mathbf{v}\rangle = \langle T\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, T^*\mathbf{v}\rangle = \langle \mathbf{u}, \overline{\mu}\mathbf{v}\rangle = \mu\langle\mathbf{u}, \mathbf{v}\rangle.$$

Since $\lambda \neq \mu$, $\langle\mathbf{u}, \mathbf{v}\rangle = 0$.

Finally, choose an orthonormal basis $\mathcal{O}_j$ of each E_{λ_j} (Theorem 6.4.20), and let $\mathcal{O}$ be the concatenation $\mathcal{O}_1, \ldots, \mathcal{O}_k$. Vectors within one $\mathcal{O}_j$ are orthonormal, and vectors from different $\mathcal{O}_j$ are orthogonal by the previous paragraph, so $\mathcal{O}$ is orthonormal (Definition 6.3.15), hence linearly independent (Proposition 6.3.13). It spans $E_{\lambda_1} + \cdots + E_{\lambda_k} = V$. So $\mathcal{O}$ is an orthonormal eigenbasis. □

Remark 8.2.6 (Comparing the three proofs). The first proof is the shortest, because Schur's theorem has already done the work of finding a basis, and it isolates the role of normality in a single matrix lemma. The second is self-contained apart from the existence of one eigenvalue, and it is the template for the real theorem below, where Schur's theorem is not available in the same form. The third proves more than it states: it shows directly that V is the orthogonal direct sum of the eigenspaces, which is the spectral decomposition of Section 8.2.4. It also shows a pattern worth remembering: to prove that some subspace E is everything, show that $E^{\perp}$ is invariant and that any eigenvector in it would have to lie in E.

The lecture states the spectral theorem for self-adjoint operators on complex spaces. It is a special case, and Schur's theorem proves it without the lemma.

Corollary 8.2.7 (Hermitian matrices). Let V be a finite-dimensional complex inner product space with $\dim V \geq 1$, and $T \in \mathrm{L}(V)$. Then T is self-adjoint if and only if V has an orthonormal eigenbasis for T and every eigenvalue of T is real. In matrix form: $A \in \mathbb{C}^{n\times n}$ is Hermitian if and only if $A = UDU^*$ with U unitary and D diagonal with real entries.

Proof (from the complex spectral theorem). ROUTINE

> *Uses.* Theorem 8.2.5 (Complex spectral theorem), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Proposition 8.2.2 (Necessary conditions)
> *Hint 1.* A self-adjoint operator is normal.
> *Hint 2.* Which result says the eigenvalues of a self-adjoint operator are real?
> *Sketch.* If $T^* = T$ then $T^*T = T^2 = TT^*$, so Theorem 8.2.5 gives an orthonormal eigenbasis, and the eigenvalues are real by Proposition 8.1.1. The converse is Proposition 8.2.2(b).

If T is self-adjoint, then $T^*T = TT = TT^*$, so T is normal and has an orthonormal eigenbasis by Theorem 8.2.5; its eigenvalues are real by Proposition 8.1.1. The converse is Proposition 8.2.2(b). The matrix form follows as in the first proof of Theorem 8.2.5: A is Hermitian if and only if T_A is self-adjoint (Example 6.6.10), and $U^*AU = D$ is real diagonal exactly when the columns of U form an orthonormal eigenbasis with real eigenvalues. □

Proof (directly from Schur, as in lecture). ROUTINE

> *Uses.* Theorem 7.5.18 (Schur's theorem), Corollary 6.6.28 (Self-adjointness in an orthonormal basis), Proposition 8.2.2 (Necessary conditions)
> *Hint 1.* Take a Schur basis $\mathcal{O}$. If T is self-adjoint, what is the relation between $[T]_{\mathcal{O}}$ and its conjugate transpose?
> *Hint 2.* A matrix that equals its conjugate transpose and is upper triangular is also lower triangular.
> *Sketch.* For a Schur basis, $R = [T]_{\mathcal{O}}$ is upper triangular and $R^* = R$ (Corollary 6.6.28). So R is also lower triangular, hence diagonal, and $r_{jj} = \overline{r_{jj}}$ is real.

Let T be self-adjoint, and let $\mathcal{O}$ be an orthonormal basis with $R = [T]_{\mathcal{O}}$ upper triangular (Theorem 7.5.18). By Corollary 6.6.28, $R^* = R$, that is, $r_{ij} = \overline{r_{ji}}$ for all i, j. For $i > j$ we have $r_{ij} = 0$, so $r_{ji} = \overline{r_{ij}} = 0$: the entries above the diagonal vanish too, and R is diagonal. Its diagonal entries satisfy $r_{jj} = \overline{r_{jj}}$, so they are real, and they are the eigenvalues of the vectors of $\mathcal{O}$. The converse is Proposition 8.2.2(b). □

Example 8.2.8 (Two unitary diagonalizations).

(a) The real matrix $N = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = I_2 + J$ is normal, since J commutes with $J^{\mathsf{T}} = -J$. It is not symmetric, so it is not orthogonally diagonalizable over $\mathbb{R}$ (and indeed $p_N = (x-1)^2 + 1$ has no real root). Over $\mathbb{C}$ its eigenvectors are those of J from Example 8.2.3, with eigenvalues $1+i$ and $1-i$:

$$N = U\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}U^*, \quad U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}.$$

(b) The Hermitian matrix $H = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}$ of Example 6.6.27 has $p_H = (x-2)(x-3) - (1-i)(1+i) = x^2 - 5x + 4 = (x-1)(x-4)$. Row reducing $H - I_2$ and $H - 4I_2$ gives the eigenvectors $(-1+i, 1)$ for 1 and $(1, 1+i)$ for 4. Each has squared norm 3, and they are orthogonal, as Corollary 8.2.7 predicts: $(-1+i)\cdot\overline{1} + 1\cdot\overline{1+i} = -1+i+1-i = 0$. So

$$H = U\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}U^*, \quad U = \frac{1}{\sqrt{3}}\begin{pmatrix} -1+i & 1 \\ 1 & 1+i \end{pmatrix}.$$

Exercise 8.2.9. ROUTINE For each matrix, decide whether it is unitarily diagonalizable (over $\mathbb{C}$), and for the real ones also whether it is orthogonally diagonalizable (over $\mathbb{R}$).

(a) $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$
(b) $\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$
(c) $\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$
(d) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$
(e) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

Solution on page 1641

8.2.3 The real spectral theorem

Over $\mathbb{R}$, Example 8.2.3(b) shows that normality is not enough, and Proposition 8.2.2(b) shows that self-adjointness is necessary. To prove it is sufficient, we want to run the second proof of Theorem 8.2.5 again. Self-adjointness takes care of the invariance of W for free (Exercise 6.6.38), so the only step that used $\mathbb{C}$ is the existence of an eigenvector. That is the content of the next lemma.

Lemma 8.2.10 (Self-adjoint operators have eigenvalues). Let V be a finite-dimensional real inner product space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$ be self-adjoint. Then T has an eigenvalue. In fact every complex root of p_T is real, so p_T splits over $\mathbb{R}$.

The first proof borrows the eigenvalue from $\mathbb{C}$. It turns T into a real symmetric matrix, regards that matrix as a complex Hermitian matrix, and uses that Hermitian matrices have only real eigenvalues.

Proof (via a Hermitian matrix). STANDARD

Uses. Theorem 6.4.20 (Orthonormal bases exist), Corollary 6.6.28 (Self-adjointness in an orthonormal basis), Example 6.6.10 (The adjoint of a matrix), Corollary 7.1.35 (Factorization over $\mathbb{C}$), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Definition 7.3.12 (Characteristic polynomial of an operator), Corollary 7.3.13 (Eigenvalues of an operator)
Hint 1. Choose an orthonormal basis and let $A = [T]_{\mathcal{O}} \in \mathbb{R}^{n\times n}$. What do you know about A?
Hint 2. A is symmetric, so as a complex matrix it is Hermitian. Where are the complex roots of p_A?
Sketch. $A = [T]_{\mathcal{O}}$ is real symmetric, hence Hermitian as a complex matrix. Over $\mathbb{C}$, p_A factors into linear factors, and each root is an eigenvalue of the self-adjoint T_A on $\mathbb{C}^n$, hence real. So $p_T = p_A$ splits over $\mathbb{R}$, and any of its roots is an eigenvalue of T.

Let $\mathcal{O}$ be an orthonormal basis of V (Theorem 6.4.20) and $A = [T]_{\mathcal{O}} \in \mathbb{R}^{n\times n}$, $n = \dim V$. By Corollary 6.6.28, $A^{\mathsf{T}} = A$. Regard A as a matrix in $\mathbb{C}^{n\times n}$. Its entries are real, so $A^* = \overline{A}^{\mathsf{T}} = A^{\mathsf{T}} = A$, and T_A is a self-adjoint operator on $\mathbb{C}^n$ with the dot product (Example 6.6.10). By Corollary 7.1.35, $p_A = (x-\lambda_1)\cdots(x-\lambda_n)$ for some $\lambda_i \in \mathbb{C}$. Each λ_i is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ (Theorem 7.3.6), that is, of T_A, so $\lambda_i \in \mathbb{R}$ by Proposition 8.1.1. Hence p_A splits over $\mathbb{R}$. (The determinant $\det(xI_n - A)$ is the same polynomial whether A is viewed as real or complex.) Since $p_T = p_A$ (Definition 7.3.12), the real number λ_1 is a root of p_T, hence an eigenvalue of T by Corollary 7.3.13. □

The second proof stays inside V. The idea comes from factoring polynomials over $\mathbb{R}$: some nonzero polynomial p satisfies $p(T)\mathbf{v} = \mathbf{0}$, and p factors into linear factors and irreducible quadratics (Theorem 7.1.39). The rotation by $\pi/2$ shows that a quadratic factor such as x^2+1 can be what kills $\mathbf{v}$. Self-adjointness rules this out: for self-adjoint T, an irreducible quadratic q gives an invertible $q(T)$, because $\langle q(T)\mathbf{v}, \mathbf{v}\rangle$ is bounded below by a quadratic in $\|T\mathbf{v}\|$ with negative discriminant.

Proof (without complex numbers). HARD

> *Uses.* Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.1.39 (Factorization over $\mathbb{R}$), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 6.2.1 (Cauchy–Schwarz inequality), Definition 6.6.26 (Self-adjoint operator), Theorem 7.2.9 (Eigenvalues and invertibility)
> *Techniques.* A nonnegative quadratic has nonpositive discriminant (Technique 6.2.7), Move an operator across the inner product (Technique 6.6.20)
> *Hint 1.* Take $\mathbf{v} \neq \mathbf{0}$. Some monic p of degree $d \geq 1$ has $p(T)\mathbf{v} = \mathbf{0}$ (Proposition 7.2.35). Factor p over $\mathbb{R}$.
> *Hint 2.* Show that $q(T) = T^2 + bT + c\,\mathrm{id}$ is injective when $b^2 < 4c$: for $\mathbf{v} \neq \mathbf{0}$, show $\langle q(T)\mathbf{v}, \mathbf{v}\rangle > 0$. Use $\langle T^2\mathbf{v}, \mathbf{v}\rangle = \|T\mathbf{v}\|^2$ and Cauchy–Schwarz on $\langle T\mathbf{v}, \mathbf{v}\rangle$.
> *Hint 3.* Complete the square: $\|T\mathbf{v}\|^2 - |b|\|T\mathbf{v}\|\|\mathbf{v}\| + c\|\mathbf{v}\|^2 = \left(\|T\mathbf{v}\| - \frac{|b|}{2}\|\mathbf{v}\|\right)^2 + \left(c - \frac{b^2}{4}\right)\|\mathbf{v}\|^2$. Then $p(T)$ is a product of injective quadratic factors and factors $T - \lambda_i\,\mathrm{id}$, and it is not injective.
> *Sketch.* For $b^2 < 4c$ and $\mathbf{v} \neq \mathbf{0}$, self-adjointness and Cauchy–Schwarz give $\left\langle (T^2 + bT + c\,\mathrm{id})\mathbf{v}, \mathbf{v}\right\rangle \geq \left(\|T\mathbf{v}\| - \frac{|b|}{2}\|\mathbf{v}\|\right)^2 + \left(c - \frac{b^2}{4}\right)\|\mathbf{v}\|^2 > 0$, so $T^2 + bT + c\,\mathrm{id}$ is injective. Pick a monic p of degree ≥ 1 with $p(T)\mathbf{v} = \mathbf{0}$ for some $\mathbf{v} \neq \mathbf{0}$ and factor it over $\mathbb{R}$. The product $p(T)$ of the factors is not injective, so some factor is not; it cannot be a quadratic one, so it is some $T - \lambda\,\mathrm{id}$, and λ is an eigenvalue.

We prove that T has an eigenvalue. First let $b, c \in \mathbb{R}$ with $b^2 < 4c$; we show that $q(T) = T^2 + bT + c\,\mathrm{id}$ is injective. Let $\mathbf{v} \neq \mathbf{0}$. Since T is self-adjoint, $\langle T^2\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$, and by Cauchy–Schwarz, $b\langle T\mathbf{v}, \mathbf{v}\rangle \geq -|b|\|T\mathbf{v}\|\|\mathbf{v}\|$. Hence

$$\begin{aligned}\langle q(T)\mathbf{v}, \mathbf{v}\rangle &= \|T\mathbf{v}\|^2 + b\langle T\mathbf{v}, \mathbf{v}\rangle + c\|\mathbf{v}\|^2 \\ &\geq \|T\mathbf{v}\|^2 - |b|\|T\mathbf{v}\|\|\mathbf{v}\| + c\|\mathbf{v}\|^2 \\ &= \left(\|T\mathbf{v}\| - \frac{|b|}{2}\|\mathbf{v}\|\right)^2 + \left(c - \frac{b^2}{4}\right)\|\mathbf{v}\|^2 > 0,\end{aligned}$$

because $c - b^2/4 > 0$ and $\|\mathbf{v}\| > 0$. So $q(T)\mathbf{v} \neq \mathbf{0}$.

Now fix $\mathbf{v} \neq \mathbf{0}$. By Proposition 7.2.35 there is a monic $p \in \mathbb{R}[x]$ of degree $d \geq 1$ with $p(T)\mathbf{v} = \mathbf{0}$. By Theorem 7.1.39,

$$p = (x-\lambda_1)\cdots(x-\lambda_k)q_1\cdots q_l, \quad q_j = x^2 + b_j x + c_j, \quad b_j^2 < 4c_j,$$

and by Proposition 7.2.27,

$$p(T) = (T - \lambda_1\,\mathrm{id})\cdots(T - \lambda_k\,\mathrm{id})q_1(T)\cdots q_l(T).$$

A composition of injective maps is injective, and $p(T)$ is not injective since it kills $\mathbf{v}$. So at least one factor is not injective. The factors $q_j(T)$ are injective by the first paragraph, so some $T - \lambda_i\,\mathrm{id}$ is not injective (in particular $k \geq 1$), and λ_i is an eigenvalue of T by Theorem 7.2.9.

The second statement of the lemma (all complex roots of p_T are real) follows from the first proof of Theorem 8.2.12 below, which uses only the first statement: in an orthonormal eigenbasis, $[T]_{\mathcal{O}} = \operatorname{diag}(\mu_1, \ldots, \mu_n)$ with real μ_j, so $p_T = (x - \mu_1)\cdots(x - \mu_n)$. □

Remark 8.2.11 (A third proof, by maximizing). There is a third proof, which finds the eigenvector as the solution of an optimization problem. On the unit sphere $S = \{\mathbf{v} \in V \mid \|\mathbf{v}\| = 1\}$, the function $\mathbf{v} \mapsto \langle T\mathbf{v}, \mathbf{v}\rangle$ is continuous, and S is closed and bounded, so the function attains a maximum at some $\mathbf{u} \in S$. Perturbing a maximizer (Technique 6.5.6) along any direction $\mathbf{w} \perp \mathbf{u}$ shows that $\langle T\mathbf{u}, \mathbf{w}\rangle = 0$ (the derivative of the perturbed value is $2\langle T\mathbf{u}, \mathbf{w}\rangle$ because $\langle T\mathbf{w}, \mathbf{u}\rangle = \langle \mathbf{w}, T\mathbf{u}\rangle$, which is where self-adjointness enters), that is, $T\mathbf{u} \in \operatorname{span}(\mathbf{u})$, so $\mathbf{u}$ is an eigenvector, and its eigenvalue is the maximum value. The existence of the maximum is a fact from analysis, which is why this book postpones the argument to Section 12.3. That section also explains what the other eigenvalues maximize (the Courant–Fischer theorem).

Theorem 8.2.12 (Real spectral theorem). Let V be a finite-dimensional real inner product space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. The following are equivalent:

(I) T is self-adjoint;

(II) V has an orthonormal basis consisting of eigenvectors of T;

(III) $[T]_{\mathcal{O}}$ is diagonal for some orthonormal basis $\mathcal{O}$ of V.

In matrix form: $A \in \mathbb{R}^{n\times n}$ is symmetric if and only if $A = QDQ^{\mathsf{T}}$ for some orthogonal $Q \in \mathbb{R}^{n\times n}$ and some diagonal $D \in \mathbb{R}^{n\times n}$, that is, if and only if A is orthogonally diagonalizable.

Again (II) ⇔ (III) is Theorem 7.4.2 and (II) ⇒ (I) is Proposition 8.2.2(b), so we prove (I) ⇒ (II).

Proof (induction with invariant complements). STANDARD

Uses. Lemma 8.2.10 (Self-adjoint operators have eigenvalues), Exercise 6.6.38, Definition 6.6.26 (Self-adjoint operator), Theorem 6.4.24 (Orthogonal decomposition), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 4.4.13 (Bases and dimension of a sum), Definition 6.3.15 (Orthonormal list, orthonormal basis), Proposition 8.2.2 (Necessary conditions), Theorem 4.5.17 (Change of basis), Definition 6.3.33 (Unitary and orthogonal matrices)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Proving TFAE (Technique 2.7.1)
Hint 1. Follow the second proof of Theorem 8.2.5. Which step used $\mathbb{C}$, and what replaces it?
Hint 2. Lemma 8.2.10 provides a unit eigenvector $\mathbf{u}$. Why is $W = \operatorname{span}(\mathbf{u})^{\perp}$ invariant this time, and why is $T{\restriction_W}$ self-adjoint?
Sketch. Induct on $\dim V$. By Lemma 8.2.10, T has a unit eigenvector $\mathbf{u}$. Since $T^* = T$ maps $\operatorname{span}(\mathbf{u})$ into itself, T maps $W = \operatorname{span}(\mathbf{u})^{\perp}$ into itself (Exercise 6.6.38), and $T{\restriction_W}$ is self-adjoint because $\langle T\mathbf{w}, \mathbf{w}'\rangle = \langle \mathbf{w}, T\mathbf{w}'\rangle$ for $\mathbf{w}, \mathbf{w}' \in W$. By induction W has an orthonormal eigenbasis; add $\mathbf{u}$. For matrices, put the eigenbasis of T_A into the columns of Q.

We prove (I) ⇒ (II) by induction on $n = \dim V$. If $n = 1$, any unit vector is an orthonormal eigenbasis, as in the complex case. Let $n \geq 2$, assume the claim in dimension $n - 1$, and let T be self-adjoint. By Lemma 8.2.10, T has an eigenvector, and dividing by its norm gives $\mathbf{u}$ with $\|\mathbf{u}\| = 1$ and $T\mathbf{u} = \lambda\mathbf{u}$. Let $U = \operatorname{span}(\mathbf{u})$ and $W = U^{\perp}$. Since $T^* = T$ maps U into U, Exercise 6.6.38 (applied to T^*, whose adjoint is T) shows that T maps W into W. The operator $S = T{\restriction_W} \in \mathrm{L}(W)$ is self-adjoint for the inner product of V restricted to W, since $\langle S\mathbf{w}, \mathbf{w}'\rangle = \langle T\mathbf{w}, \mathbf{w}'\rangle = \langle \mathbf{w}, T\mathbf{w}'\rangle = \langle \mathbf{w}, S\mathbf{w}'\rangle$ for $\mathbf{w}, \mathbf{w}' \in W$ (Definition 6.6.26). By Theorem 6.4.24 and Corollary 6.4.25, $V = U \oplus W$ with $\dim W = n - 1 \geq 1$. By the induction hypothesis, W has an orthonormal basis $\mathbf{u}_2, \ldots, \mathbf{u}_n$ of eigenvectors of S, hence of T. As in the complex case, $\mathbf{u}, \mathbf{u}_2, \ldots, \mathbf{u}_n$ is an orthonormal basis of V (Definition 6.3.15, Theorem 4.4.13) consisting of eigenvectors of T.

For the matrix form, $\left(QDQ^{\mathsf{T}}\right)^{\mathsf{T}} = QD^{\mathsf{T}}Q^{\mathsf{T}} = QDQ^{\mathsf{T}}$, so orthogonally diagonalizable matrices are symmetric. Conversely, if A is symmetric, T_A is self-adjoint on $\mathbb{R}^n$ with the dot product, so $\mathbb{R}^n$ has an orthonormal eigenbasis $\mathbf{q}_1, \ldots, \mathbf{q}_n$. The matrix Q with these columns is orthogonal (Definition 6.3.33), and $Q^{\mathsf{T}}AQ = Q^{-1}AQ = D$ is diagonal by Theorem 4.5.17. Hence $A = QDQ^{\mathsf{T}}$. □

The second proof follows the first proof of the complex theorem instead, with a real Schur basis.

Proof (triangularize, then use symmetry). STANDARD

> *Uses.* Lemma 8.2.10 (Self-adjoint operators have eigenvalues), Theorem 7.5.17 (Triangularizable means the characteristic polynomial splits), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 6.4.15 (Gram–Schmidt is correct), Corollary 6.6.28 (Self-adjointness in an orthonormal basis)
> *Techniques.* Build a flag of invariant subspaces (Technique 7.5.4), Prove it for triangular matrices first (Technique 7.5.23)
> *Hint 1.* By Lemma 8.2.10, p_T splits over $\mathbb{R}$. What does Theorem 7.5.17 give?
> *Hint 2.* Gram–Schmidt preserves the flag of a triangularizing basis, exactly as in the first proof of Theorem 7.5.18. Then compare $[T]_{\mathcal{O}}$ with its transpose.
> *Sketch.* p_T splits over $\mathbb{R}$ (Lemma 8.2.10), so some basis $\mathcal{B}$ triangularizes T; Gram–Schmidt turns it into an orthonormal $\mathcal{O}$ with the same invariant flag, so $R = [T]_{\mathcal{O}}$ is upper triangular. It is also symmetric (Corollary 6.6.28), hence diagonal.

By Lemma 8.2.10, p_T splits over $\mathbb{R}$, so by Theorem 7.5.17 there is a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of V with $[T]_{\mathcal{B}}$ upper triangular; by Proposition 7.5.2, every $V_j = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$ is T-invariant. Gram–Schmidt applied to $\mathcal{B}$ gives an orthonormal basis $\mathcal{O} = \mathbf{u}_1, \ldots, \mathbf{u}_n$ with $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j) = V_j$ for every j (Theorem 6.4.15(d)), so $R = [T]_{\mathcal{O}}$ is upper triangular by Proposition 7.5.2. (This is the first proof of Theorem 7.5.18, which used nothing about $\mathbb{C}$ beyond the triangularizing basis.) By Corollary 6.6.28, $R^{\mathsf{T}} = R$. For $i < j$, $r_{ij} = r_{ji} = 0$, so R is diagonal, and $\mathcal{O}$ is an orthonormal eigenbasis. □

Non-example 8.2.13 (Symmetric matrices over other fields). The real spectral theorem is about *real* symmetric matrices, and the inner product on $\mathbb{R}^n$ is what makes it work. Over other fields the transpose still exists, but symmetric matrices need not be diagonalizable.

(a) $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ satisfies $A^{\mathsf{T}} = A$, and $A^2 = \begin{pmatrix} 1+i^2 & i-i \\ i-i & i^2+1 \end{pmatrix} = 0$. So A is nilpotent and nonzero, and is not diagonalizable (its only eigenvalue is 0, and $A \neq 0$). Over $\mathbb{C}$, the right condition is $A^* = A$, not $A^{\mathsf{T}} = A$.

(b) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$ is symmetric with $A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = 0$ in $\mathbb{F}_2$. Again a nonzero nilpotent matrix, so not diagonalizable.

In both cases the proofs above fail at the same point: the "inner product" $\mathbf{x}^{\mathsf{T}}\mathbf{y}$ has nonzero vectors with $\mathbf{x}^{\mathsf{T}}\mathbf{x} = 0$, such as $(1, i)$ and $(1, 1)$, and every argument above that divided by $\|\mathbf{u}\|$ or used $\|\mathbf{w}\| = 0 \Rightarrow \mathbf{w} = \mathbf{0}$ breaks. Section 12.4 studies what survives.

Computing an orthogonal diagonalization

The proofs suggest the computation. Eigenspaces for different eigenvalues are automatically orthogonal (third proof of Theorem 8.2.5, which applies verbatim to self-adjoint operators on real spaces), so orthogonality only has to be arranged inside each eigenspace.

Algorithm 8.2.14 (Orthogonal diagonalization)

Input A symmetric $A \in \mathbb{R}^{n\times n}$.
Output An orthogonal Q and a diagonal D with $A = QDQ^{\mathsf{T}}$.

1. Compute p_A and its roots $\lambda_1, \ldots, \lambda_k$ (all real, by Lemma 8.2.10).
2. For each λ_j, find a basis of $E_{\lambda_j} = \ker(A - \lambda_j I_n)$ by row reduction.
3. Apply Gram–Schmidt to each of these bases separately, to get an orthonormal basis of each E_{λ_j}.
4. Let Q have these orthonormal bases as its columns, and let D be diagonal with the eigenvalue of each column in the matching position.

The columns found in step 3 are orthonormal within each eigenspace, and orthogonal across eigenspaces by the computation in the third proof of Theorem 8.2.5. There are n of them, since A is diagonalizable and so the eigenspace dimensions add up to n (Theorem 7.4.16). Only step 3 is new compared with Algorithm 7.4.18. For a simple eigenvalue it just normalizes the eigenvector; Gram–Schmidt does real work only for repeated eigenvalues.

Example 8.2.15 (A repeated eigenvalue). Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = I_3 + J_3$, where J_3 is the all-ones matrix. Since $J_3\mathbf{x} = (x_1 + x_2 + x_3)(1, 1, 1)$, the eigenvalues of J_3 are 3 (on $\operatorname{span}((1, 1, 1))$) and 0 (on the plane $x_1 + x_2 + x_3 = 0$), so A has eigenvalues 4 and 1, with

$$E_4 = \operatorname{span}((1, 1, 1)), \quad E_1 = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 0\} = \operatorname{span}((1, -1, 0), (1, 0, -1)).$$

The basis of E_1 from row reduction is not orthogonal. Gram–Schmidt replaces $(1, 0, -1)$ by $(1, 0, -1) - \frac{1}{2}(1, -1, 0) = \frac{1}{2}(1, 1, -2)$. Normalizing,

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}, \quad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = QDQ^{\mathsf{T}}.$$

Any orthonormal basis of the plane E_1 would do in the last two columns; step 3 only picks one. The first column is orthogonal to the plane E_1 automatically.

Exercise 8.2.16. ROUTINE Find an orthogonal Q and a diagonal D with $\begin{pmatrix} 7 & -2 \\ -2 & 4 \end{pmatrix} = QDQ^{\mathsf{T}}$.

Solution on page 1641

Symmetric matrices and ellipses

Geometrically, the real spectral theorem says that a real symmetric matrix stretches $\mathbb{R}^n$ by the factors $\lambda_1, \ldots, \lambda_n$ along n perpendicular axes, with a flip when $\lambda_j < 0$. In the plane, this is visible on the unit circle.

Example 8.2.17 (The image of the unit circle). Let $A \in \mathbb{R}^{2\times 2}$ be symmetric and invertible, and write $A = QDQ^{\mathsf{T}}$ with $Q = (\mathbf{q}_1 \mid \mathbf{q}_2)$ orthogonal and $D = \operatorname{diag}(\lambda_1, \lambda_2)$. A unit vector is $\mathbf{x} = y_1\mathbf{q}_1 + y_2\mathbf{q}_2$ with $y_1^2 + y_2^2 = 1$ (Proposition 6.3.18, Theorem 6.3.22), and $A\mathbf{x} = \lambda_1 y_1\mathbf{q}_1 + \lambda_2 y_2\mathbf{q}_2$. So the image of the unit circle is

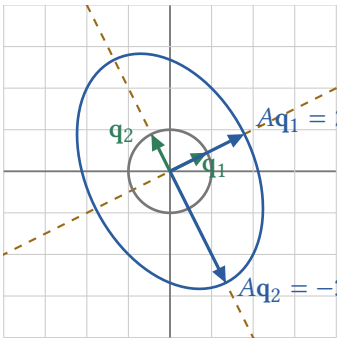
$$\left\{ z_1\mathbf{q}_1 + z_2\mathbf{q}_2 \;\middle|\; \frac{z_1^2}{\lambda_1^2} + \frac{z_2^2}{\lambda_2^2} = 1 \right\},$$

an ellipse whose axes point along the eigenvectors, with semi-axes $|\lambda_1|$ and $|\lambda_2|$.

For instance, $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ has $p_A = x^2 + x - 6 = (x - 2)(x + 3)$, with $E_2 = \operatorname{span}((2, 1))$ and $E_{-3} = \operatorname{span}((-1, 2))$. So $A = QDQ^{\mathsf{T}}$ with

$$Q = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix},$$

and the image of the unit circle is an ellipse with semi-axis 2 along $(2, 1)$ and 3 along $(-1, 2)$. A non-symmetric matrix also maps the circle to an ellipse, but its axes are generally not along its eigenvectors; that is the subject of Section 9.1.



The unit circle and its image under the symmetric matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$. The eigenvectors $\mathbf{q}_1, \mathbf{q}_2$ are perpendicular, and A scales them by 2 and -3; the images lie along the axes of the ellipse. In the interactive version the two off-diagonal entries move together, so A stays symmetric. Move the point $\mathbf{x}$ around the circle: $A\mathbf{x}$ is farthest from $\mathbf{0}$ when $\mathbf{x} = \pm\mathbf{q}$ for the eigenvalue of largest absolute value. Make an eigenvalue 0 to flatten the ellipse to a segment, and make A a multiple of I, where every direction is an eigenvector. (Interactive in the web version.)

8.2.4 The spectral decomposition

An orthonormal eigenbasis is not unique: in Example 8.2.15 any orthonormal basis of the plane E_1 would do. What is unique is the splitting of V into eigenspaces, and the orthogonal projections onto them. Writing T in terms of these projections gives a basis-free form of the spectral theorems.

For the rest of the section, we say that T satisfies the **spectral hypothesis** if T is normal and $\mathbb{F} = \mathbb{C}$, or T is self-adjoint and $\mathbb{F} = \mathbb{R}$. By Theorem 8.2.5 and Theorem 8.2.12, these are exactly the cases in which V has an orthonormal eigenbasis. (Self-adjoint operators on complex spaces are included, being normal.)

Theorem 8.2.18 (Spectral decomposition). Let T satisfy the spectral hypothesis, let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T, and let $P_j = \mathrm{Pr}_{E_{\lambda_j}}$ be the orthogonal projection onto E_{λ_j}.

(a) The eigenspaces are pairwise orthogonal, and $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

(b) $P_i P_j = 0$ for $i \neq j$, and $P_1 + \cdots + P_k = \mathrm{id}_V$.

(c) $T = \lambda_1 P_1 + \cdots + \lambda_k P_k$.

In matrix form: if $A = UDU^*$ with U unitary (orthogonal, if $\mathbb{F} = \mathbb{R}$) with columns $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $D = \mathrm{diag}(\mu_1, \dots, \mu_n)$, then

$$A = \mu_1 \mathbf{u}_1 \mathbf{u}_1^* + \cdots + \mu_n \mathbf{u}_n \mathbf{u}_n^*,$$

and grouping the terms with equal μ_i gives $A = \sum_j \lambda_j P_j$, where P_j is the matrix of $\mathrm{Pr}_{E_{\lambda_j}}$.

Proof. STANDARD

Uses. Theorem 8.2.5 (Complex spectral theorem), Theorem 8.2.12 (Real spectral theorem), Proposition 6.3.18 (Coordinates in an orthonormal basis), Definition 6.4.11 (Orthogonal projection), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 6.4.31 (The matrix of an orthogonal projection), Proposition 2.4.8 (Four views of the product)
Techniques. Pass to orthonormal coordinates (Technique 6.7.2)
Hint 1. Take an orthonormal eigenbasis $\mathbf{u}_1, \dots, \mathbf{u}_n$ and group its vectors by eigenvalue. Show that the group belonging to λ_j spans E_{λ_j}.
Hint 2. If $\mathbf{v} = \sum c_i \mathbf{u}_i$ is in E_{λ_j}, compare the coordinates of $T\mathbf{v}$ and $\lambda_j \mathbf{v}$.
Hint 3. With an orthonormal basis of E_{λ_j} in hand, $P_j \mathbf{v}$ is the part of the expansion $\mathbf{v} = \sum \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i$ coming from that group.

Sketch. Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be an orthonormal eigenbasis with $T\mathbf{u}_i = \mu_i \mathbf{u}_i$, and let $\mathcal{O}_j$ consist of the $\mathbf{u}_i$ with $\mu_i = \lambda_j$. Comparing coordinates shows $E_{\lambda_j} = \operatorname{span}(\mathcal{O}_j)$. So the eigenspaces are spanned by disjoint pieces of one orthonormal basis, which gives (a), and $P_j\mathbf{v} = \sum_{\mathbf{u}_i \in \mathcal{O}_j} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i$. Summing over j gives $\mathbf{v}$, and applying T to the expansion of $\mathbf{v}$ gives $\sum_j \lambda_j P_j \mathbf{v}$.

By Theorem 8.2.5 or Theorem 8.2.12, V has an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ with $T\mathbf{u}_i = \mu_i \mathbf{u}_i$, and every μ_i is one of $\lambda_1, \dots, \lambda_k$. For each j, let $\mathcal{O}_j$ be the list of those $\mathbf{u}_i$ with $\mu_i = \lambda_j$.

We claim $E_{\lambda_j} = \operatorname{span}(\mathcal{O}_j)$. Each vector of $\mathcal{O}_j$ lies in E_{λ_j}, which gives $\supseteq$. Conversely, let $\mathbf{v} \in E_{\lambda_j}$ and write $\mathbf{v} = \sum_i c_i \mathbf{u}_i$. Then

$$\sum_i c_i \mu_i \mathbf{u}_i = T\mathbf{v} = \lambda_j \mathbf{v} = \sum_i c_i \lambda_j \mathbf{u}_i,$$

so $c_i(\mu_i - \lambda_j) = 0$ for every i, by uniqueness of coordinates. Hence $c_i = 0$ unless $\mu_i = \lambda_j$, and $\mathbf{v} \in \operatorname{span}(\mathcal{O}_j)$.

(a) For $i \neq j$, $\mathcal{O}_i$ and $\mathcal{O}_j$ have no vector in common, so every vector of $\mathcal{O}_i$ is orthogonal to every vector of $\mathcal{O}_j$, and hence every vector of $\operatorname{span}(\mathcal{O}_i) = E_{\lambda_i}$ is orthogonal to every vector of E_{λ_j}. The lists $\mathcal{O}_1, \dots, \mathcal{O}_k$ together are a reordering of a basis of V, so $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ by Theorem 4.4.13.

(b) Since $\mathcal{O}_j$ is an orthonormal basis of E_{λ_j}, Definition 6.4.11 gives

$$P_j\mathbf{v} = \sum_{\mathbf{u}_i \in \mathcal{O}_j} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i. \tag{8.2.1}$$

Summing over j gives $\sum_j P_j\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i = \mathbf{v}$ by Proposition 6.3.18, so $\sum_j P_j = \mathrm{id}_V$. For $i \neq j$, $P_j\mathbf{v} \in E_{\lambda_j} \subseteq E_{\lambda_i}^{\perp} = \ker P_i$ by (a) and Proposition 6.4.29(a), so $P_iP_j = 0$.

(c) By Proposition 6.3.18 and linearity,

$$T\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{u}_i \rangle T\mathbf{u}_i = \sum_{i=1}^n \mu_i \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i = \sum_{j=1}^k \lambda_j \sum_{\mathbf{u}_i \in \mathcal{O}_j} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i = \sum_{j=1}^k \lambda_j P_j \mathbf{v},$$

by (8.2.1).

For the matrix form, $A = UDU^* = (UD)U^*$, and the column–row expansion of a product (Proposition 2.4.8) gives $\sum_i (\mu_i \mathbf{u}_i)\mathbf{u}_i^*$. By Theorem 6.4.31, $\sum_{\mathbf{u}_i \in \mathcal{O}_j} \mathbf{u}_i \mathbf{u}_i^*$ is the matrix of $\mathrm{Pr}_{E_{\lambda_j}}$, where E_{λ_j} is now the eigenspace of A in $\mathbb{F}^n$. □

So a normal operator is a linear combination of orthogonal projections onto mutually perpendicular subspaces that fill V. The converse also holds, and the decomposition is unique, which is what makes it useful for defining things.

Proposition 8.2.19 (Uniqueness of the spectral decomposition). Let $\mu_1, \dots, \mu_m \in \mathbb{F}$ be distinct, and let $Q_1, \dots, Q_m \in \mathrm{L}(V)$ be nonzero orthogonal projections with $Q_iQ_j = 0$ for $i \neq j$ and $Q_1 + \cdots + Q_m = \mathrm{id}_V$. Let $T = \mu_1 Q_1 + \cdots + \mu_m Q_m$. Then:

(a) T is normal, and self-adjoint if every μ_j is real;

(b) the eigenvalues of T are exactly $\mu_1, \dots, \mu_m$, and $Q_j = \mathrm{Pr}_{E_{\mu_j}}$ for every j.

Proof. STANDARD

Uses. Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.6.18 (Properties of the adjoint), Definition 7.2.13 (Eigenspace), Proposition 6.4.29 (The orthogonal projection as a linear map)
Hint 1. Each Q_j is self-adjoint (Corollary 6.6.29). Compute T^*.
Hint 2. For (b), first show that T acts as μ_j on $\operatorname{im} Q_j$. Then take an eigenvector $\mathbf{v}$ with eigenvalue λ, apply Q_i to $T\mathbf{v} = \lambda\mathbf{v}$, and use $Q_iQ_j = 0$.

Sketch. (a) $T^* = \sum \overline{\mu_j} Q_j$, and $TT^* = \sum_j |\mu_j|^2 Q_j = T^*T$ because $Q_iQ_j = 0$ for $i \neq j$ and $Q_j^2 = Q_j$. (b) If $\mathbf{v} \in \operatorname{im} Q_j$ then $Q_i\mathbf{v} = Q_iQ_j\mathbf{v} = 0$ for $i \neq j$, so $T\mathbf{v} = \mu_j\mathbf{v}$: $\operatorname{im} Q_j \subseteq E_{\mu_j}$. If $T\mathbf{v} = \lambda\mathbf{v}$, applying Q_i gives $\mu_i Q_i\mathbf{v} = \lambda Q_i\mathbf{v}$, so $Q_i\mathbf{v} = \mathbf{0}$ unless $\mu_i = \lambda$; as $\mathbf{v} = \sum Q_i\mathbf{v} \neq \mathbf{0}$, λ is some μ_j and $\mathbf{v} = Q_j\mathbf{v} \in \operatorname{im} Q_j$.

Each Q_j is an orthogonal projection, so $Q_j^2 = Q_j = Q_j^*$ (Corollary 6.6.29).

(a) By Proposition 6.6.18, $T^* = \sum_j \overline{\mu_j} Q_j$. Expanding the products and using $Q_iQ_j = 0$ for $i \neq j$ and $Q_j^2 = Q_j$,

$$TT^* = \sum_{i,j} \mu_i \overline{\mu_j} Q_iQ_j = \sum_j |\mu_j|^2 Q_j = \sum_{i,j} \overline{\mu_j} \mu_i Q_jQ_i = T^*T.$$

If every μ_j is real, then $T^* = T$.

(b) Let $\mathbf{v} \in \operatorname{im} Q_j$, say $\mathbf{v} = Q_j\mathbf{x}$. Then $Q_j\mathbf{v} = Q_j^2\mathbf{x} = \mathbf{v}$ and $Q_i\mathbf{v} = Q_iQ_j\mathbf{x} = \mathbf{0}$ for $i \neq j$, so $T\mathbf{v} = \mu_j\mathbf{v}$. Hence $\operatorname{im} Q_j \subseteq E_{\mu_j}$, and since $Q_j \neq 0$, μ_j is an eigenvalue.

Conversely, let $T\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. For each i, $Q_iT = \sum_j \mu_j Q_iQ_j = \mu_i Q_i$, so applying Q_i to $T\mathbf{v} = \lambda\mathbf{v}$ gives $\mu_i Q_i\mathbf{v} = \lambda Q_i\mathbf{v}$, that is, $(\mu_i - \lambda)Q_i\mathbf{v} = \mathbf{0}$. Since $\mathbf{v} = \sum_i Q_i\mathbf{v} \neq \mathbf{0}$, some $Q_j\mathbf{v} \neq \mathbf{0}$, and then $\lambda = \mu_j$. Moreover $Q_i\mathbf{v} = \mathbf{0}$ for all $i \neq j$ (as $\mu_i \neq \mu_j = \lambda$), so $\mathbf{v} = Q_j\mathbf{v} \in \operatorname{im} Q_j$. This shows that the eigenvalues are exactly the μ_j and that $E_{\mu_j} = \operatorname{im} Q_j$. Finally, an orthogonal projection is the orthogonal projection onto its image (Corollary 6.6.29), so $Q_j = \operatorname{Pr}_{\operatorname{im} Q_j} = \operatorname{Pr}_{E_{\mu_j}}$. □

The projections P_j can be computed without finding any eigenvectors. On E_{λ_i}, a polynomial $p(T)$ acts as the scalar $p(\lambda_i)$ (Proposition 7.2.28), so a polynomial that is 1 at λ_j and 0 at the other eigenvalues acts as P_j. These are the Lagrange polynomials.

Proposition 8.2.20 (Spectral projections are polynomials in T). Let T satisfy the spectral hypothesis, with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and $P_j = \operatorname{Pr}_{E_{\lambda_j}}$. Then

$$P_j = L_j(T), \quad \text{where} \quad L_j = \prod_{i \neq j} \frac{x - \lambda_i}{\lambda_j - \lambda_i}.$$

Consequently:

(a) every operator $S \in \mathrm{L}(V)$ with $ST = TS$ commutes with every P_j and maps every eigenspace E_{λ_j} into itself;

(b) if $\mathbb{F} = \mathbb{C}$, then $T^* = p(T)$ for the polynomial $p = \sum_j \overline{\lambda_j} L_j$.

Proof. STANDARD

Uses. Theorem 8.2.18 (Spectral decomposition), Proposition 7.1.29 (Lagrange polynomials in closed form), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 7.2.27 (Evaluating polynomials at an operator), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.6.18 (Properties of the adjoint)
Techniques. Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. Both P_j and $L_j(T)$ are linear, and V is the sum of the eigenspaces. Compare them on a vector of E_{λ_i}.
Hint 2. For (a), S commutes with T, hence with every polynomial in T.
Hint 3. For (b), take adjoints in $T = \sum_j \lambda_j P_j$.
Sketch. On $\mathbf{u} \in E_{\lambda_i}$, $L_j(T)\mathbf{u} = L_j(\lambda_i)\mathbf{u}$, which is $\mathbf{u}$ if $i = j$ and $\mathbf{0}$ otherwise, and P_j does the same. Since $V = \oplus_i E_{\lambda_i}$, $P_j = L_j(T)$. Then (a): S commutes with T^m for all m, hence with $L_j(T) = P_j$, and $SP_j\mathbf{v} = P_jS\mathbf{v} \in \operatorname{im} P_j$. (b) $T^* = \sum \overline{\lambda_j} P_j^* = \sum \overline{\lambda_j} L_j(T)$.

The formula for L_j is the Lagrange polynomial for the points $\lambda_1, \ldots, \lambda_k$ (Proposition 7.1.29), so $L_j(\lambda_j) = 1$ and $L_j(\lambda_i) = 0$ for $i \neq j$. (If $k = 1$, $L_1 = 1$ and the claim is $P_1 = \mathrm{id}_V$.) Let $\mathbf{u} \in E_{\lambda_i}$. By Proposition 7.2.28, $L_j(T)\mathbf{u} = L_j(\lambda_i)\mathbf{u}$. On the other hand, $P_j\mathbf{u} = \mathbf{u}$ if $i = j$, and $P_j\mathbf{u} = \mathbf{0}$ if $i \neq j$, since then $\mathbf{u} \in E_{\lambda_j}^{\perp} = \ker P_j$ by Theorem 8.2.18(a). So P_j and $L_j(T)$ agree on every eigenspace. Every $\mathbf{v} \in V$ is a sum of vectors from the eigenspaces (Theorem 8.2.18(a)), so $P_j = L_j(T)$ by linearity.

(a) If $ST = TS$, then $ST^m = T^m S$ for every $m \geq 0$ by induction on m, and so $Sq(T) = q(T)S$ for every polynomial q. In particular $SP_j = P_j S$, and for $\mathbf{v} \in E_{\lambda_j} = \operatorname{im} P_j$ we get $S\mathbf{v} = SP_j\mathbf{v} = P_j S\mathbf{v} \in \operatorname{im} P_j = E_{\lambda_j}$.

(b) By Theorem 8.2.18(c), Proposition 6.6.18, and $P_j^* = P_j$ (Corollary 6.6.29), $T^* = \sum_j \overline{\lambda_j} P_j = \sum_j \overline{\lambda_j} L_j(T) = p(T)$ by Proposition 7.2.27. □

Part (b) gives another description of normality: on a finite-dimensional complex space, T is normal if and only if T^* is a polynomial in T. (The converse holds because polynomials in T commute with T.)

Example 8.2.21 (A spectral decomposition). For $A = I_3 + J_3$ of Example 8.2.15, the eigenvalues are 4 and 1, so

$$P_4 = L_4(A) = \frac{A - I_3}{4 - 1} = \frac{1}{3}J_3, \quad P_1 = L_1(A) = \frac{A - 4I_3}{1 - 4} = I_3 - \frac{1}{3}J_3.$$

Indeed $\frac{1}{3}J_3 = \mathbf{q}_1\mathbf{q}_1^{\mathsf{T}}$ for $\mathbf{q}_1 = \frac{1}{\sqrt{3}}(1,1,1)$, as Theorem 6.4.31 predicts, and P_1 is the projection onto the plane $x_1 + x_2 + x_3 = 0$. The decomposition reads

$$A = 4 \cdot \frac{1}{3}J_3 + 1 \cdot \left(I_3 - \frac{1}{3}J_3\right).$$

We found P_1 without choosing a basis of the plane, which is exactly the choice that was not unique.

Exercise 8.2.22. ROUTINE Let $H = \begin{pmatrix} 3 & i \\ -i & 3 \end{pmatrix}$. Find the spectral decomposition $H = \lambda_1 P_1 + \lambda_2 P_2$, once with Lagrange polynomials and once from unit eigenvectors, and check that $P_1 P_2 = 0$.

Solution on page 1641

8.2.5 Functions of an operator

On an eigenspace E_λ, the operator T is just multiplication by λ, so T^2 is multiplication by λ^2, T^{-1} by λ^{-1}, and so on. The spectral decomposition lets us do this with any function of λ, not only with powers.

Definition 8.2.23 (Functions of an operator). Let T satisfy the spectral hypothesis, with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and $P_j = \operatorname{Pr}_{E_{\lambda_j}}$. For a function $f : \{\lambda_1, \ldots, \lambda_k\} \to \mathbb{F}$, define

$$f(T) = f(\lambda_1)P_1 + \cdots + f(\lambda_k)P_k.$$

For a matrix $A = UDU^*$ as in Theorem 8.2.18, with $D = \operatorname{diag}(\mu_1, \ldots, \mu_n)$, this is

$$f(A) = U \operatorname{diag}(f(\mu_1), \ldots, f(\mu_n))U^*.$$

The two formulas agree by the matrix form of Theorem 8.2.18, grouping equal μ_i. In particular, $f(A)$ does not depend on the choice of U, because the projections P_j do not. Note also that f only needs to be defined on the eigenvalues. When $\mathbb{F} = \mathbb{R}$, f is real-valued and $f(T)$ is an operator on the real space V.

Proposition 8.2.24 (Properties of $f(T)$). Let T satisfy the spectral hypothesis, with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, and let f, g be functions from $\{\lambda_1, \ldots, \lambda_k\}$ to $\mathbb{F}$, and $c \in \mathbb{F}$.

(a) $f(T)\mathbf{v} = f(\lambda_j)\mathbf{v}$ for every $\mathbf{v} \in E_{\lambda_j}$.

(b) If $p \in \mathbb{F}[x]$ and $f(\lambda_j) = p(\lambda_j)$ for every j, then $f(T) = p(T)$. In particular, $f(T) = p(T)$ for the interpolating polynomial $p = \sum_j f(\lambda_j) L_j$ of Proposition 8.2.20, and $f(T)$ commutes with every operator that commutes with T.

(c) $(f+g)(T) = f(T) + g(T)$, $(cf)(T) = cf(T)$, and $(fg)(T) = f(T)g(T)$.

(d) $f(T)^* = \overline{f}(T)$, where $\overline{f}(\lambda) = \overline{f(\lambda)}$. So $f(T)$ is normal, and it is self-adjoint if and only if every $f(\lambda_j)$ is real.

(e) The eigenvalues of $f(T)$ are the values $f(\lambda_1), \ldots, f(\lambda_k)$.

Proof. STANDARD

Uses. Definition 8.2.23 (Functions of an operator), Theorem 8.2.18 (Spectral decomposition), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 8.2.20 (Spectral projections are polynomials in T), Proposition 6.6.18 (Properties of the adjoint), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 8.2.19 (Uniqueness of the spectral decomposition)
Techniques. Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. Prove (a) first, from $P_i\mathbf{v} = \mathbf{0}$ for $i \neq j$ and $P_j\mathbf{v} = \mathbf{v}$.
Hint 2. Then every identity between operators built from T can be tested on each eigenspace, where everything is a scalar. V is the sum of the eigenspaces.
Hint 3. For (d), take adjoints term by term. For (e), use Proposition 8.2.19, after grouping equal values of f.
Sketch. (a) is immediate from $P_iP_j = 0$ and $P_j\mathbf{v} = \mathbf{v}$ on E_{λ_j}. Since $V = \oplus E_{\lambda_j}$, two operators agree as soon as they agree on each E_{λ_j}; on E_{λ_j}, $p(T)$, $f(T) + g(T)$, and $f(T)g(T)$ act as $p(\lambda_j)$, $f(\lambda_j) + g(\lambda_j)$, and $f(\lambda_j)g(\lambda_j)$, which gives (b) and (c). (d) $f(T)^* = \sum \overline{f(\lambda_j)} P_j$, and $f(T)$ commutes with $\overline{f}(T)$ by (c). (e) Group the P_j by the value of f and apply Proposition 8.2.19.

By Theorem 8.2.18(a), $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$, so two linear maps on V are equal as soon as they agree on each E_{λ_j}.

(a) Let $\mathbf{v} \in E_{\lambda_j}$. Then $P_j\mathbf{v} = \mathbf{v}$, and $P_i\mathbf{v} = \mathbf{0}$ for $i \neq j$ since $\mathbf{v} \in E_{\lambda_i}^{\perp}$. So $f(T)\mathbf{v} = f(\lambda_j)\mathbf{v}$.

(b) On E_{λ_j}, $p(T)$ acts as $p(\lambda_j)$ (Proposition 7.2.28) and $f(T)$ as $f(\lambda_j)$ by (a). These agree, so $f(T) = p(T)$. The interpolating polynomial satisfies $p(\lambda_i) = \sum_j f(\lambda_j) L_j(\lambda_i) = f(\lambda_i)$. If $ST = TS$, then S commutes with $p(T) = f(T)$, as in the proof of Proposition 8.2.20(a).

(c) On E_{λ_j}, $f(T) + g(T)$ and $cf(T)$ act as $f(\lambda_j) + g(\lambda_j)$ and $cf(\lambda_j)$. For $\mathbf{v} \in E_{\lambda_j}$, $g(T)\mathbf{v} = g(\lambda_j)\mathbf{v} \in E_{\lambda_j}$, so $f(T)g(T)\mathbf{v} = g(\lambda_j)f(T)\mathbf{v} = f(\lambda_j)g(\lambda_j)\mathbf{v} = (fg)(T)\mathbf{v}$.

(d) By Proposition 6.6.18 and $P_j^* = P_j$ (Corollary 6.6.29), $f(T)^* = \sum_j \overline{f(\lambda_j)} P_j = \overline{f}(T)$. By (c), $f(T)f(T)^* = (f\overline{f})(T) = (\overline{f}f)(T) = f(T)^* f(T)$. If every $f(\lambda_j)$ is real, $\overline{f} = f$ and $f(T)^* = f(T)$. Conversely, if $f(T)^* = f(T)$, then applying both sides to a nonzero $\mathbf{v} \in E_{\lambda_j}$ and using (a) gives $\overline{f(\lambda_j)} = f(\lambda_j)$.

(e) Let $\nu_1, \ldots, \nu_m$ be the distinct values among $f(\lambda_1), \ldots, f(\lambda_k)$, and let Q_i be the sum of the P_j with $f(\lambda_j) = \nu_i$. Then $f(T) = \sum_i \nu_i Q_i$. Each Q_i is nonzero, and by Theorem 8.2.18(b), $Q_i^2 = Q_i = Q_i^*$, $Q_iQ_{i'} = 0$ for $i \neq i'$, and $\sum_i Q_i = \mathrm{id}_V$. So by Corollary 6.6.29 the Q_i are orthogonal projections, and Proposition 8.2.19 says that the eigenvalues of $f(T)$ are $\nu_1, \ldots, \nu_m$. □

Technique 8.2.25 (Apply a function to the eigenvalues). To build an operator that does to T what a scalar function does to a number (a square root, an inverse, an exponential, the sign, a cutoff that keeps only large eigenvalues), keep the eigenspaces of T and apply the function to the eigenvalues: $f(T) = \sum_j f(\lambda_j) \operatorname{Pr}_{E_{\lambda_j}}$, or $f(A) = Uf(D)U^*$. Identities between such operators reduce to identities between numbers, one eigenvalue at a time (Proposition 8.2.24), which is Technique 7.4.25 in its strongest form.

Reach for it when the hypothesis says normal or self-adjoint and the problem asks for an operator with prescribed algebraic behavior: "find S with $S^2 = T$", "show that T has a self-adjoint inverse square root", "solve $\mathbf{x}' = A\mathbf{x}$". For a diagonalizable but not normal operator the same recipe works with SDS^{-1} in place

of UDU^* (Technique 7.4.22), but the result is not normal in general. The canonical example is the square root (Example 8.2.26); Section 8.3 shows that this square root is the only self-adjoint square root with nonnegative eigenvalues.

Used in. Exercise 8.2.39, Exercise 8.2.40, Exercise 8.2.47, Notation 8.3.10, Exercise 8.3.12, Exercise 8.3.37, Exercise 8.4.38, Exercise 8.2.33, Exercise 8.2.34, Exercise 8.2.38.

Example 8.2.26 (A square root). Let T be self-adjoint with every eigenvalue $\lambda_j \geq 0$, and let $f(\lambda) = \sqrt{\lambda}$. By Proposition 8.2.24, $\sqrt{T} = \sum_j \sqrt{\lambda_j} P_j$ is self-adjoint, its eigenvalues $\sqrt{\lambda_j}$ are nonnegative, and $\left(\sqrt{T}\right)^2 = \left(f^2\right)(T) = T$, since $f^2(\lambda) = \lambda$ on the eigenvalues.

For $A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$, the eigenvalues are 9 and 1, with $E_9 = \operatorname{span}((1,1))$ and $E_1 = \operatorname{span}((1,-1))$, so

$$P_9 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad P_1 = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \sqrt{A} = 3P_9 + 1P_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Check: $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$. The choice of $3 = \sqrt{9}$ and $1 = \sqrt{1}$ was ours; $3P_9 - P_1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is another symmetric square root of A, with eigenvalues 3 and -1.

Similarly, if no eigenvalue is 0, then $f(\lambda) = 1/\lambda$ gives $f(T) = \sum_j \lambda_j^{-1} P_j$ with $Tf(T) = (\lambda \cdot 1/\lambda)(T) = \mathrm{id}_V$, so $T^{-1} = \sum_j \lambda_j^{-1} P_j$.

BEYOND 21-242: The matrix exponential is not covered in 21-242.

Example 8.2.27 (The exponential). For a real symmetric A with $A = \sum_j \lambda_j P_j$, define $e^{tA} = \sum_j e^{t\lambda_j} P_j$ for $t \in \mathbb{R}$. By Proposition 8.2.24, $e^{sA}e^{tA} = e^{(s+t)A}$, and $e^{0A} = \mathrm{id}$. For $\mathbf{x}_0 \in \mathbb{R}^n$, the function

$$\mathbf{x}(t) = e^{tA}\mathbf{x}_0 = \sum_j e^{t\lambda_j} P_j \mathbf{x}_0$$

satisfies $\mathbf{x}'(t) = \sum_j \lambda_j e^{t\lambda_j} P_j \mathbf{x}_0 = A\mathbf{x}(t)$, using $AP_j = \lambda_j P_j$, and $\mathbf{x}(0) = \mathbf{x}_0$. This is the solution of Proposition 7.4.34 written with orthogonal projections: the component of $\mathbf{x}_0$ in each eigenspace grows or decays at its own rate $e^{t\lambda_j}$, independently of the others. One can show that $e^{tA} = \sum_{m \geq 0} t^m A^m / m!$, the series that defines e^{tA} for every square matrix, but we do not need this.

8.2.6 Quadratic forms and principal axes

The historical origin of the spectral theorem is the problem of finding the axes of an ellipse or ellipsoid given by an equation such as $5x^2 - 4xy + 8y^2 = 36$. A homogeneous quadratic polynomial in $x_1, \ldots, x_n$ can always be written as $\mathbf{x}^{\mathsf{T}} A \mathbf{x}$ with A symmetric: put the coefficient of x_i^2 at a_{ii}, and split the coefficient of $x_i x_j$ $(i \neq j)$ evenly between a_{ij} and a_{ji}. Section 12.2 develops quadratic forms systematically; here is the one fact the spectral theorem gives immediately.

Corollary 8.2.28 (Principal axes). Let $A \in \mathbb{R}^{n \times n}$ be symmetric, and write $A = QDQ^{\mathsf{T}}$ with Q orthogonal with columns $\mathbf{q}_1, \ldots, \mathbf{q}_n$ and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ (Theorem 8.2.12). For $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{y} = Q^{\mathsf{T}}\mathbf{x}$, so that $y_i = \langle \mathbf{x}, \mathbf{q}_i \rangle$. Then

$$\mathbf{x}^{\mathsf{T}} A \mathbf{x} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \quad \text{and} \quad \|\mathbf{y}\| = \|\mathbf{x}\|.$$

Proof. ROUTINE

Uses. Theorem 8.2.12 (Real spectral theorem), Proposition 2.4.23 (Transpose of a product), Definition 6.3.33 (Unitary and orthogonal matrices), Theorem 6.3.22 (Parseval's identity)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint. Substitute $A = QDQ^{\mathsf{T}}$ and regroup.
Sketch. $\mathbf{x}^{\mathsf{T}} QDQ^{\mathsf{T}} \mathbf{x} = \left(Q^{\mathsf{T}}\mathbf{x}\right)^{\mathsf{T}} D \left(Q^{\mathsf{T}}\mathbf{x}\right) = \mathbf{y}^{\mathsf{T}} D \mathbf{y} = \sum \lambda_i y_i^2$. And $\|\mathbf{y}\|^2 = \mathbf{x}^{\mathsf{T}} QQ^{\mathsf{T}} \mathbf{x} = \|\mathbf{x}\|^2$.

By Proposition 2.4.23, $\mathbf{x}^{\mathsf{T}}Q = \left(Q^{\mathsf{T}}\mathbf{x}\right)^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}$, so

$$\mathbf{x}^{\mathsf{T}}A\mathbf{x} = \left(\mathbf{x}^{\mathsf{T}}Q\right)D\left(Q^{\mathsf{T}}\mathbf{x}\right) = \mathbf{y}^{\mathsf{T}}D\mathbf{y} = \sum_{i=1}^{n}\lambda_i y_i^2.$$

The i-th entry of $Q^{\mathsf{T}}\mathbf{x}$ is $\mathbf{q}_i^{\mathsf{T}}\mathbf{x} = \langle \mathbf{x}, \mathbf{q}_i\rangle$. Since Q is orthogonal, $QQ^{\mathsf{T}} = I_n$, so $\|\mathbf{y}\|^2 = \mathbf{x}^{\mathsf{T}}QQ^{\mathsf{T}}\mathbf{x} = \|\mathbf{x}\|^2$ (alternatively, this is Theorem 6.3.22 in the basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$). □

In the coordinates y_i along the eigenvectors, the cross terms disappear. The lines $\operatorname{span}(\mathbf{q}_i)$ are the **principal axes** of the quadratic form.

Example 8.2.29 (An ellipse). The equation $5x^2 - 4xy + 8y^2 = 36$ is $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = 36$ with $A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$; the cross term $-4xy$ is split as $-2xy - 2yx$. Here $p_A = x^2 - 13x + 36 = (x-4)(x-9)$, with $E_4 = \operatorname{span}((2,1))$ and $E_9 = \operatorname{span}((-1,2))$, so $\mathbf{q}_1 = \frac{1}{\sqrt{5}}(2,1)$ and $\mathbf{q}_2 = \frac{1}{\sqrt{5}}(-1,2)$. In the coordinates $y_1 = (2x+y)/\sqrt{5}$, $y_2 = (-x+2y)/\sqrt{5}$, Corollary 8.2.28 turns the equation into

$$4y_1^2 + 9y_2^2 = 36, \quad \text{that is,} \quad \frac{y_1^2}{9} + \frac{y_2^2}{4} = 1.$$

This is an ellipse with semi-axis 3 along $\mathbf{q}_1$ and semi-axis 2 along $\mathbf{q}_2$. Its vertices are $\pm 3\mathbf{q}_1 = \pm\frac{3}{\sqrt{5}}(2,1)$ and $\pm 2\mathbf{q}_2 = \pm\frac{2}{\sqrt{5}}(-1,2)$; for instance $\mathbf{x} = \frac{3}{\sqrt{5}}(2,1)$ gives $\frac{1}{5}(5\cdot 36 - 4\cdot 18 + 8\cdot 9) = 36$.

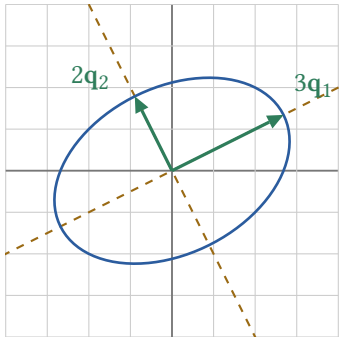


Figure 8.2.1. The ellipse $5x^2 - 4xy + 8y^2 = 36$ and its principal axes (dashed), along the eigenvectors $\mathbf{q}_1 = \frac{1}{\sqrt{5}}(2,1)$ and $\mathbf{q}_2 = \frac{1}{\sqrt{5}}(-1,2)$ of $A = \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix}$. The semi-axes are $\sqrt{36/4} = 3$ and $\sqrt{36/9} = 2$.

Remark 8.2.30 (Classifying conics). The same computation classifies every curve $ax^2 + 2bxy + cy^2 = 1$. Let $\lambda_1 \geq \lambda_2$ be the eigenvalues of $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, so the curve is $\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1$ in the rotated coordinates of Corollary 8.2.28.

- If $\lambda_1 \geq \lambda_2 > 0$, it is an ellipse with semi-axes $1/\sqrt{\lambda_1}$ and $1/\sqrt{\lambda_2}$ along $\mathbf{q}_1$ and $\mathbf{q}_2$ (a circle if $\lambda_1 = \lambda_2$).
- If $\lambda_1 > 0 > \lambda_2$, it is a hyperbola whose vertices $\pm\mathbf{q}_1/\sqrt{\lambda_1}$ lie on the axis of the positive eigenvalue.
- If $\lambda_1 > 0 = \lambda_2$, it is the pair of parallel lines $y_1 = \pm 1/\sqrt{\lambda_1}$.
- If $\lambda_1 \leq 0$, the left side is never positive and the curve is empty.

Since $\lambda_1\lambda_2 = \det A = ac - b^2$ and $\lambda_1 + \lambda_2 = \operatorname{tr} A = a + c$, the type can be read off without computing eigenvalues: an ellipse when $ac - b^2 > 0$ and $a > 0$, a hyperbola when $ac - b^2 < 0$. This is the discriminant test from analytic geometry, now with a proof.

BEYOND 21-242: Infinite-dimensional spaces are not covered in 21-242.

Remark 8.2.31 (Infinite dimensions). The spectral theorems need V to be finite-dimensional. On $C([0,1],\mathbb{C})$ with $\langle f, g\rangle = \int_0^1 f\overline{g}$, the multiplication operator $(Mf)(t) = tf(t)$ is self-adjoint (Example 6.6.14), but it has no eigenvalues at all: if $tf(t) = \lambda f(t)$ for all t, then $f(t) = 0$ for every $t \neq \lambda$, and by

continuity $f = 0$. So no eigenvector exists, let alone an orthonormal eigenbasis. The infinite-dimensional theory replaces the sum $\sum \lambda_j P_j$ by an integral over a "continuous spectrum"; Section 15.3 gives a first look.

Exercises

Exercise 8.2.32. STANDARD Find an orthogonal Q and a diagonal D with

$$\begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix} = QDQ^{\mathsf{T}}.$$

Solution on page 1642

Exercise 8.2.33. STANDARD Let V be a finite-dimensional complex inner product space and $T \in \mathrm{L}(V)$ normal. Show:

(a) T is self-adjoint if and only if every eigenvalue of T is real;
(b) $T^* = -T$ if and only if every eigenvalue of T is purely imaginary (in $i\mathbb{R}$);
(c) T is an orthogonal projection if and only if every eigenvalue of T is 0 or 1;
(d) $T^*T = \mathrm{id}_V$ if and only if every eigenvalue of T has absolute value 1.

Then show by an example that (a) fails for operators that are not normal.

Solution on page 1643

Exercise 8.2.34. STANDARD Prove or disprove each statement. Throughout, V is a finite-dimensional inner product space.

(a) If V is complex, $T \in \mathrm{L}(V)$ is normal, and $T^m = 0$ for some $m \geq 1$, then $T = 0$.
(b) If V is real, $T \in \mathrm{L}(V)$ is self-adjoint, and $T^3 + T = 2\mathrm{id}_V$, then $T = \mathrm{id}_V$.
(c) The statement in (b) remains true without the hypothesis that T is self-adjoint.
(d) If $A \in \mathbb{R}^{n\times n}$ is symmetric, then every basis of $\mathbb{R}^n$ consisting of eigenvectors of A is orthogonal.
(e) If $A \in \mathbb{R}^{n\times n}$ is diagonalizable over $\mathbb{R}$ and its eigenspaces are pairwise orthogonal, then A is symmetric.

Solution on page 1643

Exercise 8.2.35. STANDARD Find the error in the following "proof" that every normal operator T on a finite-dimensional real inner product space V with $\dim V \geq 1$ has an orthonormal eigenbasis. Then give a counterexample to the claim, and name a hypothesis on T that repairs the proof.

"Proof." Induct on $n = \dim V$; the case $n = 1$ is clear. Let $n \geq 2$, and let $\mathbf{u}$ be a unit eigenvector of T, with $T\mathbf{u} = \lambda\mathbf{u}$. Since T is normal, $\|(T^* - \lambda\,\mathrm{id})\mathbf{u}\| = \|(T - \lambda\,\mathrm{id})\mathbf{u}\| = 0$, so $\mathbf{u}$ is an eigenvector of T^* as well. Hence $W = \mathrm{span}(\mathbf{u})^{\perp}$ is invariant under T and T^*, and $T{\restriction}_W$ is normal. By the induction hypothesis, W has an orthonormal eigenbasis for $T{\restriction}_W$, and adding $\mathbf{u}$ gives one for T.

Solution on page 1643

Exercise 8.2.36. STANDARD Let V be a finite-dimensional inner product space over $\mathbb{F}$ and let $S, T \in \mathrm{L}(V)$ be self-adjoint. Show that $ST = TS$ if and only if V has an orthonormal basis consisting of vectors that are eigenvectors of both S and T.

Solution on page 1644

Exercise 8.2.37. STANDARD Let V be a finite-dimensional complex inner product space, $T \in \mathrm{L}(V)$ normal, and $S \in \mathrm{L}(V)$ with $ST = TS$.

(a) Show that $ST^* = T^*S$.
(b) Show by an example that (a) can fail if T is not normal.
(c) Suppose that S is normal as well. Show that S, S^*, T, T^* pairwise commute, and deduce that ST and $S + T$ are normal.

Solution on page 1644

Exercise 8.2.38. STANDARD

(a) Find the symmetric $A \in \mathbb{R}^{2\times 2}$ with eigenvalues 1 and 3 such that $(1, 2)$ is an eigenvector for 3, and show that it is unique.
(b) Find a matrix $B \in \mathbb{R}^{2\times 2}$, not symmetric, with eigenvalues 1 and 3 such that $(1, 2)$ is an eigenvector for 3.

Solution on page 1645

Exercise 8.2.39. STANDARD Let $A = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}$.

(a) Find the symmetric B with nonnegative eigenvalues such that $B^2 = A$.
(b) Find all $B \in \mathbb{R}^{2\times 2}$ with $B^2 = A$.
(c) Explain why I_2, in contrast, has infinitely many square roots in $\mathbb{R}^{2\times 2}$.

Solution on page 1645

Exercise 8.2.40. STANDARD BEYOND 21-242 Using Example 8.2.27 (and its analogue $e^{tN} = \sum_j e^{t\lambda_j} P_j$ for normal $N \in \mathbb{C}^{n\times n}$), compute e^{tA} for $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and e^{tJ} for $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, where $t \in \mathbb{R}$. Interpret e^{tJ} geometrically.

Solution on page 1646

Exercise 8.2.41. STANDARD Identify each curve in $\mathbb{R}^2$ (ellipse, hyperbola, pair of lines, or empty), and find its principal axes.

(a) $x^2 + 4xy + y^2 = 1$
(b) $3x^2 + 2xy + 3y^2 = 8$
(c) $2x^2 + 4xy + 2y^2 = 1$

Solution on page 1646

Exercise 8.2.42. STANDARD Let $T \in \mathrm{L}(V)$ be self-adjoint, and let $\lambda_{\min}$ and $\lambda_{\max}$ be its smallest and largest eigenvalues.

(a) Show that $\lambda_{\min}\|\mathbf{v}\|^2 \le \langle T\mathbf{v}, \mathbf{v}\rangle \le \lambda_{\max}\|\mathbf{v}\|^2$ for every $\mathbf{v} \in V$, with equality on the right if and only if $\mathbf{v} \in E_{\lambda_{\max}}$.
(b) Deduce that for a real symmetric $A \in \mathbb{R}^{n\times n}$, $\lambda_{\min} \le a_{ii} \le \lambda_{\max}$ for every i.
(c) Show that $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}$ has an eigenvalue greater than 3.

Solution on page 1647

Exercise 8.2.43. HARD Let $A \in \mathbb{R}^{n\times n}$ be symmetric and nonzero. Show that

$$\operatorname{rank} A \geq \frac{(\operatorname{tr} A)^2}{\operatorname{tr}(A^2)},$$

and determine when equality holds.

Solution on page 1647

Exercise 8.2.44. HARD Let V be a finite-dimensional complex inner product space and $T \in \mathrm{L}(V)$. Show that T is normal if and only if every eigenvector of T is also an eigenvector of T^*.

Solution on page 1648

Exercise 8.2.45. STANDARD

(a) Show that $A \in \mathbb{R}^{2\times 2}$ satisfies $A^{\mathsf{T}}A = AA^{\mathsf{T}}$ if and only if A is symmetric or $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ for some $a, b \in \mathbb{R}$.

(b) Deduce that a real normal 2×2 matrix that is not symmetric has no real eigenvalue, and interpret $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ geometrically.

Solution on page 1648

Exercise 8.2.46. STANDARD Adapt the third proof of Theorem 8.2.5 (the sum of the eigenspaces is everything) to give another proof of (I) ⇒ (II) in Theorem 8.2.12. Point out exactly where self-adjointness is used.

Solution on page 1649

Exercise 8.2.47. HARD Let $H \in \mathbb{C}^{n\times n}$ be Hermitian.

(a) Show that $I_n + iH$ is invertible.

(b) Show that $C = (I_n - iH)(I_n + iH)^{-1}$ is unitary, and that -1 is not an eigenvalue of C.

(c) Conversely, show that every unitary C without the eigenvalue -1 arises in this way from a unique Hermitian H.

Solution on page 1649

Exercise 8.2.48. CHALLENGE BEYOND 21-242 Let $A \in \mathbb{R}^{n\times n}$ be normal ($A^{\mathsf{T}}A = AA^{\mathsf{T}}$). Show that there is an orthogonal Q such that $Q^{\mathsf{T}}AQ$ is block diagonal, with each block either 1×1 or of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b > 0$.

Solution on page 1650

8.3 Positive operators

Among complex numbers, the nonnegative reals stand out: they are the numbers of the form $\overline{w}w$, and also the real numbers that are squares of real numbers. In the analogy between operators and complex numbers (Section 6.6.5), the adjoint plays the role of complex conjugation and self-adjoint operators play the role of real numbers. This section finds the operators that play the role of nonnegative numbers. We have already met several: orthogonal projections, the matrices A^*A of the normal equations (Theorem 6.5.5), and Gram matrices (Definition 6.3.28). Positive operators have nonnegative eigenvalues and a unique positive square root, and the positive definite ones are exactly the matrices of inner products. In Section 9.1 the positive operator T^*T produces the singular value decomposition.

Throughout, V is a finite-dimensional inner product space over $\mathbb{F}$ with $\dim V \geq 1$, and $T \in \mathrm{L}(V)$; only the definitions make sense more generally. Matrices in $\mathbb{F}^{n\times n}$ act on $\mathbb{F}^n$ with the dot product, so that $\langle A\mathbf{x}, \mathbf{x}\rangle = \mathbf{x}^* A\mathbf{x}$. For an orthonormal basis $\mathcal{O}$ we write $[T]_{\mathcal{O}}$ for $[T]_{\mathcal{O}\mathcal{O}}$, as in Section 8.2.

8.3.1 Positive operators

Definition 8.3.1 (Positive operator). Let V be an inner product space. An operator $T \in \mathrm{L}(V)$ is **positive** if T is self-adjoint and

$$\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0 \quad \text{for every } \mathbf{v} \in V.$$

A matrix $A \in \mathbb{F}^{n\times n}$ is **positive** if T_A is positive on $\mathbb{F}^n$ with the dot product, that is, if $A^* = A$ and $\mathbf{x}^* A\mathbf{x} \geq 0$ for every $\mathbf{x} \in \mathbb{F}^n$.

The name is traditional but a little misleading: "positive" here means "nonnegative", in the sense in which $0 \geq 0$. For matrices the same notion is often called **positive semidefinite**. The strict version, $\langle T\mathbf{v}, \mathbf{v}\rangle > 0$ for $\mathbf{v} \neq \mathbf{0}$, is positive definiteness (Definition 8.3.14).

Intuition (What the condition says). In a real space, $\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ says that $T\mathbf{v}$ never makes an obtuse angle with $\mathbf{v}$: T may stretch, shrink, and turn a vector, but never by more than a right angle. For a self-adjoint T the spectral theorem says more: T stretches V along perpendicular axes, and the stretch factors are its eigenvalues. On the axis of a unit eigenvector $\mathbf{u}$ with $T\mathbf{u} = \lambda\mathbf{u}$ we have $\langle T\mathbf{u}, \mathbf{u}\rangle = \lambda$, so positivity says at least that no axis is flipped. Theorem 8.3.5 shows that this is all it says.

Example 8.3.2 (Positive operators).

(a) Every orthogonal projection $P = \mathrm{Pr}_U$ is positive. It is self-adjoint and $P^2 = P$ (Corollary 6.6.29), so

$$\langle P\mathbf{v}, \mathbf{v}\rangle = \langle P^2\mathbf{v}, \mathbf{v}\rangle = \langle P\mathbf{v}, P\mathbf{v}\rangle = \|P\mathbf{v}\|^2 \geq 0.$$

(b) Let W be an inner product space and let $S \in \mathrm{L}(V, W)$ have an adjoint. Then $S^*S \in \mathrm{L}(V)$ is positive: $(S^*S)^* = S^*(S^*)^* = S^*S$ by Proposition 6.6.18, and $\langle S^*S\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, S\mathbf{v}\rangle = \|S\mathbf{v}\|^2 \geq 0$. In particular A^*A is positive for every $A \in \mathbb{F}^{m\times n}$; it is the coefficient matrix of the normal equations (Theorem 6.5.5).

(c) The Gram matrix G of a list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V (Definition 6.3.28) is positive. Its entries satisfy $g_{ji} = \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \overline{\langle \mathbf{v}_j, \mathbf{v}_i\rangle} = \overline{g_{ij}}$, so $G^* = G$, and $\mathbf{c}^* G\mathbf{c} = \|c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k\|^2 \geq 0$ by Proposition 6.3.30. For matrices, (b) is the special case of the columns of A, whose Gram matrix is A^*A.

(d) $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is positive. It is Hermitian, and for $\mathbf{x} \in \mathbb{C}^2$,

$$\mathbf{x}^* A\mathbf{x} = 2|x_1|^2 - 2\,\mathrm{Re}(\overline{x_1}x_2) + 2|x_2|^2 = |x_1 - x_2|^2 + |x_1|^2 + |x_2|^2 \geq 0.$$

Non-example 8.3.3 (Matrices that are not positive).

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is symmetric with positive entries, but for $\mathbf{x} = (1, -1)$ we get $\mathbf{x}^{\mathsf{T}} A\mathbf{x} = 1 - 2 - 2 + 1 = -2 < 0$. Together with Example 8.3.2(d), this shows that the signs of the entries say little about positivity.

(b) On $\mathbb{R}^2$, let $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Then $\mathbf{x}^{\mathsf{T}} B\mathbf{x} = x_1^2 - x_1x_2 + x_2x_1 + x_2^2 = \|\mathbf{x}\|^2 \geq 0$ for every $\mathbf{x} \in \mathbb{R}^2$, but B is not symmetric, so it is not positive. Geometrically, B is $\sqrt{2}$ times the rotation by $\pi/4$: it turns every vector by an angle of $\pi/4$, less than a right angle, and its characteristic polynomial $(x-1)^2 + 1$ has no real roots.

Remark 8.3.4 (Why self-adjointness is part of the definition). Over $\mathbb{C}$ the self-adjointness in Definition 8.3.1 comes for free. If $\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ for every $\mathbf{v}$, then in particular every value $\langle T\mathbf{v}, \mathbf{v}\rangle$ is real, and T is self-adjoint by Proposition 6.6.30. So on a complex space, T is positive if and only if $\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ for every $\mathbf{v} \in V$.

Over $\mathbb{R}$ it does not come for free, and Non-example 8.3.3(b) shows why we must ask for it. On a real space $\langle T^*\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \langle T\mathbf{v}, \mathbf{v}\rangle$, so

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \left\langle \frac{1}{2}(T + T^*)\mathbf{v}, \mathbf{v}\right\rangle \quad \text{for every } \mathbf{v} \in V.$$

The values $\langle T\mathbf{v}, \mathbf{v}\rangle$ only see the self-adjoint part $(T + T^*)/2$ of T (Exercise 6.1.52), and the condition $\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ says nothing about the rest. For B, the self-adjoint part is I_2 and the rest is the rotation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Without the self-adjointness requirement, B would count as positive although it has no real eigenvalues, it is not of the form S^*S (such operators are self-adjoint), and it has no self-adjoint square root (the square of a self-adjoint operator is self-adjoint). With the requirement, the real and complex notions agree: a real matrix is positive in $\mathbb{R}^{n\times n}$ if and only if it is positive in $\mathbb{C}^{n\times n}$ (Exercise 8.3.30), and the characterizations below hold over both fields with the same proofs.

Intuition (Positive operators as nonnegative numbers). A complex number z is a nonnegative real number exactly when it can be written in any one of the forms $z = \overline{w}w$ with $w \in \mathbb{C}$, $z = r^2$ with $r \in \mathbb{R}$, or $z = r^2$ with $r \geq 0$. Each form has an operator analogue: $T = S^*S$, $T = R^2$ with R self-adjoint, and $T = R^2$ with R positive. One more description makes sense only for operators. By the spectral theorem, a self-adjoint operator is described by real numbers, its eigenvalues, attached to perpendicular axes, and we can ask that these numbers be nonnegative. The next theorem says that all of these descriptions agree with the definition.

8.3.2 Characterizations

Theorem 8.3.5 (Characterizations of positive operators). Let V be a finite-dimensional inner product space with $\dim V \geq 1$, and let $T \in \mathrm{L}(V)$. The following are equivalent:

(I) T is positive;

(II) T is self-adjoint and every eigenvalue of T is nonnegative;

(III) $[T]_{\mathcal{O}}$ is diagonal with nonnegative diagonal entries for some orthonormal basis $\mathcal{O}$ of V;

(IV) $T = R^2$ for some positive $R \in \mathrm{L}(V)$;

(V) $T = R^2$ for some self-adjoint $R \in \mathrm{L}(V)$;

(VI) $T = S^*S$ for some $S \in \mathrm{L}(V, W)$, where W is a finite-dimensional inner product space;

(VII) for some orthonormal basis $\mathcal{O}$ of V, $[T]_{\mathcal{O}}$ is the Gram matrix of a list of vectors in some inner product space.

In matrix form: $A \in \mathbb{F}^{n\times n}$ is positive if and only if $A = UDU^*$ with U unitary and D diagonal with nonnegative entries, if and only if $A = B^*B$ for some $B \in \mathbb{F}^{m\times n}$, if and only if A is a Gram matrix.

Over $\mathbb{C}$, condition (I) may be replaced by "$\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ for every $\mathbf{v}$" (Remark 8.3.4).

Intuition (How the cycle is arranged). Seven conditions need seven implications if we arrange them in a cycle (Technique 2.7.1). Most of the cycle runs downhill, each condition handing the next one a witness: (II) $\Rightarrow$ (III) is the spectral theorem, (III) $\Rightarrow$ (IV) takes square roots of the diagonal entries, and (IV) $\Rightarrow$ (V) $\Rightarrow$ (VI) forget information. The cycle closes with one computation. Once $T = S^*S$, we have $\langle T\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2$, and the Gram condition (VII) is the same computation written in coordinates. The only deep step is the spectral theorem.

Proof. STANDARD

Uses. Definition 8.3.1 (Positive operator), Theorem 8.2.12 (Real spectral theorem), Corollary 8.2.7 (Hermitian matrices), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 2.2.1 (Linear maps are determined by a basis),

Corollary 6.6.28 (Self-adjointness in an orthonormal basis), Proposition 6.3.18 (Coordinates in an orthonormal basis), Definition 6.3.28 (Gram matrix), Proposition 6.3.30 (Gram matrices detect independence), Theorem 6.6.15 (Existence of adjoints), Definition 8.2.1 (Unitarily and orthogonally diagonalizable)
Techniques. Proving TFAE (Technique 2.7.1), Gram matrices (Technique 6.3.31), Pass to orthonormal coordinates (Technique 6.7.2)
Hint 1. Use the cycle (I) ⇒ (II) ⇒ ⋯ ⇒ (VII) ⇒ (I). For (I) ⇒ (II), put an eigenvector into $\langle T\mathbf{v}, \mathbf{v}\rangle$.
Hint 2. (II) ⇒ (III) is a spectral theorem. For (III) ⇒ (IV), define R on the orthonormal eigenbasis by taking square roots of the eigenvalues.
Hint 3. For (VI) ⇒ (VII), compute the entries $\langle T\mathbf{e}_j, \mathbf{e}_i\rangle$ of $[T]_{\mathcal{O}}$ when $T = S^*S$. For (VII) ⇒ (I), expand $\langle T\mathbf{v}, \mathbf{v}\rangle$ in the coordinates $\mathbf{c} = [\mathbf{v}]_{\mathcal{O}}$ and use Proposition 6.3.30.
Sketch. (I) ⇒ (II): $\lambda\|\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ for an eigenvector. (II) ⇒ (III): spectral theorem. (III) ⇒ (IV): $R\mathbf{e}_i = \sqrt{d_i}\mathbf{e}_i$ on the eigenbasis. (IV) ⇒ (V) ⇒ (VI): R is self-adjoint, so $T = R^*R$. (VI) ⇒ (VII): $\langle T\mathbf{e}_j, \mathbf{e}_i\rangle = \langle S\mathbf{e}_j, S\mathbf{e}_i\rangle$, so $[T]_{\mathcal{O}}$ is the Gram matrix of $S\mathbf{e}_1, \ldots, S\mathbf{e}_n$. (VII) ⇒ (I): a Gram matrix is Hermitian, and $\langle T\mathbf{v}, \mathbf{v}\rangle = \mathbf{c}^*G\mathbf{c} = \left\|\sum_j c_j\mathbf{w}_j\right\|^2$.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (IV) ⇒ (V) ⇒ (VI) ⇒ (VII) ⇒ (I).

(I) ⇒ (II). T is self-adjoint by definition. Let λ be an eigenvalue of T, with eigenvector $\mathbf{v} \neq \mathbf{0}$. Then

$$\lambda\|\mathbf{v}\|^2 = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, \mathbf{v}\rangle \geq 0,$$

and since $\|\mathbf{v}\|^2 > 0$, $\lambda = \langle T\mathbf{v}, \mathbf{v}\rangle / \|\mathbf{v}\|^2$ is a nonnegative real number.

(II) ⇒ (III). By Theorem 8.2.12 if $\mathbb{F} = \mathbb{R}$, or Corollary 8.2.7 if $\mathbb{F} = \mathbb{C}$, V has an orthonormal basis $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ of eigenvectors of T, say $T\mathbf{e}_i = d_i\mathbf{e}_i$. Then $[T]_{\mathcal{O}} = \operatorname{diag}(d_1, \ldots, d_n)$ (Theorem 7.4.2), and each d_i is an eigenvalue of T, so $d_i \geq 0$.

(III) ⇒ (IV). Let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis with $[T]_{\mathcal{O}} = \operatorname{diag}(d_1, \ldots, d_n)$ and every $d_i \geq 0$, so that $T\mathbf{e}_i = d_i\mathbf{e}_i$. By Theorem 2.2.1 there is $R \in \mathrm{L}(V)$ with $R\mathbf{e}_i = \sqrt{d_i}\mathbf{e}_i$ for every i. Its matrix $[R]_{\mathcal{O}} = \operatorname{diag}\left(\sqrt{d_1}, \ldots, \sqrt{d_n}\right)$ is real and diagonal, hence Hermitian, so R is self-adjoint (Corollary 6.6.28). For $\mathbf{v} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$, orthonormality gives

$$\langle R\mathbf{v}, \mathbf{v}\rangle = \left\langle \sum_i \sqrt{d_i}c_i\mathbf{e}_i, \sum_j c_j\mathbf{e}_j \right\rangle = \sum_i \sqrt{d_i}|c_i|^2 \geq 0,$$

so R is positive. Finally $R^2\mathbf{e}_i = \sqrt{d_i}\sqrt{d_i}\mathbf{e}_i = d_i\mathbf{e}_i = T\mathbf{e}_i$ for every i, so $R^2 = T$ by the uniqueness part of Theorem 2.2.1.

(IV) ⇒ (V). A positive operator is self-adjoint.

(V) ⇒ (VI). Take $W = V$ and $S = R$. Then $S^*S = RR = T$.

(VI) ⇒ (VII). Let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be *any* orthonormal basis of V. By Proposition 6.3.18, the i-th coordinate of $T\mathbf{e}_j$ is $\langle T\mathbf{e}_j, \mathbf{e}_i\rangle$, so the (i, j) entry of $[T]_{\mathcal{O}}$ is

$$\langle T\mathbf{e}_j, \mathbf{e}_i\rangle = \langle S^*S\mathbf{e}_j, \mathbf{e}_i\rangle = \langle S\mathbf{e}_j, S\mathbf{e}_i\rangle.$$

This is the (i, j) entry of the Gram matrix of the list $S\mathbf{e}_1, \ldots, S\mathbf{e}_n$ in W (Definition 6.3.28).

(VII) ⇒ (I). Let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis such that $G = [T]_{\mathcal{O}}$ is the Gram matrix of a list $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in an inner product space W. As in Example 8.3.2(c), $G^* = G$, so $[T^*]_{\mathcal{O}} = G^* = [T]_{\mathcal{O}}$ by Theorem 6.6.15, and $T^* = T$. Let $\mathbf{v} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$ and $\mathbf{c} = (c_1, \ldots, c_n)$. Since $g_{ij} = \langle T\mathbf{e}_j, \mathbf{e}_i\rangle$,

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \sum_{i,j} c_j\overline{c_i}\langle T\mathbf{e}_j, \mathbf{e}_i\rangle = \sum_{i,j} \overline{c_i}g_{ij}c_j = \mathbf{c}^*G\mathbf{c} = \|c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n\|^2 \geq 0$$

by Proposition 6.3.30. So T is positive.

For the matrix form, apply the theorem to T_A on $\mathbb{F}^n$; its matrix in the standard basis $\mathcal{E}$ is A. Condition (III) for T_A says that $\mathbb{F}^n$ has an orthonormal eigenbasis for A with nonnegative eigenvalues, which means $A = UDU^*$ with U unitary (its columns are the eigenbasis) and D diagonal with nonnegative entries, as

explained after Definition 8.2.1. If $A = B^*B$ with $B \in \mathbb{F}^{m\times n}$, then $T_A = (T_B)^* T_B$, because the adjoint of T_B is T_{B^*} (Theorem 6.6.15); this is (VI). Conversely, (V) gives $A = B^*B$ with $B = [R]_{\mathcal{E}}$. Finally, if A is a Gram matrix, then (VII) holds with $\mathcal{O} = \mathcal{E}$, so A is positive; and if A is positive, the proof of (VI) $\Rightarrow$ (VII), which worked for every orthonormal basis, shows that $A = [T_A]_{\mathcal{E}}$ is a Gram matrix. □

The proof of (VI) $\Rightarrow$ (VII) did not depend on the orthonormal basis. So for a positive T the matrix in *every* orthonormal basis is a Gram matrix, and in particular its diagonal entries $\langle T\mathbf{e}_i, \mathbf{e}_i\rangle$ are nonnegative. For bases that are not orthonormal this fails (Exercise 8.3.29). Note also that conditions (V) and (VI) look weaker than (IV). A self-adjoint square root may have negative eigenvalues, and S need not be self-adjoint or even map V to itself, yet each of them forces T to be positive.

Example 8.3.6 (Three certificates of positivity). Let

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Each of three conditions of Theorem 8.3.5 can be checked directly.

Eigenvalues (II). A is symmetric, and expanding $\det(xI_3 - A)$ along the first row gives $p_A = (x-2)^3 - 2(x-2) = (x-2)\big(x^2 - 4x + 2\big)$. The eigenvalues 2 and $2 \pm \sqrt{2}$ are all positive.

The form (I). For $\mathbf{x} \in \mathbb{R}^3$,

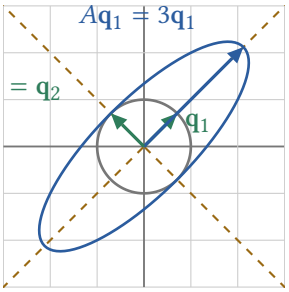
$$\mathbf{x}^{\mathsf{T}} A\mathbf{x} = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 \geq 0.$$

A factorization (VI). $A = B^{\mathsf{T}}B$ for the 4×3 difference matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \qquad B\mathbf{x} = (x_1,\, x_2 - x_1,\, x_3 - x_2,\, -x_3),$$

so A is the Gram matrix of the columns of B.

The last two certificates are the same computation, since $\mathbf{x}^{\mathsf{T}}B^{\mathsf{T}}B\mathbf{x} = \|B\mathbf{x}\|^2$ is exactly the sum of squares. The factorization is the most useful in practice: a sum of squares or a product B^*B certifies positivity without computing any eigenvalue. (The matrix A is the second-difference matrix, a close relative of the graph Laplacians of Section 16.5.)



A symmetric 2×2 matrix A maps the unit circle to an ellipse with axes along the orthonormal eigenvectors $\mathbf{q}_1, \mathbf{q}_2$. A is positive exactly when neither $\lambda_1\mathbf{q}_1$ nor $\lambda_2\mathbf{q}_2$ is reversed, which is the same as saying that $A\mathbf{x}$ never points more than a right angle away from $\mathbf{x}$. Drag $A\mathbf{e}_1$ and $A\mathbf{e}_2$, or type entries, until an eigenvalue turns negative, then move $\mathbf{x}$ to that eigenvector. (Interactive in the web version.)

Exercise 8.3.7. ROUTINE Decide which of the following matrices are positive, and justify each answer by a condition of Theorem 8.3.5 or by a vector $\mathbf{x}$ with $\mathbf{x}^* A\mathbf{x} < 0$.

(a) $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

(b) $\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

(c) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$.

(d) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

(e) $\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

Solution on page 1651

The last characterization we need concerns vectors on which the form $\langle T\mathbf{v}, \mathbf{v}\rangle$ vanishes. For a positive number t, $tx^2 = 0$ forces $x = 0$; for a positive operator, the analogue is that $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ forces $T\mathbf{v} = \mathbf{0}$.

Lemma 8.3.8 (Where the form vanishes). Let $T \in \mathrm{L}(V)$ be positive and $\mathbf{v} \in V$. Then $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ if and only if $T\mathbf{v} = \mathbf{0}$. Consequently

$$\ker T = \{\mathbf{v} \in V \mid \langle T\mathbf{v}, \mathbf{v}\rangle = 0\}.$$

Proof (in an eigenbasis). ROUTINE

Uses. Theorem 8.3.5 (Characterizations of positive operators), Proposition 6.3.18 (Coordinates in an orthonormal basis)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Pass to orthonormal coordinates (Technique 6.7.2)
Hint 1. Only one direction needs proof. Take an orthonormal basis as in Theorem 8.3.5(III) and expand $\langle T\mathbf{v}, \mathbf{v}\rangle$ in it.
Hint 2. You get $\sum_i d_i |c_i|^2$ with every $d_i \geq 0$. When is such a sum 0?
Sketch. In an orthonormal eigenbasis with $T\mathbf{e}_i = d_i\mathbf{e}_i$, $d_i \geq 0$, and $\mathbf{v} = \sum c_i\mathbf{e}_i$: $\langle T\mathbf{v}, \mathbf{v}\rangle = \sum_i d_i |c_i|^2$. If this is 0, every term is 0, so $d_i c_i = 0$ for every i, and $T\mathbf{v} = \sum_i d_i c_i \mathbf{e}_i = \mathbf{0}$.

If $T\mathbf{v} = \mathbf{0}$, then $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$. Conversely, by Theorem 8.3.5(III) there is an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ with $T\mathbf{e}_i = d_i\mathbf{e}_i$ and every $d_i \geq 0$. Write $\mathbf{v} = \sum_i c_i\mathbf{e}_i$. Then

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \left\langle \sum_i d_i c_i \mathbf{e}_i, \sum_j c_j \mathbf{e}_j \right\rangle = \sum_i d_i |c_i|^2.$$

If this is 0, then, being a sum of nonnegative terms, each term $d_i|c_i|^2$ is 0. So for each i, either $d_i = 0$ or $c_i = 0$, and in both cases $d_i c_i = 0$. Hence $T\mathbf{v} = \sum_i d_i c_i \mathbf{e}_i = \mathbf{0}$. □

Proof (with a square root). ROUTINE

Uses. Theorem 8.3.5 (Characterizations of positive operators), Definition 6.1.26 (Norm)
Techniques. Show a vector is zero via its norm (Technique 6.1.36), Move an operator across the inner product (Technique 6.6.20)
Hint. Write $T = R^2$ with R positive (Theorem 8.3.5(IV)) and move one factor R across the inner product.
Sketch. With $T = R^2$ and R self-adjoint, $\langle T\mathbf{v}, \mathbf{v}\rangle = \langle R\mathbf{v}, R\mathbf{v}\rangle = \|R\mathbf{v}\|^2$. If this is 0, then $R\mathbf{v} = \mathbf{0}$ and $T\mathbf{v} = R(R\mathbf{v}) = \mathbf{0}$.

By Theorem 8.3.5(IV), $T = R^2$ for some positive, hence self-adjoint, R. Then

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \langle R(R\mathbf{v}), \mathbf{v}\rangle = \langle R\mathbf{v}, R\mathbf{v}\rangle = \|R\mathbf{v}\|^2.$$

If this is 0, then $R\mathbf{v} = \mathbf{0}$, and $T\mathbf{v} = R(R\mathbf{v}) = R\mathbf{0} = \mathbf{0}$. □

The second proof is the general pattern: once the form $\langle T\mathbf{v}, \mathbf{v}\rangle$ is a squared norm $\|R\mathbf{v}\|^2$, its vanishing kills $R\mathbf{v}$ and therefore $T\mathbf{v}$. Without positivity the lemma fails. On $\mathbb{R}^2$ the rotation $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\langle J\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$ (Example 6.1.39), but $\ker J = \{\mathbf{0}\}$; and the symmetric matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ gives 0 at $\mathbf{v} = (1, 1)$, which is not in its kernel.

8.3.3 The positive square root

Every nonnegative real number has exactly one nonnegative square root, although a positive number has two real square roots. Operators have many more square roots. Every reflection squares to the identity (Exercise 8.2.39), so when $\dim V \geq 2$ the identity has infinitely many self-adjoint square roots. Asking for a *positive* square root restores uniqueness.

Theorem 8.3.9 (The positive square root). Let $T \in \mathrm{L}(V)$ be positive, with spectral decomposition $T = \lambda_1 P_1 + \cdots + \lambda_k P_k$ (Theorem 8.2.18). Then

$$R = \sqrt{\lambda_1} P_1 + \cdots + \sqrt{\lambda_k} P_k$$

is the only positive operator $R \in \mathrm{L}(V)$ with $R^2 = T$. Moreover $R = p(T)$ for some polynomial $p \in \mathbb{R}[x]$, so R commutes with every operator that commutes with T.

Notation 8.3.10. For a positive T, the operator of Theorem 8.3.9 is the **positive square root** of T, written $\sqrt{T}$. In the notation of Definition 8.2.23, $\sqrt{T} = f(T)$ for $f(\lambda) = \sqrt{\lambda}$. For a positive matrix A, $\sqrt{A}$ is the positive matrix whose square is A; it is the matrix of $\sqrt{T_A}$.

A positive operator is self-adjoint, so it satisfies the spectral hypothesis of Section 8.2.4 (over $\mathbb{C}$ because self-adjoint operators are normal), and its eigenvalues λ_j are nonnegative by Theorem 8.3.5. So the formula for R makes sense.

Proof, part 1 of 2: existence. ROUTINE

> *Uses.* Definition 8.2.23 (Functions of an operator), Proposition 8.2.24 (Properties of $f(T)$), Theorem 8.2.18 (Spectral decomposition), Theorem 8.3.5 (Characterizations of positive operators)
> *Techniques.* Apply a function to the eigenvalues (Technique 8.2.25)
> *Hint 1.* This is Example 8.2.26. Apply Proposition 8.2.24 to $f(\lambda) = \sqrt{\lambda}$ on the eigenvalues.
> *Hint 2.* Which parts of Proposition 8.2.24 give self-adjointness, the eigenvalues of $f(T)$, the identity $f(T)^2 = T$, and the polynomial?
> *Sketch.* $R = f(T)$ for $f = \sqrt{\cdot}$ on $\{\lambda_1, \ldots, \lambda_k\}$. It is self-adjoint because f is real, its eigenvalues $\sqrt{\lambda_j}$ are nonnegative, so it is positive, and $R^2 = (f^2)(T) = T$ because $f^2(\lambda_j) = \lambda_j$. It equals $p(T)$ for the real interpolating polynomial of Proposition 8.2.24(b).

Let $f(\lambda) = \sqrt{\lambda}$ on $\{\lambda_1, \ldots, \lambda_k\}$, so that $R = f(T)$ (Definition 8.2.23). By Proposition 8.2.24(d), R is self-adjoint, since every value $f(\lambda_j)$ is real, and by Proposition 8.2.24(e), its eigenvalues $\sqrt{\lambda_j}$ are nonnegative. So R is positive by Theorem 8.3.5. By Proposition 8.2.24(c),

$$R^2 = (f^2)(T) = \lambda_1 P_1 + \cdots + \lambda_k P_k = T,$$

because $f^2(\lambda_j) = \lambda_j$ and $T = \sum_j \lambda_j P_j$ (Theorem 8.2.18(c)). Finally, by Proposition 8.2.24(b), $R = p(T)$ for $p = \sum_j \sqrt{\lambda_j} L_j$, where the Lagrange polynomials L_j have real coefficients because the λ_j are real; and R commutes with every operator that commutes with T.

> *Intuition (Three ways to see uniqueness).* Let R be any positive operator with $R^2 = T$. We must show that R acts on each eigenspace E_λ of T as $\sqrt{\lambda}$. The three proofs below look at different objects. The first looks at the eigenspaces of T: R commutes with T, so it preserves them, and on each one it is a positive operator whose square is a scalar. The second looks at the eigenspaces of R: squaring the spectral decomposition of R gives a spectral decomposition of T, and such decompositions are unique. The third compares two square roots R and R' directly and shows that their difference has no nonzero eigenvalue. In the first two proofs the nonnegativity of the eigenvalues of R is used exactly once, to rule out $-\sqrt{\lambda}$; in the third, positivity enters through the signs of $\langle R\mathbf{v}, \mathbf{v}\rangle$ and $\langle R'\mathbf{v}, \mathbf{v}\rangle$.

Proof (via the eigenspaces of T), part 2 of 2: uniqueness. STANDARD

Uses. Proposition 8.2.20 (Spectral projections are polynomials in T), Theorem 8.3.5 (Characterizations of positive operators), Theorem 8.2.18 (Spectral decomposition), Definition 6.6.26 (Self-adjoint operator)
Techniques. Restrict to a subspace (Technique 4.2.5), Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. R commutes with $R^2 = T$. What does Proposition 8.2.20(a) say about R and the eigenspaces of T?
Hint 2. Let R_λ be the restriction of R to E_λ. Check that R_λ is a positive operator on E_λ and that $R_\lambda^2 = \lambda \mathrm{id}_{E_\lambda}$.
Hint 3. Apply Theorem 8.3.5 to R_λ: its eigenvalues μ satisfy $\mu^2 = \lambda$ and $\mu \geq 0$.
Sketch. $RT = R^3 = TR$, so R maps each E_λ into itself. The restriction R_λ is self-adjoint with $\langle R_\lambda \mathbf{v}, \mathbf{v}\rangle \geq 0$, so E_λ has an orthonormal basis of eigenvectors of R_λ with eigenvalues $\mu \geq 0$, and $\mu^2 \mathbf{u} = R^2\mathbf{u} = T\mathbf{u} = \lambda\mathbf{u}$ gives $\mu = \sqrt{\lambda}$. So $R = \sqrt{\lambda}\,\mathrm{id}$ on every E_λ, and $V = \oplus E_\lambda$ gives $R = \sum \sqrt{\lambda_j} P_j$.

Let $R \in \mathrm{L}(V)$ be positive with $R^2 = T$. Then $RT = R^3 = TR$, so by Proposition 8.2.20(a), R maps each eigenspace E_{λ_j} of T into itself. Fix j, write $\lambda = \lambda_j$, and let $R_\lambda \in \mathrm{L}(E_\lambda)$ be the restriction of R to E_λ, which is an inner product space with the inner product of V. For $\mathbf{v}, \mathbf{w} \in E_\lambda$,

$$\langle R_\lambda \mathbf{v}, \mathbf{w}\rangle = \langle R\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, R\mathbf{w}\rangle = \langle \mathbf{v}, R_\lambda \mathbf{w}\rangle \quad \text{and} \quad \langle R_\lambda \mathbf{v}, \mathbf{v}\rangle = \langle R\mathbf{v}, \mathbf{v}\rangle \geq 0,$$

so R_λ is positive. By Theorem 8.3.5(III), applied on E_λ, there is an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ of E_λ with $R\mathbf{u}_i = \mu_i \mathbf{u}_i$ and $\mu_i \geq 0$. Then

$$\lambda \mathbf{u}_i = T\mathbf{u}_i = R^2 \mathbf{u}_i = \mu_i^2 \mathbf{u}_i,$$

so $\mu_i^2 = \lambda$, and since $\mu_i \geq 0$, $\mu_i = \sqrt{\lambda}$. Thus R and $\sqrt{\lambda}\,\mathrm{id}$ agree on a basis of E_λ, so $R\mathbf{v} = \sqrt{\lambda_j}\mathbf{v}$ for every $\mathbf{v} \in E_{\lambda_j}$, for every j. Finally, every $\mathbf{v} \in V$ is $\mathbf{v} = \sum_j P_j \mathbf{v}$ with $P_j \mathbf{v} \in E_{\lambda_j}$ (Theorem 8.2.18), so

$$R\mathbf{v} = \sum_j R P_j \mathbf{v} = \sum_j \sqrt{\lambda_j} P_j \mathbf{v},$$

that is, $R = \sum_j \sqrt{\lambda_j} P_j$. □

The step $\mu_i^2 = \lambda \Rightarrow \mu_i = \sqrt{\lambda}$ is where positivity of R entered. A self-adjoint R with $R^2 = T$ also preserves each E_λ, but on E_λ it only has to satisfy $R_\lambda^2 = \lambda\,\mathrm{id}$, and when $\lambda > 0$ and $\dim E_\lambda \geq 2$ there are infinitely many such R_λ (Exercise 8.3.38).

Proof (via the spectral decomposition of R), part 2 of 2: uniqueness. STANDARD

Uses. Theorem 8.2.18 (Spectral decomposition), Proposition 8.2.19 (Uniqueness of the spectral decomposition), Theorem 8.3.5 (Characterizations of positive operators)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Write $R = \mu_1 Q_1 + \cdots + \mu_m Q_m$, its own spectral decomposition, and square it.
Hint 2. $R^2 = \sum_i \mu_i^2 Q_i$. Why are the numbers μ_i^2 distinct? Now use Proposition 8.2.19.
Sketch. $R = \sum_i \mu_i Q_i$ with distinct $\mu_i \geq 0$, so $T = R^2 = \sum_i \mu_i^2 Q_i$, and the μ_i^2 are distinct because squaring is injective on $[0, \infty)$. By Proposition 8.2.19, the μ_i^2 are the eigenvalues of T and $Q_i = \mathrm{Pr}_{E_{\mu_i^2}}$. So $R = \sum \sqrt{\lambda_j} P_j$.

Let $R \in \mathrm{L}(V)$ be positive with $R^2 = T$. By Theorem 8.2.18, $R = \mu_1 Q_1 + \cdots + \mu_m Q_m$, where $\mu_1, \ldots, \mu_m$ are the distinct eigenvalues of R, all nonnegative by Theorem 8.3.5, and Q_i is the orthogonal projection onto the eigenspace of R for μ_i. The Q_i are nonzero, $Q_i Q_l = 0$ for $i \neq l$, $Q_i^2 = Q_i$, and $\sum_i Q_i = \mathrm{id}_V$. Expanding,

$$T = R^2 = \sum_{i,l} \mu_i \mu_l Q_i Q_l = \mu_1^2 Q_1 + \cdots + \mu_m^2 Q_m.$$

The numbers μ_i^2 are distinct, because $\mu_i \geq 0$ and $t \mapsto t^2$ is injective on $[0, \infty)$. So Proposition 8.2.19 applies to this expression of T: the eigenvalues of T are exactly $\mu_1^2, \ldots, \mu_m^2$, and Q_i is the orthogonal projection onto $E_{\mu_i^2}$, the eigenspace of T. Relabelling so that $\lambda_i = \mu_i^2$ (and $k = m$), we get $\mu_i = \sqrt{\lambda_i}$ and $Q_i = P_i$, so $R = \sum_i \sqrt{\lambda_i} P_i$. □

Proof (by comparing two square roots), part 2 of 2: uniqueness. STANDARD

Uses. Lemma 8.3.8 (Where the form vanishes), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Theorem 8.2.12 (Real spectral theorem), Corollary 8.2.7 (Hermitian matrices), Proposition 6.6.18 (Properties of the adjoint)

Techniques. A sum of nonnegative terms (Technique 6.1.9), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Let R and R' be positive square roots of T and $D = R - R'$. D is self-adjoint, so it suffices to show that every eigenvalue of D is 0.
Hint 2. Check the identity $R^2 - R'^2 = RD + DR'$. Apply it to an eigenvector $\mathbf{v}$ of D, $D\mathbf{v} = \mu\mathbf{v}$, and take the inner product with $\mathbf{v}$.
Hint 3. You get $0 = \mu(\langle R\mathbf{v}, \mathbf{v}\rangle + \langle R'\mathbf{v}, \mathbf{v}\rangle)$. If $\mu \neq 0$, use Lemma 8.3.8.
Sketch. Let $D = R - R'$, self-adjoint, with $D\mathbf{v} = \mu\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, $\mu \in \mathbb{R}$. Since $0 = R^2 - R'^2 = RD + DR'$, we get $0 = \langle (RD + DR')\mathbf{v}, \mathbf{v}\rangle = \mu(\langle R\mathbf{v}, \mathbf{v}\rangle + \langle R'\mathbf{v}, \mathbf{v}\rangle)$. If $\mu \neq 0$, both nonnegative terms vanish, so $R\mathbf{v} = R'\mathbf{v} = \mathbf{0}$ by Lemma 8.3.8, and $\mu\mathbf{v} = D\mathbf{v} = \mathbf{0}$, a contradiction. So all eigenvalues of D are 0, and a self-adjoint operator with only the eigenvalue 0 is 0.

Let $R, R' \in \mathrm{L}(V)$ be positive with $R^2 = R'^2 = T$, and let $D = R - R'$. Then D is self-adjoint (Proposition 6.6.18), and

$$RD + DR' = R^2 - RR' + RR' - R'^2 = R^2 - R'^2 = 0.$$

Let μ be an eigenvalue of D with eigenvector $\mathbf{v} \neq \mathbf{0}$; μ is real by Proposition 8.1.1. Using $D^* = D$,

$$0 = \langle (RD + DR')\mathbf{v}, \mathbf{v}\rangle = \langle R(\mu\mathbf{v}), \mathbf{v}\rangle + \langle R'\mathbf{v}, D\mathbf{v}\rangle = \mu\langle R\mathbf{v}, \mathbf{v}\rangle + \overline{\mu}\langle R'\mathbf{v}, \mathbf{v}\rangle,$$

and since μ is real, $0 = \mu(\langle R\mathbf{v}, \mathbf{v}\rangle + \langle R'\mathbf{v}, \mathbf{v}\rangle)$. Suppose $\mu \neq 0$. Then the two nonnegative numbers $\langle R\mathbf{v}, \mathbf{v}\rangle$ and $\langle R'\mathbf{v}, \mathbf{v}\rangle$ add up to 0, so both are 0, and by Lemma 8.3.8, $R\mathbf{v} = R'\mathbf{v} = \mathbf{0}$. Then $\mu\mathbf{v} = D\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$, so $\mu = 0$, a contradiction. Hence every eigenvalue of D is 0. By Theorem 8.2.12 or Corollary 8.2.7, D has an orthonormal eigenbasis, in which its matrix is diagonal with the eigenvalues on the diagonal, that is, zero. So $D = 0$ and $R = R'$. Since $\sqrt{T}$ is a positive square root of T, every positive square root of T equals $\sqrt{T}$. □

The third proof never mentions the eigenspaces of T and uses only Lemma 8.3.8 and the identity $R^2 - R'^2 = RD + DR'$. That makes it the one that adapts to inequalities: the same computation shows that $A \mapsto \sqrt{A}$ preserves the order between positive operators (Exercise 8.3.40).

Example 8.3.11 (A square root as a polynomial). For $A = I_3 + J_3 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ of Example 8.2.21, the eigenvalues are 4 and 1, with $P_4 = (A - I_3)/3$ and $P_1 = (A - 4I_3)/(-3)$. So

$$\sqrt{A} = 2P_4 + 1P_1 = \frac{2(A - I_3) - (A - 4I_3)}{3} = \frac{A + 2I_3}{3} = \begin{pmatrix} 4/3 & 1/3 & 1/3 \\ 1/3 & 4/3 & 1/3 \\ 1/3 & 1/3 & 4/3 \end{pmatrix}.$$

The polynomial is $p(x) = (x + 2)/3$, the line through $(1, 1)$ and $(4, 2)$: it agrees with $\sqrt{x}$ at the two eigenvalues, which is all that matters. Check: $(I_3 + J_3/3)^2 = I_3 + 2J_3/3 + J_3^2/9 = I_3 + J_3$, since $J_3^2 = 3J_3$.

Exercise 8.3.12. ROUTINE Let $A = \begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.

(a) Compute $\sqrt{A}$, and check that $B^2 = A$.

(b) Compute $\sqrt{A^2}$ and $\sqrt{B^2}$. Which of $\sqrt{A^2} = A$ and $\sqrt{B^2} = B$ holds, and why?

Solution on page 1651

Remark 8.3.13 (The absolute value of an operator). Let W be a finite-dimensional inner product space and $S \in \mathrm{L}(V, W)$. Then S^*S is positive (Example 8.3.2(b)), so it has a positive square root

$$|S| = \sqrt{S^*S} \in \mathrm{L}(V).$$

It satisfies $\||S|\mathbf{v}\|^2 = \langle |S|^2\mathbf{v}, \mathbf{v}\rangle = \langle S^*S\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2$, so $|S|$ moves every vector exactly as far as S does. For a complex number, $|z| = \sqrt{\overline{z}z}$ is the size of z with its direction forgotten, and $|S|$ plays the same role. In Section 9.1 the eigenvalues of $|S|$ become the singular values of S.

8.3.4 Positive definite operators

Definition 8.3.14 (Positive definite). An operator $T \in \mathrm{L}(V)$ is **positive definite** if T is self-adjoint and

$$\langle T\mathbf{v}, \mathbf{v}\rangle > 0 \quad \text{for every nonzero } \mathbf{v} \in V.$$

A matrix $A \in \mathbb{F}^{n\times n}$ is **positive definite** if $A^* = A$ and $\mathbf{x}^* A\mathbf{x} > 0$ for every nonzero $\mathbf{x} \in \mathbb{F}^n$.

The lecture notes define a positive definite matrix as a matrix whose eigenvalues are all positive. For Hermitian matrices this is the same notion, by (II) of the next proposition. Every positive definite operator is positive, and by Lemma 8.3.8 a positive operator is positive definite exactly when its kernel is $\{\mathbf{0}\}$. So positive definite operators are to positive operators what positive numbers are to nonnegative ones.

Proposition 8.3.15 (Characterizations of positive definite operators). Let $T \in \mathrm{L}(V)$. The following are equivalent:

(I) T is positive definite;

(II) T is self-adjoint and every eigenvalue of T is positive;

(III) T is positive and invertible;

(IV) $T = S^*S$ for some injective $S \in \mathrm{L}(V, W)$, where W is a finite-dimensional inner product space;

(V) for some orthonormal basis $\mathcal{O}$ of V, $[T]_{\mathcal{O}}$ is the Gram matrix of a linearly independent list.

In matrix form: $A \in \mathbb{F}^{n\times n}$ is positive definite if and only if $A = B^*B$ for some $B \in \mathbb{F}^{m\times n}$ with $\operatorname{rank} B = n$, if and only if A is the Gram matrix of a linearly independent list, if and only if A is positive and $\det A \neq 0$.

Proof. STANDARD

Uses. Theorem 8.3.5 (Characterizations of positive operators), Theorem 8.3.9 (The positive square root), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 6.3.30 (Gram matrices detect independence), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 5.4.7 (Invertibility and the determinant), Corollary 4.2.18 (Injective if and only if surjective)
Techniques. Proving TFAE (Technique 2.7.1), Gram matrices (Technique 6.3.31)
Hint 1. Follow the proof of Theorem 8.3.5 with strict inequalities: (I) ⇒ (II) ⇒ (III) ⇒ (IV) ⇒ (V) ⇒ (I).
Hint 2. For (III) ⇒ (IV), take $S = \sqrt{T}$; why is it injective? For (IV) ⇒ (V), injective maps send independent lists to independent lists.
Hint 3. For (V) ⇒ (I), the computation $\langle T\mathbf{v}, \mathbf{v}\rangle = \left\|\sum_j c_j \mathbf{w}_j\right\|^2$ from Theorem 8.3.5 is still available. When is $\sum_j c_j \mathbf{w}_j = \mathbf{0}$?
Sketch. (I) ⇒ (II): $\lambda\|\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle > 0$. (II) ⇒ (III): positive by Theorem 8.3.5, and 0 is not an eigenvalue. (III) ⇒ (IV): $S = \sqrt{T}$; $S\mathbf{v} = \mathbf{0}$ gives $T\mathbf{v} = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$. (IV) ⇒ (V): $[T]_{\mathcal{O}}$ is the Gram matrix of $S\mathbf{e}_1, \ldots, S\mathbf{e}_n$, which is independent as S is injective. (V) ⇒ (I): T is positive, and $\langle T\mathbf{v}, \mathbf{v}\rangle = \left\|\sum c_j \mathbf{w}_j\right\|^2 > 0$ for $\mathbf{c} \neq \mathbf{0}$ by independence.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (IV) ⇒ (V) ⇒ (I).

(I) ⇒ (II). If $T\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\lambda\|\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle > 0$, so $\lambda > 0$.

(II) ⇒ (III). T is positive by Theorem 8.3.5. Since 0 is not an eigenvalue of T, T is invertible by Theorem 7.2.9.

(III) ⇒ (IV). Let $W = V$ and $S = \sqrt{T}$ (Theorem 8.3.9). Then $S^*S = S^2 = T$. If $S\mathbf{v} = \mathbf{0}$, then $T\mathbf{v} = S(S\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$ since T is invertible. Hence S is injective.

(IV) ⇒ (V). Let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis. By the proof of Theorem 8.3.5 ((VI) ⇒ (VII)), $[T]_{\mathcal{O}}$ is the Gram matrix of $S\mathbf{e}_1, \ldots, S\mathbf{e}_n$. This list is independent: if $\sum_i c_i S\mathbf{e}_i = \mathbf{0}$, then $S\left(\sum_i c_i \mathbf{e}_i\right) = \mathbf{0}$, so $\sum_i c_i \mathbf{e}_i = \mathbf{0}$ because S is injective, and every $c_i = 0$.

(V) ⇒ (I). T is positive by Theorem 8.3.5 ((VII) ⇒ (I)), in particular self-adjoint. Let $[T]_{\mathcal{O}}$ be the Gram matrix of the independent list $\mathbf{w}_1, \ldots, \mathbf{w}_n$, and let $\mathbf{v} \neq \mathbf{0}$ have coordinates $\mathbf{c} = [\mathbf{v}]_{\mathcal{O}} \neq \mathbf{0}$. As in the proof of Theorem 8.3.5, $\langle T\mathbf{v}, \mathbf{v}\rangle = \|c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n\|^2$, and the vector inside the norm is nonzero because the list is independent. So $\langle T\mathbf{v}, \mathbf{v}\rangle > 0$.

For the matrix form, apply this to T_A as in Theorem 8.3.5: T_B is injective if and only if $\operatorname{rank} B = n$ (Theorem 3.5.12), and the Gram matrix statement is (V) with the standard basis. Finally, if A is positive then A is invertible exactly when $\det A \neq 0$ (Theorem 5.4.7), which is (III). □

Corollary 8.3.16 (Inner products on $\mathbb{F}^n$). The inner products on $\mathbb{F}^n$ are exactly the functions

$$\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{y}^* A \mathbf{x}$$

with $A \in \mathbb{F}^{n\times n}$ positive definite. For such an inner product, A is the Gram matrix of the standard basis: $a_{ij} = \langle \mathbf{e}_j, \mathbf{e}_i \rangle_A$.

Proof. ROUTINE

Uses. Proposition 6.1.22 (Inner products on $\mathbb{F}^n$), Definition 8.3.14 (Positive definite)
Hint. Compare Definition 8.3.14 with the condition in Proposition 6.1.22(b).
Sketch. Proposition 6.1.22(b) says that $\mathbf{y}^* M \mathbf{x}$ is an inner product if and only if $M^* = M$ and $\mathbf{x}^* M \mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$, that is, M is positive definite. Part (a) says every inner product has this form with $m_{ij} = \langle \mathbf{e}_j, \mathbf{e}_i \rangle$.

By Proposition 6.1.22(b), the formula $\langle \mathbf{x}, \mathbf{y} \rangle_M = \mathbf{y}^* M \mathbf{x}$ defines an inner product if and only if $M^* = M$ and $\mathbf{x}^* M \mathbf{x} > 0$ for every nonzero $\mathbf{x}$, that is, if and only if M is positive definite. By Proposition 6.1.22(a), every inner product on $\mathbb{F}^n$ is of this form, with $m_{ij} = \langle \mathbf{e}_j, \mathbf{e}_i \rangle$. □

So positive definite matrices and inner products on $\mathbb{F}^n$ are the same objects, and Example 6.1.23 was a positive definite matrix in disguise. Questions about one inner product compared with another become questions about a positive definite matrix compared with I_n.

Example 8.3.17 (Positive, and positive definite). The matrix A of Example 8.3.6 is positive definite: its eigenvalues are positive, and equivalently the difference matrix B has independent columns (it has rank 3). The matrix $C = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ of Exercise 8.3.7 is positive but not positive definite: $C = B^* B$ only for matrices B with dependent columns, such as $B = \begin{pmatrix} 1 & i \end{pmatrix}$, whose columns $1, i$ are dependent in $\mathbb{C}^1$. Accordingly $\det C = 0$ and $C(-i, 1) = \mathbf{0}$, where $\mathbf{x} = (-i, 1)$ has $\mathbf{x}^* C \mathbf{x} = |-i + i|^2 = 0$.

If T is positive definite, so are $\sqrt{T}$ and T^{-1}: their eigenvalues $\sqrt{\lambda_j}$ and $1/\lambda_j$ are positive (Proposition 8.2.24(e), with $T^{-1} = \sum_j \lambda_j^{-1} P_j$ as in Section 8.2.5). Applying Proposition 8.2.24(c) to $f(\lambda) = 1/\sqrt{\lambda}$ shows that $f(T)\sqrt{T} = \mathrm{id}_V$ and $f(T)^2 = T^{-1}$; since $f(T)$ is positive, uniqueness of square roots gives $\left(\sqrt{T}\right)^{-1} = f(T) = \sqrt{T^{-1}}$. We write $T^{-1/2}$ for this operator.

Exercise 8.3.18. ROUTINE For which real numbers t is $A_t = \begin{pmatrix} 1 & t \\ t & 4 \end{pmatrix}$ positive? For which is it positive definite?

Solution on page 1651

Quadratic forms

For a real symmetric matrix A, the function $q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}} A \mathbf{x}$ on $\mathbb{R}^n$ is the **quadratic form** of A (Section 8.2.6). By definition, A is positive definite exactly when $q(\mathbf{x}) > 0$ for every nonzero $\mathbf{x}$, and this is how positive definite matrices usually enter: through quadratic forms. By Corollary 8.2.28, in the coordinates $\mathbf{y} = Q^{\mathsf{T}} \mathbf{x}$ along an orthonormal eigenbasis $\mathbf{q}_1, \ldots, \mathbf{q}_n$,

$$q(\mathbf{x}) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

If A is positive definite, the level set $q(\mathbf{x}) = 1$ is therefore an ellipsoid with semi-axes $1/\sqrt{\lambda_i}$ along the $\mathbf{q}_i$ (in $\mathbb{R}^2$, the ellipses of Remark 8.2.30). If $A \neq 0$ is positive but some $\lambda_i = 0$, the level set is unbounded in the direction of $\mathbf{q}_i$, and if some $\lambda_i < 0$, then q takes negative values. For example, the form $q(\mathbf{x}) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3$ of Example 8.3.6 is positive definite, and $q(\mathbf{x}) = 1$ is an ellipsoid with semi-axes $1/\sqrt{2+\sqrt{2}}$, $1/\sqrt{2}$, and $1/\sqrt{2-\sqrt{2}}$.

This is the linear algebra behind the second-derivative test. At a critical point $\mathbf{x}_0$ of a smooth $f:\mathbb{R}^n\to\mathbb{R}$, $f(\mathbf{x}_0+\mathbf{h})=f(\mathbf{x}_0)+\frac{1}{2}\mathbf{h}^{\mathsf{T}}H\mathbf{h}+(\cdots)$, where H is the symmetric matrix of second partial derivatives and the omitted terms are small compared with $\|\mathbf{h}\|^2$; a positive definite H makes $\mathbf{x}_0$ a strict local minimum. Section 12.2 studies quadratic forms in general, and Section 12.3 gives further tests for positive definiteness, among them Sylvester's criterion on the leading minors, together with the min–max description of eigenvalues.

8.3.5 Products of positive operators

The product of two nonnegative numbers is nonnegative. The product of two positive operators need not even be self-adjoint: for $A=\begin{pmatrix}2&1\\1&1\end{pmatrix}$ and the projection $B=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, the product $AB=\begin{pmatrix}2&0\\1&0\end{pmatrix}$ is not symmetric. The obstruction is that A and B need not commute. What survives is a statement about eigenvalues, and the proof introduces a move that we will use repeatedly.

Technique 8.3.19 (Sandwich between square roots). Two positive operators A and B do not commute in general, so AB need not be positive. Split $A=\sqrt{A}\sqrt{A}$ and move one factor to the other side: the operator $\sqrt{A}B\sqrt{A}$ is always positive, since $\langle\sqrt{A}B\sqrt{A}\mathbf{v},\mathbf{v}\rangle=\langle B\sqrt{A}\mathbf{v},\sqrt{A}\mathbf{v}\rangle\geq 0$, and it is close to AB. It has the same characteristic polynomial (by Proposition 7.3.22, applied to $\sqrt{A}$ and $\sqrt{A}B$), and if A is invertible it is similar to AB. In the same way, when B is positive definite, the congruence $X\mapsto B^{-1/2}XB^{-1/2}$ turns the pair (A,B) into (C,id) with $C=B^{-1/2}AB^{-1/2}$ positive, and a question about two operators becomes a question about the eigenvalues of one.

Reach for it when a statement involves two positive operators (a product, a comparison $A\preceq B$, $\det(A+B)$, $\operatorname{tr}(AB)$) and the case of a single operator is easy by Technique 8.2.25. The canonical example is Proposition 8.3.20.

Used in. Proposition 8.3.20, Exercise 8.3.34, Exercise 8.3.39, Exercise 8.3.42.

Proposition 8.3.20 (Products of positive operators). Let $A,B\in\mathrm{L}(V)$ be positive.

(a) AB is positive if and only if $AB=BA$.

(b) Every root of p_{AB} in $\mathbb{C}$ is real and nonnegative. In particular every eigenvalue of AB is nonnegative.

(c) If A is invertible, then AB is diagonalizable.

Proof. STANDARD

Uses. Theorem 8.3.9 (The positive square root), Proposition 6.6.18 (Properties of the adjoint), Proposition 7.3.22 (AB and BA), Definition 7.3.12 (Characteristic polynomial of an operator), Theorem 2.4.2 (The matrix of a composite), Theorem 8.3.5 (Characterizations of positive operators), Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Corollary 7.3.13 (Eigenvalues of an operator)
Techniques. Sandwich between square roots (Technique 8.3.19), Compare AB and BA (Technique 7.3.21)
Hint 1. For (a), a positive operator is self-adjoint, and $(AB)^*=BA$. For the converse, B commutes with $\sqrt{A}$ (Theorem 8.3.9).
Hint 2. For (b), write $AB=XY$ with $X=\sqrt{A}$ and $Y=\sqrt{A}B$. What is YX? Use Proposition 7.3.22 on the matrices of X and Y.
Hint 3. For (c), $AB=\sqrt{A}\big(\sqrt{A}B\sqrt{A}\big)\big(\sqrt{A}\big)^{-1}$. Push an eigenbasis of the positive operator $\sqrt{A}B\sqrt{A}$ forward by $\sqrt{A}$.
Sketch. (a) If AB is positive it is self-adjoint, so $AB=(AB)^*=BA$. If $AB=BA$, then B commutes with $\sqrt{A}$, so $AB=\sqrt{A}B\sqrt{A}$, which is positive. (b) AB and $\sqrt{A}B\sqrt{A}$ are XY and YX for $X=\sqrt{A}$, $Y=\sqrt{A}B$, so they have the same characteristic polynomial, which for the positive operator $\sqrt{A}B\sqrt{A}$ is $\prod(x-d_i)$ with $d_i\geq 0$. (c) If $\sqrt{A}B\sqrt{A}\mathbf{u}_i=d_i\mathbf{u}_i$ for an eigenbasis $\mathbf{u}_i$, then $AB\big(\sqrt{A}\mathbf{u}_i\big)=d_i\sqrt{A}\mathbf{u}_i$, and $\sqrt{A}\mathbf{u}_1,\dots,\sqrt{A}\mathbf{u}_n$ is a basis because $\sqrt{A}$ is invertible.

First, $C=\sqrt{A}B\sqrt{A}$ is positive: it is self-adjoint by Proposition 6.6.18, and $\langle C\mathbf{v},\mathbf{v}\rangle=\langle B\sqrt{A}\mathbf{v},\sqrt{A}\mathbf{v}\rangle\geq 0$.

(a) If AB is positive, it is self-adjoint, so $AB=(AB)^*=B^*A^*=BA$. Conversely, if $AB=BA$, then B commutes with $\sqrt{A}$ by Theorem 8.3.9, and $AB=\sqrt{A}\sqrt{A}B=\sqrt{A}B\sqrt{A}=C$, which is positive.

(b) Fix an orthonormal basis $\mathcal{O}$ and let $X = \left[\sqrt{A}\right]_{\mathcal{O}}$ and $Y = \left[\sqrt{AB}\right]_{\mathcal{O}}$. By Theorem 2.4.2, $XY = \left[\sqrt{A}\sqrt{AB}\right]_{\mathcal{O}} = [AB]_{\mathcal{O}}$ and $YX = \left[\sqrt{AB}\sqrt{A}\right]_{\mathcal{O}} = [C]_{\mathcal{O}}$. So by Definition 7.3.12 and Proposition 7.3.22, $p_{AB} = p_{XY} = p_{YX} = p_C$. Since C is positive, Theorem 8.3.5(III) gives an orthonormal basis in which C is $\operatorname{diag}(d_1, \ldots, d_n)$ with every $d_i \geq 0$, so $p_C = (x - d_1)\cdots(x - d_n)$. Hence every root of p_{AB} is one of the d_i, and every eigenvalue of AB, being a root of p_{AB} (Corollary 7.3.13), is nonnegative.

(c) If A is invertible, then 0 is not an eigenvalue of A, so the eigenvalues $\sqrt{\lambda_j}$ of $\sqrt{A}$ are positive and $\sqrt{A}$ is invertible (Theorem 7.2.9). Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be an eigenbasis of the positive operator C, with $C\mathbf{u}_i = d_i\mathbf{u}_i$ (Theorem 8.3.5(III)). Then

$$AB\left(\sqrt{A}\mathbf{u}_i\right) = \sqrt{A}\left(\sqrt{A}B\sqrt{A}\right)\mathbf{u}_i = \sqrt{A}C\mathbf{u}_i = d_i\sqrt{A}\mathbf{u}_i,$$

and $\sqrt{A}\mathbf{u}_1, \ldots, \sqrt{A}\mathbf{u}_n$ is a basis of V, being the image of a basis under an invertible map. So V has a basis of eigenvectors of AB, and AB is diagonalizable (Theorem 7.4.2). □

For the example above, $AB = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}$ has eigenvalues 2 and 0, and it is diagonalizable, as (c) predicts since $\det A = 1$. Its eigenvectors $(2, 1)$ and $(0, 1)$ are not orthogonal, so AB is diagonalizable but not unitarily diagonalizable. When A is invertible, AB is similar to a positive operator, $AB = \sqrt{A}C\left(\sqrt{A}\right)^{-1}$, but similarity does not preserve orthogonality.

8.3.6 The Löwner order

BEYOND 21-242

The Löwner order is not covered in 21-242.

For real numbers, $x \leq y$ means that $y - x \geq 0$. Positivity gives the same kind of comparison between self-adjoint operators.

Definition 8.3.21 (Löwner order). For self-adjoint $S, T \in \mathrm{L}(V)$, write $S \preceq T$ if $T - S$ is positive, that is, if $\langle S\mathbf{v}, \mathbf{v}\rangle \leq \langle T\mathbf{v}, \mathbf{v}\rangle$ for every $\mathbf{v} \in V$. In particular $0 \preceq T$ means that T is positive.

The two descriptions agree because $T - S$ is automatically self-adjoint. For matrices, $A \preceq B$ is defined in the same way for Hermitian A and B.

Proposition 8.3.22 (Properties of the Löwner order). Let $S, T, U \in \mathrm{L}(V)$ be self-adjoint.

(a) $\preceq$ is a partial order: $T \preceq T$; if $S \preceq T$ and $T \preceq S$, then $S = T$; and if $S \preceq T$ and $T \preceq U$, then $S \preceq U$.

(b) If $S \preceq T$, then $S + U \preceq T + U$, and $cS \preceq cT$ for every real $c \geq 0$.

(c) If $S \preceq T$, V' is a finite-dimensional inner product space, and $X \in \mathrm{L}(V', V)$, then $X^*SX \preceq X^*TX$.

(d) For $c \in \mathbb{R}$, $T \preceq c\mathrm{id}_V$ if and only if every eigenvalue of T is at most c, and $c\mathrm{id}_V \preceq T$ if and only if every eigenvalue of T is at least c.

Proof. ROUTINE

Uses. Definition 8.3.21 (Löwner order), Lemma 8.3.8 (Where the form vanishes), Proposition 6.6.18 (Properties of the adjoint), Theorem 8.3.5 (Characterizations of positive operators)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Each part is a statement about one positive operator, such as $T - S$. For antisymmetry, $P = T - S$ and $-P$ are both positive; use Lemma 8.3.8.
Hint 2. For (c), move X across the inner product. For (d), compare the eigenvalues of $c\mathrm{id}_V - T$ with those of T.
Sketch. (a) $T - T = 0$. If $P = T - S$ and $-P$ are positive, then $\langle P\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$, so $P = 0$ by Lemma 8.3.8. $U - S = (U - T) + (T - S)$ is a sum of positive operators. (b) The differences are $T - S$ and $c(T - S)$. (c) $\langle X^*(T - S)X\mathbf{u}, \mathbf{u}\rangle = \langle (T - S)X\mathbf{u}, X\mathbf{u}\rangle \geq 0$. (d) $c\mathrm{id}_V - T$ is self-adjoint with eigenvalues $c - \lambda$; apply Theorem 8.3.5(II).

(a) $T - T = 0$ is positive. If $P = T - S$ and $-P = S - T$ are both positive, then $\langle P\mathbf{v}, \mathbf{v}\rangle \geq 0$ and $-\langle P\mathbf{v}, \mathbf{v}\rangle \geq 0$, so $\langle P\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$, and $P\mathbf{v} = \mathbf{0}$ for every $\mathbf{v}$ by Lemma 8.3.8. So $S = T$. If $S \preceq T \preceq U$, then $U - S = (U - T) + (T - S)$ is self-adjoint and $\langle (U - S)\mathbf{v}, \mathbf{v}\rangle = \langle (U - T)\mathbf{v}, \mathbf{v}\rangle + \langle (T - S)\mathbf{v}, \mathbf{v}\rangle \geq 0$.

(b) $(T+U)-(S+U)=T-S$ is positive, and $cT-cS=c(T-S)$ is self-adjoint (c is real) with $\langle c(T-S)\mathbf{v},\mathbf{v}\rangle = c\langle (T-S)\mathbf{v},\mathbf{v}\rangle \geq 0$.

(c) $X^*TX - X^*SX = X^*(T-S)X$ is self-adjoint by Proposition 6.6.18, and for $\mathbf{u}\in V'$, $\langle X^*(T-S)X\mathbf{u},\mathbf{u}\rangle = \langle (T-S)X\mathbf{u}, X\mathbf{u}\rangle \geq 0$.

(d) $T\mathbf{v}=\lambda\mathbf{v}$ if and only if $(c\mathrm{id}_V - T)\mathbf{v} = (c-\lambda)\mathbf{v}$, so the eigenvalues of the self-adjoint operator $c\mathrm{id}_V - T$ are the numbers $c-\lambda$ with λ an eigenvalue of T. By Theorem 8.3.5(II), $c\mathrm{id}_V - T$ is positive if and only if $c-\lambda\geq 0$ for all of them. The second statement is the same argument for $T - c\mathrm{id}_V$. □

For example, every orthogonal projection satisfies $0 \preceq \mathrm{Pr}_U \preceq \mathrm{id}_V$, by (d), because its eigenvalues are 0 and 1. Unlike the order on $\mathbb{R}$, the Löwner order is not total: neither of $\mathrm{diag}(1,0)$ and $\mathrm{diag}(0,1)$ is $\preceq$ the other, since their difference has eigenvalues 1 and -1. Some familiar inequalities survive and some do not. Inverses reverse the order and square roots preserve it, while squares need not preserve it (Exercise 8.3.39, Exercise 8.3.40). The inverse inequality is proved with Technique 8.3.19, which compares an operator with id_V, where (d) applies; the square-root inequality uses the eigenvector trick of the third uniqueness proof in Theorem 8.3.9.

8.3.7 Cholesky factorization BEYOND 21-242

Cholesky factorization is standard in numerical linear algebra (LADR 7D), but it is not part of the course.

By Theorem 8.3.5 a positive matrix can be written as B^*B in many ways, for example with $B=\sqrt{A}$. One of these factorizations is triangular, and it can be computed by a short algorithm that needs no eigenvalues. It is the standard way to solve $A\mathbf{x}=\mathbf{b}$ when A is positive definite, and it is also the cheapest test of positive definiteness.

Theorem 8.3.23 (Cholesky factorization). Let $A\in\mathbb{F}^{n\times n}$ be positive definite. Then there is exactly one upper triangular $R\in\mathbb{F}^{n\times n}$ with positive real diagonal entries such that

$$A = R^*R.$$

Proof (via QR), part 1 of 2: existence. ROUTINE

Uses. Proposition 8.3.15 (Characterizations of positive definite operators), Theorem 6.4.38 (QR factorization), Theorem 6.3.29 (Orthonormal columns)
Hint 1. Write $A=B^*B$ with B invertible, and factor B.
Hint 2. Apply QR factorization (Theorem 6.4.38) to B. What is Q^*Q?
Sketch. $A=B^*B$ with B invertible (for instance $B=\sqrt{A}$). QR gives $B=QR$ with orthonormal columns in Q, so $A = R^*Q^*QR = R^*R$.

By Proposition 8.3.15, $A=B^*B$ for some $B\in\mathbb{F}^{n\times n}$ with $\mathrm{rank}\, B = n$ (for instance $B=\sqrt{A}$). The columns of B are independent, so Theorem 6.4.38 gives $B=QR$ with $Q\in\mathbb{F}^{n\times n}$ having orthonormal columns and R upper triangular with positive diagonal entries. Since $Q^*Q=I_n$ (Theorem 6.3.29),

$$A = B^*B = R^*Q^*QR = R^*R.$$

The first proof hides the work inside QR. The second builds R one row at a time, and it is the proof behind the algorithm below.

Proof (by induction on n), part 1 of 2: existence. STANDARD

Uses. Definition 2.4.33 (Block matrix), Definition 8.3.14 (Positive definite)
Techniques. Induction on matrix size (Technique 5.5.17), Block decomposition (Technique 2.4.35), Plug in clever test vectors (Technique 6.1.40)
Hint 1. Write $A=\begin{pmatrix} a & \mathbf{w}^* \\ \mathbf{w} & A'\end{pmatrix}$ with $a\in\mathbb{R}$. Why is $a>0$? Guess the first row of R from the $(1,1)$ and $(1,2)$ blocks of R^*R.
Hint 2. The first row must be $(\sqrt{a}, \mathbf{w}^*/\sqrt{a})$. The rest of R must then be a Cholesky factor of $A'' = A' - \mathbf{w}\mathbf{w}^*/a$.
Hint 3. To see that A'' is positive definite, evaluate $\mathbf{x}^*A\mathbf{x}$ at $\mathbf{x}=(t,\mathbf{y})$ with $t=-\mathbf{w}^*\mathbf{y}/a$.

Sketch. Write $A = \begin{pmatrix} a & \mathbf{w}^* \\ \mathbf{w} & A' \end{pmatrix}$ with $a = \mathbf{e}_1^* A \mathbf{e}_1 > 0$, and let $A'' = A' - \mathbf{w}\mathbf{w}^*/a$. For $\mathbf{y} \neq \mathbf{0}$ and $t = -\mathbf{w}^*\mathbf{y}/a$, $\mathbf{y}^* A'' \mathbf{y} = (t, \mathbf{y})^* A (t, \mathbf{y}) > 0$, so A'' is positive definite and $A'' = R''^* R''$ by induction. Then $R = \begin{pmatrix} \sqrt{a} & \mathbf{w}^*/\sqrt{a} \\ \mathbf{0} & R'' \end{pmatrix}$ works, by block multiplication.

We induct on n. If $n = 1$, then $A = (a)$ with $a = \mathbf{e}_1^* A \mathbf{e}_1 > 0$, and $R = \left(\sqrt{a}\right)$. Let $n \geq 2$, and write in block form (Definition 2.4.33)

$$A = \begin{pmatrix} a & \mathbf{w}^* \\ \mathbf{w} & A' \end{pmatrix}, \quad a \in \mathbb{R}, \quad \mathbf{w} \in \mathbb{F}^{n-1}, \quad A' \in \mathbb{F}^{(n-1)\times(n-1)},$$

where a is real and $A'^* = A'$ because $A^* = A$, and $a = \mathbf{e}_1^* A \mathbf{e}_1 > 0$. Let $A'' = A' - \mathbf{w}\mathbf{w}^*/a$, which is Hermitian. For $t \in \mathbb{F}$ and $\mathbf{y} \in \mathbb{F}^{n-1}$, block multiplication gives

$$\begin{pmatrix} t \\ \mathbf{y} \end{pmatrix}^* A \begin{pmatrix} t \\ \mathbf{y} \end{pmatrix} = a|t|^2 + \bar{t}\mathbf{w}^*\mathbf{y} + t\mathbf{y}^*\mathbf{w} + \mathbf{y}^* A' \mathbf{y}.$$

Take $\mathbf{y} \neq \mathbf{0}$ and $t = -\mathbf{w}^*\mathbf{y}/a$. Since $\mathbf{y}^*\mathbf{w} = \overline{\mathbf{w}^*\mathbf{y}}$, the first three terms are $|\mathbf{w}^*\mathbf{y}|^2/a - 2|\mathbf{w}^*\mathbf{y}|^2/a = -|\mathbf{w}^*\mathbf{y}|^2/a = -\mathbf{y}^*\mathbf{w}\mathbf{w}^*\mathbf{y}/a$, so the whole expression equals $\mathbf{y}^* A'' \mathbf{y}$. The vector $(t, \mathbf{y})$ is nonzero, so $\mathbf{y}^* A'' \mathbf{y} > 0$, and A'' is positive definite. By the induction hypothesis, $A'' = R''^* R''$ with R'' upper triangular with positive diagonal entries. Let

$$R = \begin{pmatrix} \sqrt{a} & \mathbf{w}^*/\sqrt{a} \\ \mathbf{0} & R'' \end{pmatrix}.$$

It is upper triangular with positive diagonal entries, and

$$R^* R = \begin{pmatrix} \sqrt{a} & \mathbf{0}^* \\ \mathbf{w}/\sqrt{a} & R''^* \end{pmatrix} \begin{pmatrix} \sqrt{a} & \mathbf{w}^*/\sqrt{a} \\ \mathbf{0} & R'' \end{pmatrix} = \begin{pmatrix} a & \mathbf{w}^* \\ \mathbf{w} & \mathbf{w}\mathbf{w}^*/a + R''^* R'' \end{pmatrix} = \begin{pmatrix} a & \mathbf{w}^* \\ \mathbf{w} & A' \end{pmatrix} = A.$$

Proof, part 2 of 2: uniqueness. STANDARD

Uses. Lemma 3.3.32 (Products and inverses of triangular matrices), Definition 3.3.31 (Triangular matrices), Proposition 5.3.15 (Triangular matrices), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Suppose $R^* R = S^* S$, with R, S as in the theorem. Rearrange to $SR^{-1} = (S^*)^{-1} R^*$.
Hint 2. The left side is upper triangular and the right side lower triangular, so both are diagonal. Compare their diagonal entries.
Sketch. R and S are invertible, and $SR^{-1} = (S^*)^{-1} R^* = \left(RS^{-1}\right)^*$ is both upper and lower triangular, hence diagonal, with diagonal entries s_{ii}/r_{ii} and r_{ii}/s_{ii}. These agree, so $s_{ii}^2 = r_{ii}^2$ and $s_{ii} = r_{ii}$. Thus $SR^{-1} = I_n$.

Suppose $A = R^* R = S^* S$ with R and S upper triangular with positive diagonal entries. Both are invertible, since their determinants are the products of the diagonal entries (Proposition 5.3.15, Theorem 5.4.7). An invertible upper triangular matrix R has an upper triangular inverse whose diagonal entries are $1/r_{ii}$: write $R = DU$ with $D = \operatorname{diag}(r_{11}, \ldots, r_{nn})$ and $U = D^{-1}R$ unit upper triangular, so that $R^{-1} = U^{-1}D^{-1}$, and apply Lemma 3.3.32. Multiplying $S^* S = R^* R$ on the left by $(S^*)^{-1}$ and on the right by R^{-1} gives

$$SR^{-1} = (S^*)^{-1} R^* = \left(RS^{-1}\right)^*.$$

By Lemma 3.3.32, SR^{-1} and RS^{-1} are upper triangular, with diagonal entries s_{ii}/r_{ii} and r_{ii}/s_{ii}. So the matrix $M = SR^{-1}$ is upper triangular, and also lower triangular, being the conjugate transpose of an upper triangular matrix. Hence M is diagonal, and comparing diagonal entries, $s_{ii}/r_{ii} = \overline{r_{ii}/s_{ii}} = r_{ii}/s_{ii}$. So $s_{ii}^2 = r_{ii}^2$, and since both are positive, $s_{ii} = r_{ii}$. Therefore $M = I_n$ and $S = R$. □

The equations behind the algorithm come from reading $A = R^* R$ entry by entry. Since $r_{ik} = 0$ for $i > k$, for $k \leq j$ the (k, j) entry is

$$a_{kj} = \sum_{i \le k} \overline{r_{ik}} r_{ij} = \sum_{i<k} \overline{r_{ik}} r_{ij} + r_{kk} r_{kj}.$$

The last term is the only one that involves row k of R. So if rows $1, \dots, k-1$ are known, the equation with $j = k$ determines r_{kk}, and then the equations with $j > k$ determine the rest of row k.

Algorithm 8.3.24 (Cholesky factorization)

Input A Hermitian matrix $A \in \mathbb{F}^{n \times n}$.
Output The Cholesky factor R of A, or the conclusion that A is not positive definite.

1. Set $k = 1$.
2. Compute $d = a_{kk} - \sum_{i<k} |r_{ik}|^2$. If $d \le 0$, stop: A is not positive definite. Otherwise set $r_{kk} = \sqrt{d}$.
3. For $j = k+1, \dots, n$, set $r_{kj} = \left(a_{kj} - \sum_{i<k} \overline{r_{ik}} r_{ij}\right) / r_{kk}$.
4. If $k < n$, increase k by 1 and return to step 2. Otherwise output $R = \left(r_{ij}\right)$, with $r_{ij} = 0$ for $i > j$.

Proposition 8.3.25 (Correctness of the Cholesky algorithm). Let $A \in \mathbb{F}^{n \times n}$ be Hermitian. If A is positive definite, Algorithm 8.3.24 outputs the matrix R of Theorem 8.3.23. If A is not positive definite, Algorithm 8.3.24 stops in step 2.

Proof. STANDARD

Uses. Theorem 8.3.23 (Cholesky factorization), Proposition 8.3.15 (Characterizations of positive definite operators), Proposition 5.3.15 (Triangular matrices)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5), Compute entry by entry (Technique 2.4.15)
Hint 1. If A is positive definite, the Cholesky factor exists. Show that its entries satisfy exactly the equations that the algorithm solves, in the same order.
Hint 2. If the algorithm runs to the end, it has produced an upper triangular R with positive diagonal and $a_{kj} = (R^* R)_{kj}$ for $k \le j$. Why does this give $A = R^* R$, and why is A then positive definite?
Sketch. The entries of the Cholesky factor satisfy $r_{kk}^2 = a_{kk} - \sum_{i<k} |r_{ik}|^2$ and $r_{kk} r_{kj} = a_{kj} - \sum_{i<k} \overline{r_{ik}} r_{ij}$, so by induction on the order of computation the algorithm reproduces them, with $d = r_{kk}^2 > 0$. Conversely, if the algorithm finishes, A and $R^* R$ agree on and above the diagonal, hence everywhere, and $A = R^* R$ with R invertible is positive definite. So for A not positive definite it cannot finish.

Suppose A is positive definite, and let R be its Cholesky factor (Theorem 8.3.23). By the entry formula above, for each k,

$$r_{kk}^2 = a_{kk} - \sum_{i<k} |r_{ik}|^2, \quad r_{kk} r_{kj} = a_{kj} - \sum_{i<k} \overline{r_{ik}} r_{ij} \quad (j > k).$$

We show by induction on k that the algorithm computes row k of R. If it has computed rows $1, \dots, k-1$ correctly, then the number d of step 2 is $r_{kk}^2 > 0$, so the algorithm does not stop and sets $r_{kk} = \sqrt{d}$, the correct (positive) value; and step 3 solves the second equation for r_{kj}. So the output is R.

Conversely, suppose the algorithm runs to the end and outputs R. Then R is upper triangular with positive diagonal entries, and the equations it solved say that $a_{kj} = (R^* R)_{kj}$ for all $k \le j$. Both A and $R^* R$ are Hermitian, so they also agree below the diagonal, and $A = R^* R$. Since $\det R$ is the product of its diagonal entries (Proposition 5.3.15), R has rank n, and A is positive definite by Proposition 8.3.15. So if A is not positive definite, the algorithm cannot run to the end, and it stops in step 2. □

Example 8.3.26 (A Cholesky factorization). Let

$$A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 10 & 2 \\ -2 & 2 & 6 \end{pmatrix}.$$

For $k=1$: $r_{11}=\sqrt{4}=2$, $r_{12}=2/2=1$, $r_{13}=-2/2=-1$. For $k=2$: $d=10-1^2=9$, so $r_{22}=3$, and $r_{23}=(2-1\cdot(-1))/3=1$. For $k=3$: $d=6-(-1)^2-1^2=4$, so $r_{33}=2$. Hence

$$A=R^{\mathsf{T}}R,\quad R=\begin{pmatrix}2&1&-1\\0&3&1\\0&0&2\end{pmatrix},$$

and A is positive definite. If the corner entry 6 is replaced by 2, the last step gives $d=2-1-1=0$, and the algorithm stops: the new matrix A' is not positive definite. Indeed $A'(2,-1,3)=\mathbf{0}$, so $(2,-1,3)^{\mathsf{T}}A'(2,-1,3)=0$.

Remark 8.3.27 (What Cholesky adds). For a real positive definite $A=R^{\mathsf{T}}R$, the matrices $L=R^{\mathsf{T}}\operatorname{diag}(1/r_{11},\dots,1/r_{nn})$, which is unit lower triangular, and $D=\operatorname{diag}\bigl(r_{11}^2,\dots,r_{nn}^2\bigr)$ satisfy $LDL^{\mathsf{T}}=R^{\mathsf{T}}R=A$. This is the factorization of Theorem 3.3.37, and $R=\sqrt{D}L^{\mathsf{T}}$: Cholesky is LDL^{T} with the pivots split evenly between the two factors. Because of the symmetry it needs about $n^3/6$ multiplications, half as many as LU.

The induction proof never used eigenvalues. It shows directly that a positive definite matrix has the form B^*B, which is the hard implication of Theorem 8.3.5, by elimination alone. It also shows that a Hermitian matrix is positive definite if and only if the pivots r_{kk}^2 produced along the way are all positive, a criterion that Section 12.3 rephrases in terms of determinants (Sylvester's criterion).

Exercises

Exercise 8.3.28. ROUTINE Let $S,T\in\mathrm{L}(V)$ be positive, let $c\geq 0$ be real, and let $X\in\mathrm{L}(U,V)$, where U is a finite-dimensional inner product space. Show that:

(a) $S+T$ and cT are positive;

(b) $X^*TX\in\mathrm{L}(U)$ is positive;

(c) T^m is positive for every integer $m\geq 0$;

(d) if T is invertible, then T^{-1} is positive.

Solution on page 1652

Exercise 8.3.29. STANDARD Prove or disprove each statement. Throughout, V is a finite-dimensional inner product space.

(a) If $T\in\mathrm{L}(V)$ is positive, then for every basis $\mathcal{B}$ of V, the diagonal entries of $[T]_{\mathcal{B}\mathcal{B}}$ are nonnegative.

(b) If $T\in\mathrm{L}(V)$ is self-adjoint and T^3 is positive, then T is positive.

(c) If $T\in\mathrm{L}(V)$ is self-adjoint and T^2 is positive, then T is positive.

(d) If $S,T\in\mathrm{L}(V)$ are positive, then $ST+TS$ is positive.

(e) If $A\in\mathbb{C}^{n\times n}$ and $\mathbf{x}^*A\mathbf{x}\geq 0$ for every $\mathbf{x}\in\mathbb{R}^n$, then A is positive.

Solution on page 1652

Exercise 8.3.30. STANDARD Let $A\in\mathbb{R}^{n\times n}$. Show that the following are equivalent:

(i) A is positive as an element of $\mathbb{R}^{n\times n}$: $A^{\mathsf{T}}=A$ and $\mathbf{x}^{\mathsf{T}}A\mathbf{x}\geq 0$ for every $\mathbf{x}\in\mathbb{R}^n$;

(ii) $\mathbf{x}^*A\mathbf{x}\geq 0$ for every $\mathbf{x}\in\mathbb{C}^n$.

Then show that (ii) does not follow from "$\mathbf{x}^{\mathsf{T}}A\mathbf{x}\geq 0$ for every $\mathbf{x}\in\mathbb{R}^n$" alone. This is why Definition 8.3.1 requires self-adjointness over $\mathbb{R}$ (Remark 8.3.4).

Solution on page 1653

Exercise 8.3.31. STANDARD Let $A = \begin{pmatrix} a & b \\ \overline{b} & c \end{pmatrix} \in \mathbb{F}^{2\times 2}$ with $a, c \in \mathbb{R}$.

(a) Show that A is positive if and only if $a \geq 0$, $c \geq 0$, and $ac - |b|^2 \geq 0$.

(b) Show that A is positive definite if and only if $a > 0$ and $ac - |b|^2 > 0$.

(c) Show by an example that $a \geq 0$ and $\det A \geq 0$ do not imply that A is positive.

Solution on page 1653

Solution on page 1654

Exercise 8.3.32. STANDARD Let $T \in \mathrm{L}(V)$ be positive, let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis of V, and let $A = [T]_{\mathcal{O}}$.

(a) Show that $|a_{ij}|^2 \leq a_{ii}a_{jj}$ for all i, j.

(b) Deduce that if $a_{ii} = 0$, then row i and column i of A are zero.

(c) Deduce that the largest of the numbers $|a_{ij}|$ occurs on the diagonal.

Solution on page 1654

Solution on page 1655

Exercise 8.3.33. STANDARD Let $T \in \mathrm{L}(V)$ be positive.

(a) Show that $|\langle T\mathbf{v}, \mathbf{w}\rangle|^2 \leq \langle T\mathbf{v}, \mathbf{v}\rangle\langle T\mathbf{w}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$.

(b) Use (a) to give a third proof of Lemma 8.3.8.

(c) Show that $\ker\sqrt{T} = \ker T$ and $\operatorname{im}\sqrt{T} = \operatorname{im} T$.

Solution on page 1655

Exercise 8.3.34. STANDARD Let $S, T \in \mathrm{L}(V)$ be positive.

(a) Show that $\operatorname{tr} T \geq 0$ and $\det T \geq 0$, and that $\operatorname{tr} T = 0$ only if $T = 0$.

(b) Show that $\operatorname{tr}(ST) \geq 0$.

(c) Show that $\operatorname{tr}(ST) = 0$ if and only if $ST = 0$.

Solution on page 1656

Exercise 8.3.35. STANDARD

(a) Show that the $n \times n$ **Hilbert matrix** H_n, with entries $h_{ij} = 1/(i + j - 1)$, is positive definite. Deduce that $\det H_n > 0$.

(b) Show that the $n \times n$ matrix K_n with entries $k_{ij} = 1/(i + j)$ is positive definite.

Solution on page 1656

Exercise 8.3.36. STANDARD Let $A \in \mathbb{F}^{2\times 2}$ be positive and nonzero, and let $\delta = \sqrt{\det A}$.

(a) Show that

$$\sqrt{A} = \frac{A + \delta I_2}{\sqrt{\operatorname{tr} A + 2\delta}}.$$

(b) Use (a) to compute $\sqrt{A}$ for $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ and for $A = \begin{pmatrix} 5 & 4i \\ -4i & 5 \end{pmatrix}$.

Solution on page 1656

Exercise 8.3.37. STANDARD Let $T \in \mathrm{L}(V)$ be self-adjoint, with spectral decomposition $T = \sum_j \lambda_j P_j$. Define $|T| = \sqrt{T^2}$, which makes sense because $T^2 = T^*T$ is positive (Remark 8.3.13).

(a) Show that $|T| = \sum_j |\lambda_j| P_j$.

(b) Show that $|T| = T$ if and only if T is positive.

(c) Let $T_+ = (|T| + T)/2$ and $T_- = (|T| - T)/2$. Show that T_+ and T_- are positive, $T = T_+ - T_-$, and $T_+ T_- = 0$.

Solution on page 1657

Exercise 8.3.38. STANDARD Consider the following argument.

"*Claim.* A positive operator T has exactly one self-adjoint square root. *Proof.* Let R be self-adjoint with $R^2 = T$. Then R commutes with T, so by Proposition 8.2.20(a) it maps each eigenspace E_λ of T into itself. On E_λ we have $R^2 = T = \lambda\,\mathrm{id}$, so $R = \sqrt{\lambda}\,\mathrm{id}$ on E_λ. Hence $R = \sum_j \sqrt{\lambda_j} P_j = \sqrt{T}$."

(a) Find the error, and show that the claim is false.

(b) Suppose $\dim V = n$ and T is positive with n distinct eigenvalues, k of which are nonzero. How many self-adjoint square roots does T have?

Solution on page 1658

Exercise 8.3.39. HARD BEYOND 21-242 Let $A, B \in \mathbb{F}^{n\times n}$ be positive definite with $A \preceq B$ (Definition 8.3.21).

(a) Show that $B^{-1} \preceq A^{-1}$.

(b) Show that $\det A \le \det B$.

(c) Show that for positive matrices, $A \preceq B$ does not imply $A^2 \preceq B^2$: take $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Solution on page 1658

Exercise 8.3.40. HARD BEYOND 21-242 Let $A, B \in \mathrm{L}(V)$ with $0 \preceq A \preceq B$. Show that $\sqrt{A} \preceq \sqrt{B}$. Then show that the converse fails: $\sqrt{A} \preceq \sqrt{B}$ does not imply $A \preceq B$.

Solution on page 1659

Exercise 8.3.41. STANDARD BEYOND 21-242

(a) Find the Cholesky factorization of $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 13 & 8 \\ 1 & 8 & 14 \end{pmatrix}$ with Algorithm 8.3.24.

(b) Use Algorithm 8.3.24 to decide whether $B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 6 \end{pmatrix}$ is positive definite.

(c) From the computation in (b), find a vector $\mathbf{x}$ with $\mathbf{x}^{\mathsf{T}} B \mathbf{x} < 0$.

Solution on page 1659

Exercise 8.3.42. HARD BEYOND 21-242

(a) (Hadamard's inequality.) Show that every positive $A \in \mathbb{F}^{n\times n}$ satisfies $\det A \le a_{11} a_{22} \cdots a_{nn}$.

(b) Deduce that every $B \in \mathbb{F}^{n\times n}$ with columns $\mathbf{b}_1, \ldots, \mathbf{b}_n$ satisfies $|\det B| \le \|\mathbf{b}_1\| \|\mathbf{b}_2\| \cdots \|\mathbf{b}_n\|$. If B is invertible, show that equality holds if and only if the columns of B are orthogonal.

Solution on page 1660

Solution on page 1660

Exercise 8.3.43. CHALLENGE BEYOND 21-242 Let $A \in \mathbb{F}^{n\times n}$ be positive definite. For $C \in \mathbb{F}^{n\times n}$ consider the equation

$$AX + XA = C \quad (X \in \mathbb{F}^{n\times n}).$$

(a) Show that the equation has exactly one solution X.

(b) Show that if C is Hermitian, then so is X.

(c) Show that if C is positive, then so is X.

(d) Deduce that for real numbers $a_1, \ldots, a_n > 0$, the matrix with entries $1/(a_i + a_j)$ is positive.

Solution on page 1661

8.4 Isometries and unitary operators

A rotation of the plane moves almost every vector but changes no length and no angle. Section 6.3 met the matrices that behave like this, the unitary matrices, through a condition on their columns, $U^*U = I_n$. This section starts from the other end, from linear maps that preserve lengths, and shows that preserving lengths forces the rest: inner products are preserved, $S^*S = \mathrm{id}$, and the columns of every matrix in orthonormal bases are orthonormal. For operators on a finite-dimensional space these maps are the unitary operators of Table 8.1.1, the analogues of the complex numbers on the unit circle, and over $\mathbb{C}$ the spectral theorem describes them completely. Over $\mathbb{R}$ the description is the classical one, by rotations in perpendicular planes and reflections. We derive it in the second half of the section, together with the distance-preserving maps of $\mathbb{R}^n$ that are not linear.

Throughout, V and W are inner product spaces over $\mathbb{F}$, and inner products are linear in the first slot.

8.4.1 Isometries

Definition 8.4.1 (Isometry). Let V and W be inner product spaces over $\mathbb{F}$. A linear map $S \in \mathrm{L}(V, W)$ is an **isometry** if

$$\|S\mathbf{v}\| = \|\mathbf{v}\| \quad \text{for every } \mathbf{v} \in V.$$

Since S is linear, $\|S\mathbf{u} - S\mathbf{v}\| = \|S(\mathbf{u} - \mathbf{v})\| = \|\mathbf{u} - \mathbf{v}\|$, so an isometry preserves distances as well as lengths. In metric geometry "isometry" means any distance-preserving map, linear or not, and translations are the basic nonlinear examples. Our isometries are always linear; Section 8.4.6 shows that, in a real inner product space, the general distance-preserving maps differ from them only by a translation.

Example 8.4.2 (Isometries).

(a) Every unitary matrix $U \in \mathbb{F}^{n\times n}$ gives an isometry $\mathbf{x} \mapsto U\mathbf{x}$ of $\mathbb{F}^n$ (Exercise 6.3.37). In particular, the rotations R_θ and the reflections S_θ of Example 6.3.35 are isometries of $\mathbb{R}^2$.

(b) More generally, let $A \in \mathbb{F}^{m\times n}$ have orthonormal columns. Then $A^*A = I_n$ (Theorem 6.3.29), so for every $\mathbf{x} \in \mathbb{F}^n$,

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2.$$

For instance, $A = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$ maps $\mathbb{R}^2$ isometrically onto a plane in $\mathbb{R}^3$; it sends $(1,1)$ to $(1,1,0)$, and both have length $\sqrt{2}$.

(c) If U is a subspace of V, the inclusion $U \to V$, $\mathbf{u} \mapsto \mathbf{u}$, is an isometry.

(d) Let $\mathcal{O} = \mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis of V. By Theorem 6.3.22(b), the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_\mathcal{O}$ is an isometry from V to $\mathbb{F}^n$ with the dot product, and so is its inverse $(a_1, \ldots, a_n) \mapsto a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n$.

Non-example 8.4.3 (Maps that are not isometries).

- The shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ sends $\mathbf{e}_2$ to $(1,1)$, which has length $\sqrt{2}$.
- An orthogonal projection Pr_U with $U \neq V$ is not an isometry. For $\mathbf{v} \notin U$, the vector $\mathbf{v} - \mathrm{Pr}_U \mathbf{v}$ is nonzero and orthogonal to $\mathrm{Pr}_U \mathbf{v}$, so $\|\mathrm{Pr}_U \mathbf{v}\| < \|\mathbf{v}\|$ by Theorem 6.3.3.
- The columns of $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ are orthogonal but have length $\sqrt{2}$. The matrix is $\sqrt{2}$ times an orthogonal matrix, so it multiplies every length by $\sqrt{2}$. It preserves right angles without being an isometry (compare Exercise 8.4.34).

Intuition (Why lengths control angles). Over $\mathbb{R}$, the law of cosines

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$$

computes the inner product of $\mathbf{u}$ and $\mathbf{v}$ from the three side lengths of the triangle with vertices $\mathbf{0}$, $\mathbf{u}$, $\mathbf{v}$. A linear map S sends this triangle to the triangle with vertices $\mathbf{0}$, $S\mathbf{u}$, $S\mathbf{v}$, whose third side is $S\mathbf{u} - S\mathbf{v} = S(\mathbf{u} - \mathbf{v})$. If S preserves lengths, the two triangles have the same side lengths, hence the same angle at $\mathbf{0}$. Over $\mathbb{C}$ the polarization identity (Proposition 6.1.32) plays the role of the law of cosines. Linearity is used once, to know that the third side of the image is the image of the third side.

The main theorem lists the equivalent forms of the definition. Parts of it appeared as exercises in Section 6.6 (Exercise 6.1.51 and Exercise 6.6.40).

Theorem 8.4.4 (Characterizations of isometries). Let V and W be finite-dimensional inner product spaces with $n = \dim V$, and let $S \in \mathrm{L}(V, W)$. The following are equivalent.

(I) S is an isometry.

(II) S preserves inner products: $\langle S\mathbf{u}, S\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ for all $\mathbf{u}, \mathbf{v} \in V$.

(III) $S^*S = \mathrm{id}_V$.

(IV) S maps every orthonormal list in V to an orthonormal list in W.

(V) S maps some orthonormal basis of V to an orthonormal list in W.

(VI) For some orthonormal bases $\mathcal{B}$ of V and $\mathcal{C}$ of W, the columns of $A = [S]_{\mathcal{C}\mathcal{B}}$ are orthonormal, that is, $A^*A = I_n$.

(VII) For all orthonormal bases $\mathcal{B}$ of V and $\mathcal{C}$ of W, the columns of $[S]_{\mathcal{C}\mathcal{B}}$ are orthonormal.

Proof (by polarization). STANDARD

Uses. Proposition 6.1.32 (Polarization identities), Definition 6.6.8 (Adjoint), Proposition 6.1.37 (Inner products distinguish vectors), Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.3.18 (Coordinates in an orthonormal basis), Corollary 6.3.6 (Pythagoras for orthogonal lists), Theorem 6.6.15 (Existence of adjoints), Theorem 2.4.2 (The matrix of a composite), Theorem 2.3.20 (Linear maps correspond to matrices), Theorem 6.3.29 (Orthonormal columns)
Techniques. Proving TFAE (Technique 2.7.1), Polarization (Technique 6.1.33), Move an operator across the inner product (Technique 6.6.20), Translate between maps and matrices (Technique 2.3.27)
Hint 1. For (I) ⇒ (II), express $\langle S\mathbf{u}, S\mathbf{v}\rangle$ through lengths of the vectors $S\mathbf{u} + cS\mathbf{v}$, and use linearity to pull S out.
Hint 2. For (II) ⇒ (III), rewrite $\langle S\mathbf{u}, S\mathbf{v}\rangle$ with the adjoint as $\langle \mathbf{u}, S^*S\mathbf{v}\rangle$. For (V) ⇒ (I), expand $\mathbf{v}$ in the orthonormal basis and apply Pythagoras to $\mathbf{v}$ and to $S\mathbf{v}$.
Hint 3. For the matrix conditions, $[S^*]_{\mathcal{BC}} = A^*$ in orthonormal bases, so $[S^*S]_{\mathcal{BB}} = A^*A$.
Sketch. Polarization writes $\langle S\mathbf{u}, S\mathbf{v}\rangle$ as a combination of $\|S(\mathbf{u} + c\mathbf{v})\|^2 = \|\mathbf{u} + c\mathbf{v}\|^2$, which gives (II). Then $\langle \mathbf{u}, S^*S\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ for all $\mathbf{u}$ forces $S^*S\mathbf{v} = \mathbf{v}$. From (III), $\langle S\mathbf{e}_j, S\mathbf{e}_i\rangle = \langle \mathbf{e}_j, S^*S\mathbf{e}_i\rangle = \delta_{ij}$. If $S\mathbf{e}_1, \dots, S\mathbf{e}_n$ is orthonormal, Pythagoras gives $\|S\mathbf{v}\|^2 = \sum |a_j|^2 = \|\mathbf{v}\|^2$. In orthonormal bases S^*S has matrix A^*A, so (III) holds iff $A^*A = I_n$ in one, or in all, such pairs of bases.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (IV) ⇒ (V) ⇒ (I), and then (III) ⇒ (VII) ⇒ (VI) ⇒ (III).

(I) ⇒ (II). Let $\mathbf{u}, \mathbf{v} \in V$. For every scalar c, linearity gives $S\mathbf{u} + cS\mathbf{v} = S(\mathbf{u} + c\mathbf{v})$, so by (I)

$$\|S\mathbf{u} + cS\mathbf{v}\| = \|\mathbf{u} + c\mathbf{v}\|.$$

If $\mathbb{F} = \mathbb{R}$, Proposition 6.1.32(a) with $c = \pm 1$ gives

$$\langle S\mathbf{u}, S\mathbf{v}\rangle = \frac{1}{4}\Big(\|S\mathbf{u} + S\mathbf{v}\|^2 - \|S\mathbf{u} - S\mathbf{v}\|^2\Big) = \frac{1}{4}\Big(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2\Big) = \langle \mathbf{u}, \mathbf{v}\rangle.$$

If $\mathbb{F} = \mathbb{C}$, the same computation with the four values $c = i^k$ in Proposition 6.1.32(b) gives $\langle S\mathbf{u}, S\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$.

(II) ⇒ (III). Let $\mathbf{v} \in V$. By the definition of the adjoint (Definition 6.6.8) and (II), for every $\mathbf{u} \in V$

$$\langle \mathbf{u}, S^*S\mathbf{v}\rangle = \langle S\mathbf{u}, S\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle.$$

Conjugating both sides, $\langle S^*S\mathbf{v}, \mathbf{u}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$ for all $\mathbf{u}$, so $S^*S\mathbf{v} = \mathbf{v}$ by Proposition 6.1.37. Hence $S^*S = \mathrm{id}_V$.

(III) ⇒ (IV). Let $\mathbf{e}_1, \dots, \mathbf{e}_k$ be orthonormal in V. Then

$$\langle S\mathbf{e}_j, S\mathbf{e}_i\rangle = \langle \mathbf{e}_j, S^*S\mathbf{e}_i\rangle = \langle \mathbf{e}_j, \mathbf{e}_i\rangle = \delta_{ij},$$

so $S\mathbf{e}_1, \dots, S\mathbf{e}_k$ is orthonormal.

(IV) ⇒ (V). V has an orthonormal basis (Theorem 6.4.20(a)), and by (IV) its image is orthonormal.

(V) ⇒ (I). Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be an orthonormal basis of V with $S\mathbf{e}_1, \dots, S\mathbf{e}_n$ orthonormal, and let $\mathbf{v} \in V$. By Proposition 6.3.18, $\mathbf{v} = \sum_j a_j \mathbf{e}_j$ with $a_j = \langle \mathbf{v}, \mathbf{e}_j\rangle$, so $S\mathbf{v} = \sum_j a_j S\mathbf{e}_j$. Applying Corollary 6.3.6 to both orthonormal lists,

$$\|S\mathbf{v}\|^2 = \sum_{j=1}^{n} |a_j|^2 \|S\mathbf{e}_j\|^2 = \sum_{j=1}^{n} |a_j|^2 = \sum_{j=1}^{n} |a_j|^2 \|\mathbf{e}_j\|^2 = \|\mathbf{v}\|^2.$$

(III) ⇒ (VII). Let $\mathcal{B}, \mathcal{C}$ be orthonormal bases of V and W, and $A = [S]_{\mathcal{CB}}$. By Theorem 6.6.15(b), $[S^*]_{\mathcal{BC}} = A^*$, and by Theorem 2.4.2,

$$A^*A = [S^*]_{\mathcal{BC}}[S]_{\mathcal{CB}} = [S^*S]_{\mathcal{BB}} = [\mathrm{id}_V]_{\mathcal{BB}} = I_n.$$

By Theorem 6.3.29, $A^*A = I_n$ says exactly that the columns of A are orthonormal.

(VII) ⇒ (VI). Orthonormal bases of V and W exist (Theorem 6.4.20(a)).

(VI) ⇒ (III). The computation in (III) ⇒ (VII), read backwards, gives $[S^*S]_{\mathcal{BB}} = A^*A = I_n = [\mathrm{id}_V]_{\mathcal{BB}}$, and a linear map is determined by its matrix (Theorem 2.3.20). So $S^*S = \mathrm{id}_V$. □

The polarization step is the only place where an identity for single vectors becomes one for pairs. There is a second route from lengths to S^*S that hides the pairs inside a result about self-adjoint operators.

Proof (through the quadratic form of $S^*S - \mathrm{id}_V$). STANDARD

> *Uses.* Proposition 6.6.18 (Properties of the adjoint), Definition 6.6.8 (Adjoint), Proposition 8.1.4 (Self-adjoint operators are determined by their diagonal values)
> *Techniques.* Move an operator across the inner product (Technique 6.6.20)
> *Hint 1.* Let $P = S^*S - \mathrm{id}_V$. Is P self-adjoint? Compute $\langle P\mathbf{v}, \mathbf{v}\rangle$.
> *Hint 2.* $\langle P\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2 - \|\mathbf{v}\|^2$. Which result says that a self-adjoint operator with $\langle P\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$ is zero?
> *Sketch.* $P = S^*S - \mathrm{id}_V$ is self-adjoint, and $\langle P\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2 - \|\mathbf{v}\|^2$. So (I) says that the quadratic form of P vanishes, and Proposition 8.1.4 gives $P = 0$. Then (III) ⇒ (II) is one line, and the rest is as in the first proof.

We prove the cycle (I) ⇒ (III) ⇒ (II) ⇒ (I); the remaining implications are as in the first proof.

Let $P = S^*S - \mathrm{id}_V$. By Proposition 6.6.18, $(S^*S)^* = S^*(S^*)^* = S^*S$, so P is self-adjoint. For $\mathbf{v} \in V$, the definition of the adjoint gives $\langle S^*S\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, S\mathbf{v}\rangle$ (apply Definition 6.6.8 to S^*, whose adjoint is S), so

$$\langle P\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, S\mathbf{v}\rangle - \langle \mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2 - \|\mathbf{v}\|^2.$$

If (I) holds, this vanishes for every $\mathbf{v}$, and $P = 0$ by Proposition 8.1.4; that is (III). Conversely, if (III) holds, then for all $\mathbf{u}, \mathbf{v}$,

$$\langle S\mathbf{u}, S\mathbf{v}\rangle = \langle \mathbf{u}, S^*S\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle,$$

which is (II), and taking $\mathbf{u} = \mathbf{v}$ in (II) gives (I). □

The first proof recovers inner products from lengths; the second recovers the operator $S^*S - \mathrm{id}_V$ from its values $\langle P\mathbf{v}, \mathbf{v}\rangle$. They rest on the same fact (the proof of Proposition 8.1.4 is a polarization argument), but the second shows the shape of the condition. An isometry is a map whose "defect" $S^*S - \mathrm{id}_V$ has zero quadratic form, and Section 8.3 studies maps whose S^*S is merely compared with id_V.

Remark 8.4.5 (Without finite dimension). The implications (I) ⇒ (II) ⇒ (I) use neither bases nor adjoints, and neither does (II) ⇒ (IV), since $\langle S\mathbf{e}_j, S\mathbf{e}_i\rangle = \langle \mathbf{e}_j, \mathbf{e}_i\rangle = \delta_{ij}$ for an orthonormal list; so they hold for a linear map between any two inner product spaces. So does (II) ⇒ (III), whenever S has an adjoint. In particular, every isometry preserves inner products, orthogonality, and angles (Definition 6.2.16).

Corollary 8.4.6 (Properties of isometries). Let $S \in \mathrm{L}(V, W)$ be an isometry.

(a) S is injective, and it preserves inner products, distances, and orthogonality.

(b) If $R \in \mathrm{L}(W, X)$ is also an isometry, so is RS.

(c) If V and W are finite-dimensional, then $\dim V \le \dim W$, and SS^* is the orthogonal projection $\mathrm{Pr}_{\mathrm{im}\, S}$.

(d) If $\dim V = \dim W < \infty$, then S is bijective, $S^{-1} = S^*$, and S^{-1} is an isometry.

Proof. ROUTINE

> *Uses.* Remark 8.4.5 (Without finite dimension), Theorem 8.4.4 (Characterizations of isometries), Proposition 4.2.6 (Injective maps preserve dimension), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.6.18 (Properties of the adjoint), Theorem 4.2.11 (Rank–nullity)
> *Techniques.* Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
> *Hint 1.* For injectivity, test at $\mathbf{0}$: what is $\|\mathbf{v}\|$ if $S\mathbf{v} = \mathbf{0}$?
> *Hint 2.* For (c), check that $P = SS^*$ satisfies $P^2 = P = P^*$, and compare $\mathrm{im}\, P$ with $\mathrm{im}\, S$ using $S = S(S^*S)$.
> *Hint 3.* For (d), S is injective between spaces of equal finite dimension, and S^* is a left inverse.
> *Sketch.* If $S\mathbf{v} = \mathbf{0}$ then $\|\mathbf{v}\| = \|S\mathbf{v}\| = 0$. Composites: $\|RS\mathbf{v}\| = \|S\mathbf{v}\| = \|\mathbf{v}\|$. Injectivity gives $\dim V = \dim \mathrm{im}\, S \le \dim W$. $P = SS^*$ has $P^2 = S(S^*S)S^* = P$ and $P^* = P$, and $\mathrm{im}\, P = \mathrm{im}\, S$ because $S = PS$. If $\dim V = \dim W$, S is bijective, and $S^* = S^*\left(SS^{-1}\right) = S^{-1}$.

(a) If $S\mathbf{v} = \mathbf{0}$, then $\|\mathbf{v}\| = \|S\mathbf{v}\| = 0$, so $\mathbf{v} = \mathbf{0}$; hence S is injective. Inner products are preserved by Remark 8.4.5, distances because $S\mathbf{u} - S\mathbf{v} = S(\mathbf{u} - \mathbf{v})$, and orthogonality because it is a statement about inner products.

(b) $\|RS\mathbf{v}\| = \|S\mathbf{v}\| = \|\mathbf{v}\|$ for every $\mathbf{v}$.

(c) Since S is injective, $\dim V = \dim \operatorname{im} S \le \dim W$ (Proposition 4.2.6). Let $P = SS^*$. By (III) of Theorem 8.4.4 and Proposition 6.6.18,

$$P^2 = S(S^*S)S^* = SS^* = P, \quad P^* = (S^*)^*S^* = SS^* = P,$$

so P is the orthogonal projection onto $\operatorname{im} P$ (Corollary 6.6.29). Clearly $\operatorname{im} P \subseteq \operatorname{im} S$, and $S = S(S^*S) = PS$ gives $\operatorname{im} S \subseteq \operatorname{im} P$. So $P = \operatorname{Pr}_{\operatorname{im} S}$.

(d) By (a) and rank–nullity (Theorem 4.2.11), $\dim \operatorname{im} S = \dim V = \dim W$, so S is also surjective. Then $S^* = S^*\big(SS^{-1}\big) = (S^*S)S^{-1} = S^{-1}$, and for $\mathbf{w} \in W$, $\big\|S^{-1}\mathbf{w}\big\| = \big\|S\big(S^{-1}\mathbf{w}\big)\big\| = \|\mathbf{w}\|$. □

For example, the matrix $A = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$ of Example 8.4.2(b) has

$$A^{\mathsf{T}}A = I_2, \quad AA^{\mathsf{T}} = \frac{1}{9}\begin{pmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{pmatrix},$$

and AA^{T} is the projection onto the plane $\operatorname{im} A$, whose unit normal is $\frac{1}{3}(2,-2,1)$. Indeed $AA^{\mathsf{T}} = I_3 - \mathbf{n}\mathbf{n}^{\mathsf{T}}$ with $\mathbf{n} = \frac{1}{3}(2,-2,1)$. So A^{T} is a left inverse of A but not a right inverse (compare Exercise 6.6.41).

The next corollary makes precise the remark in Section 6.3 that an n-dimensional inner product space "looks like" $\mathbb{F}^n$ with the dot product.

Corollary 8.4.7 (Inner product spaces are classified by dimension). Let V and W be finite-dimensional inner product spaces over $\mathbb{F}$. There is a bijective isometry $V \to W$ if and only if $\dim V = \dim W$. In particular, for every orthonormal basis $\mathcal{O}$ of an n-dimensional V, the coordinate map $\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{O}}$ is a bijective isometry $V \to \mathbb{F}^n$.

Proof. ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Theorem 2.2.1 (Linear maps are determined by a basis), Theorem 8.4.4 (Characterizations of isometries), Corollary 8.4.6 (Properties of isometries), Example 8.4.2 (Isometries)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. A bijective linear map preserves dimension. For the converse, which map should send what where?
Hint 2. Choose orthonormal bases of V and W, and define S on the first basis by sending it to the second.
Sketch. A bijective linear map sends a basis to a basis, so the dimensions agree. Conversely, pick orthonormal bases $\mathbf{e}_1, \dots, \mathbf{e}_n$ and $\mathbf{f}_1, \dots, \mathbf{f}_n$, and let $S\mathbf{e}_j = \mathbf{f}_j$; by (V) of Theorem 8.4.4 S is an isometry, and it is bijective by Corollary 8.4.6(d).

If $S : V \to W$ is a bijective isometry, it is a bijective linear map, so it sends a basis of V to a basis of W and $\dim V = \dim W$. Conversely, suppose $\dim V = \dim W = n$. By Theorem 6.4.20(a) there are orthonormal bases $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V and $\mathbf{f}_1, \dots, \mathbf{f}_n$ of W, and by Theorem 2.2.1 there is a linear $S : V \to W$ with $S\mathbf{e}_j = \mathbf{f}_j$ for each j. It maps an orthonormal basis to an orthonormal list, so it is an isometry by (V) ⇒ (I) of Theorem 8.4.4, and it is bijective by Corollary 8.4.6(d). The last sentence is Example 8.4.2(d): the coordinate map sends $\mathbf{e}_j$ to the standard basis vector $\mathbf{e}_j$ of $\mathbb{F}^n$. □

So everything about lengths and angles in an n-dimensional inner product space can be computed in $\mathbb{F}^n$ with the dot product, once an orthonormal basis is fixed. What does depend on the space is how particular operators sit inside it.

Exercise 8.4.8. ROUTINE Let V, W be finite-dimensional and $S \in \mathrm{L}(V, W)$.

(a) Prove or disprove: if $\|S\mathbf{e}_j\| = 1$ for every vector of some orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V, then S is an isometry.

(b) Prove or disprove: if $\|S\mathbf{v}\| = 1$ for every unit vector $\mathbf{v} \in V$, then S is an isometry.

Solution on page 1661

8.4.2 Unitary operators

An isometry $V \to V$ is injective. In finite dimensions it is therefore invertible, and its inverse is again an isometry; in infinite dimensions it need not be (Example 8.4.12). The invertible ones get a name.

Definition 8.4.9 (Unitary operator). Let V be an inner product space. An operator $S \in \mathrm{L}(V)$ is **unitary** if it is an isometry and invertible. When $\mathbb{F} = \mathbb{R}$, unitary operators are also called **orthogonal operators**.

In Table 8.1.1 the unitary row was described by $T^*T = TT^* = \mathrm{id}$. The two descriptions agree in every dimension.

Proposition 8.4.10 (Unitary means $S^* = S^{-1}$). Let V be an inner product space and $S \in \mathrm{L}(V)$. Then S is unitary if and only if S has an adjoint and

$$S^*S = SS^* = \mathrm{id}_V.$$

In that case $S^* = S^{-1}$, and S^* is unitary as well.

Proof. ROUTINE

Uses. Remark 8.4.5 (Without finite dimension), Definition 6.6.8 (Adjoint), Proposition 6.6.9 (Adjoints are unique and linear), Definition 2.5.12 (Invertible)
Techniques. Find the adjoint by isolating the input (Technique 6.6.11)
Hint 1. To show that S^* exists, guess it: which map could satisfy $\langle S\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, S^*\mathbf{w}\rangle$?
Hint 2. Write $\mathbf{w} = S\big(S^{-1}\mathbf{w}\big)$ and use that S preserves inner products.
Sketch. If S is unitary, $\langle S\mathbf{v}, \mathbf{w}\rangle = \big\langle S\mathbf{v}, SS^{-1}\mathbf{w}\big\rangle = \big\langle \mathbf{v}, S^{-1}\mathbf{w}\big\rangle$, so S^{-1} is the adjoint. Conversely $S^*S = SS^* = \mathrm{id}$ says that S is invertible with inverse S^*, and $\|S\mathbf{v}\|^2 = \langle \mathbf{v}, S^*S\mathbf{v}\rangle = \|\mathbf{v}\|^2$.

Suppose S is unitary. By Remark 8.4.5, S preserves inner products, so for all $\mathbf{v}, \mathbf{w} \in V$,

$$\langle S\mathbf{v}, \mathbf{w}\rangle = \big\langle S\mathbf{v}, S\big(S^{-1}\mathbf{w}\big)\big\rangle = \big\langle \mathbf{v}, S^{-1}\mathbf{w}\big\rangle.$$

So S^{-1} is an adjoint of S (Definition 6.6.8), and by Proposition 6.6.9, $S^* = S^{-1}$. Hence $S^*S = SS^* = \mathrm{id}_V$.

Conversely, suppose $S^*S = SS^* = \mathrm{id}_V$. Then S is invertible with $S^{-1} = S^*$ (Definition 2.5.12), and for every $\mathbf{v}$,

$$\|S\mathbf{v}\|^2 = \langle S\mathbf{v}, S\mathbf{v}\rangle = \langle \mathbf{v}, S^*S\mathbf{v}\rangle = \|\mathbf{v}\|^2.$$

So S is unitary. Finally, $S^* = S^{-1}$ is invertible, and $\big\|S^{-1}\mathbf{v}\big\| = \big\|S\big(S^{-1}\mathbf{v}\big)\big\| = \|\mathbf{v}\|$, so S^* is unitary. □

Theorem 8.4.11 (Characterizations of unitary operators). Let V be a finite-dimensional inner product space and $S \in \mathrm{L}(V)$. The following are equivalent.

(I) S is an isometry.
(II) S is unitary.
(III) $S^*S = \mathrm{id}_V$.
(IV) $SS^* = \mathrm{id}_V$.
(V) S^* is an isometry.
(VI) S maps some orthonormal basis of V to an orthonormal basis of V.
(VII) S maps every orthonormal basis of V to an orthonormal basis of V.
(VIII) $[S]_{\mathcal{O}}$ is a unitary matrix for some orthonormal basis $\mathcal{O}$ of V, or equivalently for every one.

Proof. ROUTINE

Uses. Theorem 8.4.4 (Characterizations of isometries), Corollary 8.4.6 (Properties of isometries), Proposition 8.4.10 (Unitary means $S^* = S^{-1}$), Remark 4.2.20 (One-sided inverses of operators), Proposition 6.6.18 (Properties of the adjoint), Corollary 6.3.14 (Counting orthogonal vectors), Definition 6.3.33 (Unitary and orthogonal matrices)

Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Most of the work is in Theorem 8.4.4 with $W = V$. What does finite dimension add?
Hint 2. An injective operator on a finite-dimensional space is invertible, and a one-sided inverse of such an operator is two-sided.
Hint 3. For (V), apply Theorem 8.4.4 to S^*, whose adjoint is S. For (VI) and (VII), an orthonormal list of length $\dim V$ is a basis.
Sketch. (I) ⇔ (III) is Theorem 8.4.4, and an isometry of a finite-dimensional space is invertible (Corollary 8.4.6(d)), which gives (II). Unitary operators satisfy (IV) by Proposition 8.4.10, and (IV) ⇒ (III) because one-sided inverses of operators are two-sided. (V) is (IV) by Theorem 8.4.4 applied to S^*. An orthonormal list of length n is an orthonormal basis, so (VI) and (VII) are (V) and (IV) of Theorem 8.4.4; and (VIII) is its (VI)–(VII) with $\mathcal{B} = \mathcal{C} = \mathcal{O}$.

We prove (I) ⇒ (II) ⇒ (IV) ⇒ (III) ⇒ (I), then (IV) ⇔ (V), and finally (I) ⇔ (VI) ⇔ (VII) and (III) ⇔ (VIII).

(I) ⇒ (II). By Corollary 8.4.6(d) with $W = V$, S is bijective, with $S^{-1} = S^*$; so S is an invertible isometry.

(II) ⇒ (IV). Proposition 8.4.10.

(IV) ⇒ (III). S^* is a right inverse of the operator S on the finite-dimensional space V, hence also a left inverse (Remark 4.2.20).

(III) ⇒ (I). (III) ⇒ (I) of Theorem 8.4.4.

(IV) ⇔ (V). By Proposition 6.6.18, $(S^*)^* = S$. So by (I) ⇔ (III) of Theorem 8.4.4 applied to $S^* \in \mathrm{L}(V)$, S^* is an isometry if and only if $SS^* = \mathrm{id}_V$.

(I) ⇔ (VI) ⇔ (VII). Let $n = \dim V$. An orthonormal list of length n in V is a basis (Corollary 6.3.14). So (VI) says the same as (V) of Theorem 8.4.4, and (VII) says the same as its (IV) restricted to lists of length n. Thus (I) ⇒ (VII) by (I) ⇒ (IV) there, (VII) ⇒ (VI) is clear, and (VI) ⇒ (I) is (V) ⇒ (I) there.

(III) ⇔ (VIII). With $\mathcal{B} = \mathcal{C} = \mathcal{O}$, (VI) and (VII) of Theorem 8.4.4 say that $[S]_{\mathcal{O}}^* [S]_{\mathcal{O}} = I_n$ for some, respectively every, orthonormal basis $\mathcal{O}$, which is the definition of a unitary matrix (Definition 6.3.33). Both are equivalent to (III). □

So for operators on finite-dimensional spaces "isometry", "unitary", and "unitary matrix in an orthonormal basis" mean the same thing, and we use them interchangeably. For a real matrix Q, the operator $\mathbf{x} \mapsto Q\mathbf{x}$ on $\mathbb{R}^n$ with the dot product is orthogonal exactly when Q is an orthogonal matrix.

BEYOND 21-242: Infinite-dimensional spaces are not covered in 21-242.

Example 8.4.12 (The shift). On $\mathbb{F}[x]$, define

$$\left\langle a_0 + a_1 x + \cdots + a_d x^d, b_0 + b_1 x + \cdots + b_d x^d \right\rangle = \sum_{k=0}^{d} a_k \overline{b_k},$$

padding the shorter polynomial with zero coefficients. This is the dot product of coefficient vectors, so it is an inner product, and $1, x, x^2, \ldots$ is an orthonormal list. Let $M \in \mathrm{L}(\mathbb{F}[x])$ be multiplication by x, which shifts the coefficients one step to the right. Then $\|xp\| = \|p\|$ for every p, so M is an isometry, but it is not surjective: every xp has constant term 0, so $1 \notin \operatorname{im} M$. So M is an isometry of an inner product space into itself that is not unitary, and (I) ⇒ (II) of Theorem 8.4.11 fails without finite dimension. Exercise 8.4.40 computes its adjoint, which is a left inverse but not a right inverse.

Isometries have two more invariants that we can read off right away.

Corollary 8.4.13 (Eigenvalues and determinant).

(a) If $S \in \mathrm{L}(V)$ is an isometry and λ is an eigenvalue of S, then $|\lambda| = 1$.

(b) If V is finite-dimensional and $S \in \mathrm{L}(V)$ is unitary, then $|\det S| = 1$. In particular $\det S = \pm 1$ when $\mathbb{F} = \mathbb{R}$.

Proof (by norms and by $\det(U^*) = \overline{\det U}$). ROUTINE

> *Uses.* Definition 7.2.6 (Eigenvalue, eigenvector), Definition 5.4.18 (Determinant of an operator), Theorem 8.4.11 (Characterizations of unitary operators), Theorem 5.4.14 (The transpose), Theorem 5.4.10 (Multiplicativity), Definition 5.3.6 (Determinant)
> *Hint 1.* For (a), take norms in $S\mathbf{v} = \lambda\mathbf{v}$.
> *Hint 2.* For (b), pass to the unitary matrix $U = [S]_{\mathcal{O}}$ and take the determinant of $U^*U = I_n$. You need $\det(U^*) = \overline{\det U}$.
> *Sketch.* (a) $\|\mathbf{v}\| = \|S\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$ with $\mathbf{v} \neq \mathbf{0}$. (b) $\det S = \det U$ for the unitary matrix $U = [S]_{\mathcal{O}}$, and $1 = \det(U^*U) = \overline{\det U}\det U = |\det U|^2$, since conjugation commutes with the polynomial det and det is unchanged by transposing.

(a) Let $S\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Then $\|\mathbf{v}\| = \|S\mathbf{v}\| = \|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$, and dividing by $\|\mathbf{v}\| \neq 0$ gives $|\lambda| = 1$.

(b) Let $\mathcal{O}$ be an orthonormal basis of V and $U = [S]_{\mathcal{O}}$, which is a unitary matrix by Theorem 8.4.11; $\det S = \det U$ by Definition 5.4.18. The determinant is a polynomial in the entries with integer coefficients (Definition 5.3.6), and complex conjugation respects sums and products, so $\det(\overline{U}) = \overline{\det U}$, where $\overline{U}$ conjugates every entry. With Theorem 5.4.14, $\det(U^*) = \det\left(\overline{U}^{\mathsf{T}}\right) = \det(\overline{U}) = \overline{\det U}$. Now by Theorem 5.4.10,

$$1 = \det I_n = \det(U^*U) = \det(U^*)\det U = \overline{\det U}\det U = |\det U|^2.$$

So $|\det S| = 1$. Over $\mathbb{R}$, $\det S$ is a real number of absolute value 1. □

Proof (through the eigenvalues, over $\mathbb{C}$). ROUTINE

> *Uses.* Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem 8.4.11 (Characterizations of unitary operators), Definition 5.4.18 (Determinant of an operator)
> *Hint 1.* Over $\mathbb{C}$, the determinant is the product of the eigenvalues counted with multiplicity.
> *Hint 2.* For a real orthogonal operator, regard its matrix as a complex unitary matrix.
> *Sketch.* For (b): $\det U = \lambda_1 \cdots \lambda_n$, and each $|\lambda_j| = 1$ by (a). A real orthogonal matrix is also a complex unitary matrix, with the same determinant.

Part (a) is as in the first proof. For (b), let $U = [S]_{\mathcal{O}}$ for an orthonormal basis $\mathcal{O}$. Over $\mathbb{R}$, U has real entries, and $U^*U = I_n$ also holds in $\mathbb{C}^{n\times n}$, so in either case we may regard U as a complex unitary matrix with the same determinant. By Theorem 7.5.22, $\det U = \lambda_1 \cdots \lambda_n$, where $\lambda_1, \ldots, \lambda_n$ are the complex eigenvalues of U counted with multiplicity. Each is an eigenvalue of the isometry $\mathbf{x} \mapsto U\mathbf{x}$ of $\mathbb{C}^n$, so $|\lambda_j| = 1$ by (a), and $|\det U| = \prod_j |\lambda_j| = 1$. □

For unitary operators, (a) is also Corollary 8.1.20(c), which was proved through normality. The norm argument here needs neither normality nor an adjoint, so it applies to every isometry, including the shift of Example 8.4.12 (which turns out to have no eigenvalues at all). Neither condition is sufficient for unitarity: the shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has determinant 1 and the single eigenvalue 1, and it is not an isometry.

Invariant subspaces and the spectral form

For a normal operator, the orthogonal complement of an invariant subspace spanned by eigenvectors is invariant; this is what made the induction in the second proof of Theorem 8.2.5 work. For a unitary operator the same conclusion holds for every invariant subspace, with a proof that needs no eigenvectors. (By the spectral theorem it holds for normal operators on complex spaces too, but that is a consequence of the theorem, not a step towards it.)

> **Lemma 8.4.14** (Invariant complements). Let V be finite-dimensional, let $S \in \mathrm{L}(V)$ be unitary, and let U be an S-invariant subspace. Then $S(U) = U$, the complement $U^\perp$ is S-invariant, and $S|_U$ and $S|_{U^\perp}$ are unitary operators on U and on $U^\perp$.

Proof (directly). ROUTINE

> *Uses.* Corollary 4.2.18 (Injective if and only if surjective), Corollary 8.4.6 (Properties of isometries), Theorem 8.4.11 (Characterizations of unitary operators), Definition 4.2.1 (Restriction)
> *Hint 1.* $S|_U$ is an injective operator on the finite-dimensional space U. So every $\mathbf{u} \in U$ is $S\mathbf{u}'$ for some $\mathbf{u}' \in U$.
> *Hint 2.* For $\mathbf{w} \in U^\perp$, compute $\langle S\mathbf{w}, S\mathbf{u}'\rangle$.

Sketch. $S\restriction_U$ is injective, hence onto U. For $\mathbf{w} \in U^\perp$ and $\mathbf{u} = S\mathbf{u}' \in U$, $\langle S\mathbf{w}, \mathbf{u}\rangle = \langle S\mathbf{w}, S\mathbf{u}'\rangle = \langle \mathbf{w}, \mathbf{u}'\rangle = 0$. Both restrictions are isometries of finite-dimensional spaces into themselves, hence unitary.

Since S is injective, so is $S\restriction_U : U \to U$, and U is finite-dimensional, so $S\restriction_U$ is surjective (Corollary 4.2.18). That is, $S(U) = U$. Now let $\mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$. Write $\mathbf{u} = S\mathbf{u}'$ with $\mathbf{u}' \in U$; since S preserves inner products (Corollary 8.4.6(a)),

$$\langle S\mathbf{w}, \mathbf{u}\rangle = \langle S\mathbf{w}, S\mathbf{u}'\rangle = \langle \mathbf{w}, \mathbf{u}'\rangle = 0.$$

So $S\mathbf{w} \in U^\perp$, and $U^\perp$ is S-invariant. The restrictions of S to U and to $U^\perp$ preserve norms, so they are isometries of finite-dimensional spaces into themselves, hence unitary by Theorem 8.4.11. □

Proof (through the adjoint). ROUTINE

Uses. Proposition 8.4.10 (Unitary means $S^* = S^{-1}$), Exercise 6.6.38, Corollary 4.2.18 (Injective if and only if surjective), Theorem 8.4.11 (Characterizations of unitary operators)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. T maps U into U if and only if T^* maps $U^\perp$ into $U^\perp$ (Exercise 6.6.38). Apply this with $T = S^*$.
Hint 2. Why does $S^* = S^{-1}$ map U into U?
Sketch. $S(U) = U$ as before, so $S^{-1}(U) = U$, that is, U is invariant under $S^* = S^{-1}$. By Exercise 6.6.38 applied to $T = S^*$, $U^\perp$ is invariant under $(S^*)^* = S$.

As in the first proof, $S(U) = U$, so S^{-1} maps U onto U. By Proposition 8.4.10, $S^* = S^{-1}$, so U is S^*-invariant. By Exercise 6.6.38 applied to $T = S^*$, whose adjoint is S, the complement $U^\perp$ is S-invariant. The restrictions are unitary as in the first proof. □

The first proof uses only that S preserves inner products and maps U onto itself. The second places the lemma in the pattern of Section 6.6: invariance passes from T on U to T^* on $U^\perp$, and for unitary S the adjoint is the inverse, which has the same invariant subspaces. The lemma fails for isometries that are not onto: the shift M of Example 8.4.12 maps $U = \operatorname{span}(x, x^2, \ldots)$ (the polynomials with constant term 0) into itself, but $U^\perp = \operatorname{span}(1)$ is not M-invariant.

Over $\mathbb{C}$ every operator has an eigenvector, so the lemma splits off one eigenline at a time. Alternatively, unitary operators are normal, and the complex spectral theorem applies. Either way:

Theorem 8.4.15 (Spectral form of unitary operators). Let V be a finite-dimensional complex inner product space and $S \in \mathrm{L}(V)$. The following are equivalent.

(I) S is unitary.

(II) S is normal and every eigenvalue of S has absolute value 1.

(III) V has an orthonormal basis consisting of eigenvectors of S, and every eigenvalue of S has absolute value 1.

In matrix form: $U \in \mathbb{C}^{n\times n}$ is unitary if and only if $U = WDW^*$ for some unitary W and some $D = \operatorname{diag}\left(e^{i\theta_1}, \ldots, e^{i\theta_n}\right)$ with $\theta_1, \ldots, \theta_n \in \mathbb{R}$.

Proof (from the complex spectral theorem). ROUTINE

Uses. Proposition 8.4.10 (Unitary means $S^* = S^{-1}$), Corollary 8.4.13 (Eigenvalues and determinant), Theorem 8.2.5 (Complex spectral theorem), Theorem 8.4.11 (Characterizations of unitary operators), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Exercise 6.3.37, Example 6.3.35 (Unitary and orthogonal matrices)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Unitary operators satisfy $S^*S = SS^* = \mathrm{id}$, so they are normal. Which theorem gives orthonormal eigenbases of normal operators?
Hint 2. For (III) ⇒ (I), show that S maps the orthonormal eigenbasis to an orthonormal basis.
Sketch. (I) ⇒ (II): $S^*S = SS^* = \mathrm{id}$, and eigenvalues of isometries have absolute value 1. (II) ⇒ (III) is Theorem 8.2.5. (III) ⇒ (I): $\langle S\mathbf{e}_j, S\mathbf{e}_k\rangle = \lambda_j\overline{\lambda_k}\delta_{jk} = \delta_{jk}$, so S maps an orthonormal basis to one. For matrices, apply this to $\mathbf{x} \mapsto U\mathbf{x}$ and put the eigenbasis into the columns of W.

(I) ⇒ (II). By Proposition 8.4.10, $S^*S = \mathrm{id}_V = SS^*$, so S is normal, and its eigenvalues have absolute value 1 by Corollary 8.4.13(a).

(II) ⇒ (III). Theorem 8.2.5 gives an orthonormal eigenbasis, and the eigenvalues have absolute value 1 by hypothesis.

(III) ⇒ (I). Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis with $S\mathbf{e}_j = \lambda_j \mathbf{e}_j$ and $|\lambda_j| = 1$. Then

$$\langle S\mathbf{e}_j, S\mathbf{e}_k \rangle = \lambda_j \overline{\lambda_k} \langle \mathbf{e}_j, \mathbf{e}_k \rangle = \lambda_j \overline{\lambda_k} \delta_{jk} = |\lambda_j|^2 \delta_{jk} = \delta_{jk},$$

so S maps an orthonormal basis to an orthonormal basis, and S is unitary by (VI) ⇒ (II) of Theorem 8.4.11.

Matrices. If U is unitary, apply (I) ⇒ (III) to $\mathbf{x} \mapsto U\mathbf{x}$ on $\mathbb{C}^n$ with the dot product. Let W have the orthonormal eigenvectors as columns; it is unitary, and by Theorem 7.4.2, $U = WDW^{-1} = WDW^*$ with D diagonal, whose entries have absolute value 1, so they are $e^{i\theta_j}$ for some real θ_j. Conversely, D is unitary (Example 6.3.35(c)), and products of unitary matrices are unitary (Exercise 6.3.37), so WDW^* is unitary. □

Proof ((I) ⇒ (III) by induction, without the spectral theorem). STANDARD

> *Uses.* Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Corollary 8.4.13 (Eigenvalues and determinant), Lemma 8.4.14 (Invariant complements), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 6.4.24 (Orthogonal decomposition)
> *Techniques.* Induction on dimension via an invariant subspace (Technique 7.5.12)
> *Hint 1.* Over $\mathbb{C}$, S has an eigenvector. Normalize it.
> *Hint 2.* The eigenline is S-invariant. Apply Lemma 8.4.14 and induct on $\dim V$.
> *Sketch.* Take a unit eigenvector $\mathbf{e}$; $\mathrm{span}(\mathbf{e})$ is invariant, so its complement is invariant and S restricted to it is unitary (Lemma 8.4.14). By induction the complement has an orthonormal eigenbasis; add $\mathbf{e}$. Eigenvalues have absolute value 1 by Corollary 8.4.13(a).

We prove (I) ⇒ (III) by induction on $n = \dim V$; the other implications are as in the first proof. Every eigenvalue has absolute value 1 by Corollary 8.4.13(a), so we only need the orthonormal eigenbasis. If $n = 0$ the empty list works. Let $n \geq 1$. By Theorem 7.3.27, S has an eigenvector, and we scale it to a unit vector $\mathbf{e}$. The line $L = \mathrm{span}(\mathbf{e})$ is S-invariant, so by Lemma 8.4.14 $L^\perp$ is S-invariant and $S{\restriction}_{L^\perp}$ is unitary. Since $\dim L^\perp = n - 1$ (Corollary 6.4.25), the induction hypothesis gives an orthonormal basis $\mathbf{e}_2, \ldots, \mathbf{e}_n$ of $L^\perp$ consisting of eigenvectors of $S{\restriction}_{L^\perp}$, hence of S. Since $V = L \oplus L^\perp$ (Theorem 6.4.24), the list $\mathbf{e}, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is an orthonormal basis of V consisting of eigenvectors of S. □

The first proof inherits everything from normal operators. The second shows why the unitary case is easier to prove than the normal one: the invariance of the complement comes from Lemma 8.4.14, which needs neither eigenvectors nor adjoint identities. Over $\mathbb{R}$ the second proof breaks at the first step, because a rotation of $\mathbb{R}^2$ has no eigenvector. Repairing it with two-dimensional invariant subspaces gives the normal form of Section 8.4.4.

Example 8.4.16 (Two unitary diagonalizations).

(a) The rotation $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ with $\theta \notin \pi\mathbb{Z}$ has no real eigenvalue, but as a complex matrix

$$R_\theta \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} \cos\theta + i\sin\theta \\ \sin\theta - i\cos\theta \end{pmatrix} = e^{i\theta}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad R_\theta \begin{pmatrix} 1 \\ i \end{pmatrix} = e^{-i\theta}\begin{pmatrix} 1 \\ i \end{pmatrix}.$$

With $W = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, which is unitary, $R_\theta = W \operatorname{diag}\left(e^{i\theta}, e^{-i\theta}\right) W^*$. The same W works for every θ, which reflects the fact that rotations of the plane commute with each other.

(b) Let $C \in \mathbb{R}^{n\times n}$ be the cyclic shift, $C\mathbf{e}_k = \mathbf{e}_{k+1}$, with indices $0, \ldots, n-1$ read modulo n as in Definition 6.3.40. It is a permutation matrix, hence orthogonal, and $(C\mathbf{x})_j = x_{j-1}$. For the Fourier vector $\mathbf{f}_k = \left(\omega^{jk}\right)_{j=0}^{n-1}$,

$$(C\mathbf{f}_k)_j = \omega^{(j-1)k} = \omega^{-k}(\mathbf{f}_k)_j,$$

so $C\mathbf{f}_k = \omega^{-k}\mathbf{f}_k$. Since $\Phi_n/\sqrt{n}$ is unitary (Proposition 6.3.41), $C = WDW^*$ with $W = \Phi_n/\sqrt{n}$ and $D = \operatorname{diag}\left(1, \omega^{-1}, \ldots, \omega^{-(n-1)}\right)$. The eigenvalues of C are all the nth roots of unity. For $n = 3$, C is the rotation of $\mathbb{R}^3$ by $2\pi/3$ about the line through $(1,1,1)$ (Example 8.4.28).

Exercise 8.4.17. ROUTINE Find the matrix $U \in \mathbb{C}^{2\times 2}$ with $U\begin{pmatrix} 1 \\ 1 \end{pmatrix} = i\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $U\begin{pmatrix} 1 \\ -1 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Show that U is unitary, and compute U^4 and U^*.

Solution on page 1662

Remark 8.4.18 (The unitary group). The identity is unitary, and composites and inverses of unitary operators are unitary, by Corollary 8.4.6(b) and Proposition 8.4.10. So the unitary operators on V form a group under composition. For matrices, the **unitary group** $\mathrm{U}(n)$ consists of the unitary $n \times n$ matrices, the **orthogonal group** $\mathrm{O}(n)$ of the real orthogonal ones, and the **special orthogonal group** $\mathrm{SO}(n)$ of the orthogonal matrices with determinant 1; $\mathrm{U}(1)$ is the unit circle in $\mathbb{C}$. Groups of matrices are the subject of Section 14.1. Under the analogy of Table 8.1.1, $\mathrm{U}(1)$ is the set of z with $\bar{z}z = 1$, and just as every complex number is $z = e^{i\theta}|z|$, every operator on a finite-dimensional space is a unitary operator times a positive one (Definition 8.3.1), the **polar decomposition** of Section 9.2.

Unitary operators can move any orthonormal list to any other of the same length. This is the precise form of the statement that an inner product space "looks the same" in all directions.

Proposition 8.4.19 (Moving orthonormal lists). Let V be finite-dimensional.

(a) If $\mathbf{e}_1, \ldots, \mathbf{e}_k$ and $\mathbf{f}_1, \ldots, \mathbf{f}_k$ are orthonormal lists in V, there is a unitary $S \in \mathrm{L}(V)$ with $S\mathbf{e}_j = \mathbf{f}_j$ for $j = 1, \ldots, k$.

(b) If U is a subspace of V and $S_0 \in \mathrm{L}(U, V)$ is an isometry, there is a unitary $S \in \mathrm{L}(V)$ with $S|_U = S_0$.

Proof. ROUTINE

Uses. Theorem 6.4.20 (Orthonormal bases exist), Theorem 2.2.1 (Linear maps are determined by a basis), Theorem 8.4.11 (Characterizations of unitary operators), Theorem 8.4.4 (Characterizations of isometries)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Extend a basis of a subspace (Technique 4.1.18)
Hint 1. Extend both lists to orthonormal bases of V. How long are the extensions?
Hint 2. Define S on the first basis. For (b), apply (a) to an orthonormal basis of U and its image.
Sketch. Both lists extend to orthonormal bases of length $n = \dim V$; the linear map sending the first basis to the second, term by term, maps an orthonormal basis to one and is unitary. For (b), S_0 maps an orthonormal basis of U to an orthonormal list; extend as in (a), and two linear maps agreeing on a basis of U agree on U.

(a) By Theorem 6.4.20(b), the lists extend to orthonormal bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V, both of length $n = \dim V$. By Theorem 2.2.1 there is $S \in \mathrm{L}(V)$ with $S\mathbf{e}_j = \mathbf{f}_j$ for all $j \le n$. It maps an orthonormal basis to an orthonormal basis, so it is unitary by Theorem 8.4.11.

(b) Let $\mathbf{e}_1, \dots, \mathbf{e}_k$ be an orthonormal basis of U. By (IV) of Theorem 8.4.4, $S_0\mathbf{e}_1, \dots, S_0\mathbf{e}_k$ is orthonormal in V, so (a) gives a unitary S with $S\mathbf{e}_j = S_0\mathbf{e}_j$ for all j. The linear maps $S|_U$ and S_0 agree on a basis of U, so they are equal by the uniqueness in Theorem 2.2.1. □

In particular, any unit vector can be moved to any other by a unitary operator. When a problem is unchanged by unitary operators (it involves only lengths and inner products), we may therefore assume that a given unit vector is $\mathbf{e}_1$. Section 8.4.3 writes down an explicit operator that does this.

8.4.3 Householder reflections

The simplest unitary operators other than scalars are reflections across hyperplanes.

Example 8.4.20 (Householder reflections). Let $\mathbf{u} \in \mathbb{F}^n$ be a unit vector, and let

$$H_{\mathbf{u}} = I_n - 2\mathbf{u}\mathbf{u}^*.$$

Then $H_{\mathbf{u}}\mathbf{x} = \mathbf{x} - 2(\mathbf{u}^*\mathbf{x})\mathbf{u} = \mathbf{x} - 2(\mathbf{x}\cdot\mathbf{u})\mathbf{u}$. So $H_{\mathbf{u}}\mathbf{u} = \mathbf{u} - 2\mathbf{u} = -\mathbf{u}$, and $H_{\mathbf{u}}\mathbf{x} = \mathbf{x}$ for every $\mathbf{x} \perp \mathbf{u}$. In words, $H_{\mathbf{u}}$ fixes the hyperplane $\mathbf{u}^\perp$ and reverses the normal direction: it is the reflection across $\mathbf{u}^\perp$. By Theorem 6.4.31, $P = \mathbf{u}\mathbf{u}^*$ is the orthogonal projection onto $\operatorname{span}(\mathbf{u})$, and

$$H_{\mathbf{u}} = (I_n - P) - P,$$

which keeps the component of $\mathbf{x}$ in $\mathbf{u}^\perp$ and negates the component along $\mathbf{u}$. This is the operator $-(2P - I_n)$ of Exercise 8.1.30, so $H_{\mathbf{u}}$ is Hermitian and unitary, and $H_{\mathbf{u}}^2 = I_n$. Directly: $H_{\mathbf{u}}^* = I_n - 2(\mathbf{u}\mathbf{u}^*)^* = H_{\mathbf{u}}$, and

$$H_{\mathbf{u}}^2 = I_n - 4\mathbf{u}\mathbf{u}^* + 4\mathbf{u}(\mathbf{u}^*\mathbf{u})\mathbf{u}^* = I_n.$$

Extending $\mathbf{u}$ to an orthonormal basis $\mathbf{u}, \mathbf{u}_2, \dots, \mathbf{u}_n$ of eigenvectors, with eigenvalues $-1, 1, \dots, 1$, gives $\det H_{\mathbf{u}} = -1$ (Theorem 7.5.22).

For instance, for $\mathbf{u} = \frac{1}{\sqrt{3}}(1,-1,-1)$,

$$H_{\mathbf{u}} = I_3 - \frac{2}{3}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix},$$

the orthogonal matrix of Example 6.3.35(d). So that matrix is a reflection, across the plane $x_1 - x_2 - x_3 = 0$.

The point of Householder reflections is that they can move a given vector to any other vector of the same length.

> **Proposition 8.4.21** (A reflection exchanging two vectors). Let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ with $\mathbf{x} \neq \mathbf{y}$, $\|\mathbf{x}\| = \|\mathbf{y}\|$, and $\mathbf{x}\cdot\mathbf{y} \in \mathbb{R}$. Let $\mathbf{u} = (\mathbf{x}-\mathbf{y})/\|\mathbf{x}-\mathbf{y}\|$. Then
>
> $$H_{\mathbf{u}}\mathbf{x} = \mathbf{y} \quad \text{and} \quad H_{\mathbf{u}}\mathbf{y} = \mathbf{x}.$$

Proof (by computation). ROUTINE

> *Uses.* Example 8.4.20 (Householder reflections), Proposition 6.1.29 (Expanding a norm), Proposition 6.1.4 (Sesquilinearity)
> *Hint 1.* $H_{\mathbf{u}}\mathbf{x} = \mathbf{x} - 2(\mathbf{x}\cdot\mathbf{u})\mathbf{u}$. Write everything in terms of $\mathbf{d} = \mathbf{x} - \mathbf{y}$.
> *Hint 2.* You need $2(\mathbf{x}\cdot\mathbf{d}) = \|\mathbf{d}\|^2$. Expand both sides using $\|\mathbf{x}\| = \|\mathbf{y}\|$ and $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\mathbf{x}$.
> *Sketch.* With $\mathbf{d} = \mathbf{x} - \mathbf{y}$: $H_{\mathbf{u}}\mathbf{x} = \mathbf{x} - \big(2(\mathbf{x}\cdot\mathbf{d})/\|\mathbf{d}\|^2\big)\mathbf{d}$. Both $2(\mathbf{x}\cdot\mathbf{d})$ and $\|\mathbf{d}\|^2$ equal $2\|\mathbf{x}\|^2 - 2(\mathbf{x}\cdot\mathbf{y})$, so the coefficient is 1 and $H_{\mathbf{u}}\mathbf{x} = \mathbf{y}$. Then $H_{\mathbf{u}}\mathbf{y} = H_{\mathbf{u}}^2\mathbf{x} = \mathbf{x}$.

Let $\mathbf{d} = \mathbf{x} - \mathbf{y} \neq \mathbf{0}$, so $\mathbf{u} = \mathbf{d}/\|\mathbf{d}\|$ and

$$H_{\mathbf{u}}\mathbf{x} = \mathbf{x} - 2(\mathbf{x}\cdot\mathbf{u})\mathbf{u} = \mathbf{x} - \frac{2(\mathbf{x}\cdot\mathbf{d})}{\|\mathbf{d}\|^2}\mathbf{d}.$$

Since $\mathbf{x}\cdot\mathbf{y}$ is real, $\mathbf{y}\cdot\mathbf{x} = \overline{\mathbf{x}\cdot\mathbf{y}} = \mathbf{x}\cdot\mathbf{y}$. Using $\|\mathbf{y}\| = \|\mathbf{x}\|$,

$$\mathbf{x}\cdot\mathbf{d} = \|\mathbf{x}\|^2 - \mathbf{x}\cdot\mathbf{y},$$
$$\|\mathbf{d}\|^2 = \|\mathbf{x}\|^2 - \mathbf{x}\cdot\mathbf{y} - \mathbf{y}\cdot\mathbf{x} + \|\mathbf{y}\|^2 = 2\Big(\|\mathbf{x}\|^2 - \mathbf{x}\cdot\mathbf{y}\Big).$$

So $2(\mathbf{x}\cdot\mathbf{d}) = \|\mathbf{d}\|^2$, and $H_{\mathbf{u}}\mathbf{x} = \mathbf{x} - \mathbf{d} = \mathbf{y}$. Since $H_{\mathbf{u}}^2 = I_n$ (Example 8.4.20), $H_{\mathbf{u}}\mathbf{y} = H_{\mathbf{u}}^2\mathbf{x} = \mathbf{x}$. □

Proof (by decomposing along the mirror). ROUTINE

> *Uses.* Example 8.4.20 (Householder reflections), Proposition 6.1.4 (Sesquilinearity)
> *Techniques.* Orthogonal decomposition (Technique 6.2.4)
> *Hint 1.* Write $\mathbf{x} = \frac{1}{2}(\mathbf{x}+\mathbf{y}) + \frac{1}{2}(\mathbf{x}-\mathbf{y})$. Where does each piece lie relative to the mirror $\mathbf{u}^\perp$?
> *Hint 2.* Show that $\mathbf{x}+\mathbf{y} \perp \mathbf{x}-\mathbf{y}$; this is where both hypotheses are used.
> *Sketch.* $(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 + (\mathbf{y}\cdot\mathbf{x} - \mathbf{x}\cdot\mathbf{y}) = 0$, so $\mathbf{x}+\mathbf{y}$ lies in the mirror and is fixed, while $\mathbf{x}-\mathbf{y}$ is a multiple of $\mathbf{u}$ and is negated. Hence $H_{\mathbf{u}}\mathbf{x} = \frac{1}{2}(\mathbf{x}+\mathbf{y}) - \frac{1}{2}(\mathbf{x}-\mathbf{y}) = \mathbf{y}$.

First,

$$(\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = \|\mathbf{x}\|^2 - \mathbf{x}\cdot\mathbf{y} + \mathbf{y}\cdot\mathbf{x} - \|\mathbf{y}\|^2 = 0,$$

since $\|\mathbf{x}\| = \|\mathbf{y}\|$ and $\mathbf{y}\cdot\mathbf{x} = \overline{\mathbf{x}\cdot\mathbf{y}} = \mathbf{x}\cdot\mathbf{y}$. So $\mathbf{x}+\mathbf{y} \in \mathbf{u}^\perp$, and $H_{\mathbf{u}}$ fixes it, while $\mathbf{x}-\mathbf{y} = \|\mathbf{x}-\mathbf{y}\|\mathbf{u}$ is negated (Example 8.4.20). Hence

$$H_{\mathbf{u}}\mathbf{x} = H_{\mathbf{u}}\left(\frac{1}{2}(\mathbf{x}+\mathbf{y}) + \frac{1}{2}(\mathbf{x}-\mathbf{y})\right) = \frac{1}{2}(\mathbf{x}+\mathbf{y}) - \frac{1}{2}(\mathbf{x}-\mathbf{y}) = \mathbf{y},$$

and in the same way $H_{\mathbf{u}}\mathbf{y} = \frac{1}{2}(\mathbf{x}+\mathbf{y}) + \frac{1}{2}(\mathbf{x}-\mathbf{y}) = \mathbf{x}$. □

The second proof is the picture in Figure 8.4.1: the mirror is the perpendicular bisector of the segment from $\mathbf{x}$ to $\mathbf{y}$, which passes through $\mathbf{0}$ because $\mathbf{x}$ and $\mathbf{y}$ are equally long. The condition $\mathbf{x}\cdot\mathbf{y} \in \mathbb{R}$ cannot be dropped over $\mathbb{C}$: since $H_{\mathbf{u}}$ is Hermitian, $(H_{\mathbf{u}}\mathbf{x})\cdot\mathbf{x}$ is always real (Proposition 6.6.30), so no Householder reflection maps $\mathbf{x}$ to a $\mathbf{y}$ with $\mathbf{y}\cdot\mathbf{x} \notin \mathbb{R}$. Multiplying $\mathbf{y}$ by a suitable scalar of absolute value 1 first repairs this (??).

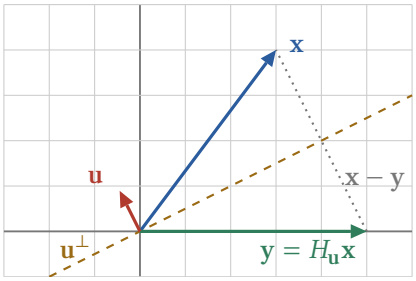


Figure 8.4.1. The Householder reflection with $\mathbf{x} = (3,4)$ and $\mathbf{y} = (5,0)$, both of length 5. The unit normal is $\mathbf{u} = (\mathbf{x}-\mathbf{y})/\|\mathbf{x}-\mathbf{y}\| = \frac{1}{\sqrt{5}}(-1,2)$, and the mirror $\mathbf{u}^\perp$ (dashed) is the line through $\mathbf{0}$ and $\mathbf{x}+\mathbf{y} = (8,4)$, the perpendicular bisector of the dotted segment.

Take $\mathbf{y} = -c\|\mathbf{x}\|\mathbf{e}_1$, where $c = x_1/|x_1|$ (and $c = 1$ if $x_1 = 0$). Then $\mathbf{x}\cdot\mathbf{y} = -\bar{c}x_1\|\mathbf{x}\| = -|x_1|\|\mathbf{x}\|$ is real, so unless $\mathbf{x} = \mathbf{y}$ already, one reflection turns $\mathbf{x}$ into $\mathbf{y}$, whose entries after the first are 0. Applied to the first column of a matrix $A \in \mathbb{F}^{m\times n}$, and then to the first column of the lower right block, and so on (skipping a step when the column is already in the required form), this produces a unitary $Q \in \mathbb{F}^{m\times m}$, a product of reflections, with $R_0 = Q^*A$ upper triangular. This is a full QR factorization $A = QR_0$; when the columns of A are independent, keeping the first n columns of Q and the top $n\times n$ block of R_0, and rescaling by scalars of absolute value 1 to make the diagonal positive, gives the factorization of Theorem 6.4.38. This is how QR factorizations are computed in practice; ?? carries it out for a 3×2 matrix.

Exercise 8.4.22. ROUTINE Find a Householder reflection H with $H(2,1,2) = (3,0,0)$, and one with $H(2,1,2) = (-3,0,0)$. Check that the first is symmetric and orthogonal.

Solution on page 1662

8.4.4 Real orthogonal operators BEYOND 21-242

The normal form of real orthogonal operators, orientation, and rigid motions are not covered in lecture; see LADW Chapters 5 and 6 and LADR 7D.

Over $\mathbb{R}$ an orthogonal operator need not have an eigenvector: a rotation of the plane has none. So the spectral form of Theorem 8.4.15 has no direct real analogue. Instead, we will show that every orthogonal operator is assembled from rotations of perpendicular planes and from ± 1 on perpendicular lines. We first list the orthogonal operators of the plane.

Lemma 8.4.23 (Orthogonal 2×2 matrices). Let $Q \in \mathbb{R}^{2\times 2}$ be orthogonal. Then there is exactly one $\theta \in [0, 2\pi)$ such that Q is one of

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},$$

and $\det R_\theta = 1$, $\det S_\theta = -1$. The matrix S_θ is the reflection across the line spanned by $\mathbf{c} = (\cos(\theta/2), \sin(\theta/2))$: it fixes $\mathbf{c}$ and negates $\mathbf{c}' = (-\sin(\theta/2), \cos(\theta/2))$, so it has the eigenvalues 1 and -1.

Proof. ROUTINE

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.18 (Coordinates in an orthonormal basis), Definition 5.1.14 (The 2×2 determinant)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. The first column is a unit vector, so it is $(\cos\theta, \sin\theta)$ for exactly one $\theta \in [0, 2\pi)$.
Hint 2. The second column is a unit vector orthogonal to the first. Expand it in the orthonormal basis $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$.
Sketch. The first column is $(\cos\theta, \sin\theta)$. The second column is orthogonal to it, so its coordinates in the orthonormal basis $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ are $(0, t)$, and being a unit vector forces $t = \pm 1$. The sign is the determinant. For S_θ, the angle-subtraction formulas give $S_\theta \mathbf{c} = \mathbf{c}$ and $S_\theta \mathbf{c}' = -\mathbf{c}'$.

The columns (a, b) and (c, d) of Q are orthonormal (Definition 6.3.33). Since $a^2 + b^2 = 1$, there is exactly one $\theta \in [0, 2\pi)$ with $(a, b) = (\cos\theta, \sin\theta)$. The vectors $(\cos\theta, \sin\theta)$ and $(-\sin\theta, \cos\theta)$ are orthonormal, hence an orthonormal basis of $\mathbb{R}^2$ (Corollary 6.3.14), so by Proposition 6.3.18

$$(c, d) = (c\cos\theta + d\sin\theta)(\cos\theta, \sin\theta) + t(-\sin\theta, \cos\theta), \quad t = -c\sin\theta + d\cos\theta.$$

The first coefficient is $(c, d) \cdot (a, b) = 0$, and $1 = \|(c, d)\| = |t|$, so $t = \pm 1$. For $t = 1$ we get $Q = R_\theta$ and for $t = -1$ we get $Q = S_\theta$; the determinants are $\cos^2\theta + \sin^2\theta = 1$ and $-\cos^2\theta - \sin^2\theta = -1$. Since θ is determined by the first column and the sign by the determinant, the description is unique.

Write $\varphi = \theta/2$. By the angle-subtraction formulas,

$$S_\theta \mathbf{c} = (\cos\theta\cos\varphi + \sin\theta\sin\varphi, \sin\theta\cos\varphi - \cos\theta\sin\varphi) = (\cos\varphi, \sin\varphi) = \mathbf{c},$$
$$S_\theta \mathbf{c}' = (-\cos\theta\sin\varphi + \sin\theta\cos\varphi, -\sin\theta\sin\varphi - \cos\theta\cos\varphi) = (\sin\varphi, -\cos\varphi) = -\mathbf{c}'. \qquad \square$$

So an orthogonal operator of the plane is a rotation if its determinant is 1 and a reflection across a line if it is -1. Rotations by 0 and π are $\pm I_2$, the only orthogonal 2×2 matrices that are scalars. In higher dimensions we need one more lemma, the block version of "a triangular unitary matrix is diagonal" (Exercise 6.3.57(c)).

Lemma 8.4.24 (Block-triangular unitary matrices). Let $R \in \mathbb{F}^{n\times n}$ be unitary and block upper triangular, with square diagonal blocks $B_1, \dots, B_m$. Then all blocks above the diagonal are 0, and each B_i is unitary.

Proof. STANDARD

Uses. Proposition 6.3.34 (Characterizations of unitary matrices), Definition 6.3.33 (Unitary and orthogonal matrices), Definition 2.4.33 (Block matrix)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Induction on matrix size (Technique 5.5.17)
Hint 1. Count the sum of $|r_{ij}|^2$ over the first block of columns, and then over the first block of rows.
Hint 2. Both the columns and the rows of a unitary matrix are orthonormal (Proposition 6.3.34). The first k columns are zero outside B_1.
Sketch. Let B_1 be $k \times k$. The first k columns are orthonormal and vanish below B_1, so the squared entries of B_1 add up to k. The first k rows are orthonormal, so the squared entries in those rows also add up to k. Hence every entry to the right of B_1 is 0. Then $R = \left(\begin{smallmatrix} B_1 & 0 \\ 0 & R' \end{smallmatrix}\right)$ with B_1 and R' unitary, and induction on m finishes.

We induct on m. If $m = 1$, then $R = B_1$. Let $m \geq 2$, let B_1 be $k \times k$, and write

$$R = \begin{pmatrix} B_1 & X \\ 0 & R' \end{pmatrix}$$

with $X \in \mathbb{F}^{k\times(n-k)}$ and R' block upper triangular with diagonal blocks $B_2, \dots, B_m$. The first k columns of R are orthonormal and have zero entries below B_1, so adding their squared lengths gives $\sum |b|^2 = k$, the sum over all entries b of B_1. By Proposition 6.3.34, the first k rows of R are orthonormal too, so the sum of $|r|^2$ over all entries r of these rows is also k. These rows consist of the entries of B_1 and of X, so

$$k = \sum_{b \text{ in } B_1} |b|^2 + \sum_{x \text{ in } X} |x|^2 = k + \sum_{x \text{ in } X} |x|^2.$$

A sum of nonnegative terms vanishes only if each term does, so $X = 0$. Now $R = \left(\begin{smallmatrix} B_1 & 0 \\ 0 & R' \end{smallmatrix}\right)$ and

$$I_n = R^* R = \begin{pmatrix} B_1^* B_1 & 0 \\ 0 & R'^* R' \end{pmatrix},$$

so B_1 and R' are unitary. By the induction hypothesis applied to R', its blocks above the diagonal vanish and $B_2, \dots, B_m$ are unitary. □

Theorem 8.4.25 (Normal form of orthogonal operators). Let V be a real inner product space with $1 \leq \dim V < \infty$, and let $S \in \mathrm{L}(V)$ be orthogonal. Then V has an orthonormal basis $\mathcal{O}$ in which $[S]_{\mathcal{O}}$ is block diagonal with diagonal blocks

$$R_{\theta_1}, \dots, R_{\theta_k}, I_p, -I_q,$$

where $k, p, q \geq 0$, $2k + p + q = \dim V$, and $0 < \theta_1 \leq \cdots \leq \theta_k < \pi$. Equivalently, every $Q \in \mathrm{O}(n)$ can be written as $Q = PBP^{\mathsf{T}}$ with P orthogonal and B block diagonal of this form.

Intuition. Read geometrically, the theorem says that V splits into perpendicular planes, each rotated by its own angle, and perpendicular lines, each fixed or reversed. Over $\mathbb{C}$ we could diagonalize; over $\mathbb{R}$ the planes are what is left of pairs of conjugate eigenvalues $e^{\pm i\theta_j}$. The two proofs below are the two routes to real block forms we have seen: triangularize first and then use orthogonality to clear the upper part, or split off one small invariant subspace at a time.

Proof (from the real Schur form). HARD

Uses. Theorem 7.5.27 (Real block upper-triangular form), Theorem 8.4.11 (Characterizations of unitary operators), Lemma 8.4.24 (Block-triangular unitary matrices), Lemma 8.4.23 (Orthogonal 2×2 matrices), Theorem 4.5.17 (Change of basis)

Techniques. Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. Theorem 7.5.27(b) gives an orthonormal basis in which $[S]_{\mathcal{O}}$ is block upper triangular with blocks of size 1 and 2. What does orthogonality of this matrix force?
Hint 2. By Lemma 8.4.24 the matrix is block diagonal with orthogonal blocks. Identify the 1×1 and 2×2 orthogonal blocks; recall that the 2×2 blocks have no real eigenvalue.
Hint 3. To move $\theta\in(\pi,2\pi)$ into $(0,\pi)$, replace one vector of the pair by its negative.
Sketch. The real Schur form gives an orthonormal $\mathcal{O}$ with $R=[S]_{\mathcal{O}}$ block upper triangular, diagonal blocks of size 1 or 2, the latter without real eigenvalues. R is orthogonal, so it is block diagonal with orthogonal blocks. A 1×1 block is ± 1; a 2×2 block without real eigenvalues is a rotation R_θ with $\theta\notin\{0,\pi\}$, since reflections have eigenvalues ± 1. Flipping one basis vector replaces R_θ by $R_{-\theta}$, and reordering the basis puts the blocks in order.

By Theorem 7.5.27(b), V has an orthonormal basis $\mathcal{O}$ such that $R=[S]_{\mathcal{O}}$ is block upper triangular with diagonal blocks of size 1×1 or 2×2, where no 2×2 diagonal block has a real eigenvalue. By Theorem 8.4.11, R is orthogonal, so by Lemma 8.4.24 it is block diagonal and each diagonal block is orthogonal.

A 1×1 orthogonal matrix (a) has $a^2=1$, so it is (1) or (-1). A 2×2 diagonal block B is orthogonal without real eigenvalues. By Lemma 8.4.23, it is R_θ or S_θ for some $\theta\in[0,2\pi)$; the reflection S_θ has the eigenvalue 1, and $R_0=I_2$, $R_\pi=-I_2$ have real eigenvalues, so $B=R_\theta$ with $\theta\in(0,\pi)\cup(\pi,2\pi)$. If $\theta\in(\pi,2\pi)$, let $\mathbf{a},\mathbf{b}$ be the two vectors of $\mathcal{O}$ belonging to this block, and replace $\mathbf{b}$ by $-\mathbf{b}$. The new list is still orthonormal, and by Theorem 4.5.17 the block becomes

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}R_\theta\begin{pmatrix}1&0\\0&-1\end{pmatrix}=\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}=R_{2\pi-\theta},$$

with $2\pi-\theta\in(0,\pi)$. Finally, reordering $\mathcal{O}$ (keeping the two vectors of each rotation block together) permutes the diagonal blocks, so we can list the rotation blocks first, in increasing order of angle, then the blocks (1), then the blocks (-1).

For the matrix form, apply this to $\mathbf{x}\mapsto Q\mathbf{x}$ on $\mathbb{R}^n$ with the dot product, and let P have the vectors of $\mathcal{O}$ as columns. Then P is orthogonal and $Q=PBP^{-1}=PBP^{\mathsf{T}}$ with $B=[S]_{\mathcal{O}}$. □

Proof (by induction on dimension). HARD

Uses. Theorem 7.3.31 (Small invariant subspaces), Lemma 8.4.14 (Invariant complements), Lemma 8.4.23 (Orthogonal 2×2 matrices), Corollary 8.4.13 (Eigenvalues and determinant), Theorem 6.4.24 (Orthogonal decomposition), Remark 7.2.5 (Invariant subspaces and block matrices)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Theorem 7.3.31 gives an S-invariant subspace U of dimension 1 or 2. What does Lemma 8.4.14 say about $U^\perp$?
Hint 2. On U, S is orthogonal of dimension 1 or 2. Use Lemma 8.4.23, and note that a reflection of a plane splits it into two eigenlines.
Sketch. Pick an invariant U with $\dim U\in\{1,2\}$. Then $U^\perp$ is invariant and both restrictions are orthogonal. On a line, $S=\pm\,\mathrm{id}$. On a plane, S is a rotation R_θ in an orthonormal basis (normalize θ as in the first proof; $\theta\in\{0,\pi\}$ gives $\pm I_2$) or a reflection, which is $\mathrm{diag}(1,-1)$ in the basis of its eigenvectors. Apply induction to $U^\perp$, concatenate, and reorder.

We induct on $n=\dim V$, proving the statement without the ordering of the blocks; reordering the basis as in the first proof then arranges them. By Theorem 7.3.31, S has an invariant subspace U with $\dim U\in\{1,2\}$. By Lemma 8.4.14, $U^\perp$ is S-invariant and $S{\restriction}_U$, $S{\restriction}_{U^\perp}$ are orthogonal.

The piece U. If $\dim U=1$, then U is spanned by an eigenvector, whose eigenvalue is real of absolute value 1 (Corollary 8.4.13(a)), so $S{\restriction}_U=\pm\mathrm{id}_U$, a block (1) or (-1) in a unit basis vector of U. If $\dim U=2$, let $\mathbf{a},\mathbf{b}$ be an orthonormal basis of U; the matrix of $S{\restriction}_U$ in it is orthogonal, so by Lemma 8.4.23 it is R_θ or S_θ. If it is S_θ, the unit eigenvectors $\mathbf{c},\mathbf{c}'$ of the lemma (taken as combinations of $\mathbf{a},\mathbf{b}$) form an orthonormal basis of U in which the matrix is $\mathrm{diag}(1,-1)$. If it is R_θ with $\theta\in\{0,\pi\}$, it is $\pm I_2$. Otherwise it is a rotation block, and replacing $\mathbf{b}$ by $-\mathbf{b}$ if necessary makes $\theta\in(0,\pi)$, as in the first proof.

The piece $U^\perp$. If $U=V$ we are done. Otherwise $1\le\dim U^\perp<n$, and the induction hypothesis gives an orthonormal basis of $U^\perp$ in which $S{\restriction}_{U^\perp}$ has the required block form.

By Theorem 6.4.24, $V = U \oplus U^\perp$, so the two bases together form an orthonormal basis of V. Both summands are invariant, so in this basis $[S]$ is block diagonal, with the blocks of $S|_U$ followed by those of $S|_{U^\perp}$ (Remark 7.2.5). □

The first proof spends its effort in the real Schur form and needs only bookkeeping afterwards; the second needs only a small invariant subspace and does the splitting by hand, using that invariant complements come for free for orthogonal operators. A third proof complexifies and pairs conjugate eigenvectors, as in Proposition 7.3.34; ?? carries it out.

The numbers in the normal form are determined by S. The complex eigenvalues of $[S]_{\mathcal{O}}$, counted with multiplicity, are $e^{\pm i\theta_1}, \dots, e^{\pm i\theta_k}$ together with 1 (p times) and -1 (q times), by Proposition 7.3.15, since R_θ has characteristic polynomial $x^2 - 2\cos\theta x + 1 = \big(x - e^{i\theta}\big)\big(x - e^{-i\theta}\big)$. In particular $\det S = (-1)^q$ (Proposition 5.3.17).

8.4.5 Orientation and rotations of $\mathbb{R}^3$ BEYOND 21-242

By Corollary 8.4.13(b), every orthogonal operator has determinant 1 or -1. The sign records whether S preserves orientation: whether it maps a right-handed orthonormal basis to a right-handed one, or reverses the handedness as a mirror does. In the plane, Lemma 8.4.23 says the two kinds are the rotations and the reflections, and the "F" in the picture below shows the difference.

Definition 8.4.26 (Rotation). Let V be a finite-dimensional real inner product space. An orthogonal operator $S \in \mathrm{L}(V)$ is a **rotation** (or **proper**) if $\det S = 1$, and **improper** if $\det S = -1$. The rotations of $\mathbb{R}^n$ form the group $\mathrm{SO}(n)$.

$$S = \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{pmatrix}, \quad \det S = -1, \quad S\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad S\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

The reflection $S = \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{pmatrix} = S_\theta$ with $\cos\theta = 0.6$, $\sin\theta = 0.8$. It fixes the line through $(2, 1)$ and reverses $(-1, 2)$, and $\det S = -1$: the "F" comes out mirrored. In the interactive version, drag $S\mathbf{e}_2$ to $(-0.8, 0.6)$ to get the rotation R_θ, with determinant 1 and an unmirrored "F". Every other position of $S\mathbf{e}_2$ leaves the orthogonal matrices. (Interactive in the web version.)

The determinant is a continuous function of the entries that takes only the values ± 1 on $\mathrm{O}(n)$, so no path of orthogonal matrices leads from a rotation to a reflection. Conversely, by Theorem 8.4.25 every rotation is PBP^{T} with an even number of blocks -1, which pair up into blocks $-I_2 = R_\pi$; letting all the angles shrink to 0 moves B continuously to I_n inside $\mathrm{SO}(n)$. This is the precise sense in which the rotations are the orthogonal operators that can be "performed physically".

In three dimensions a rotation has more structure than the word suggests: it always turns about a fixed axis.

Theorem 8.4.27 (Euler's rotation theorem). Let V be a real inner product space with $\dim V = 3$, and let $S \in \mathrm{L}(V)$ be a rotation. Then S has an **axis**: a unit vector $\mathbf{a}$ with $S\mathbf{a} = \mathbf{a}$. Moreover, for every orthonormal basis $\mathcal{O} = \mathbf{a}, \mathbf{b}, \mathbf{c}$ of V starting with $\mathbf{a}$,

$$[S]_{\mathcal{O}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

for some $\theta \in [0, 2\pi)$, and $\operatorname{tr} S = 1 + 2\cos\theta$.

Intuition. We need 1 to be an eigenvalue. Three different reasons are available: a determinant identity that says $\det(S - \mathrm{id})$ equals its own negative, a count of complex eigenvalues on the unit circle with product 1, and the normal form, in which an odd dimension leaves at least one 1×1 block. Once an axis exists, the plane perpendicular to it is invariant, and Lemma 8.4.23 identifies what happens there.

Proof (by a determinant identity). STANDARD

Uses. Theorem 8.4.11 (Characterizations of unitary operators), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Definition 5.1.1 (Determinant function), Theorem 7.2.9 (Eigenvalues and invertibility), Lemma 8.4.14 (Invariant complements), Lemma 8.4.23 (Orthogonal 2×2 matrices), Proposition 5.3.17 (Block-diagonal matrices)
Hint 1. Let A be the matrix of S in an orthonormal basis. Write $A - I_3 = A - AA^{\mathsf{T}}$ and factor out A.
Hint 2. $\det(A - I_3) = \det A \cdot \det\left(I_3 - A^{\mathsf{T}}\right) = \det(I_3 - A)$. Compare $\det(I_3 - A)$ with $\det(A - I_3)$ in dimension 3.
Hint 3. For the matrix, the plane $\mathbf{a}^{\perp}$ is invariant. What is the determinant of S on it?
Sketch. $\det(A - I) = \det\left(A\left(I - A^{\mathsf{T}}\right)\right) = \det A \det\left((I - A)^{\mathsf{T}}\right) = \det(I - A) = (-1)^3 \det(A - I)$, so $\det(A - I) = 0$ and 1 is an eigenvalue. With a unit fixed vector $\mathbf{a}$, $\mathbf{a}^{\perp}$ is invariant and S is orthogonal on it with determinant $\det S/1 = 1$, so it is a rotation R_θ of that plane. The trace is $1 + 2\cos\theta$.

An axis. Let $A = [S]_{\mathcal{B}}$ for some orthonormal basis $\mathcal{B}$; it is orthogonal with $\det A = \det S = 1$. Using $AA^{\mathsf{T}} = I_3$, Theorem 5.4.10, and Theorem 5.4.14,

$$\det(A - I_3) = \det\left(A - AA^{\mathsf{T}}\right) = \det(A)\det\left(I_3 - A^{\mathsf{T}}\right) = \det\left((I_3 - A)^{\mathsf{T}}\right) = \det(I_3 - A).$$

Since $I_3 - A = (-1)(A - I_3)$ and the determinant is multilinear in the three columns (Definition 5.1.1), $\det(I_3 - A) = (-1)^3 \det(A - I_3)$. So $\det(A - I_3) = -\det(A - I_3)$, that is, $\det(A - I_3) = 0$. Hence 1 is an eigenvalue of A, and so of S (Theorem 7.2.9); a unit eigenvector $\mathbf{a}$ is an axis.

The matrix. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be an orthonormal basis. The line $\operatorname{span}(\mathbf{a})$ is S-invariant, so by Lemma 8.4.14 the plane $P = \operatorname{span}(\mathbf{b}, \mathbf{c}) = \mathbf{a}^{\perp}$ is invariant and $S{\restriction}_P$ is orthogonal. So $[S]_{\mathcal{O}}$ is block diagonal with blocks (1) and the orthogonal matrix B of $S{\restriction}_P$ in $\mathbf{b}, \mathbf{c}$, and $1 = \det S = 1 \cdot \det B$ (Proposition 5.3.17). By Lemma 8.4.23, $B = R_\theta$ for some $\theta \in [0, 2\pi)$. Finally $\operatorname{tr} S = 1 + 2\cos\theta$. □

Proof (by counting eigenvalues). STANDARD

Uses. Theorem 7.1.32 (Fundamental Theorem of Algebra), Proposition B.1.4 (Roots of real polynomials come in conjugate pairs), Corollary 7.1.43 (Odd degree gives a real root), Theorem 7.5.22 (Trace and determinant from eigenvalues), Corollary 8.4.13 (Eigenvalues and determinant), Corollary 7.3.13 (Eigenvalues of an operator)
Techniques. Complexify, then take real and imaginary parts (Technique 7.3.35)
Hint 1. Regard the matrix A of S as a complex unitary matrix. Its three eigenvalues lie on the unit circle and multiply to $\det A = 1$.
Hint 2. The characteristic polynomial is real of degree 3: its nonreal roots come in conjugate pairs, and it has a real root.
Sketch. The eigenvalues $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ of A have absolute value 1 and product 1. If two are a conjugate pair, their product is $|\lambda|^2 = 1$, so the third is 1. If all are real, they are ± 1 with product 1, so they are not all -1. Either way 1 is a root of p_S, hence an eigenvalue of S.

We show that 1 is an eigenvalue; the rest is as in the first proof. Let A be the matrix of S in an orthonormal basis, regarded as a complex unitary matrix, and let $\lambda_1, \lambda_2, \lambda_3$ be its eigenvalues counted with multiplicity, the roots of p_A (Theorem 7.1.32). Each satisfies $|\lambda_j| = 1$ (Corollary 8.4.13(a)), and $\lambda_1\lambda_2\lambda_3 = \det A = 1$ (Theorem 7.5.22). Since p_A is a real polynomial of odd degree, it has a real root (Corollary 7.1.43), say $\lambda_3 \in \{1, -1\}$, and its nonreal roots come in conjugate pairs (Proposition B.1.4).

If $\lambda_1 \notin \mathbb{R}$, then $\lambda_2 = \overline{\lambda_1}$, so $\lambda_1\lambda_2 = |\lambda_1|^2 = 1$ and $\lambda_3 = 1$. If all three are real, each is ± 1, and their product is 1, so they are not all -1. In both cases 1 is a root of $p_A = p_S$, hence an eigenvalue of S (Corollary 7.3.13).□

Proof (from the normal form). ROUTINE

Uses. Theorem 8.4.25 (Normal form of orthogonal operators), Proposition 5.3.17 (Block-diagonal matrices)
Hint 1. In the normal form, $2k + p + q = 3$ and $\det S = (-1)^q$.
Hint 2. The number q of blocks -1 is even. Can p be 0?

Sketch. $\det S = (-1)^q = 1$ forces q even. If $k = 1$, then $p + q = 1$, so $q = 0$ and $p = 1$. If $k = 0$, then $p + q = 3$ with $q \in \{0, 2\}$, so $p \geq 1$. A block (1) is a fixed unit vector.

By Theorem 8.4.25, $[S]_{\mathcal{O}}$ is block diagonal with k rotation blocks, p blocks (1), and q blocks (-1), where $2k + p + q = 3$. By Proposition 5.3.17, $\det S = (-1)^q$ (each R_θ has determinant 1), so q is even. If $k = 1$, then $p + q = 1$ forces $q = 0$ and $p = 1$. If $k = 0$, then $p + q = 3$ with $q \in \{0, 2\}$, so $p \geq 1$. Either way some vector of $\mathcal{O}$ is fixed by S. The matrix and trace follow as in the first proof. □

The first proof is the shortest and shows what matters: in any odd dimension n, $\det(I_n - A) = (-1)^n \det(A - I_n)$, so every rotation of $\mathbb{R}^n$ with n odd has a fixed vector (Exercise 8.4.42). In even dimensions this fails: $R_\theta \in \mathrm{SO}(2)$ with $\theta \neq 0$, and $\begin{pmatrix} R_\theta & 0 \\ 0 & R_\varphi \end{pmatrix} \in \mathrm{SO}(4)$ with $\theta, \varphi \neq 0$, fix only $\mathbf{0}$. The second proof is a computation with the spectrum, and the third reads the answer off the classification.

The trace formula identifies a rotation of $\mathbb{R}^3$ quickly. Its axis is a kernel vector of $S - \mathrm{id}$, and the angle θ satisfies $\cos\theta = (\operatorname{tr} S - 1)/2$. The cosine determines θ only up to sign; replacing $\mathbf{b}, \mathbf{c}$ by $\mathbf{c}, \mathbf{b}$ replaces θ by $2\pi - \theta$, so the sense of the rotation depends on an orientation of the plane $\mathbf{a}^\perp$.

Example 8.4.28 (Axis and angle). Let

$$Q = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}.$$

Its columns have squared lengths $(4 + 4 + 1)/9 = 1$ and pairwise dot products $(-2 + 4 - 2)/9$, $(4 - 2 - 2)/9$, $(-2 - 2 + 4)/9$, all 0, so Q is orthogonal. Expanding along the first row, $\det(3Q) = 2(4 + 2) + (4 - 1) + 2(4 + 2) = 27$, so $\det Q = 1$ and Q is a rotation. Each row of Q sums to 1, so $Q(1, 1, 1) = (1, 1, 1)$, and the axis is $\mathbf{a} = \frac{1}{\sqrt{3}}(1, 1, 1)$. Since $\operatorname{tr} Q = 2 = 1 + 2\cos\theta$, we get $\cos\theta = 1/2$: Q turns the plane $x_1 + x_2 + x_3 = 0$ through $\pi/3$. For instance $Q(1, -1, 0) = (1, 0, -1)$, and these two vectors of length $\sqrt{2}$ have dot product $1 = \sqrt{2} \cdot \sqrt{2} \cdot \cos(\pi/3)$.

Squaring, $Q^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, the cyclic shift C of Example 8.4.16(b) for $n = 3$: it turns about the same axis through $2\pi/3$, and its trace is $0 = 1 + 2\cos(2\pi/3)$.

Since $\mathrm{SO}(3)$ is a group, the composite of two rotations of $\mathbb{R}^3$, about any two axes, is again a rotation about some axis, which is not obvious from the pictures (??).

8.4.6 Rigid motions BEYOND 21-242

We now drop linearity and ask for maps that preserve distances.

Definition 8.4.29 (Rigid motion). Let V be a real inner product space. A function $f : V \to V$, not assumed linear, is a **rigid motion** if

$$\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Translations $\mathbf{x} \mapsto \mathbf{x} + \mathbf{b}$ and orthogonal operators are rigid motions, and so are their composites $\mathbf{x} \mapsto S\mathbf{x} + \mathbf{b}$. These are affine maps (Definition 4.7.35) with orthogonal linear part. The theorem of this subsection is that there are no others. In Problem 6.8 the linearity of $f - f(\mathbf{0})$ was proved by polarization; here we give a different proof that uses distances only, through the observation that in an inner product space a point is determined by its distances to the two ends of a segment that it lies on.

Lemma 8.4.30 (Points between two points). Let V be an inner product space, let $\mathbf{a}, \mathbf{b}, \mathbf{z} \in V$, and let $t \in [0, 1]$. If

$$\|\mathbf{z}-\mathbf{a}\| = t\|\mathbf{b}-\mathbf{a}\| \quad \text{and} \quad \|\mathbf{b}-\mathbf{z}\| = (1-t)\|\mathbf{b}-\mathbf{a}\|,$$

then $\mathbf{z} = (1-t)\mathbf{a} + t\mathbf{b}$.

Proof. STANDARD

Uses. Theorem 6.2.11 (Triangle inequality), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. Let $\mathbf{x} = \mathbf{z}-\mathbf{a}$ and $\mathbf{y} = \mathbf{b}-\mathbf{z}$. Then $\mathbf{x}+\mathbf{y} = \mathbf{b}-\mathbf{a}$. What do the hypotheses say about $\|\mathbf{x}\| + \|\mathbf{y}\|$?
Hint 2. Equality holds in the triangle inequality, so one of $\mathbf{x}, \mathbf{y}$ is a nonnegative multiple of the other (Theorem 6.2.11(b)). Deduce that $\mathbf{x}$ is a nonnegative multiple of $\mathbf{b}-\mathbf{a}$, and find the factor from $\|\mathbf{x}\|$.
Sketch. With $\mathbf{x} = \mathbf{z}-\mathbf{a}$, $\mathbf{y} = \mathbf{b}-\mathbf{z}$: $\|\mathbf{x}\| + \|\mathbf{y}\| = \|\mathbf{x}+\mathbf{y}\|$, so $\mathbf{x}$ and $\mathbf{y}$ point the same way, and then $\mathbf{x} = s(\mathbf{b}-\mathbf{a})$ with $s \geq 0$. Taking norms, $s = t$ (if $\mathbf{b} \neq \mathbf{a}$; if $\mathbf{b} = \mathbf{a}$, then $\mathbf{z} = \mathbf{a}$ directly).

If $\mathbf{b} = \mathbf{a}$, the first hypothesis gives $\|\mathbf{z}-\mathbf{a}\| = 0$, so $\mathbf{z} = \mathbf{a} = (1-t)\mathbf{a} + t\mathbf{b}$. Suppose $\mathbf{b} \neq \mathbf{a}$, and let $\mathbf{x} = \mathbf{z}-\mathbf{a}$, $\mathbf{y} = \mathbf{b}-\mathbf{z}$, so that $\mathbf{x}+\mathbf{y} = \mathbf{b}-\mathbf{a}$ and

$$\|\mathbf{x}\| + \|\mathbf{y}\| = t\|\mathbf{b}-\mathbf{a}\| + (1-t)\|\mathbf{b}-\mathbf{a}\| = \|\mathbf{x}+\mathbf{y}\|.$$

By Theorem 6.2.11(b), one of $\mathbf{x}, \mathbf{y}$ is a nonnegative real multiple of the other. If $\mathbf{x} = c\mathbf{y}$ with $c \geq 0$, then $\mathbf{b}-\mathbf{a} = (c+1)\mathbf{y}$, so $\mathbf{x} = c/(c+1)(\mathbf{b}-\mathbf{a})$. If $\mathbf{y} = c\mathbf{x}$ with $c \geq 0$, then $\mathbf{b}-\mathbf{a} = (1+c)\mathbf{x}$, so $\mathbf{x} = 1/(1+c)(\mathbf{b}-\mathbf{a})$. Either way $\mathbf{x} = s(\mathbf{b}-\mathbf{a})$ for some real $s \geq 0$, and taking norms gives $s\|\mathbf{b}-\mathbf{a}\| = \|\mathbf{x}\| = t\|\mathbf{b}-\mathbf{a}\|$, so $s = t$. Hence $\mathbf{z} = \mathbf{a} + t(\mathbf{b}-\mathbf{a}) = (1-t)\mathbf{a} + t\mathbf{b}$. □

Theorem 8.4.31 (Distance-preserving maps are affine). Let V and W be real inner product spaces, and let $f : V \to W$ satisfy $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x}-\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in V$. Then $S(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$ defines an isometry $S \in \mathrm{L}(V, W)$, so $f = S + f(\mathbf{0})$ is affine.

In particular, if V is finite-dimensional, every rigid motion of V has the form $\mathbf{x} \mapsto S\mathbf{x} + \mathbf{b}$ with S orthogonal, and is bijective. The rigid motions of $\mathbb{R}^n$ are the maps $\mathbf{x} \mapsto Q\mathbf{x} + \mathbf{b}$ with $Q \in \mathrm{O}(n)$ and $\mathbf{b} \in \mathbb{R}^n$.

Proof (by betweenness). HARD

Uses. Lemma 8.4.30 (Points between two points), Theorem 2.1.3 (Checking linearity), Definition 4.7.35 (Affine map), Theorem 8.4.11 (Characterizations of unitary operators), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. Use Lemma 8.4.30 to show that $f((1-t)\mathbf{a} + t\mathbf{b}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b})$ for $t \in [0,1]$: the point on the left has the right distances to $f(\mathbf{a})$ and $f(\mathbf{b})$.
Hint 2. Extend to all real t: for $t > 1$, the point $\mathbf{b}$ lies between $\mathbf{a}$ and $(1-t)\mathbf{a} + t\mathbf{b}$.
Hint 3. Now $S(t\mathbf{x}) = f((1-t)\mathbf{0} + t\mathbf{x}) - f(\mathbf{0})$, and $\mathbf{x}+\mathbf{y}$ is twice the midpoint of $\mathbf{x}$ and $\mathbf{y}$.
Sketch. Distances from $f(\mathbf{z})$ to $f(\mathbf{a})$ and $f(\mathbf{b})$ equal those from $\mathbf{z} = (1-t)\mathbf{a} + t\mathbf{b}$ to $\mathbf{a}$ and $\mathbf{b}$, so Lemma 8.4.30 forces $f(\mathbf{z}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b})$ for $t \in [0,1]$, and a relabelling gives all $t \in \mathbb{R}$. Then $S = f - f(\mathbf{0})$ satisfies $S(t\mathbf{x}) = tS(\mathbf{x})$ and $S(\mathbf{x}+\mathbf{y}) = 2S((\mathbf{x}+\mathbf{y})/2) = S(\mathbf{x}) + S(\mathbf{y})$, and $\|S\mathbf{x}\| = \|\mathbf{x}\|$. In finite dimension S is orthogonal, hence bijective.

Step 1: segments go to segments. Let $\mathbf{a}, \mathbf{b} \in V$, $t \in [0,1]$, and $\mathbf{z} = (1-t)\mathbf{a} + t\mathbf{b}$. Then $\mathbf{z}-\mathbf{a} = t(\mathbf{b}-\mathbf{a})$ and $\mathbf{b}-\mathbf{z} = (1-t)(\mathbf{b}-\mathbf{a})$, so, as f preserves distances,

$$\|f(\mathbf{z}) - f(\mathbf{a})\| = t\|f(\mathbf{b}) - f(\mathbf{a})\|, \quad \|f(\mathbf{b}) - f(\mathbf{z})\| = (1-t)\|f(\mathbf{b}) - f(\mathbf{a})\|.$$

By Lemma 8.4.30 in W, $f(\mathbf{z}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b})$.

Step 2: lines go to lines. We show that $f((1-t)\mathbf{a} + t\mathbf{b}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b})$ for every $t \in \mathbb{R}$. Let $t > 1$ and $\mathbf{c} = (1-t)\mathbf{a} + t\mathbf{b}$. With $s = 1/t \in (0,1)$, a direct check gives $\mathbf{b} = (1-s)\mathbf{a} + s\mathbf{c}$, so Step 1 gives $f(\mathbf{b}) = (1-s)f(\mathbf{a}) + sf(\mathbf{c})$. Solving for $f(\mathbf{c})$,

$$f(\mathbf{c}) = tf(\mathbf{b}) - (t-1)f(\mathbf{a}) = (1-t)f(\mathbf{a}) + tf(\mathbf{b}).$$

For $t < 0$, write $(1-t)\mathbf{a} + t\mathbf{b} = (1-t')\mathbf{b} + t'\mathbf{a}$ with $t' = 1-t > 1$, and apply the case just proved with $\mathbf{a}$ and $\mathbf{b}$ exchanged.

Step 3: S is linear. Let $S(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$. For $t \in \mathbb{R}$, Step 2 with $\mathbf{a} = \mathbf{0}$, $\mathbf{b} = \mathbf{x}$ gives

$$S(t\mathbf{x}) = f((1-t)\mathbf{0} + t\mathbf{x}) - f(\mathbf{0}) = (1-t)f(\mathbf{0}) + tf(\mathbf{x}) - f(\mathbf{0}) = tS(\mathbf{x}).$$

For $\mathbf{x}, \mathbf{y} \in V$, Step 1 with $t = 1/2$ gives $f\left(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}\right) = \frac{1}{2}f(\mathbf{x}) + \frac{1}{2}f(\mathbf{y})$, so, using homogeneity,

$$S(\mathbf{x}+\mathbf{y}) = 2S\left(\frac{1}{2}(\mathbf{x}+\mathbf{y})\right) = 2\left(\frac{1}{2}f(\mathbf{x}) + \frac{1}{2}f(\mathbf{y}) - f(\mathbf{0})\right) = S(\mathbf{x}) + S(\mathbf{y}).$$

So S is linear (Theorem 2.1.3), and $\|S\mathbf{x}\| = \|f(\mathbf{x}) - f(\mathbf{0})\| = \|\mathbf{x} - \mathbf{0}\| = \|\mathbf{x}\|$, so S is an isometry. Thus $f = S + f(\mathbf{0})$ is affine (Definition 4.7.35).

Finite dimension. If $W = V$ is finite-dimensional, S is orthogonal by (I) $\Rightarrow$ (II) of Theorem 8.4.11, so it is bijective, and so is $f = S + \mathbf{b}$, with inverse $\mathbf{y} \mapsto S^{-1}(\mathbf{y} - \mathbf{b})$. For $V = \mathbb{R}^n$, S is $\mathbf{x} \mapsto Q\mathbf{x}$ for a matrix Q (Proposition 2.3.17), which is orthogonal by Theorem 8.4.11. □

The proof used the inner product only through Lemma 8.4.30, that is, through the equality case of the triangle inequality: the "midpoint" of $\mathbf{a}$ and $\mathbf{b}$, defined by distances alone, is unique. This is the idea behind the Mazur–Ulam theorem, which says that every surjective distance-preserving map between real normed spaces is affine. In a normed space such as $\mathbb{R}^2$ with $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|)$ there are many points at distance 1 from both $(0,0)$ and $(2,0)$, and Exercise 8.4.45 shows that the conclusion of Theorem 8.4.31 then fails for maps that are not onto. Over $\mathbb{C}$, Lemma 8.4.30 still holds and Steps 1–3 go through with real t, so a distance-preserving map of complex spaces is additive and real-homogeneous after subtracting $f(\mathbf{0})$; complex conjugation on $\mathbb{C}^n$ shows that it need not be complex-linear (Problem 6.8(d)).

In homogeneous coordinates (Section 4.7.6), the rigid motion $f(\mathbf{x}) = Q\mathbf{x} + \mathbf{b}$ of $\mathbb{R}^n$ is the matrix

$$M_f = \begin{pmatrix} Q & \mathbf{b} \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)},$$

and by Theorem 4.7.42 composites and inverses correspond to products and inverses:

$$\begin{pmatrix} Q_1 & \mathbf{b}_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} Q_2 & \mathbf{b}_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_1Q_2 & Q_1\mathbf{b}_2 + \mathbf{b}_1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} Q & \mathbf{b} \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} Q^{\mathsf{T}} & -Q^{\mathsf{T}}\mathbf{b} \\ 0 & 1 \end{pmatrix}.$$

So the rigid motions of $\mathbb{R}^n$ form a group, the **Euclidean group**. A rigid motion is **proper** if $\det Q = 1$; these are the motions of a solid body in space, and in the plane they are the translations and the rotations about a point (Exercise 4.7.62).

Exercises

Exercise 8.4.32. ROUTINE Decide which of the following linear maps are isometries, and which of the isometries are unitary.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = \frac{1}{3}(x + 2y, 2x + y, 2x - 2y)$.

(b) $T : \mathbb{C}^2 \to \mathbb{C}^2$, $T(z,w) = \frac{1}{\sqrt{2}}(z + iw, iz + w)$.

(c) $T : V \to V$, $(Tp)(x) = p(1-x)$, where $V = \mathbb{R}[x]_{\le 1}$ with $\langle p, q\rangle = p(0)q(0) + p(1)q(1)$.

(d) $T : V \to V$, $(Tp)(x) = p(x/2)$, with V as in (c).

(e) $T : \mathbb{C}^{2\times 2} \to \mathbb{C}^{2\times 2}$, $T(A) = A^{\mathsf{T}}$, with the Frobenius inner product.

Solution on page 1663

Exercise 8.4.33. STANDARD Prove or disprove each statement. Here V and W are finite-dimensional.

(a) If $S \in \mathrm{L}(V, W)$ is an isometry, then so is S^*.

(b) If V is complex and every eigenvalue of $T \in \mathrm{L}(V)$ has absolute value 1, then T is unitary.

(c) If V is complex, $T \in \mathrm{L}(V)$ is normal, and $\|T\mathbf{v}\| = \|\mathbf{v}\|$ for every eigenvector $\mathbf{v}$ of T, then T is unitary.

(d) Statement (c) with "complex" replaced by "real".

Solution on page 1663

Exercise 8.4.34. STANDARD Let V and W be finite-dimensional, and let $T \in \mathrm{L}(V, W)$ preserve orthogonality: $\langle T\mathbf{u}, T\mathbf{v}\rangle = 0$ whenever $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Show that either $T = 0$ or $T = cS$ for a real number $c > 0$ and an isometry S. (So the maps that preserve right angles are exactly the scalar multiples of isometries.)

Solution on page 1664

Exercise 8.4.35. STANDARD

(a) Show that an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ with integer entries is a **signed permutation matrix**: each row and each column has exactly one nonzero entry, which is 1 or -1.

(b) Count the orthogonal $n \times n$ matrices with integer entries, and count those with determinant 1 (for $n \geq 1$).

Solution on page 1664

Exercise 8.4.36. STANDARD Let $Q = \left(q_{ij}\right) \in \mathbb{R}^{n\times n}$ be orthogonal.

(a) Show that $\left|\sum_{i,j} q_{ij}\right| \leq n$, and that the sum equals n if and only if every row of Q sums to 1.

(b) Show that $\sum_{i,j}|q_{ij}| \leq n\sqrt{n}$, with equality if and only if every $|q_{ij}| = 1/\sqrt{n}$. Give an example of equality for $n = 2$.

Solution on page 1664

Exercise 8.4.37. STANDARD Let $U \in \mathbb{C}^{n\times n}$ be unitary. Show that $|\mathrm{tr}\, U| \leq n$, with equality if and only if $U = cI_n$ for some $c \in \mathbb{C}$ with $|c| = 1$. Deduce that a real orthogonal Q with $\mathrm{tr}\, Q = n$ is I_n.

Solution on page 1665

Exercise 8.4.38. HARD Let V be a finite-dimensional complex inner product space and $S \in \mathrm{L}(V)$ unitary.

(a) Show that there is a self-adjoint $A \in \mathrm{L}(V)$, with all eigenvalues in $[0, 2\pi)$, such that $S = f(A)$ for $f(t) = e^{it}$ (Definition 8.2.23).

(b) Show that S has a unitary square root: a unitary R with $R^2 = S$.

(c) Suppose $\dim V = n$ and S has n distinct eigenvalues. Show that S has exactly 2^n square roots $R \in \mathrm{L}(V)$, and that all of them are unitary.

Solution on page 1665

Exercise 8.4.39. STANDARD A student argues: "Let S be an orthogonal operator on a finite-dimensional real inner product space. Then $S^*S = SS^* = \mathrm{id}$, so S is normal, and by the spectral theorem V has an orthonormal basis of eigenvectors of S."

(a) Find the error, and give a counterexample to the conclusion.

(b) Show that an orthogonal S has an orthonormal eigenbasis if and only if $S^2 = \mathrm{id}$, and describe such operators geometrically.

Solution on page 1666

Exercise 8.4.40. STANDARD BEYOND 21-242 Let M be multiplication by x on $\mathbb{F}[x]$, with the inner product of Example 8.4.12, and define $L \in \mathrm{L}(\mathbb{F}[x])$ by

$$L\left(a_0 + a_1 x + \cdots + a_d x^d\right) = a_1 + a_2 x + \cdots + a_d x^{d-1}.$$

(a) Show that L is the adjoint of M.

(b) Show that $LM = \mathrm{id}$ and $ML = \mathrm{id} - \mathrm{Pr}_{\mathrm{span}(1)}$.

(c) Show that M has no eigenvalues, and that the only eigenvalue of L is 0.

Solution on page 1666

Exercise 8.4.41. ROUTINE BEYOND 21-242 Show that

$$Q = \frac{1}{9}\begin{pmatrix} 1 & -4 & 8 \\ 8 & 4 & 1 \\ -4 & 7 & 4 \end{pmatrix}$$

is a rotation of $\mathbb{R}^3$, and find its axis and its angle.

Solution on page 1667

Exercise 8.4.42. STANDARD BEYOND 21-242 Let $Q \in \mathrm{O}(n)$.

(a) Show that if $\det Q = -1$, then -1 is an eigenvalue of Q.

(b) Show that if $\det Q = 1$ and n is odd, then 1 is an eigenvalue of Q.

(c) Give a $Q \in \mathrm{SO}(4)$ with no real eigenvalue, and a $Q \in \mathrm{O}(2)$ with $\det Q = -1$ that does not have the eigenvalue 1, or show that none exists.

Solution on page 1667

Exercise 8.4.43. STANDARD For $\alpha \in \mathbb{R}$, let $S_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$, the reflection across the line at angle $\alpha/2$ (Example 6.3.35).

(a) Show that $S_\alpha S_\beta = R_{\alpha-\beta}$. In words: reflecting across a line ℓ_2 and then across a line ℓ_1 through $\mathbf{0}$ rotates the plane by twice the angle from ℓ_2 to ℓ_1.

(b) Deduce that every rotation R_θ is a product of two reflections, and that the first of them (the one applied first) can be any reflection.

Solution on page 1667

Exercise 8.4.44. CHALLENGE BEYOND 21-242 Let $Q \in \mathrm{O}(n)$ and $r = \mathrm{rank}(Q - I_n)$. A **reflection** of $\mathbb{R}^n$ means a Householder matrix $H_{\mathbf{u}}$.

(a) Show that if $Q = H_{\mathbf{u}_1} \cdots H_{\mathbf{u}_k}$, then $k \geq r$ and $\det Q = (-1)^k$.

(b) Show that Q is a product of exactly r reflections. (This is the Cartan–Dieudonné theorem for $\mathbb{R}^n$, in its sharp form.)

(c) Deduce that every rotation of $\mathbb{R}^3$ other than I_3 is a product of two reflections, and never of an odd number of reflections.

Solution on page 1668

Exercise 8.4.45. STANDARD BEYOND 21-242 On $\mathbb{R}^2$, let $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|)$, and let $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (t, |t|)$.

(a) Show that $\|f(s) - f(t)\|_\infty = |s - t|$ for all $s, t \in \mathbb{R}$, but that f is not affine.

(b) Which step of the proof of Theorem 8.4.31 fails for $\|\cdot\|_\infty$? Find all points $\mathbf{z} \in \mathbb{R}^2$ with $\|\mathbf{z} - f(-1)\|_\infty = \|\mathbf{z} - f(1)\|_\infty = 1$.

Solution on page 1668

Exercise 8.4.46. STANDARD BEYOND 21-242 Let f, g be rigid motions of $\mathbb{R}^n$, and let $\mathbf{p}_0, \ldots, \mathbf{p}_n$ be affinely independent (Definition 4.7.18) with $f(\mathbf{p}_i) = g(\mathbf{p}_i)$ for all i. Show that $f = g$. Deduce that a rigid motion of $\mathbb{R}^2$ fixing three non-collinear points is the identity.

Solution on page 1669

8.5 Proof patterns

Chapter 9

The singular value decomposition

Singular values, the SVD and its geometry, its consequences, and a survey of matrix factorizations.

9.1 Singular values and the SVD

9.2 Consequences of the SVD BEYOND 21-242

9.3 Matrix factorizations

9.4 Proof patterns

Chapter 10

Generalized eigenspaces and Jordan form

Generalized eigenvectors, nilpotent operators, the generalized eigenspace decomposition, and Jordan canonical form.

10.1 Generalized eigenvectors and nilpotent operators

10.2 The generalized eigenspace decomposition

10.3 Jordan canonical form

10.4 Proof patterns

Chapter 11

Markov chains and positive matrices

Stochastic matrices, Markov chains and stationary distributions, Perron–Frobenius, and reversibility.

11.1 Stochastic matrices and Markov chains

11.2 Perron–Frobenius

11.3 Reversibility

11.4 Proof patterns

Chapter 12

Bilinear and quadratic forms

Bilinear forms and congruence, quadratic forms, positive definiteness and min–max, and forms over arbitrary fields.

12.1 Bilinear forms BEYOND 21-242

12.2 Quadratic forms

12.3 Positive definiteness and min–max

12.4 Inner products beyond $\mathbb{R}$ and $\mathbb{C}$ BEYOND 21-242

12.5 Proof patterns

Chapter 13

Duality and multilinear algebra

Dual spaces and dual maps, duality for inner product spaces, tensors, and the determinant revisited through alternating forms.

13.1 Dual spaces BEYOND 21-242

13.2 Duality in inner product spaces BEYOND 21-242

13.3 Multilinear maps and tensors BEYOND 21-242

13.4 Alternating forms and the determinant revisited BEYOND 21-242

13.5 Proof patterns BEYOND 21-242

Chapter 14

Linear algebra and abstract algebra

Groups through matrices, field extensions, finite fields, and modules over F[x] with rational canonical form.

14.1 Groups through matrices BEYOND 21-242

14.2 Fields and field extensions BEYOND 21-242

14.3 Finite fields BEYOND 21-242

14.4 Modules over F[x] and canonical forms BEYOND 21-242

14.5 Proof patterns BEYOND 21-242

Chapter 15

Infinite-dimensional vector spaces

Hamel bases, linear algebra inside infinite-dimensional spaces, and a first look at inner product spaces of infinite dimension.

15.1 Bases in infinite dimensions BEYOND 21-242

15.2 Linear algebra inside infinite-dimensional spaces BEYOND 21-242

15.3 Inner product spaces in infinite dimensions BEYOND 21-242

15.4 Proof patterns BEYOND 21-242

Chapter 16

Applications to data and machine learning

Regression, principal component analysis, low-rank models, quadratic optimization, spectral graph methods, and linear layers.

16.1 Regression, revisited BEYOND 21-242

16.2 Principal component analysis BEYOND 21-242

16.3 Low-rank models BEYOND 21-242

16.4 Optimization of quadratic objectives BEYOND 21-242

16.5 Graphs and spectral methods BEYOND 21-242

16.6 Linear layers BEYOND 21-242

Appendix A

Sets, functions, and proofs

Functions and inverses, induction, equivalence relations, Zorn's lemma, and an index of proof techniques.

A.1 Sets and functions

This appendix collects the set-theoretic language that the book uses without comment: functions and their inverses, finite sets, induction, equivalence relations, and Zorn's lemma. The main text cites the results here where they are needed. Several of them have linear counterparts in the main text (one-sided inverses of linear maps in Section 2.5, injectivity and surjectivity in Section 2.6, quotient spaces in Section 4.6), and comparing the two versions shows which parts of a linear-algebra argument are really about linearity.

A.1.1 Sets

Notation A.1.1 (Sets). We write $x \in X$ when x is an element of the set X, and $A \subseteq X$ when every element of A is an element of X. Two sets are equal when they have the same elements, so $A = B$ means $A \subseteq B$ and $B \subseteq A$ (Technique 1.3.14). For subsets $A, B \subseteq X$,

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}, \quad A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\}, \quad X \setminus A = \{x \in X \mid x \notin A\}.$$

The empty set is $\emptyset$, and A and B are **disjoint** if $A \cap B = \emptyset$. The **power set** $P(X)$ is the set of all subsets of X. The **Cartesian product** $X \times Y$ is the set of ordered pairs (x, y) with $x \in X$ and $y \in Y$, and X^n is the set of n-tuples of elements of X. For an integer $n \geq 0$, we write $[n] = \{1, 2, \ldots, n\}$, so that $[0] = \emptyset$. Finally, $\mathbb{N} = \{0, 1, 2, \ldots\}$ includes 0.

A.1.2 Functions, images, and preimages

Definition A.1.2 (Function). A **function** (or **map**) $f : X \to Y$ assigns to each element $x \in X$ exactly one element $f(x) \in Y$. The set X is the **domain** of f and Y is its **codomain**. Two functions $f, g : X \to Y$ are **equal** if $f(x) = g(x)$ for every $x \in X$. The **identity** of X is $\mathrm{id}_X : X \to X$, $\mathrm{id}_X(x) = x$.

Formally, f is its graph $\{(x, f(x)) \mid x \in X\} \subseteq X \times Y$, a subset that contains exactly one pair (x, y) for each $x \in X$. The codomain is part of the data: $x \mapsto x^2$ from $\mathbb{R}$ to $\mathbb{R}$ and from $\mathbb{R}$ to $[0, \infty)$ are different functions, since the second is surjective and the first is not. For every set Y there is exactly one function $\emptyset \to Y$, the empty function, and there is no function $X \to \emptyset$ when $X \neq \emptyset$.

Definition A.1.3 (Image and preimage). Let $f : X \to Y$. For $A \subseteq X$, the **image** of A under f is

$$\mathrm{im}(f|_A) = \{f(a) \mid a \in A\} \subseteq Y,$$

and the image of f is $\operatorname{im} f = \operatorname{im}(f{\restriction_X})$. For $B \subseteq Y$, the **preimage** of B under f is

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\} \subseteq X.$$

For $y \in Y$ we write $f^{-1}(y)$ for $f^{-1}(\{y\})$, the set of preimages of y, and call it the **fiber** of f over y.

The notation $\operatorname{im}(f{\restriction_A})$ reads "the image of f restricted to A", which is what it is. The notation $f^{-1}(B)$ does not presuppose that f has an inverse: every function has preimages, and a fiber $f^{-1}(y)$ can be empty or have many elements. In this language, Definition 2.6.1, Definition 2.6.2 and Definition 2.6.3 say that f is injective, surjective, or bijective when every fiber has at most one, at least one, or exactly one element. For a linear map T, the kernel is the fiber $T^{-1}(\mathbf{0})$, and the solution set of $T(\mathbf{v}) = \mathbf{w}$ is the fiber $T^{-1}(\mathbf{w})$.

Example A.1.4 (Images and preimages of $x \mapsto x^2$). Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Then $\operatorname{im}\big(f{\restriction_{[-1,2]}}\big) = [0, 4]$ and $\operatorname{im} f = [0, \infty)$. Preimages: $f^{-1}([1, 4]) = [-2, -1] \cup [1, 2]$, $f^{-1}(4) = \{-2, 2\}$, and $f^{-1}(-1) = \emptyset$. Taking the image and then the preimage can enlarge a set: $f^{-1}\big(\operatorname{im}\big(f{\restriction_{[0,1]}}\big)\big) = f^{-1}([0, 1]) = [-1, 1]$. Taking the preimage and then the image can shrink one: $\operatorname{im}\big(f{\restriction_{f^{-1}([-1,1])}}\big) = [0, 1]$.

The example shows the general pattern. Preimages respect all the set operations, images respect unions but not intersections, and the two operations undo each other exactly when f is injective (in one order) or surjective (in the other).

Proposition A.1.5 (Images and preimages). Let $f : X \to Y$, let $A, A' \subseteq X$, and let $B, B' \subseteq Y$.

(a) $f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$, $f^{-1}(B \cap B') = f^{-1}(B) \cap f^{-1}(B')$, and $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$.

(b) $\operatorname{im}(f{\restriction_{A \cup A'}}) = \operatorname{im}(f{\restriction_A}) \cup \operatorname{im}(f{\restriction_{A'}})$ and $\operatorname{im}(f{\restriction_{A \cap A'}}) \subseteq \operatorname{im}(f{\restriction_A}) \cap \operatorname{im}(f{\restriction_{A'}})$, with equality if f is injective.

(c) $A \subseteq f^{-1}(\operatorname{im}(f{\restriction_A}))$, with equality if f is injective.

(d) $\operatorname{im}\big(f{\restriction_{f^{-1}(B)}}\big) = B \cap \operatorname{im} f$. In particular $\operatorname{im}\big(f{\restriction_{f^{-1}(B)}}\big) \subseteq B$, with equality if f is surjective.

Proof. ROUTINE

Uses. Definition A.1.3 (Image and preimage), Definition 2.6.1 (Injective), Definition 2.6.2 (Surjective)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Each claim is an equality or containment of sets. Take a generic element of one side and unwind the definitions until it is visibly in the other side.
Hint 2. For preimages, $x \in f^{-1}(B)$ is the statement $f(x) \in B$, so (a) is the logic of "or", "and", and "not" applied to $f(x)$. For images, $y \in \operatorname{im}(f{\restriction_A})$ means $y = f(a)$ for some $a \in A$.
Hint 3. Injectivity enters when an element is written as a value in two ways: if $y = f(a) = f(a')$ with $a \in A$ and $a' \in A'$, then $a = a' \in A \cap A'$.
Sketch. (a) unwinds to $f(x) \in B \cup B'$ iff $f(x) \in B$ or $f(x) \in B'$, and similarly for $\cap$ and complements. In (b), $f(a)$ with $a \in A \cup A'$ lies in one of the two images; for the intersection, a common value $f(a) = f(a')$ forces $a = a'$ when f is injective. (c): $a \in A$ has $f(a) \in \operatorname{im}(f{\restriction_A})$; if $f(x) = f(a)$ and f is injective, then $x = a$. (d): the values $f(x)$ with $f(x) \in B$ are exactly the elements of B that are values.

(a) For $x \in X$, $x \in f^{-1}(B \cup B')$ means $f(x) \in B \cup B'$, that is, $f(x) \in B$ or $f(x) \in B'$, that is, $x \in f^{-1}(B)$ or $x \in f^{-1}(B')$. The same unwinding with "and" proves the second equality. For the third, $x \in f^{-1}(Y \setminus B)$ means $f(x) \notin B$, that is, $x \notin f^{-1}(B)$.

(b) An element of $\operatorname{im}(f{\restriction_{A \cup A'}})$ is $f(a)$ with $a \in A$ or $a \in A'$, so it lies in $\operatorname{im}(f{\restriction_A})$ or in $\operatorname{im}(f{\restriction_{A'}})$; conversely, $\operatorname{im}(f{\restriction_A})$ and $\operatorname{im}(f{\restriction_{A'}})$ are both contained in $\operatorname{im}(f{\restriction_{A \cup A'}})$. If $a \in A \cap A'$, then $f(a)$ lies in both images, which proves the containment. Now suppose f is injective and $y \in \operatorname{im}(f{\restriction_A}) \cap \operatorname{im}(f{\restriction_{A'}})$, say $y = f(a) = f(a')$ with $a \in A$ and $a' \in A'$. Then $a = a'$, so $a \in A \cap A'$ and $y \in \operatorname{im}(f{\restriction_{A \cap A'}})$.

(c) If $a \in A$, then $f(a) \in \operatorname{im}(f{\restriction_A})$, so $a \in f^{-1}(\operatorname{im}(f{\restriction_A}))$. Suppose f is injective and $x \in f^{-1}(\operatorname{im}(f{\restriction_A}))$. Then $f(x) = f(a)$ for some $a \in A$, so $x = a \in A$.

(d) An element of $\operatorname{im}\big(f\restriction_{f^{-1}(B)}\big)$ is $f(x)$ with $f(x) \in B$, so it lies in $B \cap \operatorname{im} f$. Conversely, if $y \in B \cap \operatorname{im} f$, write $y = f(x)$; then $x \in f^{-1}(B)$ and $y \in \operatorname{im}\big(f\restriction_{f^{-1}(B)}\big)$. If f is surjective, then $\operatorname{im} f = Y$ and $B \cap \operatorname{im} f = B$.□

The hypotheses in (b), (c), and (d) cannot be dropped. For $f(x) = x^2$, the sets $A = \{-1\}$ and $A' = \{1\}$ are disjoint, so $\operatorname{im}(f\restriction_{A\cap A'}) = \varnothing$, while $\operatorname{im}(f\restriction_A) \cap \operatorname{im}(f\restriction_{A'}) = \{1\}$; Example A.1.4 does the same for (c) and (d). Exercise A.1.24 shows that injectivity is also necessary for equality in (b).

A.1.3 Composition and inverses

Definition A.1.6 (Composition). For $f : X \to Y$ and $g : Y \to Z$, the **composition** $g \circ f : X \to Z$ is the function $(g \circ f)(x) = g(f(x))$.

Composition is defined only when the codomain of f is the domain of g, and it is read right to left: $g \circ f$ applies f first. Even when both $g \circ f$ and $f \circ g$ are defined, they are usually different (Example 2.1.19 shows this for linear maps).

Proposition A.1.7 (Associativity and identities). Let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$. Then $h \circ (g \circ f) = (h \circ g) \circ f$, and $f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f$.

Proof. ROUTINE

Uses. Definition A.1.6 (Composition), Definition A.1.2 (Function)
Hint 1. Two functions $X \to W$ are equal when they agree at every $x \in X$.
Hint 2. Evaluate both sides at x and apply the definition of composition twice on each side.
Sketch. Both sides of the first identity send x to $h(g(f(x)))$. Both $f \circ \mathrm{id}_X$ and $\mathrm{id}_Y \circ f$ send x to $f(x)$.

Both sides are functions $X \to W$. For $x \in X$,

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x).$$

Similarly $(f \circ \mathrm{id}_X)(x) = f(x) = \mathrm{id}_Y(f(x)) = (\mathrm{id}_Y \circ f)(x)$. □

So we may write $h \circ g \circ f$ without parentheses. The next proposition is used constantly, and its last two parts are easy to get backwards. If $g \circ f$ is injective, the function applied *first* must be injective, since a collision $f(x) = f(x')$ would survive into $g(f(x)) = g(f(x'))$. If $g \circ f$ is surjective, the function applied *last* must be surjective, since every value of $g \circ f$ is a value of g.

Proposition A.1.8 (Composites, injectivity, and surjectivity). Let $f : X \to Y$ and $g : Y \to Z$.

(a) If f and g are injective, so is $g \circ f$. If f and g are surjective, so is $g \circ f$.

(b) If $g \circ f$ is injective, then f is injective.

(c) If $g \circ f$ is surjective, then g is surjective.

Proof. ROUTINE

Uses. Definition 2.6.1 (Injective), Definition 2.6.2 (Surjective), Definition A.1.6 (Composition)
Hint 1. For injectivity, start from $(g \circ f)(x) = (g \circ f)(x')$, or from $f(x) = f(x')$, and see what follows.
Hint 2. For (b), apply g to both sides of $f(x) = f(x')$. For (c), every value of $g \circ f$ is g of something.
Sketch. (a) Peel off g, then f: $g(f(x)) = g(f(x'))$ gives $f(x) = f(x')$, then $x = x'$. For surjectivity, lift z through g, then through f. (b) $f(x) = f(x')$ gives $g(f(x)) = g(f(x'))$, so $x = x'$. (c) $z = g(f(x))$ is a value of g.

(a) Suppose f and g are injective, and $g(f(x)) = g(f(x'))$. Injectivity of g gives $f(x) = f(x')$, and injectivity of f gives $x = x'$. Suppose f and g are surjective, and let $z \in Z$. Then $z = g(y)$ for some $y \in Y$, and $y = f(x)$ for some $x \in X$, so $z = (g \circ f)(x)$.

(b) If $f(x) = f(x')$, then $(g \circ f)(x) = g(f(x)) = g(f(x')) = (g \circ f)(x')$, so $x = x'$ because $g \circ f$ is injective.

(c) Let $z \in Z$. Since $g \circ f$ is surjective, $z = g(f(x))$ for some $x \in X$, so z is a value of g. □

Definition A.1.9 (One-sided inverses of a function). Let $f : X \to Y$. A function $g : Y \to X$ is a **left inverse** of f if $g \circ f = \mathrm{id}_X$, and a **right inverse** of f if $f \circ g = \mathrm{id}_Y$.

Definition A.1.10 (Inverse function). A function $g : Y \to X$ is an **inverse** of $f : X \to Y$ if it is both a left and a right inverse of f, that is, $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. If f has an inverse, f is **invertible**.

These are Definition 2.5.1, Definition 2.5.2 and Definition 2.5.12 without linearity. The definitions in Section 2.5 ask for more: there, a left inverse of a linear map must itself be linear. Remark 2.5.8 explains why that matters for one-sided inverses, and Proposition 2.7.7 shows that it does not matter for two-sided ones.

Example A.1.11 (The successor function). Let $s : \mathbb{N} \to \mathbb{N}$, $s(n) = n + 1$. For any $c \in \mathbb{N}$, the function

$$g_c(n) = \begin{cases} n - 1 & \text{if } n \geq 1 \\ c & \text{if } n = 0 \end{cases}$$

satisfies $g_c(s(n)) = n$ for all n, so every g_c is a left inverse of s. The value at 0 is free because 0 is not a value of s. On the other hand, s has no right inverse: $s(g(0)) = 0$ is impossible for every function g, since $0 \notin \operatorname{im} s$. Multiplication by x on $F[x]$ (Example 2.6.4(e)) is the linear version of s: it shifts coefficients up by one place.

Theorem A.1.12 (One-sided inverses of functions). Let $f : X \to Y$.

(a) If f has a left inverse, then f is injective. Conversely, if f is injective and $X \neq \emptyset$, then f has a left inverse.

(b) f has a right inverse if and only if f is surjective.

Intuition. A left inverse g must send each value $f(x)$ back to x. This is possible exactly when no value comes from two different inputs, and then g is forced on $\operatorname{im} f$ and free elsewhere, which is where $X \neq \emptyset$ is needed: g needs somewhere to send the points outside $\operatorname{im} f$. A right inverse must choose, for each $y \in Y$, some preimage of y. This is possible exactly when every y has one, but there may be many, and g chooses one in each fiber.

Proof. ROUTINE

Uses. Definition A.1.9 (One-sided inverses of a function), Definition 2.6.1 (Injective), Definition 2.6.2 (Surjective), Remark A.1.13 (The axiom of choice)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. The two "if" directions are one line each: apply g to $f(x) = f(x')$, or apply f to $g(y)$.
Hint 2. For the converse in (a), define g on $\operatorname{im} f$ by sending $f(x)$ to x. Why is this a function there? What should g do off $\operatorname{im} f$?
Hint 3. For the converse in (b), each fiber $f^{-1}(y)$ is nonempty. Let $g(y)$ be an element of it.
Sketch. If $g \circ f = \mathrm{id}_X$, then $f(x) = f(x')$ gives $x = g(f(x)) = g(f(x')) = x'$. If f is injective, each $y \in \operatorname{im} f$ has a unique preimage; send it there, and send every $y \notin \operatorname{im} f$ to a fixed $x_0 \in X$. If $f \circ g = \mathrm{id}_Y$, then $y = f(g(y))$ is a value. If f is surjective, choose $g(y) \in f^{-1}(y)$ for each y; then $f(g(y)) = y$.

(a) Suppose $g \circ f = \mathrm{id}_X$, and $f(x) = f(x')$. Then $x = g(f(x)) = g(f(x')) = x'$, so f is injective.

Conversely, suppose f is injective, and fix some $x_0 \in X$. Each $y \in \operatorname{im} f$ has exactly one preimage, which we call x_y. Define $g : Y \to X$ by $g(y) = x_y$ if $y \in \operatorname{im} f$ and $g(y) = x_0$ otherwise. For every $x \in X$, x is the preimage of $f(x) \in \operatorname{im} f$, so $g(f(x)) = x$. Hence $g \circ f = \mathrm{id}_X$.

(b) If $f \circ g = \mathrm{id}_Y$, then every $y \in Y$ equals $f(g(y))$, so f is surjective. Conversely, suppose f is surjective. Then every fiber $f^{-1}(y)$ is nonempty, and we let $g(y)$ be an element of $f^{-1}(y)$ for each $y \in Y$ (Remark A.1.13). Then $f(g(y)) = y$ for all y, so $f \circ g = \mathrm{id}_Y$. □

Remark A.1.13 (The axiom of choice). The last step of the proof chooses an element of every fiber at once. When Y is finite, these are finitely many choices, and induction on $|Y|$ makes the construction precise. In general, the step is an instance of the **axiom of choice**: for every family of nonempty sets $(S_i)_{i \in I}$ there is a

function c on I with $c(i) \in S_i$ for every i. Conversely, the axiom of choice follows from (b): given nonempty sets S_i, the map $(i, s) \mapsto i$ from the set of pairs $\{(i, s) \mid i \in I, s \in S_i\}$ to I is surjective, and if g is a right inverse of it, then $c(i) =$ the second coordinate of $g(i)$ is a choice function. So "every surjection has a right inverse" is the axiom of choice in another form. The left inverses in (a) need no choice, because g is forced on $\operatorname{im} f$ and constant elsewhere.

The axiom is part of the standard foundations (ZFC), and we use it freely. Its most useful form for linear algebra is Zorn's lemma (Section A.4), which shows that every vector space has a basis (Section 15.1).

When f is a bijection, neither half of the construction involves a choice, and the one-sided inverses merge.

Theorem A.1.14 (Bijections and inverses). Let $f : X \to Y$.

(a) If f has a left inverse g and a right inverse h, then $g = h$. In particular, an inverse of f, if it exists, is unique. We denote it f^{-1}.

(b) f is invertible if and only if f is bijective. In that case $f^{-1}(y)$ is the unique preimage of y, and f^{-1} is bijective with $\left(f^{-1}\right)^{-1} = f$.

(c) If $f : X \to Y$ and $g : Y \to Z$ are bijective, then so is $g \circ f$, and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof. ROUTINE

Uses. Proposition A.1.7 (Associativity and identities), Theorem A.1.12 (One-sided inverses of functions), Proposition A.1.8 (Composites, injectivity, and surjectivity), Definition A.1.10 (Inverse function)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), evaluate $g \circ f \circ h$ in two ways, using associativity.
Hint 2. For (b), "$\Rightarrow$" is Theorem A.1.12. For "$\Leftarrow$", send each y to its unique preimage; no choice and no x_0 is needed.
Hint 3. For (c), guess the inverse and check both composites.
Sketch. $g = g \circ (f \circ h) = (g \circ f) \circ h = h$. An invertible f has one-sided inverses on both sides, so it is injective and surjective; a bijective f is inverted by "the unique preimage". The equations $f^{-1} \circ f = \mathrm{id}$, $f \circ f^{-1} = \mathrm{id}$ say that f is an inverse of f^{-1}. For (c), $f^{-1} \circ g^{-1}$ is an inverse of $g \circ f$ by associativity.

(a) By Proposition A.1.7,

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_X \circ h = h.$$

If g and h are both inverses of f, then g is a left inverse and h is a right inverse, so $g = h$.

(b) If f is invertible, it has a left and a right inverse, so it is injective and surjective by Theorem A.1.12. Conversely, if f is bijective, every $y \in Y$ has exactly one preimage x_y. Let $g(y) = x_y$. Then $f(g(y)) = y$, and $g(f(x)) = x$ because x is the preimage of $f(x)$. So g is an inverse of f, and by (a) it is the only one. Finally, the equations $f^{-1} \circ f = \mathrm{id}_X$ and $f \circ f^{-1} = \mathrm{id}_Y$ say that f is an inverse of f^{-1}. So f^{-1} is invertible, hence bijective, and $\left(f^{-1}\right)^{-1} = f$ by (a).

(c) By Proposition A.1.7,

$$\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f) = f^{-1} \circ \left(g^{-1} \circ g\right) \circ f = f^{-1} \circ f = \mathrm{id}_X,$$

and similarly $(g \circ f) \circ \left(f^{-1} \circ g^{-1}\right) = \mathrm{id}_Z$. So $g \circ f$ is invertible with the stated inverse, and bijective by (b).□

For a bijection f and $B \subseteq Y$, the preimage $f^{-1}(B)$ of Definition A.1.3 equals the image $\operatorname{im}\left(f^{-1}\restriction_B\right)$ of B under the inverse function, so the two meanings of $f^{-1}(B)$ agree whenever both make sense.

Exercise A.1.15. ROUTINE Prove or disprove: if $f : X \to Y$ has exactly one left inverse, then f is bijective.

Solution on page 1669

A.1.4 Finite sets and counting

A finite set is one whose elements can be listed as $x_1, \dots, x_n$ without repetition, which is a bijection $[n] \to X$. That the number n does not depend on the listing is a theorem, and the same theorem is behind every counting argument in the book.

Definition A.1.16 (Finite set and cardinality). A set X is **finite** with n **elements** if there is a bijection $[n] \to X$ for some integer $n \geq 0$. We then write $|X| = n$, the **cardinality** of X. A set that is not finite is **infinite**.

Theorem A.1.17 (Pigeonhole principle). If $m > n \geq 0$, there is no injective function $[m] \to [n]$.

Intuition. We induct on the number n of boxes. Given an injection $[m] \to [n]$, we want an injection $[m-1] \to [n-1]$, so we delete the last box and one object. The object to delete is the one in box n, if any; after swapping it into position m, the rest of the objects avoid box n.

Proof. STANDARD

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering), Proposition A.1.8 (Composites, injectivity, and surjectivity)
Hint 1. Induct on n. For $n = 0$, is there any function $[m] \to \emptyset$ with $m \geq 1$?
Hint 2. Given an injection $f : [m] \to [n]$, you want an injection $[m-1] \to [n-1]$. If $n \notin \operatorname{im} f$, restrict f to $[m-1]$.
Hint 3. If $f(k) = n$, first compose f with the bijection of $[m]$ that swaps k and m. The new function sends m to n and, being injective, sends nothing else there.
Sketch. Induction on n. For $n = 0$ there is no function $[m] \to \emptyset$. For the step, let $f : [m] \to [n]$ be injective. Compose with the swap of k and m, where $f(k) = n$ (if $n \in \operatorname{im} f$), so that only m can map to n. Then f restricted to $[m-1]$ is an injection into $[n-1]$ with $m - 1 > n - 1$, which contradicts the induction hypothesis.

We use induction on n (Theorem A.2.1). For $n = 0$ and $m \geq 1$ there is no function $[m] \to [0] = \emptyset$ at all, since $1 \in [m]$ would need a value.

Let $n \geq 1$, suppose that there is no injection $[m'] \to [n-1]$ whenever $m' > n - 1$, and suppose for contradiction that $f : [m] \to [n]$ is injective with $m > n$. We first arrange that only m can map to n. If $n = f(k)$ for some $k \in [m]$, let $\tau : [m] \to [m]$ swap k and m and fix everything else. Then τ is a bijection, so $f' = f \circ \tau$ is injective (Proposition A.1.8), and $f'(m) = f(k) = n$. If $n \notin \operatorname{im} f$, let $f' = f$. In either case, f' is injective and no $i \in [m-1]$ has $f'(i) = n$: in the first case because $f'(i) = n = f'(m)$ would force $i = m$, and in the second because n is not a value at all. So the restriction of f' to $[m-1]$ is an injective function $[m-1] \to [n-1]$. Since $m - 1 > n - 1$, this contradicts the induction hypothesis. □

Corollary A.1.18 (Cardinality is well defined). Let X and Y be finite.

(a) If there are bijections $[m] \to X$ and $[n] \to X$, then $m = n$. So $|X|$ is well defined.

(b) If there is an injection $X \to Y$, then $|X| \leq |Y|$. There is a bijection $X \to Y$ if and only if $|X| = |Y|$.

Proof. ROUTINE

Uses. Theorem A.1.17 (Pigeonhole principle), Theorem A.1.14 (Bijections and inverses), Proposition A.1.8 (Composites, injectivity, and surjectivity)
Hint 1. Transport the question to $[m]$ and $[n]$ by composing with the given bijections and their inverses.
Hint 2. In (a), $\psi^{-1} \circ \varphi : [m] \to [n]$ is a bijection; apply Theorem A.1.17 in both directions.
Sketch. With bijections $\varphi : [m] \to X$ and $\psi : [n] \to Y$, an injection $f : X \to Y$ gives the injection $\psi^{-1} \circ f \circ \varphi : [m] \to [n]$, so $m \leq n$ by pigeonhole. Part (a) is the case $Y = X$, $f = \mathrm{id}_X$, applied in both directions. A bijection $X \to Y$ exists when $m = n$ by composing $\psi \circ \varphi^{-1}$.

Let $\varphi : [m] \to X$ and $\psi : [n] \to Y$ be bijections. If $f : X \to Y$ is injective, then $\psi^{-1} \circ f \circ \varphi : [m] \to [n]$ is injective by Proposition A.1.8, since bijections are injective (Theorem A.1.14). By Theorem A.1.17, $m \leq n$.

(a) Take $Y = X$ and $f = \mathrm{id}_X$: with two bijections $\varphi : [m] \to X$ and $\psi : [n] \to X$, the argument above gives $m \leq n$, and exchanging their roles gives $n \leq m$.

(b) The first claim is the argument above. If $f : X \to Y$ is a bijection, it is injective, and so is f^{-1}, which gives $|X| \le |Y| \le |X|$. Conversely, if $m = n$, then $\psi \circ \varphi^{-1} : X \to Y$ is a bijection by Theorem A.1.14(c). □

Proposition A.1.19 (Subsets of finite sets). Let X be finite and $A \subseteq X$. Then A is finite and $|A| \le |X|$, with equality if and only if $A = X$.

Proof. STANDARD

Uses. Corollary A.1.18 (Cardinality is well defined), Theorem A.2.1 (Induction, strong induction, and well-ordering), Definition A.1.16 (Finite set and cardinality)
Hint 1. A bijection $[n] \to X$ carries subsets of $[n]$ to subsets of X of the same size, so it suffices to treat $X = [n]$.
Hint 2. Induct on n. For $A \subseteq [n]$, apply the hypothesis to $A \cap [n-1]$ and then ask whether $n \in A$.
Sketch. Reduce to $X = [n]$ via a bijection $\varphi : [n] \to X$, replacing A by its preimage. Induct on n. Write $A' = A \cap [n-1]$, with $|A'| = k \le n-1$ and equality only for $A' = [n-1]$. If $n \notin A$, then $A = A'$ and $|A| < n$. If $n \in A$, extend a bijection $[k] \to A'$ by $k+1 \mapsto n$ to get $|A| = k+1 \le n$, with equality only if $A' = [n-1]$, that is, $A = [n]$.

Let $\varphi : [n] \to X$ be a bijection and $A_0 = \varphi^{-1}(A) \subseteq [n]$. Then φ restricts to a bijection $A_0 \to A$, so A is finite with $|A| = |A_0|$ once A_0 is, and $A = X$ if and only if $A_0 = [n]$. So we may assume $X = [n]$, and we induct on n. If $n = 0$, then $A = \varnothing = [0]$.

Let $n \ge 1$ and $A \subseteq [n]$. By the induction hypothesis, $A' = A \cap [n-1]$ is finite with $k = |A'| \le n-1$, and $k = n-1$ only if $A' = [n-1]$. If $n \notin A$, then $A = A'$, so A is finite with $|A| \le n-1 < n$. If $n \in A$, then $A = A' \cup \{n\}$ with $n \notin A'$. A bijection $[k] \to A'$, extended by $k+1 \mapsto n$, is a bijection $[k+1] \to A$, so $|A| = k+1 \le n$. Equality holds if and only if $k = n-1$, that is, $A' = [n-1]$, and since $n \in A$ this means $A = [n]$. □

For finite sets of the same size, half of bijectivity is free. This is the set-theoretic model for Corollary 4.2.18, where an operator on a finite-dimensional space is injective if and only if it is surjective; there, dimension plays the role of cardinality.

Theorem A.1.20 (Injective, surjective, and bijective for finite sets). Let X and Y be finite with $|X| = |Y|$, and let $f : X \to Y$. Then f is injective if and only if f is surjective, and hence if and only if f is bijective.

Proof. STANDARD

Uses. Proposition A.1.19 (Subsets of finite sets), Corollary A.1.18 (Cardinality is well defined), Theorem A.1.12 (One-sided inverses of functions), Theorem A.1.14 (Bijections and inverses)
Hint 1. If f is injective, then f is a bijection from X onto $\operatorname{im} f$. What does that say about $|\operatorname{im} f|$?
Hint 2. Use Proposition A.1.19 to compare $\operatorname{im} f$ with Y.
Hint 3. For the converse, a surjection has a right inverse $g : Y \to X$, which is injective. Apply the first half to g.
Sketch. If f is injective, $f : X \to \operatorname{im} f$ is a bijection, so $|\operatorname{im} f| = |X| = |Y|$, and $\operatorname{im} f = Y$ by Proposition A.1.19. If f is surjective, it has a right inverse g, which is injective (it has the left inverse f), hence bijective by the first half; then $f = f \circ g \circ g^{-1} = g^{-1}$ is bijective.

Suppose f is injective. Then f, viewed as a function $X \to \operatorname{im} f$, is a bijection, so $|\operatorname{im} f| = |X| = |Y|$ by Corollary A.1.18. Since $\operatorname{im} f \subseteq Y$, Proposition A.1.19 gives $\operatorname{im} f = Y$, and f is surjective.

Suppose f is surjective. By Theorem A.1.12(b) it has a right inverse $g : Y \to X$. Since $f \circ g = \mathrm{id}_Y$, f is a left inverse of g, so g is injective by Theorem A.1.12(a). As $|Y| = |X|$, the first half applied to g shows that g is bijective. Then

$$f = f \circ \left(g \circ g^{-1}\right) = (f \circ g) \circ g^{-1} = g^{-1},$$

so f is bijective by Theorem A.1.14(b). (Here Y is finite, so the choice of g needs no axiom of choice.) □

The successor function $s(n) = n + 1$ on $\mathbb{N}$ (Example A.1.11) is injective and not surjective, so Theorem A.1.20 fails for infinite sets. It is one way of saying that $\mathbb{N}$ is infinite.

Technique A.1.21 (Pigeonhole). If a function goes from a finite set to a smaller one, two inputs share a value (Theorem A.1.17). In practice: more than n objects in n boxes put two objects in one box. The skill

is in choosing the boxes, which are often the values of some quantity that can take only n values, such as a remainder mod n or an element of a finite field or of a finite set of matrices.

Reach for it when a statement claims that some coincidence or repetition must occur, but gives no way to find it. A typical consequence is that a sequence in a finite set repeats, so an invertible operation applied repeatedly eventually returns to the start. The canonical example is Example A.1.22; the cycle decomposition of permutations (Section 5.2) uses the same idea.
Used in. Exercise A.1.29.

Example A.1.22 (Powers of a matrix over a finite field). Let F be a finite field with q elements and $A \in F^{n\times n}$ invertible. We show that $A^k = I_n$ for some k with $1 \le k \le q^{n^2}$.

The set $F^{n\times n}$ has $N = q^{n^2}$ elements (Exercise A.1.28). The function $i \mapsto A^i$ from $\{0, 1, \dots, N\}$, a set of $N+1$ elements, to $F^{n\times n}$ is not injective by Theorem A.1.17 (composed with bijections to $[N+1]$ and $[N]$). So $A^i = A^j$ for some $0 \le i < j \le N$. Multiplying by $\left(A^{-1}\right)^i$ gives $I_n = A^{j-i}$, and $k = j - i$ satisfies $1 \le k \le N$.

For example, over $\mathbb{F}_2$ the matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ has $A^2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ and $A^3 = I_2$, much sooner than the bound $2^4 = 16$. Over $\mathbb{R}$ the conclusion fails: $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^k \neq I_2$ for every $k \ge 1$.

Exercises

Exercise A.1.23. ROUTINE Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 - 2x$. Compute $\operatorname{im}\left(f\restriction_{[0,3]}\right)$, $f^{-1}([-1, 0])$, and $f^{-1}\left(\operatorname{im}\left(f\restriction_{[0,1]}\right)\right)$. Which part of Proposition A.1.5 does the last answer illustrate?

Solution on page 1670

Exercise A.1.24. STANDARD Let $f : X \to Y$. Show that f is injective if and only if $\operatorname{im}(f\restriction_{A\cap A'}) = \operatorname{im}(f\restriction_A) \cap \operatorname{im}(f\restriction_{A'})$ for all subsets $A, A' \subseteq X$.

Solution on page 1670

Exercise A.1.25. ROUTINE Let $f : X \to Y$ and $g : Y \to Z$. Prove or disprove each claim.

(a) If $g \circ f$ is injective, then g is injective.

(b) If $g \circ f$ is bijective, then f is injective and g is surjective.

(c) If $g \circ f$ is bijective, then f and g are bijective.

Solution on page 1670

Exercise A.1.26. STANDARD Let $f : X \to Y$. Show that f is injective if and only if it can be cancelled on the left: for every set W and all functions $g, h : W \to X$, $f \circ g = f \circ h$ implies $g = h$. State and prove the analogous characterization of surjectivity, with cancellation on the right.

Solution on page 1670

Exercise A.1.27. STANDARD

(a) Let X and Y be finite with $|X| = m \ge 1$ and $|Y| = n$, and let $f : X \to Y$ be injective. How many left inverses does f have?

(b) Let $f : X \to Y$ be surjective with $Y = \{y_1, \dots, y_k\}$ finite and every fiber finite. Show that f has exactly $\left|f^{-1}(y_1)\right| \cdots \left|f^{-1}(y_k)\right|$ right inverses.

(c) Find the number of left inverses of $f : [2] \to [4]$, $f(i) = i$, and the number of right inverses of $p : [4] \to [2]$, $p(i) = 1$ for i odd and $p(i) = 2$ for i even.

Solution on page 1671

Exercise A.1.28. STANDARD For sets X and Y, let Y^X be the set of all functions $X \to Y$.

(a) Show by induction on $|X|$ that if X and Y are finite, then $\left|Y^X\right| = |Y|^{|X|}$ (with $0^0 = 1$).

(b) Construct a bijection $P(X) \to \{0,1\}^X$, and deduce that $|P(X)| = 2^{|X|}$.

(c) Deduce that $|F^n| = q^n$ and $|F^{m\times n}| = q^{mn}$ for a field F with q elements.

Solution on page 1671

Exercise A.1.29. STANDARD Let $a_1, \ldots, a_n$ be integers, where $n \geq 1$. Show that some nonempty block of consecutive terms $a_{i+1} + a_{i+2} + \cdots + a_j$, with $0 \leq i < j \leq n$, is divisible by n.

Solution on page 1672

Exercise A.1.30. HARD BEYOND 21-242 Let X be any set. Show that there is no surjective function $X \to P(X)$. In particular, there is no bijection $\mathbb{N} \to P(\mathbb{N})$.

Solution on page 1672

A.2 Induction and well-ordering

Most proofs about "all n" in this book are inductions, and the measure n is rarely the obvious one: it is the length of a list, the size of a matrix, the degree of a polynomial, the number of points a permutation moves, or the dimension of a space. This section states the three equivalent forms of induction and proves that they are equivalent, then discusses how induction is set up in linear algebra, where the induction hypothesis is usually applied to a different space from the one we started with.

A.2.1 Three forms of induction

Theorem A.2.1 (Induction, strong induction, and well-ordering). For each integer $n \geq 0$ let $P(n)$ be a statement. The following three principles are equivalent.

(I) (*Induction.*) If $P(0)$ holds, and $P(n)$ implies $P(n+1)$ for every $n \geq 0$, then $P(n)$ holds for every $n \geq 0$.

(II) (*Strong induction.*) If, for every $n \geq 0$, the truth of $P(k)$ for all $0 \leq k < n$ implies $P(n)$, then $P(n)$ holds for every $n \geq 0$.

(III) (*Well-ordering.*) Every nonempty subset of $\mathbb{N}$ has a least element.

"Equivalent" means that each principle can be derived from each other one, using only the arithmetic and order of the integers (every $n \geq 1$ is $k+1$ for some $k \geq 0$, and $k < n+1$ if and only if $k \leq n$). One of them has to be taken as an axiom about $\mathbb{N}$; the usual choice is (I), one of Peano's axioms for the natural numbers. So all three are true, and we use whichever is convenient. Technique 5.2.12 describes when strong induction and well-ordering are the convenient forms. Here we prove that they are no stronger than ordinary induction.

Intuition. Strong induction looks stronger than induction, because its hypothesis assumes more. The trick in (I) $\Rightarrow$ (II) is to induct on the statement "$P(0), \ldots, P(n)$ all hold", whose single predecessor is the whole history. Well-ordering is strong induction in contrapositive form: a set with no least element cannot contain 0, then cannot contain 1, and so on, and strong induction makes "and so on" precise.

Proof. STANDARD

Uses. the order of the integers

Techniques. Proving TFAE (Technique 2.7.1), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Prove (I) ⇒ (II) ⇒ (III) ⇒ (I). For the first, apply (I) to a different statement $Q(n)$ that packages $P(0), \dots, P(n)$ together.
Hint 2. For (II) ⇒ (III), take $S \subseteq \mathbb{N}$ with no least element and prove "$n \notin S$" for all n by strong induction.
Hint 3. For (III) ⇒ (I), look at the set of n for which $P(n)$ fails, and at its least element m. Why is $m \neq 0$, and what about $m-1$?
Sketch. (I) ⇒ (II): apply induction to $Q(n)$: "$P(k)$ for all $k \leq n$". (II) ⇒ (III): if S has no least element, then $n \notin S$ whenever no $k < n$ is in S (else n would be least), so $S = \emptyset$ by strong induction. (III) ⇒ (I): a least counterexample m is not 0, and $P(m-1)$ holds by minimality, so $P(m)$ holds, a contradiction.

We prove (I) ⇒ (II) ⇒ (III) ⇒ (I).

(I) ⇒ (II). Assume the hypothesis of (II), and let $Q(n)$ be the statement "$P(k)$ holds for all $0 \leq k \leq n$". The hypothesis of (II) for $n = 0$ has an empty premise, since there is no k with $0 \leq k < 0$; so $P(0)$ holds, which is $Q(0)$. If $Q(n)$ holds, then $P(k)$ holds for all $k < n+1$, so the hypothesis of (II) gives $P(n+1)$, and $Q(n+1)$ follows. By (I), $Q(n)$ holds for all n, and in particular so does $P(n)$.

(II) ⇒ (III). Let $S \subseteq \mathbb{N}$ have no least element. We show by strong induction that $P(n)$: "$n \notin S$" holds for every n, so that $S = \emptyset$. Suppose $k \notin S$ for all $k < n$. If n were in S, then every element $s \in S$ would satisfy $s \geq n$ (as $s < n$ would give $s \notin S$), so n would be the least element of S. Hence $n \notin S$. By (II), $S = \emptyset$, which is the contrapositive of (III).

(III) ⇒ (I). Assume the hypotheses of (I), and suppose for contradiction that $S = \{n \in \mathbb{N} \mid P(n) \text{ fails}\}$ is nonempty. By (III) it has a least element m. Since $P(0)$ holds, $m \geq 1$, so $m - 1 \in \mathbb{N}$, and $m - 1 \notin S$ because $m - 1 < m$. So $P(m-1)$ holds, and the hypothesis of (I) gives $P(m)$, contradicting $m \in S$. □

Induction can start at any integer n_0: to prove $P(n)$ for all $n \geq n_0$, apply Theorem A.2.1 to the statements $P(n_0 + n)$, $n \geq 0$. Likewise, every nonempty set of integers that are all $\geq n_0$ has a least element.

Example A.2.2 (A geometric sum of matrices). For $A \in F^{n\times n}$ and every $k \geq 1$,

$$(I - A)\left(I + A + A^2 + \cdots + A^{k-1}\right) = I - A^k.$$

For $k = 1$ both sides are $I - A$. If the identity holds for k, then by distributivity

$$\begin{aligned}(I - A)\left(I + A + \cdots + A^{k}\right) &= (I - A)\left(I + A + \cdots + A^{k-1}\right) + (I - A)A^k \\ &= I - A^k + A^k - A^{k+1} = I - A^{k+1},\end{aligned}$$

so it holds for $k + 1$. In particular, if $A^k = 0$ for some k, then $I - A$ is invertible, with inverse $I + A + \cdots + A^{k-1}$ (the other product equals I by the same computation, or by Theorem 3.5.23).

Example A.2.3 (Factorization into primes). Every integer $n \geq 2$ is a product of primes (a prime counts as a product of one prime). We use strong induction, starting at 2. Let $n \geq 2$ and suppose every k with $2 \leq k < n$ is a product of primes. If n is prime, we are done. Otherwise $n = ab$ with $2 \leq a, b < n$, and a and b are products of primes by hypothesis, hence so is n.

Ordinary induction would be awkward here: the factors a and b are smaller than n, but we cannot say by how much, so knowing $P(n-1)$ alone does not help.

The well-ordering principle is the form to use when an argument begins "take the least n such that …". The book needs it once in an essential way, already in Definition 1.1.19: dividing integers with remainder.

Proposition A.2.4 (Division with remainder). Let $n \geq 1$ and $x \in \mathbb{Z}$. There are unique integers q and r with $x = qn + r$ and $0 \leq r < n$.

Proof. ROUTINE
Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering)

Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Among the numbers $x - qn$ with $q \in \mathbb{Z}$, look at the nonnegative ones, and take the least.
Hint 2. Why is the set of nonnegative values of $x - qn$ nonempty? Try $q = -|x|$.
Hint 3. If the least one, $r = x - qn$, satisfied $r \geq n$, then $x - (q+1)n$ would be smaller. For uniqueness, if $qn + r = q'n + r'$, then n divides $r - r'$, and $|r - r'| < n$.
Sketch. Let $S = \{x - qn \mid q \in \mathbb{Z}\} \cap \mathbb{N}$. It contains $x + |x|n \geq 0$, so it has a least element $r = x - qn$. If $r \geq n$, then $r - n \in S$ is smaller, so $0 \leq r < n$. If $qn + r = q'n + r'$ with $0 \leq r, r' < n$, then $(q - q')n = r' - r$ has absolute value less than n, so it is 0.

Existence. Let $S = \{x - qn \mid q \in \mathbb{Z} \text{ and } x - qn \geq 0\} \subseteq \mathbb{N}$. Taking $q = -|x|$ gives $x + |x|n \geq x + |x| \geq 0$, so S is nonempty, and by well-ordering (Theorem A.2.1(III)) it has a least element $r = x - qn$. Then $r \geq 0$. If $r \geq n$, then $x - (q+1)n = r - n \geq 0$ is an element of S smaller than r, which is impossible. So $0 \leq r < n$.

Uniqueness. If $x = qn + r = q'n + r'$ with $0 \leq r, r' < n$, then $(q - q')n = r' - r$, and $-n < r' - r < n$. A multiple of n strictly between $-n$ and n is 0, so $q = q'$ and $r = r'$. □

The remainder r is the number $x \bmod n$ of Definition 1.1.19. The polynomial version, Theorem 7.1.14, has the same shape: there the size of a remainder is its degree, and the proof uses strong induction on the degree.

A.2.2 Induction in linear algebra

An induction proof has three parts, and in linear algebra each of them needs a decision.

The first is the **measure**: the quantity n that the induction runs on. In this book it is the length of a list (Technique 1.6.9, as in Theorem 1.6.10 and Lemma 4.1.2), the number of summands in a direct sum (Technique 4.4.10), the size of a matrix (Technique 5.5.17, as in cofactor expansion), the number of points a permutation moves or the degree of a polynomial (Technique 5.2.12), and the dimension of a space (in Section 7.5 and Section 8.2). The right measure is one that a natural reduction step makes smaller.

The second is the **statement** $P(n)$, and here linear algebra has a characteristic feature. In an induction on dimension, the reduction step replaces V by a smaller space: an invariant subspace, a complement, or a quotient. So the induction hypothesis must apply to every space of smaller dimension, not just to subspaces of the original one, and $P(n)$ has to be quantified accordingly: "for every vector space V with $\dim V = n$ and every $T \in \mathrm{L}(V)$ with the stated properties, the conclusion holds". The hypothesis is then applied to some $T' \in \mathrm{L}(V')$ built from T, for instance the restriction of T to an invariant subspace (Definition 7.2.1). This T' must again be an operator on V', which is the point that Exercise A.2.13 gets wrong. Similarly, an induction on the length of a list proves a statement about every list of length n in every space, so that it can be applied to a shorter list after a vector has been removed or replaced.

The third is the **base case**, which is often $n = 0$ (the empty list, the zero space) or $n = 1$. The base case has to be checked (Exercise A.2.12), and the step from n to $n + 1$ must be valid for every n at or above the base case, including the smallest one. The classic false proof that any n objects are equal breaks exactly there: its step works for $n \geq 2$ but not from $n = 1$ to $n = 2$.

Sometimes the statement we want is too weak to prove by induction directly: the hypothesis $P(n)$ does not give enough to prove $P(n + 1)$. Then the fix is often to prove more.

Technique A.2.5 (Strengthen the induction hypothesis). If the induction step fails because $P(n)$ is too weak to imply $P(n + 1)$, replace $P(n)$ by a stronger statement $P'(n)$ that implies it and that does carry from n to $n + 1$. The stronger statement is harder to prove, but the hypothesis you get to use is stronger too, and often the step becomes easy.

Reach for it when an induction step needs information about the smaller case that $P(n)$ does not record. Typical strengthenings: replace a bound by a sharper one that depends on n; carry along an extra property

of the object being built, as Gram–Schmidt carries the equality of the spans of the first k inputs and outputs (Section 6.4); or prove a statement for all objects of size n rather than for one. The canonical example is Example A.2.6.
Used in. Exercise A.2.11.

Example A.2.6 (A bound that has to be sharpened). We show that $\sum_{k=1}^{n} 1/k^2 < 2$ for all $n \geq 1$. Induction on "$S_n = \sum_{k=1}^{n} 1/k^2 < 2$" fails: from $S_n < 2$ we only get $S_{n+1} < 2 + 1/(n+1)^2$. The trouble is that the hypothesis leaves no room for the next term. So we prove the stronger statement

$$S_n \leq 2 - \frac{1}{n} \quad \text{for all } n \geq 1.$$

For $n = 1$, $S_1 = 1 = 2 - 1$. If $S_n \leq 2 - 1/n$, then

$$S_{n+1} \leq 2 - \frac{1}{n} + \frac{1}{(n+1)^2} \leq 2 - \frac{1}{n} + \frac{1}{n(n+1)} = 2 - \frac{1}{n+1},$$

using $1/n - 1/(n+1) = 1/(n(n+1))$. Hence $S_n \leq 2 - 1/n < 2$ for every n.

Exercises

Exercise A.2.7. ROUTINE Let $\lambda \in F$ and $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Show that $J^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$ for every $n \geq 1$, where $n\lambda^{n-1}$ means λ^{n-1} added to itself n times.

Solution on page 1672

Exercise A.2.8. STANDARD Let $A, B \in F^{n \times n}$ with $AB = BA$.

(a) Show that $A^k B = BA^k$ for all $k \geq 0$.

(b) Show that $(A+B)^m = \sum_{k=0}^{m} \binom{m}{k} A^k B^{m-k}$ for all $m \geq 0$.

(c) Show that (b) can fail for $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Solution on page 1673

Exercise A.2.9. STANDARD Show that every integer $n \geq 1$ can be written as a sum of distinct powers of 2, and that this representation is unique up to the order of the terms.

Solution on page 1673

Exercise A.2.10. STANDARD Let $S \subseteq \mathbb{Z}$ be a set with $S \neq \{0\}$ and $S \neq \emptyset$ that is closed under subtraction: $a - b \in S$ whenever $a, b \in S$. Show that S has a least positive element d, and that $S = d\mathbb{Z} = \{dk \mid k \in \mathbb{Z}\}$.

Solution on page 1673

Exercise A.2.11. HARD Show that for every $n \geq 1$,

$$\frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \leq \frac{1}{\sqrt{3n}}.$$

(Induction on this inequality as stated does not work.)

Solution on page 1674

Exercise A.2.12. ROUTINE Find the error in the following argument.

Claim: $n^2 + n + 1$ is even for every integer $n \geq 0$. *Proof:* Suppose $n^2 + n + 1$ is even. Then $(n+1)^2 + (n+1) + 1 = \left(n^2 + n + 1\right) + 2(n+1)$ is a sum of two even numbers, hence even. By induction, the claim holds for every n.

Solution on page 1674

Exercise A.2.13. STANDARD Find the error in the following argument.

Claim: Every operator on a nonzero finite-dimensional real vector space has an eigenvector. *Proof:* By induction on $n = \dim V \geq 1$. If $n = 1$, then $V = \operatorname{span}(\mathbf{v})$ for any $\mathbf{v} \neq \mathbf{0}$, and $T(\mathbf{v}) = \lambda \mathbf{v}$ for some $\lambda \in \mathbb{R}$, so $\mathbf{v}$ is an eigenvector. Suppose the claim holds in dimension n, and let $\dim V = n+1$ and $T \in \mathrm{L}(V)$. Pick a subspace W of V with $\dim W = n$. By the induction hypothesis, the restriction of T to W has an eigenvector $\mathbf{w} \in W$, so $T(\mathbf{w}) = \lambda \mathbf{w}$ and $\mathbf{w}$ is an eigenvector of T.

Solution on page 1674

A.3 Equivalence relations

Much of the book classifies objects up to some notion of sameness: matrices up to row operations (Section 3.2.4), vector spaces up to isomorphism and square matrices up to change of basis (Section 4.5), vectors up to an element of a subspace (Section 4.6). Each notion is an equivalence relation, and each one cuts its set into classes. This section sets up the general language: equivalence classes, quotient sets, functions on quotient sets, and invariants.

A.3.1 Equivalence relations

Definition A.3.1 (Equivalence relation). A **relation** on a set X is a subset $R \subseteq X \times X$; we write $x \sim y$ for $(x, y) \in R$. A relation $\sim$ is an **equivalence relation** if it is

- *reflexive*: $x \sim x$ for every $x \in X$;
- *symmetric*: if $x \sim y$, then $y \sim x$;
- *transitive*: if $x \sim y$ and $y \sim z$, then $x \sim z$.

Example A.3.2 (Equivalence relations).

(a) Equality on any set.

(b) Let $n \geq 1$. **Congruence modulo** n on $\mathbb{Z}$: $x \equiv y \pmod{n}$ if n divides $x - y$. It is reflexive since $n \mid 0$, symmetric since $n \mid x - y$ implies $n \mid y - x$, and transitive since $x - z = (x - y) + (y - z)$.

(c) Any function $f : X \to Y$ defines the relation $x \underset{f}{\sim} x'$ if $f(x) = f(x')$. It inherits the three properties from equality in Y. For instance, "$A \sim B$ if $\operatorname{rank} A = \operatorname{rank} B$" is an equivalence relation on $F^{m \times n}$.

(d) Row equivalence of matrices (Proposition 3.2.22), similarity of square matrices (Proposition 4.5.24), isomorphism of vector spaces (Proposition 4.5.4), and congruence modulo a subspace U, $\mathbf{v} - \mathbf{w} \in U$ (Remark 4.6.16).

Each axiom can fail on its own, and the failures of transitivity are the instructive ones.

Non-example A.3.3 (Relations that are not equivalence relations).

(a) $\leq$ on $\mathbb{R}$ is reflexive and transitive but not symmetric.

(b) On a vector space V, let $\mathbf{v} \sim \mathbf{w}$ if one of $\mathbf{v}, \mathbf{w}$ is a scalar multiple of the other. This is reflexive and symmetric, but not transitive when $\dim V \geq 2$: for independent $\mathbf{v}, \mathbf{w}$ we have $\mathbf{v} \sim \mathbf{0}$ and $\mathbf{0} \sim \mathbf{w}$, since $\mathbf{0} = 0\mathbf{v} = 0\mathbf{w}$, but not $\mathbf{v} \sim \mathbf{w}$. On $V \setminus \{\mathbf{0}\}$, the relation becomes "$\mathbf{w} = c\mathbf{v}$ for some $c \neq 0$", which is an equivalence relation (Exercise A.3.16).

(c) On $F^{n\times n}$, let $A \sim B$ if $AB = BA$. This is reflexive and symmetric, but every A commutes with I, so transitivity would make all matrices commute, which fails for $n \geq 2$ (Example 2.1.19).

(d) On $\mathbb{R}$, let $x \sim y$ if $|x - y| \leq 1$. Then $0 \sim 1 \sim 2$ but not $0 \sim 2$. "Close to" is rarely transitive.

A.3.2 Classes and partitions

Definition A.3.4 (Equivalence class). Let $\sim$ be an equivalence relation on X. The **equivalence class** of $x \in X$ is

$$[x] = \{y \in X \mid x \sim y\}.$$

Any element of $[x]$ is called a **representative** of the class.

For congruence modulo n, the class of x is $x + n\mathbb{Z} = \{x + nk \mid k \in \mathbb{Z}\}$. For congruence modulo a subspace U, it is the coset $\mathbf{v} + U$ (Lemma 4.6.14). For $\underset{f}{\sim}$, it is the fiber $f^{-1}(f(x))$ (Definition A.1.3). The next lemma says that classes behave like cosets: two classes are equal or disjoint, and any element of a class can name it.

Lemma A.3.5 (Equal classes). Let $\sim$ be an equivalence relation on X and $x, y \in X$. The following are equivalent:

(i) $x \sim y$;
(ii) $[x] = [y]$;
(iii) $[x] \cap [y] \neq \emptyset$.

Proof. ROUTINE

Uses. Definition A.3.1 (Equivalence relation), Definition A.3.4 (Equivalence class)
Techniques. Proving TFAE (Technique 2.7.1), Double containment (Technique 1.3.14)
Hint 1. Prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i). Which axiom does each step use?
Hint 2. For (i) $\Rightarrow$ (ii), take $z \in [y]$ and chain $x \sim y \sim z$. For (iii) $\Rightarrow$ (i), take z in both classes and chain $x \sim z$ with $z \sim y$.
Sketch. (i) $\Rightarrow$ (ii): if $x \sim y$ and $y \sim z$, then $x \sim z$, so $[y] \subseteq [x]$; by symmetry $[x] \subseteq [y]$. (ii) $\Rightarrow$ (iii): $x \in [x]$ by reflexivity. (iii) $\Rightarrow$ (i): if $x \sim z$ and $y \sim z$, then $x \sim z \sim y$.

(i) $\Rightarrow$ (ii). Suppose $x \sim y$. If $z \in [y]$, then $y \sim z$, so $x \sim z$ by transitivity, and $z \in [x]$. Thus $[y] \subseteq [x]$. By symmetry $y \sim x$, and the same argument gives $[x] \subseteq [y]$.

(ii) $\Rightarrow$ (iii). By reflexivity $x \in [x] = [y]$, so $x \in [x] \cap [y]$.

(iii) $\Rightarrow$ (i). Let $z \in [x] \cap [y]$, so $x \sim z$ and $y \sim z$. By symmetry $z \sim y$, and by transitivity $x \sim y$. □

Definition A.3.6 (Partition). A **partition** of a set X is a set Π of nonempty subsets of X, called **blocks**, such that every element of X lies in exactly one block.

Theorem A.3.7 (Equivalence relations are partitions). Let X be a set.

(a) If $\sim$ is an equivalence relation on X, its equivalence classes form a partition of X.

(b) If Π is a partition of X, the relation "$x \sim y$ if x and y lie in the same block" is an equivalence relation, and its classes are the blocks of Π.

(c) These two constructions are inverse to each other. So equivalence relations on X correspond bijectively to partitions of X.

Proof. ROUTINE

> *Uses.* Lemma A.3.5 (Equal classes), Definition A.3.6 (Partition), Definition A.3.1 (Equivalence relation)
> *Hint 1.* For (a), every class is nonempty and $x \in [x]$. Two classes that share an element are equal by Lemma A.3.5.
> *Hint 2.* For (b), each axiom follows from "exactly one block". For the classes, the class of x is the block containing x.
> *Hint 3.* For (c), start from $\sim$, pass to its classes, and then to "same class"; by Lemma A.3.5, x and y are in the same class iff $x \sim y$.
> *Sketch.* (a) $x \in [x]$, and if $x \in [y] \cap [z]$, then $[y] = [x] = [z]$. (b) Same block is reflexive, symmetric and transitive because each element is in exactly one block, and the class of x is its block. (c) Partition then relation: x, y in the same class iff $x \sim y$, which recovers $\sim$. Relation then partition: by (b) the classes are the blocks.

(a) Each class $[x]$ is nonempty, since $x \in [x]$ by reflexivity, so every element lies in at least one class. If $x \in [y]$ and $x \in [z]$, then $[y] \cap [z] \neq \emptyset$, so $[y] = [z]$ by Lemma A.3.5. So x lies in exactly one class.

(b) Let $B(x)$ denote the block containing x. Then $x \sim y$ means $B(x) = B(y)$, and this is an equivalence relation by Example A.3.2(c), applied to the function $x \mapsto B(x)$. The class of x is $\{y \mid B(y) = B(x)\} = B(x)$, since $y \in B(x)$ if and only if $B(y) = B(x)$.

(c) Start with $\sim$, form the partition into classes, and let $\approx$ be "in the same class". Then $x \approx y$ means that the class containing x, which is $[x]$, is the class containing y, which is $[y]$. By Lemma A.3.5 this happens if and only if $x \sim y$, so $\approx$ is $\sim$. In the other order, starting with a partition, (b) says that the classes of the relation are the blocks, so we recover the partition. □

For congruence modulo a subspace, (a) is Corollary 4.6.15. In practice, the theorem is used to define a partition by giving a relation, which is easier to check: three axioms instead of "exactly one block".

A.3.3 Quotient sets and well-defined functions

Definition A.3.8 (Quotient set). Let $\sim$ be an equivalence relation on X. The **quotient set** $X/\sim$ is the set of equivalence classes, and the **quotient map** is

$$q : X \to X/\sim, \quad q(x) = [x].$$

The quotient map is surjective, and $q(x) = q(y)$ if and only if $x \sim y$ (Lemma A.3.5), so passing to the quotient set turns $\sim$ into equality. That is the purpose of the construction: in $X/\sim$, equivalent elements have become equal. The quotient space V/U of Definition 4.6.20 is the quotient set for congruence modulo U, with a vector space structure added.

Example A.3.9 ($\mathbb{Z}/n\mathbb{Z}$ as a quotient set). Let $n \geq 2$, and let $\sim$ be congruence modulo n. By Proposition A.2.4, each $x \in \mathbb{Z}$ is congruent to exactly one $r \in \{0, 1, \ldots, n-1\}$, namely $r = x \bmod n$. So each class contains exactly one of $0, 1, \ldots, n-1$, and

$$\mathbb{Z}/\sim = \{[0], [1], \ldots, [n-1]\},$$

with these n classes distinct. In Definition 1.1.19, $\mathbb{Z}/n\mathbb{Z}$ was defined as the set $\{0, 1, \ldots, n-1\}$ of remainders; we identify the remainder r with its class $[r]$. Under this identification, Lemma 1.1.21(a) says that $[a] + [b] = [a+b]$ and $[a][b] = [ab]$. These formulas define operations on classes, because the right side does not depend on the representatives: if $a \sim a'$ and $b \sim b'$, then n divides $(a+b) - (a'+b')$ and $ab - a'b' = a(b - b') + (a - a')b'$.

The last sentence is the general issue with quotient sets. A function on $X/\sim$ is usually given by a formula in a representative, and the formula must not depend on which representative is used (Technique 4.6.18). The theorem below says exactly when such a formula defines a function, and what properties the function then has.

Theorem A.3.10 (Functions on a quotient set). Let $\sim$ be an equivalence relation on X with quotient map $q : X \to X/\sim$, and let $g : X \to Y$.

(a) There is a function $\overline{g} : X/{\sim} \to Y$ with $\overline{g} \circ q = g$ if and only if g is constant on classes: $x \sim x'$ implies $g(x) = g(x')$. In that case $\overline{g}$ is unique, and $\overline{g}([x]) = g(x)$.

(b) Such a $\overline{g}$ is injective if and only if $g(x) = g(x')$ implies $x \sim x'$.

(c) Such a $\overline{g}$ is surjective if and only if g is surjective.

Proof. ROUTINE

Uses. Definition A.3.8 (Quotient set), Lemma A.3.5 (Equal classes), Definition A.1.2 (Function)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18), Uniqueness via two candidates (Technique 1.1.8)
Hint 1. If $\overline{g} \circ q = g$, then $\overline{g}([x]) = g(x)$ for every x, which settles uniqueness and the formula.
Hint 2. For existence, define $\overline{g}(C)$ for a class C as $g(x)$ for any $x \in C$. What has to be checked for this to be a function?
Hint 3. For (b), translate: $\overline{g}([x]) = \overline{g}([x'])$ means $g(x) = g(x')$, and $[x] = [x']$ means $x \sim x'$. For (c), $\operatorname{im}\overline{g} = \operatorname{im} g$ because q is surjective.
Sketch. If $\overline{g} \circ q = g$, then $g(x) = \overline{g}([x]) = \overline{g}([x']) = g(x')$ when $x \sim x'$, and $\overline{g}$ is forced to be $[x] \mapsto g(x)$. Conversely, if g is constant on classes, this formula is independent of the representative, so it defines $\overline{g}$. Injectivity of $\overline{g}$ reads "$g(x) = g(x') \Rightarrow [x] = [x']$". Since q is surjective, $\overline{g}$ and g have the same image.

(a) Suppose $\overline{g} \circ q = g$. Then $\overline{g}([x]) = g(x)$ for every x, which determines $\overline{g}$, since every class is $[x]$ for some x. And if $x \sim x'$, then $[x] = [x']$ by Lemma A.3.5, so $g(x) = \overline{g}([x]) = \overline{g}([x']) = g(x')$.

Conversely, suppose g is constant on classes. For a class $C \in X/{\sim}$, all $x \in C$ satisfy $x \sim x_0$ for a fixed $x_0 \in C$, so g takes the single value $g(x_0)$ on C. Setting $\overline{g}(C)$ equal to this value defines a function $X/{\sim} \to Y$, and $\overline{g}([x]) = g(x)$ because $x \in [x]$.

(b) $\overline{g}$ is injective if and only if $\overline{g}([x]) = \overline{g}([x'])$ implies $[x] = [x']$ for all $x, x' \in X$, that is, if and only if $g(x) = g(x')$ implies $x \sim x'$.

(c) Since q is surjective, the values of $\overline{g}$ are exactly the values $\overline{g}(q(x)) = g(x)$, so $\operatorname{im}\overline{g} = \operatorname{im} g$. □

Taking $\sim$ to be $\underset{f}{\sim}$ gives a factorization of every function.

Corollary A.3.11 (Factoring a function). Let $f : X \to Y$, and let $\underset{f}{\sim}$ be the relation $f(x) = f(x')$. Then $\overline{f}([x]) = f(x)$ is a well-defined bijection $X/{\sim_f} \to \operatorname{im} f$. So f is the composite

$$X \xrightarrow{q} X/{\sim_f} \xrightarrow{\overline{f}} \operatorname{im} f \xrightarrow{\iota} Y$$

of a surjection, a bijection, and the injective inclusion $\iota(y) = y$.

Proof. ROUTINE

Uses. Theorem A.3.10 (Functions on a quotient set)
Hint 1. Apply Theorem A.3.10 to f, viewed as a function $X \to \operatorname{im} f$.
Hint 2. Both conditions, in (a) and in (b) of the theorem, are the definition of $\underset{f}{\sim}$.
Sketch. For $g = f : X \to \operatorname{im} f$, "$x \underset{f}{\sim} x' \Rightarrow g(x) = g(x')$" and its converse both hold by definition of $\underset{f}{\sim}$, and g is surjective. So $\overline{f}$ exists and is injective and surjective. Composing with q and ι recovers f.

View f as a surjective function $g : X \to \operatorname{im} f$. By definition of $\underset{f}{\sim}$, $x \underset{f}{\sim} x'$ if and only if $g(x) = g(x')$. So g is constant on classes, and $\overline{f}$ exists by Theorem A.3.10(a); it is injective by (b) and surjective by (c). For every x, $\iota\big(\overline{f}(q(x))\big) = f(x)$. □

For a linear map $T : V \to W$, $T(\mathbf{v}) = T(\mathbf{v}')$ if and only if $\mathbf{v} - \mathbf{v}' \in \ker T$, so $\underset{T}{\sim}$ is congruence modulo $\ker T$ and $V/{\sim_T}$ is the quotient space $V/\ker T$. The corollary then says that $\overline{T}$ is a bijection $V/\ker T \to \operatorname{im} T$, and the first isomorphism theorem (Theorem 4.6.34) adds that it is linear.

A.3.4 Invariants and normal forms

To classify objects up to $\sim$ is to describe the quotient set $X/\sim$: decide when two given objects are equivalent, and list one object from each class.

Definition A.3.12 (Invariant, normal form). Let $\sim$ be an equivalence relation on X.

- An **invariant** of $\sim$ is a function $g : X \to Y$ that is constant on classes: $x \sim x'$ implies $g(x) = g(x')$.
- An invariant is **complete** if moreover $g(x) = g(x')$ implies $x \sim x'$.
- A set of **normal forms** for $\sim$ is a subset $N \subseteq X$ that contains exactly one element of each class.

By Theorem A.3.10, an invariant is a function on $X/\sim$, and a complete invariant is an injective one, so it identifies $X/\sim$ with the set of its values. A set of normal forms N is mapped bijectively onto $X/\sim$ by q, and "the normal form of x" is a complete invariant with values in N. The classification results of the book all have this shape:

- For row equivalence on $F^{m\times n}$, the RREF is a complete invariant, and the matrices in RREF are normal forms (Corollary 3.3.18). So is the kernel (Corollary 3.3.20).
- For isomorphism of finite-dimensional vector spaces, the dimension is a complete invariant (Theorem 4.5.6), and $F^0, F^1, F^2, \ldots$ are normal forms.
- For the relation $B = PAQ$ with P, Q invertible, the rank is a complete invariant, with normal forms $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ (Theorem 3.5.27).
- For similarity, rank, trace, and determinant are invariants (Proposition 4.5.27, Proposition 5.4.17), but even all of them together are not complete: I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ share them and are not similar. Over $\mathbb{C}$, the Jordan form, with a fixed ordering of the blocks, provides normal forms (Section 10.3).
- For congruence modulo n, the remainder $x \bmod n$ is a complete invariant, with normal forms $0, 1, \ldots, n-1$ (Example A.3.9).

Technique A.3.13 (Classify by a complete invariant). To decide when two objects are equivalent, find a quantity that equivalent objects share (an invariant), and then prove that it is complete. To prove that x and x' are *not* equivalent, one invariant on which they differ is enough (Technique 4.5.29). To prove that they *are* equivalent, completeness is needed, and it is usually proved by reducing every object to a normal form that depends only on the invariant: if both x and x' are equivalent to the same normal form, they are equivalent to each other by symmetry and transitivity.

Reach for it whenever a question asks "when are these the same up to ...?" or "how many essentially different ... are there?". The canonical example is Theorem 3.5.27, where row and column operations reduce every matrix to $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ with r the rank.

Exercises

Exercise A.3.14. ROUTINE Decide which of the following are equivalence relations. For those that are, describe the classes.

(a) On $\mathbb{R}$: $x \sim y$ if $x - y \in \mathbb{Z}$.

(b) On $\mathbb{Z}$: $x \sim y$ if $xy > 0$.

(c) On $F^{m\times n}$: $A \sim B$ if $\operatorname{rank} A = \operatorname{rank} B$.

(d) On $V \setminus \{\mathbf{0}\}$, for a vector space V: $\mathbf{v} \sim \mathbf{w}$ if $\operatorname{span}(\mathbf{v}) = \operatorname{span}(\mathbf{w})$.

(e) On $P(X)$, for a set X with at least two elements: $A \sim B$ if $A \cap B \neq \emptyset$.

Solution on page 1675

Exercise A.3.15. STANDARD Count the equivalence relations on $[3]$ and on $[4]$.

Solution on page 1675

Exercise A.3.16. STANDARD Let F be a field and $n \geq 0$. On $F^{n+1} \setminus \{\mathbf{0}\}$, let $\mathbf{v} \sim \mathbf{w}$ if $\mathbf{w} = c\mathbf{v}$ for some nonzero $c \in F$.

(a) Show that $\sim$ is an equivalence relation, and describe its classes in terms of one-dimensional subspaces of F^{n+1}.

(b) If F has q elements, show that the quotient set has $\left(q^{n+1}-1\right)/(q-1) = 1 + q + \cdots + q^n$ elements.

Solution on page 1676

Exercise A.3.17. STANDARD Let R and S be equivalence relations on X, viewed as subsets of $X \times X$. Prove or disprove: (a) $R \cap S$ is an equivalence relation; (b) $R \cup S$ is an equivalence relation.

Solution on page 1676

Exercise A.3.18. ROUTINE Find the error in the following argument.

Claim: A relation that is symmetric and transitive is reflexive. *Proof:* Let $x \in X$ and pick y with $x \sim y$. By symmetry $y \sim x$, and by transitivity, $x \sim y$ and $y \sim x$ give $x \sim x$.

Solution on page 1676

Exercise A.3.19. STANDARD Let $m, n \geq 2$, and write $[a]_n$ for the class of $a \in \mathbb{Z}$ under congruence modulo n. Show that the formula $f\left([a]_n\right) = [a]_m$ defines a function $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ if and only if m divides n. When it does, show that f is surjective, and bijective only if $m = n$.

Solution on page 1676

Exercise A.3.20. HARD Let R be any relation on X. Define $x \approx y$ if there is a finite sequence $x = x_0, x_1, \ldots, x_k = y$, with $k \geq 0$, such that for each i, $(x_{i-1}, x_i) \in R$ or $(x_i, x_{i-1}) \in R$.

(a) Show that $\approx$ is an equivalence relation containing R, and that it is contained in every equivalence relation that contains R. It is the equivalence relation **generated** by R.

(b) Let R be the relation on $F^{m \times n}$ given by "B is obtained from A by one elementary row operation". Show that the equivalence relation generated by R is row equivalence (Definition 3.2.21).

Solution on page 1677

A.4 Zorn's lemma BEYOND 21-242

In a finite-dimensional space, a basis is built by adding vectors one at a time until the list spans; the process stops because an independent list has at most $\dim V$ vectors (Theorem 4.1.16). In a space such as $F^{\mathbb{N}}$ or $\mathbb{R}$ over $\mathbb{Q}$, the process need not stop after finitely many steps, or after countably many, and there is no last step to point to. Zorn's lemma replaces the process by a single statement: if every partial result can be extended to an upper bound, then a maximal one exists. It is the standard way to prove that objects of linear algebra exist in infinite dimensions, and Section 15.1 uses it to show that every vector space has a basis.

A.4.1 Partial orders and chains

Definition A.4.1 (Partial order). A **partial order** on a set P is a relation $\leq$ that is

- *reflexive*: $p \leq p$;
- *antisymmetric*: if $p \leq p'$ and $p' \leq p$, then $p = p'$;
- *transitive*: if $p \leq p'$ and $p' \leq p''$, then $p \leq p''$.

The pair $(P, \leq)$ is a **partially ordered set**. We write $p < p'$ if $p \leq p'$ and $p \neq p'$. Two elements p, p' are **comparable** if $p \leq p'$ or $p' \leq p$. A **chain** is a subset $C \subseteq P$ in which any two elements are comparable.

Definition A.4.2 (Upper bound, maximal element). Let $(P, \leq)$ be a partially ordered set. An **upper bound** of a subset $S \subseteq P$ is an element $u \in P$ with $s \leq u$ for every $s \in S$; it need not lie in S. An element $m \in P$ is **maximal** if no element of P is strictly larger: $m \leq p$ implies $p = m$.

In linear algebra, P is almost always a collection of sets, such as subspaces or independent sets, ordered by inclusion $\subseteq$. A chain is then a family of sets that is *nested*: of any two, one contains the other. A maximal element is one that cannot be enlarged within the collection. It need not be the largest element: there may be many maximal elements, none containing the others.

Example A.4.3 (Chains and maximal elements).

(a) In $P(X)$, ordered by $\subseteq$, every chain has the upper bound X, and X is the only maximal element.

(b) Let P be the set of subspaces of $\mathbb{R}^3$ of dimension at most 2, ordered by $\subseteq$. The maximal elements are the planes through $\mathbf{0}$, and there are infinitely many of them. Every chain has an upper bound: a nonempty chain has at most three distinct members, of dimensions 0, 1, 2, and its largest member is an upper bound, while the empty chain is bounded by $\{\mathbf{0}\}$.

(c) Let P be the set of proper subspaces of $F[x]$, ordered by $\subseteq$. The subspaces $F[x]_{\leq n}$ of polynomials of degree at most n form a chain, and it has no upper bound in P: a subspace containing every $F[x]_{\leq n}$ contains every polynomial, so it is $F[x]$, which is not proper.

(d) $\mathbb{N}$ with $\leq$ is a chain with no upper bound, and it has no maximal element.

In (b), an upper bound of a chain was its largest member. In (c) and (d) the chains are infinite with no largest member, and nothing in P lies above all of their members; in (c), the natural candidate, the union, is $F[x]$, which is not in P. Zorn's lemma says that this failure is the only obstruction to a maximal element.

A.4.2 Zorn's lemma

Theorem A.4.4 (Zorn's lemma). Let $(P, \leq)$ be a partially ordered set in which every chain has an upper bound. Then P has a maximal element.

The empty set is a chain, and an upper bound of it is any element of P. So the hypothesis includes $P \neq \emptyset$, and in practice one checks that P is nonempty and that every nonempty chain has an upper bound.

Zorn's lemma is equivalent to the axiom of choice (Remark A.1.13): Exercise A.4.11 derives the axiom of choice from it, and the proof below derives it from the axiom of choice. The proof is not needed anywhere in the book, and it uses a transfinite construction that the rest of the book does not, so it is marked beyond and the reader can skip it.

Intuition (Why Zorn's lemma should be true). Suppose P has no maximal element. Then every chain C has a *strict* upper bound: it has an upper bound u, and u is not maximal, so something is larger than u. Start from any element p_0, take a strict upper bound p_1 of $\{p_0\}$, then p_2 of $\{p_0, p_1\}$, and so on. After all finite stages, the chain $\{p_0, p_1, \ldots\}$ again has a strict upper bound p_ω, and we continue. The process never gets stuck, and it never repeats an element, so it should eventually use up more elements than P has. The proof makes "continue forever, through all stages" precise without ordinals, by looking at all chains that are built by this rule at once.

Proof. CHALLENGE

Uses. Remark A.1.13 (The axiom of choice), Definition A.4.1 (Partial order), Definition A.4.2 (Upper bound, maximal element)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Suppose there is no maximal element. Show that every chain C then has a strict upper bound, and use the axiom of choice to fix one, $s(C)$, for every chain.
Hint 2. Call $W \subseteq P$ a tower if every nonempty subset of W has a least element, and every $w \in W$ equals $s(W_{<w})$, where $W_{<w} = \{x \in W \mid x < w\}$. Show that of any two towers, one is an initial segment of the other.
Hint 3. For the comparison, take the union I of all sets that are downward closed in both towers W and V. If $I \neq W$ and $I \neq V$, the least elements of $W \setminus I$ and of $V \setminus I$ are both $s(I)$.
Hint 4. Then the union U of all towers is a tower, and $U \cup \{s(U)\}$ is a strictly larger tower, a contradiction.
Sketch. Assume no maximal element; then every chain has a strict upper bound, and choice gives a function s picking one. A tower is a well-ordered W with $w = s(W_{<w})$ for all $w \in W$. Any two towers are comparable, one an initial segment of the other: the largest common downward closed set I is all of one of them, since otherwise $s(I)$ would extend it. So the union U of all towers is a tower, and so is $U \cup \{s(U)\}$, which contradicts $s(U) \notin U$.

Suppose that P has no maximal element. Let C be a chain, and let u be an upper bound of C. Since u is not maximal, there is $p \in P$ with $u < p$. Then $c < p$ for every $c \in C$: we have $c \leq u \leq p$, and $c = p$ would give $p \leq u \leq p$, so $u = p$. So the set of *strict* upper bounds of C is nonempty, and by the axiom of choice there is a function s that assigns to every chain C a strict upper bound $s(C)$.

A subset $W \subseteq P$ is **well ordered** if every nonempty subset of W has a least element; then W is a chain, since every pair $\{x, y\} \subseteq W$ has a least element. For $w \in W$ write $W_{<w} = \{x \in W \mid x < w\}$. Call W a **tower** if it is well ordered and $w = s(W_{<w})$ for every $w \in W$. (For instance $\emptyset$, $\{s(\emptyset)\}$, and $\{s(\emptyset), s(\{s(\emptyset)\})\}$ are towers.) A subset $A \subseteq W$ is **downward closed** in W if $x \in A$, $y \in W$, and $y < x$ imply $y \in A$.

Step 1. If A is downward closed in a well-ordered W and $A \neq W$, then $A = W_{<m}$, where m is the least element of $W \setminus A$. If $y \in W$ and $y < m$, then $y \in A$ by minimality of m. If $y \in A$, then $m \leq y$ is impossible, since it would put m in A (either $m = y$ or $m < y$ and A is downward closed); as W is a chain, $y < m$.

Step 2. If W and V are towers, then one of them is downward closed in the other. Let I be the union of all subsets of $W \cap V$ that are downward closed in both W and V. A union of downward closed sets is downward closed, so I is downward closed in both, and it is the largest such set. Suppose $I \neq W$ and $I \neq V$. By Step 1, $I = W_{<m}$ and $I = V_{<n}$, where m and n are the least elements of $W \setminus I$ and $V \setminus I$. Since W and V are towers,

$$m = s(W_{<m}) = s(I) = s(V_{<n}) = n.$$

Then $I \cup \{m\}$ is downward closed in W (an element of W below m lies in $W_{<m} = I$), and likewise in V. By maximality of I, $m \in I$, which is false. So $I = W$ or $I = V$; that is, W is downward closed in V, or V in W.

Step 3. The union U of all towers is a tower. Let W be a tower and $w \in W$. We claim $U_{<w} = W_{<w}$. Clearly $W_{<w} \subseteq U_{<w}$. If $u \in U$ and $u < w$, then $u \in V$ for some tower V. By Step 2, either V is downward closed in W, so $V \subseteq W$ and $u \in W$; or W is downward closed in V, and then $u \in V$, $u < w \in W$ give $u \in W$. Either way $u \in W_{<w}$.

Now U is a chain: for $x, y \in U$, say $x \in W$ and $y \in V$, Step 2 puts both in W or both in V, which are chains. It is well ordered: let $A \subseteq U$ be nonempty, pick $a \in A$ in some tower W, and let m be the least element of $A \cap W$. If $b \in A$ had $b < m$, then $b \in U_{<m} = W_{<m}$, so $b \in A \cap W$ and $b < m$, which is impossible. As U is a chain, $m \leq b$ for all $b \in A$. Finally, for $u \in U$, say $u \in W$, we have $s(U_{<u}) = s(W_{<u}) = u$. So U is a tower.

Step 4. Contradiction. Let $p = s(U)$, a strict upper bound of the chain U, so $p \notin U$. Then $U' = U \cup \{p\}$ is a tower: a nonempty subset of U' has as least element its least element in U if it meets U, and p otherwise; $U'_{<p} = U$ and $s(U) = p$; and for $u \in U$, $U'_{<u} = U_{<u}$ since $p > u$, so $s(U'_{<u}) = u$. Hence $U' \subseteq U$ by definition of U, and $p \in U$, a contradiction. So P has a maximal element. □

A.4.3 Using Zorn's lemma

Every application in linear algebra has the same three steps, and the second step rests on one small fact: a union of a chain of sets is "finitary", in that any finitely many of its elements already lie in one member of the chain. Spans and linear relations only ever involve finitely many vectors, so this is what lets linear properties pass to the union.

Technique A.4.5 (Apply Zorn's lemma). To prove that an object with some property exists in a possibly infinite-dimensional setting:

1. Let P be the set of *partial* objects (independent sets, subspaces meeting U only in $\mathbf{0}$, maps defined on part of the space, ...), ordered by inclusion or by extension. Check that $P \neq \emptyset$.
2. Show that every nonempty chain in P has an upper bound in P. The candidate is almost always the union. The work is to show that the union is still in P, and Lemma A.4.6 does most of it.
3. Let m be a maximal element. Show that m is a complete object, by contradiction: if it were not, enlarge it by one vector (or one point), and check that the enlargement is still in P.

Reach for it whenever a finite-dimensional proof proceeds by "add one more until done". The canonical example is Theorem A.4.7; Section 15.1 applies it to linearly independent sets.
Used in. Theorem A.4.7, Exercise A.4.11.

Lemma A.4.6 (Finitely many elements of a union of a chain). Let Γ be a nonempty set of sets that is a chain under $\subseteq$, and let $x_1, ..., x_k$, with $k \geq 0$, be elements of the union $\bigcup \Gamma$. Then some single member of Γ contains all of $x_1, ..., x_k$.

Proof. ROUTINE

Uses. Definition A.4.1 (Partial order), Theorem A.2.1 (Induction, strong induction, and well-ordering)
Hint 1. Each x_i lies in some member C_i. Among finitely many members of a chain, one contains all the others.
Hint 2. Prove the last statement by induction on the number of members.
Sketch. Pick $C_i \in \Gamma$ with $x_i \in C_i$. Finitely many members of a chain have a largest one, by induction: the larger of C_k and the largest of $C_1, ..., C_{k-1}$. It contains every x_i. For $k = 0$, any member works.

For $k = 0$ any member of Γ works, since $\Gamma \neq \emptyset$. For $k \geq 1$, pick $C_i \in \Gamma$ with $x_i \in C_i$ for each i. We show by induction on k that one of $C_1, ..., C_k$ contains all the others. For $k = 1$ this is clear. If C_j contains $C_1, ..., C_{k-1}$, then C_j and C_k are comparable since Γ is a chain, and the larger of the two contains all of $C_1, ..., C_k$. That member contains $x_1, ..., x_k$. □

In Section 4.4, every subspace of a finite-dimensional space was shown to have a complement by extending a basis of the subspace (Proposition 4.4.16). The same statement holds in every vector space, and the proof is a direct application of Zorn's lemma, without bases.

Theorem A.4.7 (Complements exist). Let V be any vector space over F, and U a subspace of V. Then there is a subspace W of V with $V = U \oplus W$.

Intuition. We want W as large as possible subject to $W \cap U = \{\mathbf{0}\}$, because then $U + W$ should be all of V: if some $\mathbf{v}$ were missing, we could add it to W without creating an intersection with U. In finite dimensions "as large as possible" can be achieved by adding one vector at a time. Here Zorn's lemma provides it, once we check that the union of a nested family of such subspaces is again one.

Proof. HARD

Uses. Theorem A.4.4 (Zorn's lemma), Lemma A.4.6 (Finitely many elements of a union of a chain), Theorem 1.4.3 (Subspace criterion), Theorem 4.4.7 (Criteria for direct sums)
Techniques. Apply Zorn's lemma (Technique A.4.5), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Let P be the set of subspaces W with $W \cap U = \{\mathbf{0}\}$, ordered by $\subseteq$. Why is P nonempty?

Hint 2. For a nonempty chain in P, show that the union is a subspace and still meets U only in $\mathbf{0}$. Two vectors of the union lie in one member of the chain.
Hint 3. If W is maximal and $\mathbf{v} \notin U + W$, show that $W + \operatorname{span}(\mathbf{v})$ is in P. If $\mathbf{w} + c\mathbf{v} \in U$, why must $c = 0$?
Sketch. Let $P = \{W \text{ subspace} \mid W \cap U = \{\mathbf{0}\}\}$, which contains $\{\mathbf{0}\}$. The union of a nonempty chain is a subspace, since any two of its vectors lie in a common member, and it meets U only in $\mathbf{0}$; so it is an upper bound. Zorn gives a maximal W. If $\mathbf{v} \notin U + W$, then $W + \operatorname{span}(\mathbf{v})$ still meets U trivially: $\mathbf{w} + c\mathbf{v} \in U$ with $c \neq 0$ would put $\mathbf{v}$ in $U + W$. This contradicts maximality, so $U + W = V$, and the sum is direct.

Let P be the set of subspaces W of V with $W \cap U = \{\mathbf{0}\}$, ordered by inclusion. It contains $\{\mathbf{0}\}$, so it is nonempty.

Let Γ be a nonempty chain in P, and $W^* = \bigcup \Gamma$. We show that $W^* \in P$; it is then an upper bound of Γ. It contains $\mathbf{0}$. If $\mathbf{w}, \mathbf{w'} \in W^*$ and $a \in F$, then by Lemma A.4.6 both lie in one $C \in \Gamma$, which is a subspace, so $\mathbf{w} + \mathbf{w'} \in C \subseteq W^*$ and $a\mathbf{w} \in C \subseteq W^*$. So W^* is a subspace (Theorem 1.4.3). If $\mathbf{w} \in W^* \cap U$, then $\mathbf{w} \in C$ for some $C \in \Gamma$, so $\mathbf{w} \in C \cap U = \{\mathbf{0}\}$. Hence $W^* \in P$.

By Zorn's lemma (Theorem A.4.4), P has a maximal element W. Suppose for contradiction that $U + W \neq V$, and pick $\mathbf{v} \in V \setminus (U + W)$. Let $W' = W + \operatorname{span}(\mathbf{v})$, a subspace containing W, with $W' \neq W$ since $\mathbf{v} \notin W$. We claim $W' \in P$. Let $\mathbf{w} + c\mathbf{v} \in U$ with $\mathbf{w} \in W$ and $c \in F$, say $\mathbf{w} + c\mathbf{v} = \mathbf{u}$. If $c \neq 0$, then $\mathbf{v} = c^{-1}(\mathbf{u} - \mathbf{w}) \in U + W$, which is false. So $c = 0$, and $\mathbf{w} = \mathbf{u} \in W \cap U = \{\mathbf{0}\}$. Hence $W' \cap U = \{\mathbf{0}\}$, so $W' \in P$ and $W \subsetneq W'$, contradicting maximality.

Therefore $U + W = V$ and $U \cap W = \{\mathbf{0}\}$, so $V = U \oplus W$ by Theorem 4.4.7. □

The proof gives no way to write W down. For $V = \mathbb{R}$ over $\mathbb{Q}$ and $U = \mathbb{Q}$, the theorem produces a $\mathbb{Q}$-subspace W of $\mathbb{R}$ with $\mathbb{R} = \mathbb{Q} \oplus W$, so that every real number is uniquely a rational plus an element of W, but nobody can describe such a W explicitly. Such sets are the price of the axiom of choice. In Section 4.6, this is one reason to prefer the quotient V/U to a complement: the quotient is constructed, not chosen.

Exercises

Exercise A.4.8. ROUTINE Without using Zorn's lemma, show that every finite nonempty partially ordered set has a maximal element.

Solution on page 1677

Exercise A.4.9. STANDARD Find the error in the following argument.

Claim: Every nonzero vector space V has a maximal proper subspace. *Proof:* Let P be the set of proper subspaces of V, ordered by $\subseteq$. It contains $\{\mathbf{0}\}$. If Γ is a nonempty chain in P, its union is a subspace by the argument in the proof of Theorem A.4.7, and it contains every member of Γ, so it is an upper bound. By Zorn's lemma, P has a maximal element.

Solution on page 1678

Exercise A.4.10. STANDARD Let V be any vector space, U a subspace, and $\mathbf{v} \in V \setminus U$. Show that there is a linear map $\varphi : V \to F$ with $\varphi(\mathbf{u}) = 0$ for all $\mathbf{u} \in U$ and $\varphi(\mathbf{v}) = 1$. Deduce that every nonzero vector space has a maximal proper subspace.

Solution on page 1678

Exercise A.4.11. HARD Use Zorn's lemma to prove that every surjective function $f : X \to Y$ has a right inverse, without using the axiom of choice directly. (So Zorn's lemma implies the axiom of choice, by Remark A.1.13.)

Solution on page 1678

Appendix B

Complex numbers

Arithmetic, conjugation, modulus, polar form, and roots of unity.

B.1 Complex numbers

Section 1.1.4 built $\mathbb{C}$ as $\mathbb{R}^2$ with a new multiplication and proved that it is a field. This appendix collects what else the book uses about complex numbers. Conjugation and the modulus are the raw material of inner products over $\mathbb{C}$ (Chapter 6). Polar form explains what multiplying by a complex number does to the plane. Roots of unity give the discrete Fourier basis (Section 6.3.4), and square roots give the first cases of the Fundamental Theorem of Algebra (Theorem 7.1.32).

Throughout, when we write $z = a + bi$ or $w = c + di$, the letters a, b, c, d denote real numbers.

B.1.1 Conjugation

The multiplication of $\mathbb{C}$ is forced by the single rule $i^2 = -1$ (Section 1.1.4). But $(-i)^2 = -1$ as well, and nothing in that rule, or in the arithmetic of $\mathbb{R}$, prefers i to $-i$. Exchanging the two is conjugation.

Definition B.1.1 (Complex conjugate). Let $z = a + bi \in \mathbb{C}$. The **complex conjugate** of z is $\overline{z} = a - bi$. We write $\operatorname{Re} z = a$ and $\operatorname{Im} z = b$ for the real and imaginary parts of z (Definition 1.1.29); both are real numbers, and $z = \operatorname{Re} z + i \operatorname{Im} z$. We call z **purely imaginary** if $\operatorname{Re} z = 0$.

For example, $\overline{3 + 2i} = 3 - 2i$, $\overline{i} = -i$, and $\overline{5} = 5$. In the plane, $\overline{z}$ is the mirror image of z in the real axis (Figure B.1.1). As a map $\mathbb{C} \to \mathbb{C}$, conjugation is $\mathbb{R}$-linear but not $\mathbb{C}$-linear (Non-example 2.1.14).

For $z \neq 0$ and an integer $n > 0$ we write $z^{-n} = \left(z^{-1}\right)^n$. The laws $z^m z^n = z^{m+n}$, $(z^m)^n = z^{mn}$, and $(zw)^n = z^n w^n$ then hold for all integers m, n (with $z, w \neq 0$ when an exponent is negative), by the same inductions as for real numbers, and we use them freely.

Proposition B.1.2 (Conjugation respects the field operations). Let $z, w \in \mathbb{C}$.

(a) $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z}\,\overline{w}$.

(b) $\overline{\overline{z}} = z$. In particular, conjugation is a bijection $\mathbb{C} \to \mathbb{C}$ and is its own inverse.

(c) $\overline{z} = z$ if and only if $z \in \mathbb{R}$, and $\overline{z} = -z$ if and only if z is purely imaginary.

(d) $\operatorname{Re} z = (z + \overline{z})/2$ and $\operatorname{Im} z = (z - \overline{z})/(2i)$.

(e) $\overline{-z} = -\overline{z}$. If $z \neq 0$, then $\overline{z} \neq 0$ and $\overline{z^{-1}} = \overline{z}^{-1}$. Consequently $\overline{z^n} = \overline{z}^n$ for every integer n (with $z \neq 0$ if $n < 0$).

Proof. ROUTINE

> *Uses.* Definition B.1.1 (Complex conjugate), Definition 1.1.29 (Complex numbers), Theorem 1.1.9 (Uniqueness of identities and inverses), induction on n
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* Write $z = a + bi$ and $w = c + di$, and compute both sides of (a)–(d) from Definition 1.1.29.
> *Hint 2.* For (e), avoid coordinates. Apply (a) to $z + (-z) = 0$ and to $zz^{-1} = 1$, and use the uniqueness of inverses. Then induct on n.
> *Sketch.* Parts (a)–(d) are coordinate computations; for the product, both sides equal $(ac - bd) - (ad + bc)i$. For (e), conjugating $z + (-z) = 0$ and $zz^{-1} = 1$ with (a) shows that $\overline{-z}$ and $\overline{z^{-1}}$ are the additive and multiplicative inverses of $\overline{z}$. Powers follow from (a) by induction.

Let $z = a + bi$ and $w = c + di$.

(a) $\overline{z + w} = (a + c) - (b + d)i = (a - bi) + (c - di) = \overline{z} + \overline{w}$. Since $zw = (ac - bd) + (ad + bc)i$,

$$\overline{zw} = (ac - bd) - (ad + bc)i = (a - bi)(c - di) = \overline{z}\,\overline{w},$$

where the middle equality is the multiplication rule of Definition 1.1.29 applied to $a + (-b)i$ and $c + (-d)i$: the real part of the product is $ac - (-b)(-d) = ac - bd$, and its imaginary part is $a(-d) + (-b)c = -(ad + bc)$.

(b) $\overline{\overline{z}} = \overline{a - bi} = a + bi = z$. A map that is its own inverse is a bijection.

(c) $\overline{z} = z$ means $-b = b$, that is, $b = 0$, which says that $z = a \in \mathbb{R}$. Likewise $\overline{z} = -z$ means $a = -a$, that is, $\operatorname{Re} z = 0$.

(d) $z + \overline{z} = 2a$ and $z - \overline{z} = 2bi$.

(e) Since $z + (-z) = 0$ and $\overline{0} = 0$, part (a) gives $\overline{z} + \overline{-z} = 0$, so $\overline{-z} = -\overline{z}$ by Theorem 1.1.9. If $z \neq 0$, then $zz^{-1} = 1$ and $\overline{1} = 1$, so (a) gives $\overline{z}\,\overline{z^{-1}} = 1$. Hence $\overline{z} \neq 0$, and $\overline{z^{-1}} = \overline{z}^{-1}$ by Theorem 1.1.9.

For $n \geq 0$ we show $\overline{z^n} = \overline{z}^n$ by induction on n. The case $n = 0$ is $\overline{1} = 1$, and if it holds for n, then by (a)

$$\overline{z^{n+1}} = \overline{z^n z} = \overline{z^n}\,\overline{z} = \overline{z}^n \overline{z} = \overline{z}^{n+1}.$$

For $n < 0$ and $z \neq 0$, the case $-n > 0$ applied to z^{-1} gives $\overline{z^n} = \overline{\left(z^{-1}\right)^{-n}} = \overline{z^{-1}}^{-n} = \left(\overline{z}^{-1}\right)^{-n} = \overline{z}^n$. □

A bijection $\sigma : F \to F$ of a field with $\sigma(a + b) = \sigma(a) + \sigma(b)$ and $\sigma(ab) = \sigma(a)\sigma(b)$ for all $a, b \in F$ is an **automorphism** of F. By (a) and (b), conjugation is an automorphism of $\mathbb{C}$, and by (c) the numbers it fixes are exactly the real ones. Exercise B.1.43 shows that it is the only automorphism of $\mathbb{C}$ other than the identity that fixes every real number.

> *Intuition (Conjugation as a symmetry).* Read Proposition B.1.2 as a symmetry of $\mathbb{C}$. Take any true equation built from complex numbers with sums, products, and inverses. Conjugate both sides and push the bars inward with (a) and (e): every complex number in it is replaced by its conjugate, and every real number stays as it is. So the new equation is also true. This is the precise sense in which i and $-i$ cannot be told apart.

Technique B.1.3 (Conjugate the equation). Conjugation respects sums, products, and inverses, and it fixes real numbers (Proposition B.1.2). So conjugating both sides of a true equation gives another true equation, in which each complex number is replaced by its conjugate and each real number is unchanged. Two uses come up again and again.

- To get a second solution of an equation with real coefficients from a first one, conjugate the first. The canonical example is Proposition B.1.4: the nonreal roots of a real polynomial come in pairs $z, \overline{z}$.
- To show that a complex quantity q is real, show that $\overline{q} = q$; to show that it is purely imaginary, show that $\overline{q} = -q$ (Proposition B.1.2(c)).

Reach for it when real coefficients meet complex unknowns, as when a real matrix acts on $\mathbb{C}^n$, or when a quantity built symmetrically from z and $\overline{z}$ ought to be real. The second use shows that the eigenvalues of a self-adjoint operator are real (Proposition 8.1.1).
Used in. Problem 7.18, Proposition B.1.4, Exercise B.1.40.

Proposition B.1.4 (Roots of real polynomials come in conjugate pairs). Let $p = a_0 + a_1 x + \cdots + a_d x^d$ with $a_0, \ldots, a_d \in \mathbb{R}$, and evaluate p at complex numbers (Definition 1.2.17, with $F = \mathbb{C}$). Then $p(\overline{z}) = \overline{p(z)}$ for every $z \in \mathbb{C}$. In particular, if $p(z) = 0$, then $p(\overline{z}) = 0$.

Proof. ROUTINE

Uses. Proposition B.1.2 (Conjugation respects the field operations), Definition 1.2.17 (Evaluation)
Techniques. Conjugate the equation (Technique B.1.3)
Hint 1. Conjugate the number $p(z) = \sum_k a_k z^k$ and push the bar inward.
Hint 2. The bar passes through sums and products by Proposition B.1.2(a), and $\overline{a_k} = a_k$ because a_k is real.
Sketch. $\overline{p(z)} = \sum_k \overline{a_k}\,\overline{z}^k = \sum_k a_k \overline{z}^k = p(\overline{z})$. If $p(z) = 0$, then $p(\overline{z}) = \overline{0} = 0$.

By Proposition B.1.2(a) and (e), and since $\overline{a_k} = a_k$ for real a_k (Proposition B.1.2(c)),

$$\overline{p(z)} = \overline{\sum_{k=0}^{d} a_k z^k} = \sum_{k=0}^{d} \overline{a_k}\,\overline{z}^k = \sum_{k=0}^{d} a_k \overline{z}^k = p(\overline{z}).$$

If $p(z) = 0$, then $p(\overline{z}) = \overline{0} = 0$. □

Example B.1.5 ($x^4 + 4$). Let $p = x^4 + 4$. Since $(1+i)^2 = 2i$, we get $(1+i)^4 = (2i)^2 = -4$, so $p(1+i) = 0$. By Proposition B.1.4, $p(1-i) = 0$ too. As $p(-z) = p(z)$ for every z, also $p(-1-i) = p(-1+i) = 0$. So p has the four roots $\pm 1 \pm i$, and by Proposition 7.1.22 it has no others. (Example B.1.27 finds them again as the fourth roots of -4.)

The real coefficients matter: $x - i$ has the root i but not $-i$.

Exercise B.1.6. ROUTINE Find all $z \in \mathbb{C}$ with $z^2 = \overline{z}^2$.

Solution on page 1679

B.1.2 The modulus

Definition B.1.7 (Modulus). The **modulus** (or **absolute value**) of $z = a + bi \in \mathbb{C}$ is the real number $|z| = \sqrt{a^2 + b^2}$.

So $|z|$ is the length of the vector (a, b), and $|z - w| = \sqrt{(a-c)^2 + (b-d)^2}$ is the distance between the points z and w of the plane. For a real number $a = a + 0i$, $|a| = \sqrt{a^2}$ is its usual absolute value, so the notation is consistent. For example, $|3 - 4i| = 5$, $|i| = 1$, and $|\cos\theta + i\sin\theta| = 1$ for every $\theta \in \mathbb{R}$.

Proposition B.1.8 (Properties of the modulus). Let $z, w \in \mathbb{C}$.

(a) $|z| \geq 0$, with equality if and only if $z = 0$.

(b) $|z|^2 = z\overline{z}$, and $|\overline{z}| = |-z| = |z|$.

(c) If $z \neq 0$, then $z^{-1} = \overline{z}/|z|^2$.

(d) $|zw| = |z||w|$, and if $w \neq 0$, then $|z/w| = |z|/|w|$.

(e) $|\operatorname{Re} z| \leq |z|$ and $|\operatorname{Im} z| \leq |z|$. Moreover $\operatorname{Re} z \leq |z|$, with equality if and only if z is a nonnegative real number.

(f) $|z + w|^2 = |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2$.

Most of these are one-line computations in coordinates. Parts (d) and (f) are better done without coordinates, by a move worth naming.

> **Technique B.1.9** (Square the modulus). The modulus is a square root, and square roots do not expand. Its square does: $|z|^2 = z\overline{z}$ is a product, which can be multiplied out with the field axioms and conjugated term by term with Proposition B.1.2. Nothing is lost by squaring, since moduli are nonnegative: for real $s, t \geq 0$, $s = t$ if and only if $s^2 = t^2$, and $s \leq t$ if and only if $s^2 \leq t^2$. So to prove an identity or an inequality between moduli, square both sides, write each $|u|^2$ as $u\overline{u}$, and expand.
>
> Reach for it whenever moduli of sums or products appear. The canonical example is $|zw| = |z||w|$ in Proposition B.1.8, where $|zw|^2 = zw\,\overline{z}\,\overline{w} = |z|^2|w|^2$. With $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle$ in place of $|z|^2 = z\overline{z}$, the same move expands norms in inner product spaces (Proposition 6.1.29).
>
> *Used in.* Proposition B.1.8, Theorem B.1.11, Exercise B.1.35, Exercise B.1.36, Exercise B.1.37, Exercise B.1.44.

Proof. ROUTINE

> *Uses.* Definition B.1.7 (Modulus), Proposition B.1.2 (Conjugation respects the field operations), Theorem 1.1.9 (Uniqueness of identities and inverses)
> *Techniques.* Square the modulus (Technique B.1.9)
> *Hint 1.* Parts (a), (b), and (e) are statements about $a^2 + b^2$ for $z = a + bi$. Compute in coordinates.
> *Hint 2.* For (c), multiply z by the proposed inverse and use (b). For (d), compare $|zw|^2$ with $|z|^2|w|^2$, writing each as a number times its conjugate.
> *Hint 3.* For (f), expand $(z + w)(\overline{z} + \overline{w})$ and recognize the two cross terms as a number and its conjugate.
> *Sketch.* Parts (a), (b), and (e) are coordinate facts about $a^2 + b^2$. Then $z \cdot \overline{z}/|z|^2 = 1$ gives (c), and $|zw|^2 = (zw)(\overline{z}\,\overline{w}) = |z|^2|w|^2$ gives (d). For (f), $(z + w)(\overline{z} + \overline{w}) = |z|^2 + (u + \overline{u}) + |w|^2$ with $u = z\overline{w}$, and $u + \overline{u} = 2\operatorname{Re} u$.

Let $z = a + bi$.

(a) $a^2 + b^2 \geq 0$, so the square root is defined and nonnegative. It is 0 if and only if $a^2 + b^2 = 0$, that is, $a = b = 0$.

(b) $z\overline{z} = (a + bi)(a - bi) = \left(a^2 + b^2\right) + (-ab + ba)i = a^2 + b^2 = |z|^2$. Also $|\overline{z}| = \sqrt{a^2 + (-b)^2} = |z|$ and $|-z| = \sqrt{(-a)^2 + (-b)^2} = |z|$.

(c) By (a), $|z|^2$ is a nonzero real number, and by (b), $z \cdot \overline{z}/|z|^2 = |z|^2/|z|^2 = 1$. So $\overline{z}/|z|^2$ is the inverse of z (Theorem 1.1.9).

(d) By (b) and Proposition B.1.2(a),

$$|zw|^2 = (zw)\overline{zw} = (z\overline{z})(w\overline{w}) = |z|^2|w|^2 = (|z||w|)^2.$$

Both $|zw|$ and $|z||w|$ are nonnegative, so they are equal. If $w \neq 0$, then $|z/w||w| = |z|$ by what we just proved, and dividing by $|w| > 0$ gives $|z/w| = |z|/|w|$.

(e) $|a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|$, and likewise $|b| \leq |z|$. Next, $\operatorname{Re} z = a \leq |a| \leq |z|$. The first inequality is an equality if and only if $a \geq 0$, and the second if and only if $b = 0$. So $\operatorname{Re} z = |z|$ if and only if $b = 0$ and $a \geq 0$, that is, z is a nonnegative real number.

(f) By (b) and Proposition B.1.2(a),

$$|z + w|^2 = (z + w)(\overline{z} + \overline{w}) = z\overline{z} + (z\overline{w} + \overline{z}w) + w\overline{w}.$$

By Proposition B.1.2(a) and (b), $\overline{z}w = \overline{z\overline{w}}$, so the middle term is $u + \overline{u} = 2\operatorname{Re} u$ with $u = z\overline{w}$ (Proposition B.1.2(d)). □

In coordinates, (d) says that $\left(a^2 + b^2\right)\left(c^2 + d^2\right) = (ac - bd)^2 + (ad + bc)^2$: a product of two sums of two squares is again a sum of two squares. Exercise B.1.48 puts this to use. Similarly, $\operatorname{Re}(z\overline{w}) = ac + bd$ is the dot product of (a, b) and (c, d), so (f) is the law of cosines in the plane.

Exercise B.1.10. ROUTINE Let $z = (1+2i)/(2-i)$. Find $|z|$ without computing z. Then compute z, and check your answer.

Solution on page 1679

The triangle inequality

The modulus is a length, so it should satisfy the triangle inequality. In the triangle with vertices 0, z, and $z+w$, the sides have lengths $|z|$, $|w|$, and $|z+w|$ (Figure B.1.1(b)). With more terms, the broken path $0, z_1, z_1+z_2, \ldots$ is at least as long as the straight segment from 0 to its endpoint, and the two are equally long only if the path never turns.

Theorem B.1.11 (Triangle inequality in $\mathbb{C}$). Let $n \geq 1$ and $z_1, \ldots, z_n \in \mathbb{C}$.

(a) $|z_1 + \cdots + z_n| \leq |z_1| + \cdots + |z_n|$.

(b) Equality holds in (a) if and only if there is a $u \in \mathbb{C}$ with $|u| = 1$ such that $z_k = |z_k|u$ for every k. That is, all the z_k point in the same direction u.

Intuition (Two ways in). For two terms, Technique B.1.9 applies. By Proposition B.1.8(f), $|z+w|^2 = |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2$, while $(|z|+|w|)^2 = |z|^2 + 2|z\overline{w}| + |w|^2$. So the inequality says $\operatorname{Re}(z\overline{w}) \leq |z\overline{w}|$, which is Proposition B.1.8(e), and induction handles n terms.

Alternatively, rotate the whole picture so that the sum s points along the positive real axis (Technique 6.2.8). Then $|s|$ is the real part of the rotated sum. The real part of a sum is the sum of the real parts, and each real part is at most the modulus of its term. The first argument is the one that generalizes to inner product spaces (Theorem 6.2.11). The second handles all n terms at once and makes the equality case transparent.

Proof (by squaring). STANDARD

Uses. Proposition B.1.8 (Properties of the modulus), induction on n
Techniques. Square the modulus (Technique B.1.9)
Hint 1. Prove the case $n = 2$ first. Both sides are nonnegative, so compare squares, and expand $|z+w|^2$ with Proposition B.1.8(f).
Hint 2. The case $n = 2$ reduces to $\operatorname{Re}(z\overline{w}) \leq |z\overline{w}|$, with equality exactly when $z\overline{w}$ is a nonnegative real number. For n terms, induct, splitting off z_n.
Hint 3. For the equality case with n terms, equality forces equality in both steps of $|s+z_n| \leq |s| + |z_n| \leq \sum_k |z_k|$, where $s = z_1 + \cdots + z_{n-1}$. Treat $s = 0$ separately.
Sketch. For two terms, $|z+w|^2 = |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2 \leq (|z|+|w|)^2$, with equality if and only if $z\overline{w}$ is real and nonnegative. Induction on n gives (a). For (b), "if" is a computation. For "only if", equality for n terms forces equality for $z_1, \ldots, z_{n-1}$, so they point along some u, and for the two terms $s = z_1 + \cdots + z_{n-1} = |s|u$ and z_n. If $s \neq 0$, then $u\overline{z_n} \geq 0$, which gives $z_n = |z_n|u$; if $s = 0$, then $z_1 = \cdots = z_{n-1} = 0$, and the direction of z_n works for all terms.

Two terms. Let $z, w \in \mathbb{C}$. By parts (f), (e), (d), and (b) of Proposition B.1.8,

$$\begin{aligned} |z+w|^2 &= |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\overline{w}| + |w|^2 = (|z|+|w|)^2, \end{aligned}$$

with equality if and only if $\operatorname{Re}(z\overline{w}) = |z\overline{w}|$, that is (by Proposition B.1.8(e)), if and only if $z\overline{w}$ is a nonnegative real number. Both $|z+w|$ and $|z|+|w|$ are nonnegative, so

$$|z+w| \leq |z| + |w|, \quad \text{with equality if and only if } z\overline{w} \in \mathbb{R} \text{ and } z\overline{w} \geq 0. \tag{B.1.1}$$

(a) We induct on n. For $n = 1$ there is nothing to prove. For $n \geq 2$, let $s = z_1 + \cdots + z_{n-1}$. By (B.1.1) and the induction hypothesis,

$$|s + z_n| \leq |s| + |z_n| \leq |z_1| + \cdots + |z_{n-1}| + |z_n|. \tag{B.1.2}$$

(b) Suppose $z_k = |z_k|u$ for every k, with $|u| = 1$. Then $\sum_k z_k = \big(\sum_k |z_k|\big)u$, and by Proposition B.1.8(d) its modulus is $\sum_k |z_k|$.

Conversely, suppose equality holds in (a). We find u by induction on n. For $n = 1$, take $u = z_1/|z_1|$ if $z_1 \neq 0$ and $u = 1$ if $z_1 = 0$. Let $n \geq 2$. Equality between the two ends of (B.1.2) forces equality in both of its steps:

$$|s + z_n| = |s| + |z_n| \quad \text{and} \quad |s| = |z_1| + \cdots + |z_{n-1}|.$$

By the induction hypothesis there is u with $|u| = 1$ and $z_k = |z_k|u$ for $k < n$, and then $s = |s|u$ by the computation in the previous paragraph.

- If $s = 0$, then $|z_1| + \cdots + |z_{n-1}| = 0$, so $z_1 = \cdots = z_{n-1} = 0$. Choose u' for the single number z_n as in the case $n = 1$; then $z_k = |z_k|u'$ for every k.
- If $s \neq 0$, then by the first equality and (B.1.1), $s\overline{z_n} = |s|u\overline{z_n}$ is a nonnegative real number. Dividing by $|s| > 0$, so is $u\overline{z_n}$, and hence $u\overline{z_n} = |u\overline{z_n}| = |z_n|$. Conjugating gives $\overline{u}z_n = |z_n|$, and multiplying by u, with $u\overline{u} = |u|^2 = 1$, gives $z_n = |z_n|u$. □

Proof (by rotating the phase). STANDARD

Uses. Proposition B.1.8 (Properties of the modulus), Definition 1.1.29 (Complex numbers)
Techniques. Rotate the phase (Technique 6.2.8)
Hint 1. Let $s = z_1 + \cdots + z_n$. If $s \neq 0$, multiply s by $\overline{u}$, where $u = s/|s|$. What is $\overline{u}s$?
Hint 2. $|s| = \overline{u}s = \mathrm{Re}(\overline{u}s) = \sum_k \mathrm{Re}(\overline{u}z_k)$. Bound each term with Proposition B.1.8(e).
Hint 3. Equality holds if and only if $\mathrm{Re}(\overline{u}z_k) = |\overline{u}z_k|$ for every k. When does that happen?
Sketch. If $s = 0$, (a) is clear, and equality forces every $z_k = 0$. Otherwise let $u = s/|s|$. Then $|s| = \mathrm{Re}(\overline{u}s) = \sum_k \mathrm{Re}(\overline{u}z_k) \leq \sum_k |\overline{u}z_k| = \sum_k |z_k|$, with equality if and only if each $\overline{u}z_k$ is a nonnegative real number. As $|\overline{u}z_k| = |z_k|$, that means $\overline{u}z_k = |z_k|$, that is, $z_k = |z_k|u$.

If $z_k = |z_k|u$ for all k with $|u| = 1$, then $\sum_k z_k = \big(\sum_k |z_k|\big)u$ has modulus $\sum_k |z_k|$ by Proposition B.1.8(d). This is the "if" part of (b).

Let $s = z_1 + \cdots + z_n$. If $s = 0$, then (a) holds because its right side is nonnegative, and equality means $|z_1| + \cdots + |z_n| = 0$. Then every z_k is 0, and $u = 1$ works.

Suppose $s \neq 0$, and let $u = s/|s|$. Then $|u| = 1$, and $\overline{u}s = \overline{s}s/|s| = |s|$ by Proposition B.1.8(b). The real part of a sum is the sum of the real parts, since addition in $\mathbb{C}$ is coordinatewise (Definition 1.1.29). With parts (e) and (d) of Proposition B.1.8,

$$|s| = \mathrm{Re}(\overline{u}s) = \sum_{k=1}^{n} \mathrm{Re}(\overline{u}z_k) \leq \sum_{k=1}^{n} |\overline{u}z_k| = \sum_{k=1}^{n} |z_k|.$$

This is (a). If equality holds, then $\mathrm{Re}(\overline{u}z_k) = |\overline{u}z_k|$ for every k, because a sum of nonnegative gaps vanishes only if each gap does. By Proposition B.1.8(e), each $\overline{u}z_k$ is then a nonnegative real number, and since $|\overline{u}z_k| = |z_k|$, this says $\overline{u}z_k = |z_k|$. Multiplying by u gives $z_k = |z_k|u$ for every k. □

The first proof found the direction u one term at a time; the second named it at the start, as the direction of the sum, and then checked each term against it.

Corollary B.1.12 (Two terms, and the reverse triangle inequality). Let $z, w \in \mathbb{C}$.

(a) $|z + w| = |z| + |w|$ if and only if one of z, w is a nonnegative real multiple of the other, if and only if $z\overline{w}$ is a nonnegative real number.

(b) $\big||z| - |w|\big| \leq |z - w|$.

Proof. ROUTINE

Uses. Theorem B.1.11 (Triangle inequality in $\mathbb{C}$), Proposition B.1.8 (Properties of the modulus)
Hint 1. For (a), take $n = 2$ in Theorem B.1.11(b), and translate "$z = |z|u$ and $w = |w|u$" into the other two conditions.
Hint 2. For (b), write $z = (z - w) + w$ and apply the triangle inequality; then exchange the roles of z and w.

Sketch. (a) If $z = |z|u$ and $w = |w|u$, then $z\overline{w} = |z||w| \geq 0$. If $z\overline{w} = t \geq 0$ and $w \neq 0$, then $z = \left(t/|w|^2\right)w$. If $z = tw$ with $t \geq 0$ and $w \neq 0$, then $u = w/|w|$ works. (b) $|z| \leq |z - w| + |w|$ and $|w| \leq |w - z| + |z|$.

(a) By Theorem B.1.11(b) with $n = 2$, the equality holds if and only if (i) $z = |z|u$ and $w = |w|u$ for some u with $|u| = 1$. Let (ii) say that $z\overline{w}$ is a nonnegative real number, and (iii) that one of z, w is a nonnegative real multiple of the other. We prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i).

If (i) holds, then $z\overline{w} = |z||w|u\overline{u} = |z||w| \geq 0$.

Suppose $z\overline{w} = t$ with t real and $t \geq 0$. If $w = 0$, then $w = 0 \cdot z$. If $w \neq 0$, multiplying by w gives $z|w|^2 = tw$, so $z = \left(t/|w|^2\right)w$ with $t/|w|^2 \geq 0$.

Suppose finally that $z = tw$ with $t \geq 0$; the case $w = tz$ is the same with z and w exchanged, and (i) is symmetric in z and w. If $w = 0$, then $z = 0$ as well, and $u = 1$ works. If $w \neq 0$, let $u = w/|w|$. Then $w = |w|u$ and $z = t|w|u$, where $t|w| = |tw| = |z|$ by Proposition B.1.8(d).

(b) By Theorem B.1.11, $|z| = |(z - w) + w| \leq |z - w| + |w|$, so $|z| - |w| \leq |z - w|$. Exchanging z and w, and using $|w - z| = |-(z - w)| = |z - w|$ (Proposition B.1.8(b)), gives $|w| - |z| \leq |z - w|$. Together these say $||z| - |w|| \leq |z - w|$. □

With the dot product $z\overline{w}$ on $\mathbb{C} = \mathbb{C}^1$ (Definition 6.1.8), $|z|$ is the norm, and Corollary B.1.12(a) is the case $V = \mathbb{C}$ of Theorem 6.2.11(b). The proofs above use only Proposition B.1.8, so Chapter 6 can rely on them.

B.1.3 The plane and polar form

As a real vector space, $\mathbb{C}$ is $\mathbb{R}^2$ with the basis $1, i$ (Example 1.6.5). So adding complex numbers is adding vectors, by the parallelogram rule, and multiplying by a real number scales (Figure B.1.1). What is new is multiplication by a nonreal number, and to describe it we need polar coordinates.

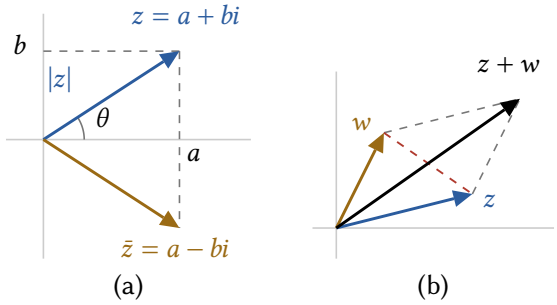


Figure B.1.1. (a) The number $z = a + bi$ as the point (a, b), with its modulus $|z|$ and an argument θ, and its conjugate $\overline{z}$, the mirror image in the real axis. (b) Addition is vector addition. The triangle $0, z, z + w$ has sides of lengths $|z|$, $|w|$, and $|z + w|$, and the red diagonal of the parallelogram has length $|z - w|$.

Remark B.1.13 (Facts about cos *and* sin*).* We use the following facts about the real functions cos and sin, from calculus.

(T1) $\cos^2\theta + \sin^2\theta = 1$, $\cos(-\theta) = \cos\theta$, $\sin(-\theta) = -\sin\theta$, and the angle addition formulas

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \quad \sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta.$$

(T2) $\cos\theta = \cos\theta'$ and $\sin\theta = \sin\theta'$ if and only if $\theta - \theta' \in 2\pi\mathbb{Z} = \{2\pi k \mid k \in \mathbb{Z}\}$.

(T3) For every $(x, y) \in \mathbb{R}^2$ with $x^2 + y^2 = 1$ there is a $\theta \in \mathbb{R}$ with $(x, y) = (\cos\theta, \sin\theta)$.

We used (T1) to show that rotations are linear (Example 2.1.11).

Definition B.1.14 (Polar form and argument). Let $z \in \mathbb{C}$ be nonzero. An **argument** of z is a real number θ with

$$z = |z|(\cos\theta + i\sin\theta).$$

Writing $z = r(\cos\theta + i\sin\theta)$ with $r > 0$ and $\theta \in \mathbb{R}$ is writing z in **polar form**.

Proposition B.1.15 (Existence and uniqueness of polar form). Let $z \in \mathbb{C}$ be nonzero.

(a) z has an argument.

(b) If $z = r(\cos\theta + i\sin\theta)$ with $r > 0$ and $\theta \in \mathbb{R}$, then $r = |z|$, so θ is an argument of z.

(c) If θ is an argument of z, then the arguments of z are exactly the numbers $\theta + 2\pi k$ with $k \in \mathbb{Z}$.

Proof. ROUTINE

Uses. Definition B.1.14 (Polar form and argument), Proposition B.1.8 (Properties of the modulus), Remark B.1.13 (Facts about cos and sin)
Hint 1. For (a), the number $z/|z|$ has modulus 1, so its coordinates (x, y) satisfy $x^2 + y^2 = 1$. Apply (T3).
Hint 2. For (b), take moduli of both sides. For (c), divide by $|z|$ and compare coordinates using (T2).
Sketch. (a) $z/|z| = x + yi$ with $x^2 + y^2 = 1$, and (T3) writes $(x, y) = (\cos\theta, \sin\theta)$. (b) Taking moduli, $|z| = r|\cos\theta + i\sin\theta| = r$. (c) θ' is an argument if and only if $\cos\theta' + i\sin\theta' = z/|z| = \cos\theta + i\sin\theta$, if and only if $\theta' - \theta \in 2\pi\mathbb{Z}$ by (T2).

(a) Let $z/|z| = x + yi$. By Proposition B.1.8(d), $x^2 + y^2 = |z/|z||^2 = 1$, so by (T3) there is a θ with $x = \cos\theta$ and $y = \sin\theta$. Then $z = |z|(\cos\theta + i\sin\theta)$.

(b) By Proposition B.1.8(d) and (T1), $|z| = r|\cos\theta + i\sin\theta| = r\sqrt{\cos^2\theta + \sin^2\theta} = r$.

(c) Dividing by $|z|$, a real number θ' is an argument of z if and only if $\cos\theta' + i\sin\theta' = z/|z| = \cos\theta + i\sin\theta$. Comparing real and imaginary parts, this holds if and only if $\cos\theta' = \cos\theta$ and $\sin\theta' = \sin\theta$, which by (T2) holds if and only if $\theta' - \theta \in 2\pi\mathbb{Z}$. □

So two nonzero complex numbers are equal if and only if they have the same modulus and their arguments differ by a multiple of 2π. The number 0 has no argument.

Example B.1.16 (Polar forms). $1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$, $-2 = 2(\cos\pi + i\sin\pi)$, $-3i = 3\left(\cos\left(-\frac{\pi}{2}\right) + i\sin\left(-\frac{\pi}{2}\right)\right)$, and $\sqrt{3} - i = 2\left(\cos\left(-\frac{\pi}{6}\right) + i\sin\left(-\frac{\pi}{6}\right)\right)$. Any argument may be changed by a multiple of 2π: also $-3i = 3\left(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right)$.

Proposition B.1.17 (Multiplication in polar form). For all $\alpha, \beta \in \mathbb{R}$,

$$(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = \cos(\alpha + \beta) + i\sin(\alpha + \beta).$$

Consequently, if $z = r(\cos\alpha + i\sin\alpha)$ and $w = s(\cos\beta + i\sin\beta)$ with $r, s > 0$, then

$$zw = rs(\cos(\alpha + \beta) + i\sin(\alpha + \beta)),$$
$$z^{-1} = r^{-1}(\cos(-\alpha) + i\sin(-\alpha)),$$
$$\overline{z} = r(\cos(-\alpha) + i\sin(-\alpha)).$$

That is, moduli multiply and arguments add, and inverting or conjugating negates the argument.

Proof. ROUTINE

Uses. Definition 1.1.29 (Complex numbers), Remark B.1.13 (Facts about cos and sin)
Hint 1. Multiply out the left side with $i^2 = -1$ and compare with the angle addition formulas (T1).
Hint 2. For z^{-1}, multiply z by the proposed inverse, using the first part with $\beta = -\alpha$.
Sketch. Expanding gives $(\cos\alpha\cos\beta - \sin\alpha\sin\beta) + (\sin\alpha\cos\beta + \cos\alpha\sin\beta)i$, which is the right side by (T1). The formula for zw follows by regrouping. The proposed inverse times z is $\cos 0 + i\sin 0 = 1$, and $\overline{z} = r(\cos\alpha - i\sin\alpha)$ is the stated number by (T1).

Expanding with $i^2 = -1$,

$$(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + (\sin\alpha\cos\beta + \cos\alpha\sin\beta)i,$$

which is $\cos(\alpha+\beta)+i\sin(\alpha+\beta)$ by (T1). For z and w as stated, $zw = rs(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)$ by commutativity and associativity, which gives the formula for zw. With $\beta = -\alpha$, the first part gives

$$z \cdot r^{-1}(\cos(-\alpha) + i\sin(-\alpha)) = \cos 0 + i\sin 0 = 1,$$

which is the formula for z^{-1}. Finally $\overline{z} = r(\cos\alpha - i\sin\alpha) = r(\cos(-\alpha) + i\sin(-\alpha))$ by (T1). □

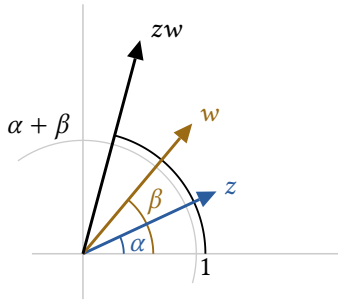


Figure B.1.2. Multiplying $z = |z|(\cos\alpha + i\sin\alpha)$ by $w = |w|(\cos\beta + i\sin\beta)$. The product zw has modulus $|z||w|$ and argument $\alpha+\beta$. The faint arc is part of the unit circle.

The computation in the proof is the one that shows that rotating by β and then by α is rotating by $\alpha+\beta$ (Exercise 2.1.43). The next proposition explains why.

B.1.4 Multiplication as rotation and scaling

Proposition B.1.18 (Multiplication by w as a map of the plane). Let $w = c + di \in \mathbb{C}$, and let $M_w : \mathbb{C} \to \mathbb{C}$, $M_w(z) = wz$. Regard $\mathbb{C}$ as a vector space over $\mathbb{R}$ with the basis $\mathcal{B} = 1, i$.

(a) M_w is $\mathbb{R}$-linear, and

$$[M_w]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}, \quad \text{where} \quad c^2 + d^2 = |w|^2.$$

(b) If $w \neq 0$ and θ is an argument of w, then

$$[M_w]_{\mathcal{B}\mathcal{B}} = |w|\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = |w|[R_\theta]_{\mathcal{E}\mathcal{E}},$$

where R_θ is the rotation of $\mathbb{R}^2$ by θ (Example 2.1.11). So, identifying $a+bi$ with (a,b), multiplication by w is the rotation by θ followed by scaling by $|w|$.

Proof. ROUTINE

Uses. Exercise 2.1.5, Definition 2.3.1 (Matrix of a linear map), Example 2.3.3 (Matrices with respect to standard bases), Definition B.1.14 (Polar form and argument)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Linearity comes from distributivity and commutativity in $\mathbb{C}$. The columns of $[M_w]_{\mathcal{B}\mathcal{B}}$ are the coordinates of $M_w(1)$ and $M_w(i)$.
Hint 2. For (b), substitute $c = |w|\cos\theta$ and $d = |w|\sin\theta$, and recall the matrix of R_θ from Example 2.3.3.
Sketch. M_w is even $\mathbb{C}$-linear (Exercise 2.1.5), hence $\mathbb{R}$-linear. $M_w(1) = c + di$ and $M_w(i) = -d + ci$ give the columns (c,d) and $(-d,c)$. In polar form $c = |w|\cos\theta$ and $d = |w|\sin\theta$, and the matrix becomes $|w|$ times the rotation matrix.

(a) By Exercise 2.1.5 with $F = \mathbb{C}$, M_w is $\mathbb{C}$-linear: $w(z+z') = wz + wz'$ and $w(\lambda z) = \lambda(wz)$ for $\lambda \in \mathbb{C}$. In particular this holds for real λ, so M_w is $\mathbb{R}$-linear. Its values on $\mathcal{B}$ are

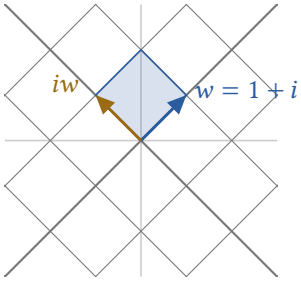
$$M_w(1) = c + di = c\cdot 1 + d\cdot i, \quad M_w(i) = ci + di^2 = (-d)\cdot 1 + c\cdot i,$$

so by Definition 2.3.1 the columns of $[M_w]_{\mathcal{B}\mathcal{B}}$ are (c,d) and $(-d,c)$.

(b) By Definition B.1.14, $c + di = |w|(\cos\theta + i\sin\theta)$, so $c = |w|\cos\theta$ and $d = |w|\sin\theta$. Substituting into (a) gives the first equality, and the second is the matrix of R_θ from Example 2.3.3. The coordinate map $a + bi \mapsto (a, b)$ sends the basis $\mathcal{B}$ to the standard basis $\mathcal{E}$ of $\mathbb{R}^2$, so under it M_w becomes the linear map of $\mathbb{R}^2$ with standard matrix $|w|[R_\theta]_{\mathcal{E}\mathcal{E}}$, namely $|w|R_\theta$. □

Part (a) is also part (a) of Exercise 2.3.35. The determinant of the matrix is $c^2 + d^2 = |w|^2$ (Definition 5.1.14): multiplication by w scales areas by $|w|^2$, as a rotation followed by scaling every direction by $|w|$ should. The shape of the matrix records $M_w(i) = iM_w(1)$: its second column is its first, turned a quarter turn counterclockwise. An $\mathbb{R}$-linear map of the plane has this property exactly when it is $\mathbb{C}$-linear, that is, multiplication by a complex number (Exercise 2.1.47), and exactly when it commutes with the quarter turn $R_{\pi/2}$ (Exercise 2.2.30).

In the plane transformer below, $A = [M_w]_{\mathcal{B}\mathcal{B}}$ for $w = 1 + i$. So $A\mathbf{e}_1 = (1, 1)$ is w and $A\mathbf{e}_2 = (-1, 1)$ is iw. The grid turns by $\pi/4$ and stretches by $|w| = \sqrt{2}$, and $\det A = 2 = |w|^2$. Drag $A\mathbf{e}_1$ to another w, and then $A\mathbf{e}_2$ to its quarter turn iw: the grid stays square, and the determinant is $|w|^2$. In any other position of $A\mathbf{e}_2$ the map is still $\mathbb{R}$-linear, but it is not multiplication by a complex number. Squares become non-square parallelograms, except at $A\mathbf{e}_2 = -iw$, where the map is $z \mapsto w\bar{z}$: it keeps squares square but reverses orientation, and $\det A = -|w|^2$.



Multiplication by $w = 1 + i$ as a map of $\mathbb{C} = \mathbb{R}^2$. The unit square turns by $\pi/4$ and grows by the factor $|w| = \sqrt{2}$ in each direction, so its area becomes $|w|^2 = \det A = 2$. (Interactive in the web version.)

The notation $e^{i\theta}$

Notation B.1.19 ($e^{i\theta}$). For $\theta \in \mathbb{R}$ we write

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

In this notation a nonzero z with argument θ is $z = |z|e^{i\theta}$, and Proposition B.1.17, Proposition B.1.8, and (T2) say that

$$e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}, \quad \overline{e^{i\theta}} = e^{-i\theta} = \left(e^{i\theta}\right)^{-1}, \quad \left|e^{i\theta}\right| = 1,$$

$$e^{i\theta} = 1 \quad \text{if and only if} \quad \theta \in 2\pi\mathbb{Z}.$$

For example, $e^{i\pi/2} = i$, $e^{i\pi} = -1$, and $e^{2\pi i} = 1$. For us $e^{i\theta}$ is an abbreviation, chosen because the first identity is a law of exponents. It is also the value of a genuine exponential function, but we never need that.

ASIDE: *Why an exponential* BEYOND 21-242

The power series $\exp(z) = \sum_{n\geq 0} z^n/n!$ converges for every $z \in \mathbb{C}$, and multiplying two such series, with the binomial theorem, gives $\exp(z + w) = \exp(z)\exp(w)$. Splitting $\exp(i\theta)$ into its even and odd terms and using $i^{2k} = (-1)^k$ gives the Taylor series of $\cos\theta$ and $i\sin\theta$. So $\exp(i\theta) = \cos\theta + i\sin\theta$ (Euler's formula), and $\exp(a + ib) = e^a e^{ib}$ for real a, b. All of this needs convergence of series, which belongs to analysis.

Powers

Theorem B.1.20 (De Moivre's formula). For every $\theta \in \mathbb{R}$ and every integer n,

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta, \quad \text{that is,} \quad \left(e^{i\theta}\right)^n = e^{in\theta}.$$

Consequently $\left(re^{i\theta}\right)^n = r^n e^{in\theta}$ for $r > 0$.

Proof. ROUTINE

Uses. Proposition B.1.17 (Multiplication in polar form), Notation B.1.19 ($e^{i\theta}$), induction on n
Hint 1. For $n \geq 0$, induct on n, using $e^{i(n+1)\theta} = e^{in\theta}e^{i\theta}$.
Hint 2. For $n < 0$, write $\left(e^{i\theta}\right)^n = \left(\left(e^{i\theta}\right)^{-1}\right)^{-n}$ and use $\left(e^{i\theta}\right)^{-1} = e^{i(-\theta)}$.
Sketch. The case $n = 0$ is $1 = 1$, and $e^{in\theta}e^{i\theta} = e^{i(n+1)\theta}$ is the induction step. For negative n, apply the nonnegative case to $-\theta$ and $-n$, since $\left(e^{i\theta}\right)^{-1} = e^{-i\theta}$.

For $n = 0$ both sides are 1. If the formula holds for some $n \geq 0$, then by Proposition B.1.17

$$\left(e^{i\theta}\right)^{n+1} = \left(e^{i\theta}\right)^n e^{i\theta} = e^{in\theta}e^{i\theta} = e^{i(n+1)\theta}.$$

So it holds for every $n \geq 0$. If $n < 0$, then $-n > 0$, and since $\left(e^{i\theta}\right)^{-1} = e^{i(-\theta)}$ (Notation B.1.19), the case already proved, applied to $-\theta$ and $-n$, gives

$$\left(e^{i\theta}\right)^n = \left(\left(e^{i\theta}\right)^{-1}\right)^{-n} = \left(e^{i(-\theta)}\right)^{-n} = e^{i(-\theta)(-n)} = e^{in\theta}.$$

Finally, $\left(re^{i\theta}\right)^n = r^n\left(e^{i\theta}\right)^n = r^n e^{in\theta}$. □

Technique B.1.21 (Use polar form for products, powers, and roots). Addition is easy in the form $a + bi$. Multiplication is easy in polar form, where moduli multiply and arguments add (Proposition B.1.17). So when a problem is about products, quotients, powers, or roots of complex numbers, or about rotations of the plane, write the numbers as $re^{i\theta}$. The problem then splits into a question about positive real numbers (the moduli) and a question about angles modulo 2π (the arguments), and each is usually easy. Converting back to the form $a + bi$ costs a cosine and a sine.

Reach for it to compute a high power, to solve $z^n = w$, or to see what a product does geometrically. The canonical example is the solution of $z^n = w$ in Theorem B.1.26.
Used in. Exercise B.1.23, Theorem B.1.25, Theorem B.1.26, Proposition B.1.29, Exercise B.1.33, Exercise B.1.34.

Example B.1.22 (Powers and multiple angles).

(a) $1 + i = \sqrt{2}e^{i\pi/4}$, so by Theorem B.1.20, $(1+i)^8 = \left(\sqrt{2}\right)^8 e^{2\pi i} = 16$. Compare the repeated squaring in Exercise 1.1.36(d).

(b) $\sqrt{3} - i = 2e^{-i\pi/6}$, so $\left(\sqrt{3} - i\right)^{-6} = 2^{-6}e^{i\pi} = -1/64$.

(c) Expand the left side of $(\cos\theta + i\sin\theta)^3 = \cos 3\theta + i\sin 3\theta$ with the binomial theorem and compare real parts:

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta = 4\cos^3\theta - 3\cos\theta,$$

using $\sin^2\theta = 1 - \cos^2\theta$. The imaginary parts give $\sin 3\theta = 3\cos^2\theta\sin\theta - \sin^3\theta = 3\sin\theta - 4\sin^3\theta$.

Exercise B.1.23. ROUTINE Write $-2 + 2i$ and $-1 - \sqrt{3}i$ in polar form, and compute $(-2+2i)^5/\left(-1-\sqrt{3}i\right)^3$.

Solution on page 1679

B.1.5 Roots of unity

Definition B.1.24 (Roots of unity). Let $n \geq 1$. An **nth root of unity** is a complex number ζ with $\zeta^n = 1$. When n is fixed we write

$$\omega = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n},$$

and ω_n when several values of n are in play.

For $n = 4$, $\omega = i$, and the 4th roots of unity are $1, i, -1, -i$. For $n = 3$, $\omega = -1/2 + \left(\sqrt{3}/2\right)i$, and $\omega^2 = \overline{\omega}$. Every nth root of unity is nonzero, and -1 is an nth root of unity exactly when n is even.

Theorem B.1.25 (The nth roots of unity). Let $n \geq 1$ and $\omega = e^{2\pi i/n}$.

(a) For every integer m, $\omega^m = e^{2\pi i m/n}$, and $\omega^m = 1$ if and only if $n \mid m$. Consequently $\omega^m = \omega^{m'}$ if and only if $n \mid m - m'$.

(b) The nth roots of unity are exactly $1, \omega, \omega^2, \ldots, \omega^{n-1}$, and these n numbers are distinct.

(c) Every nth root of unity ζ satisfies $|\zeta| = 1$ and $\overline{\zeta} = \zeta^{-1}$. Products and inverses of nth roots of unity are nth roots of unity.

(d) If n is even, then $\omega^{n/2} = -1$.

Proof. STANDARD

Uses. Theorem B.1.20 (De Moivre's formula), Notation B.1.19 ($e^{i\theta}$), Proposition B.1.15 (Existence and uniqueness of polar form), Proposition B.1.8 (Properties of the modulus)
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. For (a), apply de Moivre's formula, then ask when $e^{i\varphi} = 1$.
Hint 2. For the hard half of (b), let $\zeta^n = 1$. First show $|\zeta| = 1$, then write $\zeta = e^{i\theta}$ and use the criterion for $e^{in\theta} = 1$.
Hint 3. For (c), $|\zeta|^n = |\zeta^n| = 1$, and $\overline{\zeta}\zeta = |\zeta|^2$.
Sketch. (a) $\omega^m = e^{2\pi i m/n}$ by de Moivre, and $e^{i\varphi} = 1$ exactly when $\varphi \in 2\pi\mathbb{Z}$, which here means $m/n \in \mathbb{Z}$. (b) Each ω^k is a root, since $\left(\omega^k\right)^n = (\omega^n)^k = 1$, and they are distinct by (a). If $\zeta^n = 1$, then $|\zeta| = 1$, so $\zeta = e^{i\theta}$, and $e^{in\theta} = 1$ gives $n\theta = 2\pi k$, so $\zeta = \omega^k$; reduce k modulo n. (c) $|\zeta|^n = 1$ forces $|\zeta| = 1$, so $\overline{\zeta} = \zeta^{-1}$. (d) $\omega^{n/2} = e^{i\pi} = -1$.

(a) By Theorem B.1.20, $\omega^m = e^{i(2\pi m/n)} = e^{2\pi i m/n}$. By Notation B.1.19 this is 1 if and only if $2\pi m/n \in 2\pi\mathbb{Z}$, that is, $m/n \in \mathbb{Z}$, that is, $n \mid m$. Multiplying by $\omega^{-m'}$, $\omega^m = \omega^{m'}$ if and only if $\omega^{m-m'} = 1$, if and only if $n \mid m - m'$.

(b) For $0 \leq k \leq n-1$, $\left(\omega^k\right)^n = (\omega^n)^k = 1$ by (a), so each ω^k is an nth root of unity. If $0 \leq j < k \leq n-1$, then $0 < k - j < n$, so $n \nmid k - j$, and $\omega^j \neq \omega^k$ by (a).

Conversely, let $\zeta^n = 1$. Then $\zeta \neq 0$, and $|\zeta|^n = |\zeta^n| = 1$ by Proposition B.1.8(d). A positive real number r with $r^n = 1$ equals 1, since $r < 1$ gives $r^n < 1$ and $r > 1$ gives $r^n > 1$. So $|\zeta| = 1$, and $\zeta = e^{i\theta}$ for an argument θ of ζ (Proposition B.1.15). By Theorem B.1.20, $e^{in\theta} = \zeta^n = 1$, so $n\theta = 2\pi k$ for some $k \in \mathbb{Z}$, and $\zeta = e^{2\pi ik/n} = \omega^k$. Dividing k by n with remainder, $k = qn + r$ with $0 \leq r \leq n-1$, and $\omega^k = \omega^r$ by (a). So ζ is one of $1, \omega, \ldots, \omega^{n-1}$.

(c) As in (b), $|\zeta| = 1$, so $\overline{\zeta} = \zeta^{-1}$ by Proposition B.1.8(c). If $\zeta^n = \eta^n = 1$, then $(\zeta\eta)^n = \zeta^n\eta^n = 1$ and $\left(\zeta^{-1}\right)^n = (\zeta^n)^{-1} = 1$.

(d) By (a), $\omega^{n/2} = e^{2\pi i(n/2)/n} = e^{i\pi} = -1$. □

The roots $\omega^k = e^{2\pi ik/n}$ lie on the unit circle at the angles $2\pi k/n$, equally spaced. So for $n \geq 3$ they are the vertices of a regular n-gon inscribed in the unit circle, with one vertex at 1 (Figure B.1.3(a)). Multiplying by ω rotates the plane by $2\pi/n$ (Proposition B.1.18), which carries each vertex ω^k to the next one, ω^{k+1}, and ω^{n-1} back to $\omega^n = 1$. Parts (a)–(d) are the facts about ω behind the discrete Fourier basis (Section 6.3.4).

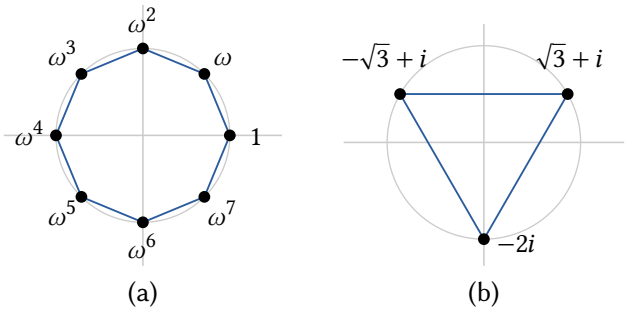


Figure B.1.3. (a) The 8th roots of unity ω^k, $\omega = e^{2\pi i/8}$, are the vertices of a regular octagon; multiplication by ω turns each vertex into the next. (b) The cube roots of $8i$ (Example B.1.27) form an equilateral triangle inscribed in the circle of radius 2.

Every nonzero number has n nth roots, by the same method.

Theorem B.1.26 (nth roots). Let $n \geq 1$, and let $w = re^{i\varphi}$ with $r > 0$ and $\varphi \in \mathbb{R}$. The equation $z^n = w$ has exactly n solutions in $\mathbb{C}$, namely

$$z_k = r^{1/n} e^{i(\varphi+2\pi k)/n}, \quad k = 0, 1, \ldots, n-1,$$

where $r^{1/n}$ is the positive real nth root of r. If z_* is any one solution, the solutions are exactly the numbers $z_*\zeta$ with ζ an nth root of unity. For $n \geq 3$ they are the vertices of a regular n-gon inscribed in the circle of radius $r^{1/n}$ about 0.

Proof. STANDARD

Uses. Theorem B.1.20 (De Moivre's formula), Theorem B.1.25 (The nth roots of unity), Corollary 1.1.14 (Cancellation), Proposition B.1.17 (Multiplication in polar form)
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. Check with de Moivre's formula that z_0 is a solution.
Hint 2. If z_* and z are both solutions, what equation does z/z_* satisfy?
Sketch. By de Moivre, $z_0^n = re^{i\varphi} = w$. If z_* is one solution, then $z^n = w$ if and only if $(z/z_*)^n = 1$, that is, $z = z_*\zeta$ for an n th root of unity ζ, and there are exactly n of these by Theorem B.1.25. With $z_* = z_0$ this is the list $z_k = z_0\omega^k$.

By Theorem B.1.20, $z_0^n = \left(r^{1/n}\right)^n e^{i\varphi} = w$, so solutions exist. Let z_* be any solution. Since $w \neq 0$, also $z_* \neq 0$. For $z \in \mathbb{C}$, $z^n = w$ if and only if $(z/z_*)^n = z^n/w = 1$, that is, z/z_* is an nth root of unity. So the solutions are the numbers $z_*\zeta$ with $\zeta^n = 1$. By Theorem B.1.25(b) there are exactly n such ζ, namely $1, \omega, \ldots, \omega^{n-1}$ with $\omega = e^{2\pi i/n}$, and the numbers $z_*\omega^k$ are distinct because $z_* \neq 0$ (Corollary 1.1.14). Taking $z_* = z_0$ and using Proposition B.1.17,

$$z_0\omega^k = r^{1/n}e^{i\varphi/n}e^{2\pi ik/n} = r^{1/n}e^{i(\varphi+2\pi k)/n} = z_k.$$

All the z_k have modulus $r^{1/n}$, and their arguments $(\varphi + 2\pi k)/n$ are spaced $2\pi/n$ apart, so for $n \geq 3$ they are the vertices of a regular n-gon. □

For $w = 0$, the only solution of $z^n = 0$ is $z = 0$, by Proposition 1.1.17.

Example B.1.27 (Cube roots of $8i$, fourth roots of -4).

(a) $8i = 8e^{i\pi/2}$, so the cube roots of $8i$ are $2e^{i(\pi/2+2\pi k)/3}$ for $k = 0, 1, 2$, with arguments $\pi/6, 5\pi/6$, and $3\pi/2$:

$$2e^{i\pi/6} = \sqrt{3} + i, \quad 2e^{5\pi i/6} = -\sqrt{3} + i, \quad 2e^{3\pi i/2} = -2i.$$

For instance, $(-2i)^3 = -8i^3 = 8i$. They form an equilateral triangle (Figure B.1.3(b)).

(b) $-4 = 4e^{i\pi}$, so the fourth roots of -4 are $\sqrt{2}e^{i(\pi+2\pi k)/4}$ for $k = 0, 1, 2, 3$. That is, they are $\sqrt{2}e^{i\pi/4} = 1 + i$ times the fourth roots of unity $1, i, -1, -i$: the numbers $\pm 1 \pm i$ of Example B.1.5.

The sum of the roots of unity

The nth roots of unity are the vertices of a regular polygon centred at 0, so their average, the centre of mass of the vertices, should be 0. Rotating by $2\pi/n$ permutes the vertices, so their sum S satisfies $\omega S = S$, and as $\omega \neq 1$, this forces $S = 0$. Here are two proofs, together with the product of the roots.

Proposition B.1.28 (Sum and product of the roots of unity). Let $n \geq 2$ and $\omega = e^{2\pi i/n}$. The nth roots of unity add up to 0, and their product is $(-1)^{n-1}$:

$$\sum_{k=0}^{n-1} \omega^k = 0, \quad \prod_{k=0}^{n-1} \omega^k = (-1)^{n-1}.$$

Proof (factor and cancel). ROUTINE

Uses. Theorem B.1.25 (The nth roots of unity), Proposition 1.1.17 (No zero divisors)
Hint 1. Multiply the sum $S = 1 + \omega + \cdots + \omega^{n-1}$ by $\omega - 1$. The products telescope.
Hint 2. For the product, add the exponents: $0 + 1 + \cdots + (n-1) = n(n-1)/2$. Treat odd and even n separately, using Theorem B.1.25(a) and (d).
Sketch. $(\omega - 1)S = \omega^n - 1 = 0$, and $\omega \neq 1$ since $n \geq 2$, so $S = 0$. The product is $\omega^{n(n-1)/2}$. For odd n, n divides $n(n-1)/2$, so the product is 1; for even n, it is $\left(\omega^{n/2}\right)^{n-1} = (-1)^{n-1}$.

Let $S = 1 + \omega + \cdots + \omega^{n-1}$. Then

$$(\omega - 1)S = \left(\omega + \omega^2 + \cdots + \omega^n\right) - \left(1 + \omega + \cdots + \omega^{n-1}\right) = \omega^n - 1 = 0.$$

Since $n \geq 2$, $n \nmid 1$, so $\omega \neq 1$ by Theorem B.1.25(a), and Proposition 1.1.17 gives $S = 0$.

The product is $\omega^{0+1+\cdots+(n-1)} = \omega^{n(n-1)/2}$. If n is odd, then $(n-1)/2$ is an integer, so n divides $n(n-1)/2$, and the product is $1 = (-1)^{n-1}$ by Theorem B.1.25(a). If n is even, then $n(n-1)/2 = (n/2)(n-1)$, and by Theorem B.1.25(d) the product is $\left(\omega^{n/2}\right)^{n-1} = (-1)^{n-1}$. □

Proof (compare coefficients). STANDARD

Uses. Theorem B.1.25 (The nth roots of unity), the factor theorem: a polynomial of degree n with leading coefficient 1 and n distinct roots $c_1, \ldots, c_n$ equals $(x - c_1)\cdots(x - c_n)$ (Section 7.1)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. The nth roots of unity are exactly the roots of the polynomial $x^n - 1$, and there are n of them. Factor $x^n - 1$.
Hint 2. In $(x - c_1)\cdots(x - c_n)$, the coefficient of x^{n-1} is $-(c_1 + \cdots + c_n)$ and the constant term is $(-1)^n c_1 \cdots c_n$.
Sketch. By Theorem B.1.25, $x^n - 1$ has the n distinct roots ω^k, so $x^n - 1 = \prod_k \left(x - \omega^k\right)$. Comparing the coefficients of x^{n-1} (which is 0 on the left, as $n \geq 2$) and the constant terms gives $-\sum_k \omega^k = 0$ and $(-1)^n \prod_k \omega^k = -1$.

The polynomial $x^n - 1 \in \mathbb{C}[x]$ has degree n and leading coefficient 1, and by Theorem B.1.25(b) it has the n distinct roots $1, \omega, \ldots, \omega^{n-1}$. By the factor theorem (Section 7.1),

$$x^n - 1 = (x - 1)(x - \omega)\cdots\left(x - \omega^{n-1}\right).$$

Multiplying out the right side, each term picks x or $-\omega^k$ from each factor. The terms of degree $n - 1$ pick $-\omega^k$ from exactly one factor, so the coefficient of x^{n-1} is $-\left(1 + \omega + \cdots + \omega^{n-1}\right)$. The constant term picks $-\omega^k$ from every factor, so it is $(-1)^n \omega^0 \omega^1 \cdots \omega^{n-1}$. On the left, the coefficient of x^{n-1} is 0, since $n - 1 \geq 1$, and the constant term is -1. Comparing,

$$\sum_{k=0}^{n-1} \omega^k = 0, \quad \prod_{k=0}^{n-1} \omega^k = (-1)^{n+1} = (-1)^{n-1}.$$

□

The first proof needs nothing about polynomials. The second reads both identities off one factorization, the way Vieta's formulas relate the coefficients of a polynomial to its roots, and the other coefficients of $x^n - 1$ give more identities of the same kind (Exercise B.1.45 uses the factorization again).

The first proof applies to any nth root of unity $\zeta \neq 1$: since $(\zeta - 1)\big(1 + \zeta + \cdots + \zeta^{n-1}\big) = \zeta^n - 1 = 0$, we get $1 + \zeta + \cdots + \zeta^{n-1} = 0$. With $\zeta = \omega^m$ this is Lemma 6.3.39, the computation behind the orthogonality of the discrete Fourier basis.

B.1.6 Square roots and quadratic equations

$\mathbb{R}$ was enlarged to $\mathbb{C}$ to give -1 a square root. Every complex number then has a square root without any further enlargement, and so every quadratic equation over $\mathbb{C}$ can be solved.

Proposition B.1.29 (Square roots in $\mathbb{C}$). Every $w \in \mathbb{C}$ has a square root in $\mathbb{C}$. If $w \neq 0$, it has exactly two square roots, z_* and $-z_*$, where z_* is any one of them. The only square root of 0 is 0.

Proof (by polar form). ROUTINE

Uses. Theorem B.1.26 (nth roots), Theorem B.1.25 (The nth roots of unity), Proposition 1.1.17 (No zero divisors)
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. Apply Theorem B.1.26 with $n = 2$.
Hint 2. The square roots of unity are 1 and $e^{i\pi} = -1$.
Sketch. For $w \neq 0$, Theorem B.1.26 with $n = 2$ gives exactly the two solutions $z_*\zeta$ with $\zeta \in \{1, -1\}$. For $w = 0$, $z^2 = 0$ forces $z = 0$.

If $w = 0$, then $z^2 = 0$ forces $z = 0$ (Proposition 1.1.17). If $w \neq 0$, Theorem B.1.26 with $n = 2$ gives exactly two solutions of $z^2 = w$, namely $z_*\zeta$ for the two square roots of unity ζ. By Theorem B.1.25 with $n = 2$, these are 1 and $\omega^{2/2} = -1$, so the solutions are z_* and $-z_*$. Explicitly, if $w = re^{i\varphi}$, then $z_* = \sqrt{r}e^{i\varphi/2}$ is one of them. □

Proof (by solving for the real and imaginary parts). STANDARD

Uses. Proposition B.1.8 (Properties of the modulus), Proposition 1.1.17 (No zero divisors), Definition 1.1.29 (Complex numbers)
Hint 1. Write $w = a + bi$ and look for $z = x + yi$. Then $z^2 = w$ says $x^2 - y^2 = a$ and $2xy = b$.
Hint 2. Taking moduli, $x^2 + y^2 = |w|$. Combine this with $x^2 - y^2 = a$ to find x^2 and y^2, then fix the signs with $2xy = b$.
Hint 3. For the count, factor $z^2 - z_*^2$.
Sketch. A solution must satisfy $x^2 - y^2 = a$ and $x^2 + y^2 = |w|$, so $x^2 = (|w| + a)/2$ and $y^2 = (|w| - a)/2$, both nonnegative because $|a| \leq |w|$. Taking the nonnegative root for x and giving y the sign of b (either sign if $b = 0$), one checks that $x^2 - y^2 = a$ and $2xy = b$. Finally, $z^2 = z_*^2$ if and only if $(z - z_*)(z + z_*) = 0$.

Let $w = a + bi$. We first find a candidate. If $z = x + yi$ satisfies $z^2 = w$, then $x^2 - y^2 = a$ and $2xy = b$, and taking moduli, $x^2 + y^2 = |z|^2 = |z^2| = |w|$ (Proposition B.1.8(d)). Hence

$$x^2 = \frac{|w| + a}{2}, \quad y^2 = \frac{|w| - a}{2}.$$

Both right sides are nonnegative, since $|a| \leq |w|$ (Proposition B.1.8(e)). So let $\sigma = 1$ if $b \geq 0$ and $\sigma = -1$ if $b < 0$, and put

$$z_* = \sqrt{\frac{|w| + a}{2}} + \sigma\sqrt{\frac{|w| - a}{2}}\, i. \tag{B.1.3}$$

We check that $z_*^2 = w$. The real part of z_*^2 is $(|w| + a)/2 - (|w| - a)/2 = a$, and its imaginary part is

$$2\sigma\sqrt{\frac{|w| + a}{2}}\sqrt{\frac{|w| - a}{2}} = \sigma\sqrt{|w|^2 - a^2} = \sigma\sqrt{b^2} = \sigma|b| = b.$$

So every w has a square root. Now if $z_*^2 = w$, then $z^2 = w$ if and only if $(z - z_*)(z + z_*) = z^2 - z_*^2 = 0$, which by Proposition 1.1.17 holds if and only if $z = z_*$ or $z = -z_*$. These differ unless $2z_* = 0$, that is, unless $z_* = 0$, which happens exactly when $w = 0$. □

The second proof needs no trigonometry: (B.1.3) computes a square root from two real square roots. For $w = 3 + 4i$ we have $|w| = 5$, so $z_* = \sqrt{4} + \sqrt{1}\,i = 2 + i$, and indeed $(2+i)^2 = 3 + 4i$. For $w = i$ it gives $z_* = (1+i)/\sqrt{2}$, as in Exercise 1.1.44(b).

Theorem B.1.30 (Quadratic equations over $\mathbb{C}$). Let $a, b, c \in \mathbb{C}$ with $a \neq 0$, let $\Delta = b^2 - 4ac$, and let s be a square root of Δ (Proposition B.1.29). Then the solutions of $az^2 + bz + c = 0$ are

$$z_1 = \frac{-b+s}{2a}, \quad z_2 = \frac{-b-s}{2a},$$

which are distinct if $\Delta \neq 0$ and equal if $\Delta = 0$. Moreover, $az^2 + bz + c = a(z - z_1)(z - z_2)$ for every $z \in \mathbb{C}$.

Proof. ROUTINE

Uses. Proposition B.1.29 (Square roots in $\mathbb{C}$), Exercise 1.1.50
Hint 1. Complete the square: multiply the equation by $4a$.
Hint 2. $4a\big(az^2 + bz + c\big) = (2az + b)^2 - \Delta$. Now use that the square roots of Δ are $\pm s$.
Hint 3. For the factorization, compute $z_1 + z_2$ and $z_1 z_2$.
Sketch. Since $4a\big(az^2 + bz + c\big) = (2az + b)^2 - \Delta$ and $4a \neq 0$, z is a solution if and only if $2az + b$ is a square root of Δ, that is, $2az + b = \pm s$. The roots coincide exactly when $s = 0$. Finally $z_1 + z_2 = -b/a$ and $z_1 z_2 = \big(b^2 - s^2\big)/\big(4a^2\big) = c/a$.

Expanding, $(2az + b)^2 - \Delta = 4a^2z^2 + 4abz + 4ac = 4a\big(az^2 + bz + c\big)$. Since $4a \neq 0$, z is a solution if and only if $(2az + b)^2 = \Delta$, that is, $2az + b$ is a square root of Δ. By Proposition B.1.29 the square roots of Δ are s and $-s$, so z is a solution if and only if $2az + b = \pm s$, that is, $z = z_1$ or $z = z_2$. These coincide if and only if $s = -s$, that is, $s = 0$, that is, $\Delta = 0$. Finally $z_1 + z_2 = -b/a$ and

$$z_1 z_2 = \frac{(-b)^2 - s^2}{4a^2} = \frac{4ac}{4a^2} = \frac{c}{a},$$

so $a(z - z_1)(z - z_2) = az^2 - a(z_1 + z_2)z + az_1z_2 = az^2 + bz + c$. □

This is Exercise 1.1.50(a) for the field $\mathbb{C}$, where the hypothesis that $b^2 - 4ac$ is a square always holds.

Example B.1.31 (A quadratic with complex coefficients). To solve $z^2 - 3z + (3 - i) = 0$, compute $\Delta = 9 - 4(3 - i) = -3 + 4i$. Here $|\Delta| = 5$, and (B.1.3) gives the square root $s = \sqrt{1} + \sqrt{4}\,i = 1 + 2i$. So the solutions are

$$z = \frac{3 \pm (1 + 2i)}{2}, \quad \text{that is,} \quad z = 2 + i \quad \text{and} \quad z = 1 - i.$$

Check: their sum is 3 and their product is $(2 + i)(1 - i) = 3 - i$. The roots are not conjugate to each other, which is allowed because the coefficients are not real (compare Proposition B.1.4).

Remark B.1.32 (The Fundamental Theorem of Algebra). Theorem B.1.26 and Theorem B.1.30 solve two kinds of polynomial equations over $\mathbb{C}$. The Fundamental Theorem of Algebra (Theorem 7.1.32) says that every nonconstant polynomial with complex coefficients has a root in $\mathbb{C}$. Over $\mathbb{R}$ this fails already for $x^2 + 1$, and $\mathbb{C}$ was built by adjoining a root of that single polynomial; the theorem says that no polynomial ever forces a further enlargement. Together with the factor theorem, it factors every polynomial in $\mathbb{C}[x]$ into linear factors, and together with Proposition B.1.4, it factors every polynomial in $\mathbb{R}[x]$ into real factors of degree at most 2 (Section 7.1). For degree 5 and higher there is no formula in radicals like Theorem B.1.30, and the proofs in Section C.1 show that a root exists without producing it. In linear algebra, the theorem is what guarantees that every operator on a nonzero finite-dimensional complex vector space has an eigenvalue (Theorem 7.3.27).

Exercises

Exercise B.1.33. ROUTINE Write each of the following in the form $a + bi$: (a) $(1 - i)^{10}$; (b) $\left(1 + \sqrt{3}i\right)^{-4}$; (c) $(\cos(\pi/9) + i\sin(\pi/9))^{12}$; (d) $((1 + i)/(1 - i))^{2027}$.

Solution on page 1680

Exercise B.1.34. ROUTINE Find all $z \in \mathbb{C}$ with $z^6 = -64$, in the form $a + bi$, and draw them. Which of them satisfy $z^2 = -4$?

Solution on page 1680

Exercise B.1.35. ROUTINE Let $z, w \in \mathbb{C}$.

(a) Show that $|z+w|^2 + |z-w|^2 = 2|z|^2 + 2|w|^2$, and interpret this in the parallelogram with vertices $0, z, w, z+w$.

(b) Show that $|z+w|^2 - |z-w|^2 = 4\,\mathrm{Re}(z\overline{w})$. Deduce that the two diagonals of the parallelogram have the same length if and only if $\mathrm{Re}(z\overline{w}) = 0$, and interpret this condition geometrically.

Solution on page 1680

Exercise B.1.36. STANDARD Let $z, w \in \mathbb{C}$.

(a) Show that $|1 - \overline{z}w|^2 - |z-w|^2 = \left(1 - |z|^2\right)\left(1 - |w|^2\right)$.

(b) Deduce that if $|z| < 1$ and $|w| < 1$, then $1 - \overline{z}w \neq 0$ and

$$\left|\frac{z-w}{1-\overline{z}w}\right| < 1.$$

(c) Show that if $|w| = 1$ and $\overline{z}w \neq 1$, then $|(z-w)/(1-\overline{z}w)| = 1$.

Solution on page 1681

Exercise B.1.37. STANDARD Describe each of the following sets geometrically.

(a) $\{z \in \mathbb{C} \mid |z-1| = |z+i|\}$.

(b) $\{z \in \mathbb{C} \mid |z-1| = 2|z+1|\}$.

(c) $\{z \in \mathbb{C} \mid z \neq -1, \mathrm{Re}((z-1)/(z+1)) = 0\}$.

Solution on page 1681

Exercise B.1.38. ROUTINE Prove or disprove each statement, for all $z, w \in \mathbb{C}$.

(a) $\mathrm{Re}(zw) = \mathrm{Re}\, z\, \mathrm{Re}\, w$.

(b) $|z^2| = |z|^2$.

(c) If z^2 is real, then z is real.

(d) If $|z| = |w|$, then $z = w$ or $z = -w$.

(e) If $|z+w| = |z-w|$, then $\mathrm{Re}(z\overline{w}) = 0$.

(f) If $z \neq 0$, then $\overline{1/z} = 1/\overline{z}$.

Solution on page 1682

Exercise B.1.39. ROUTINE Find the error in the following argument.

Claim. $1 = -1$.

Proof. By de Moivre's formula with $n = 1/2$, $(\cos\theta + i\sin\theta)^{1/2} = \cos(\theta/2) + i\sin(\theta/2)$ for every $\theta \in \mathbb{R}$. Now $1 = \cos 0 + i\sin 0 = \cos 2\pi + i\sin 2\pi$. Taking $\theta = 0$ gives $1^{1/2} = \cos 0 + i\sin 0 = 1$, and taking $\theta = 2\pi$ gives $1^{1/2} = \cos\pi + i\sin\pi = -1$. So $1 = -1$.

Solution on page 1682

Exercise B.1.40. STANDARD For $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$, write $\overline{\mathbf{z}} = (\overline{z_1}, \ldots, \overline{z_n})$. Let $A \in \mathbb{R}^{m \times n}$.

(a) Show that $A\overline{\mathbf{z}} = \overline{A\mathbf{z}}$ for every $\mathbf{z} \in \mathbb{C}^n$.

(b) Let $m = n$, $\lambda \in \mathbb{C}$, and $\mathbf{v} \in \mathbb{C}^n$ with $A\mathbf{v} = \lambda\mathbf{v}$. Show that $A\overline{\mathbf{v}} = \overline{\lambda}\,\overline{\mathbf{v}}$.

(c) Let $\mathbf{b} \in \mathbb{R}^m$. Show that the set of solutions $\mathbf{z} \in \mathbb{C}^n$ of $A\mathbf{z} = \mathbf{b}$ is closed under conjugation, and that if it is nonempty, it contains a vector in $\mathbb{R}^n$.

Solution on page 1682

Exercise B.1.41. STANDARD

(a) Find $A \in \mathbb{R}^{2\times 2}$ with $A^3 = -I_2$ but $A \neq -I_2$.

(b) For each $n \geq 2$, find $A \in \mathbb{R}^{2\times 2}$ with $A^n = I_2$ but $A^k \neq I_2$ for $1 \leq k < n$.

Solution on page 1683

Exercise B.1.42. STANDARD For $z = a + bi \in \mathbb{C}$, let $\varphi(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathbb{R}^{2\times 2}$, and let $K = \{\varphi(z) \mid z \in \mathbb{C}\}$.

(a) Show that $\varphi(z + w) = \varphi(z) + \varphi(w)$, $\varphi(zw) = \varphi(z)\varphi(w)$, and $\varphi(1) = I_2$, and that φ is injective.

(b) Show that $\varphi(\overline{z}) = \varphi(z)^{\mathsf{T}}$ and $\det \varphi(z) = |z|^2$.

(c) Show that K, with matrix addition and multiplication, is a field, and that $\varphi(z)^{-1} = \varphi\big(z^{-1}\big)$ for $z \neq 0$.

Solution on page 1683

Exercise B.1.43. STANDARD Let $\sigma : \mathbb{C} \to \mathbb{C}$ be an automorphism of $\mathbb{C}$ (Section B.1.1) with $\sigma(t) = t$ for every $t \in \mathbb{R}$. Show that σ is either the identity or conjugation.

Solution on page 1684

Exercise B.1.44. STANDARD Let $z_1, z_2, z_3 \in \mathbb{C}$ with $|z_1| = |z_2| = |z_3| = 1$ and $z_1 + z_2 + z_3 = 0$. Show that $\{z_1, z_2, z_3\} = \{u, u\omega, u\omega^2\}$ for some u with $|u| = 1$, where $\omega = e^{2\pi i/3}$. Conclude that z_1, z_2, z_3 are the vertices of an equilateral triangle.

Solution on page 1684

Exercise B.1.45. HARD Let $n \geq 2$ and $\omega = e^{2\pi i/n}$.

(a) Show that $1 + x + \cdots + x^{n-1} = (x - \omega)\big(x - \omega^2\big)\cdots\big(x - \omega^{n-1}\big)$ in $\mathbb{C}[x]$.

(b) Deduce that $(1 - \omega)\big(1 - \omega^2\big)\cdots\big(1 - \omega^{n-1}\big) = n$.

(c) Show that the product of the distances from one vertex of a regular n-gon inscribed in the unit circle to the other $n - 1$ vertices is n. Deduce that

$$\prod_{k=1}^{n-1} \sin\frac{k\pi}{n} = \frac{n}{2^{n-1}}.$$

Solution on page 1684

Exercise B.1.46. STANDARD An nth root of unity ζ is **primitive** if $\zeta^m \neq 1$ for $1 \leq m < n$. Let $\omega = e^{2\pi i/n}$.

(a) Show that for $0 \le k \le n-1$, ω^k is primitive if and only if $\gcd(k,n) = 1$.

(b) Show that if ζ is primitive, then $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$ are all the nth roots of unity.

(c) List the primitive 12th roots of unity in the form $a + bi$.

Solution on page 1685

Exercise B.1.47. HARD Let $n \ge 1$. Show that $x^2 + x + 1$ divides $x^{2n} + x^n + 1$ in $\mathbb{R}[x]$ if and only if $3 \nmid n$.

Solution on page 1685

Exercise B.1.48. STANDARD BEYOND 21-242 The **Gaussian integers** are $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$, and the **norm** of $z \in \mathbb{Z}[i]$ is $N(z) = |z|^2$.

(a) Show that $\mathbb{Z}[i]$ is closed under addition, negation, and multiplication, that $N(z)$ is a nonnegative integer, and that $N(zw) = N(z)N(w)$.

(b) A **unit** of $\mathbb{Z}[i]$ is a $z \in \mathbb{Z}[i]$ with $zw = 1$ for some $w \in \mathbb{Z}[i]$. Show that the units are $1, -1, i, -i$.

(c) Starting from $5 = N(2+i)$ and $13 = N(3+2i)$, write 65 as a sum of two squares of integers in two different ways.

Solution on page 1686

Exercise B.1.49. CHALLENGE Let $z_1, \ldots, z_n \in \mathbb{C}$. Show that there is a set $S \subseteq \{1, \ldots, n\}$ with

$$\left| \sum_{k \in S} z_k \right| \ge \frac{1}{\pi} \sum_{k=1}^{n} |z_k|.$$

Solution on page 1686

Appendix C

The Fundamental Theorem of Algebra

Two proofs: one by minimum modulus, one by linear algebra.

C.1 Proofs of the Fundamental Theorem of Algebra

BEYOND 21-242

Index of techniques

Technique 1.1.7: Proving from axioms . . . 8
Technique 1.1.8: Uniqueness via two candidates . . . 8
Technique 1.2.38: Try small cases and small fields . . . 35
Technique 1.3.13: Fix an arbitrary element and chase it . . . 43
Technique 1.3.14: Double containment . . . 43
Technique 1.4.18: Finding counterexamples and minimal examples . . . 57
Technique 1.5.4: Contrapositive . . . 66
Technique 1.5.9: Extremal choice and minimal counterexamples . . . 68
Technique 1.6.9: Induction on the length of a list . . . 82
Technique 1.7.4: Proving a disjunction . . . 97
Technique 2.1.26: Assemble linear maps from linear pieces . . . 117
Technique 2.2.4: Construct a map by choosing it on a basis . . . 124
Technique 2.2.7: Check it on a basis . . . 125
Technique 2.3.27: Translate between maps and matrices . . . 146
Technique 2.4.15: Compute entry by entry . . . 159
Technique 2.4.29: Cyclicity of trace . . . 166
Technique 2.4.35: Block decomposition . . . 169
Technique 2.5.16: Guess the inverse, then check both products . . . 180
Technique 2.6.12: Test injectivity at $\mathbf{0}$ . . . 194
Technique 2.7.1: Proving TFAE . . . 208
Technique 3.1.9: Undetermined coefficients . . . 239
Technique 3.1.18: Particular plus homogeneous . . . 241
Technique 3.2.27: Transpose to trade rows for columns . . . 259
Technique 3.3.5: Algorithms as proofs: invariants and termination . . . 269
Technique 3.3.22: Reduce to RREF . . . 280
Technique 3.4.21: Uniqueness gives existence . . . 303
Technique 4.1.18: Extend a basis of a subspace . . . 353
Technique 4.1.25: Apply an operator to a relation to kill terms . . . 356
Technique 4.1.30: Too many vectors force a dependence . . . 358
Technique 4.1.36: Double counting . . . 361
Technique 4.2.5: Restrict to a subspace . . . 367
Technique 4.2.24: Dimension counting . . . 377
Technique 4.2.29: Study a linear map on a space of matrices . . . 380
Technique 4.3.17: Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ . . . 396
Technique 4.4.8: Compare components . . . 408

Technique 4.4.10: Induction on the number of summands . . . 410
Technique 4.4.14: Glue bases of the summands . . . 412
Technique 4.4.28: Split the space with an idempotent . . . 420
Technique 4.5.7: Transport along an isomorphism . . . 428
Technique 4.5.29: Rule out similarity with an invariant . . . 440
Technique 4.5.33: Choose a convenient basis . . . 441
Technique 4.5.36: Choose a basis adapted to a preserved subspace . . . 442
Technique 4.6.18: Check that a definition does not depend on the representative . . . 454
Technique 4.6.31: Pass to a quotient . . . 460
Technique 4.7.4: Choose a base point . . . 468
Technique 4.7.21: Homogenize: add a coordinate 1 . . . 474
Technique 4.7.33: Slide along a dependence . . . 480
Technique 5.1.7: Expand by multilinearity . . . 506
Technique 5.2.12: Strong induction and well-ordering . . . 520
Technique 5.2.19: Parity via an invariant . . . 523
Technique 5.3.5: Cancel terms in pairs . . . 538
Technique 5.3.12: Read structure off the Leibniz expansion . . . 542
Technique 5.3.18: det is a polynomial in the entries . . . 546
Technique 5.3.21: Reduce modulo p . . . 548
Technique 5.4.6: Reduce to elementary matrices . . . 555
Technique 5.4.11: Recognize the determinant by its axioms . . . 557
Technique 5.4.28: Block elimination . . . 569
Technique 5.5.17: Induction on matrix size . . . 583
Technique 5.5.22: Expose a factor with row and column operations . . . 586
Technique 5.5.24: Polynomial identity arguments . . . 588
Technique 5.6.1: Factor the matrix . . . 597
Technique 6.1.9: A sum of nonnegative terms . . . 613
Technique 6.1.18: The trace of A^*A . . . 617
Technique 6.1.33: Polarization . . . 623
Technique 6.1.36: Show a vector is zero via its norm . . . 624
Technique 6.1.40: Plug in clever test vectors . . . 625
Technique 6.2.4: Orthogonal decomposition . . . 633
Technique 6.2.7: A nonnegative quadratic has nonpositive discriminant . . . 634
Technique 6.2.8: Rotate the phase . . . 635
Technique 6.2.21: Apply Cauchy–Schwarz to a well-chosen inner product . . . 642
Technique 6.3.12: Take an inner product to isolate a coefficient . . . 653
Technique 6.3.31: Gram matrices . . . 663
Technique 6.4.5: Projections shrink norms . . . 675
Technique 6.4.7: Recognize a projection by its two conditions . . . 677
Technique 6.5.6: Perturb a minimizer . . . 703
Technique 6.6.11: Find the adjoint by isolating the input . . . 727
Technique 6.6.20: Move an operator across the inner product . . . 732
Technique 6.7.2: Pass to orthonormal coordinates . . . 745
Technique 7.1.6: Compare degrees and leading coefficients . . . 758
Technique 7.2.31: Look for invariant subspaces among kernels and images of commuting operators . 793

Technique 7.3.21: Compare AB and BA . . . 810
Technique 7.3.28: Krylov sequences must become dependent . . . 814
Technique 7.3.35: Complexify, then take real and imaginary parts . . . 818
Technique 7.4.22: Diagonalize, solve, conjugate back . . . 831
Technique 7.4.25: Test an identity on eigenvectors . . . 832
Technique 7.5.4: Build a flag of invariant subspaces . . . 841
Technique 7.5.12: Induction on dimension via an invariant subspace . . . 845
Technique 7.5.23: Prove it for triangular matrices first . . . 852
Technique 7.6.13: Lagrange interpolation in T . . . 864
Technique 7.6.17: Use an annihilating polynomial . . . 866
Technique 7.6.23: Perturbation and density . . . 869
Technique 7.6.37: Look at the largest coordinate . . . 877
Technique 8.1.15: Shift by a scalar . . . 902
Technique 8.2.25: Apply a function to the eigenvalues . . . 923
Technique 8.3.19: Sandwich between square roots . . . 939
Technique A.1.21: Pigeonhole . . . 985
Technique A.2.5: Strengthen the induction hypothesis . . . 989
Technique A.3.13: Classify by a complete invariant . . . 995
Technique A.4.5: Apply Zorn's lemma . . . 999
Technique B.1.3: Conjugate the equation . . . 1002
Technique B.1.9: Square the modulus . . . 1004
Technique B.1.21: Use polar form for products, powers, and roots . . . 1011

Solutions

Section 1.1

Solution to Exercise 1.1.11

Uses. Theorem 1.1.9 (Uniqueness of identities and inverses), Definition 1.1.1 (Field)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. By Theorem 1.1.9, to show that $y = -x$ it suffices to show that $x + y = 0$.
Hint 2. For the second claim, add $(a+b)+((-a)+(-b))$ and regroup so that a meets $-a$ and b meets $-b$.
Sketch. $0 + 0 = 0$, so 0 is an additive inverse of 0. Regrouping gives $(a+b)+((-a)+(-b)) = (a+(-a))+(b+(-b)) = 0$. Both claims follow from uniqueness of additive inverses.

Since $0 + 0 = 0$ by the additive identity axiom, 0 is an additive inverse of 0, so $-0 = 0$ by Theorem 1.1.9.

For the second claim,

$$\begin{aligned}
(a+b)+((-a)+(-b)) &= a+(b+((-a)+(-b))) && \text{(associativity)}\\
&= a+((b+(-a))+(-b)) && \text{(associativity)}\\
&= a+(((-a)+b)+(-b)) && \text{(commutativity)}\\
&= a+((-a)+(b+(-b))) && \text{(associativity)}\\
&= (a+(-a))+(b+(-b)) && \text{(associativity)}\\
&= 0+0=0 && \text{(additive inverses, additive identity).}
\end{aligned}$$

So $(-a)+(-b)$ is an additive inverse of $a+b$, and by Theorem 1.1.9 it equals $-(a+b)$.

Solution to Exercise 1.1.18

Uses. Proposition 1.1.17 (No zero divisors), Proposition 1.1.16 (Sign rules), Proposition 1.1.13 (Solving $a+x=b$ and $ax=b$)
Hint 1. Move everything to one side and factor.
Hint 2. $x^2 - x = x(x-1)$. Now use Proposition 1.1.17.
Sketch. $x^2 = x$ if and only if $x(x-1) = 0$, which by Proposition 1.1.17 happens exactly when $x = 0$ or $x = 1$.

By distributivity and Proposition 1.1.16, $x(x-1) = x^2 + x(-1) = x^2 - x$. By Proposition 1.1.13, $x^2 = x$ holds if and only if $x^2 - x = 0$, that is, if and only if $x(x-1) = 0$. By Proposition 1.1.17 this happens if and only if $x = 0$ or $x - 1 = 0$, and $x - 1 = 0$ means $x = 1$ (again by Proposition 1.1.13). So the solutions are exactly $x = 0$ and $x = 1$, which are distinct since $0 \neq 1$.

Solution to Exercise 1.1.27

Uses. Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.13 (Solving $a+x=b$ and $ax=b$)
Hint 1. Multiply pairs of nonzero elements and reduce modulo 7; you want products that leave remainder 1.
Hint 2. For the equation, follow Proposition 1.1.13: subtract 5, then multiply by 3^{-1}.
Sketch. $2 \cdot 4 = 8$, $3 \cdot 5 = 15$, and $6 \cdot 6 = 36$ all leave remainder 1. Then $3x = 2 - 5 = 4$, so $x = 5 \cdot 4 = 6$.

Since $2 \cdot 4 = 8 = 7 + 1$, $3 \cdot 5 = 15 = 14 + 1$, and $6 \cdot 6 = 36 = 35 + 1$, the inverses are

$$1^{-1} = 1, \quad 2^{-1} = 4, \quad 3^{-1} = 5, \quad 4^{-1} = 2, \quad 5^{-1} = 3, \quad 6^{-1} = 6.$$

For the equation, $3x + 5 = 2$ holds if and only if $3x = 2 - 5 = -3 = 4$, which holds if and only if $x = 3^{-1} \cdot 4 = 5 \cdot 4 = 20 \bmod 7 = 6$ (Proposition 1.1.13). Check: $3 \cdot 6 + 5 = 23 = 21 + 2$, so $3 \cdot 6 + 5 = 2$ in $\mathbb{F}_7$.

Solution to Exercise 1.1.35

Uses. Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$)
Hint 1. Eliminate x as you would over $\mathbb{R}$, by subtracting 3 times the first equation from the second. Every step is field arithmetic, so it is valid over $\mathbb{F}_7$.
Hint 2. In $\mathbb{F}_7$, $2^{-1} = 4$. Over $\mathbb{F}_2$, first rewrite the coefficients: $2 = 4 = 0$ and $3 = 1$.
Sketch. Over $\mathbb{F}_7$, elimination gives $-2y = -1$, so $y = 4$ and $x = 2$; this pair checks. Over $\mathbb{F}_2$ the system collapses to $x = 1$ twice, so y is free and there are two solutions.

(a) We work in $\mathbb{F}_7$, where $2^{-1} = 4$ since $2 \cdot 4 = 8 = 7 + 1$. Suppose (x, y) is a solution. Subtracting 3 times the first equation from the second gives

$$(3x + 4y) - 3(x + 2y) = 1 - 3 \cdot 3, \quad \text{that is,} \quad -2y = -8.$$

In $\mathbb{F}_7$, $-8 = -1$, so $2y = 1$ by Proposition 1.1.16, and $y = 2^{-1} = 4$. The first equation then gives $x = 3 - 2 \cdot 4 = -5 = 2$. So $(2, 4)$ is the only possible solution, and it is one: in $\mathbb{F}_7$, $2 + 2 \cdot 4 = 10 = 3$ and $3 \cdot 2 + 4 \cdot 4 = 22 = 1$.

(b) In $\mathbb{F}_2$ we have $2 = 4 = 0$ and $3 = 1$, so the system reads $x = 1$, $x = 1$. Its solutions are $(x, y) = (1, 0)$ and $(1, 1)$. The elimination in (a) is unavailable here, because it divided by 2, which is 0 in $\mathbb{F}_2$.

Solution to Exercise 1.1.36

Uses. Proposition 1.1.30 ($\mathbb{C}$ is a field)
Hint 1. Expand products with distributivity and use $i^2 = -1$.
Hint 2. For (b) and (c), use the inverse formula of Proposition 1.1.30, or multiply by $a - bi$ where $a + bi$ is the number being inverted. For (d), square repeatedly.
Sketch. (a) $3 + 12i - 2i + 8 = 11 + 10i$. (b) $(2 - i)/5$. (c) $(1 + i)^2/2 = i$. (d) $(1 + i)^2 = 2i$, so $(1 + i)^8 = (2i)^4 = 16$.

(a) $(3 - 2i)(1 + 4i) = 3 + 12i - 2i - 8i^2 = 11 + 10i$.

(b) By Proposition 1.1.30 with $a = 2$, $b = 1$, and $a^2 + b^2 = 5$, $(2 + i)^{-1} = \frac{2}{5} - \frac{1}{5}i$. Check: $(2 + i)(2 - i) = 4 + 1 = 5$.

(c) $(1 - i)^{-1} = (1 + i)/2$, so $(1 + i)/(1 - i) = (1 + i)^2/2 = \left(1 + 2i + i^2\right)/2 = i$.

(d) $(1 + i)^2 = 2i$, so $(1 + i)^4 = (2i)^2 = -4$ and $(1 + i)^8 = (-4)^2 = 16$.

Solution to Exercise 1.1.37

Uses. Theorem 1.1.9 (Uniqueness of identities and inverses), Proposition 1.1.15 (Multiplying by zero), Proposition 1.1.17 (No zero divisors)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Before writing $\left(a^{-1}\right)^{-1}$ you must know that $a^{-1} \neq 0$, because the inverse axiom only applies to nonzero elements. What would $a^{-1} = 0$ give?
Hint 2. To identify an inverse, exhibit a candidate and invoke uniqueness: check that a is an inverse of a^{-1}, and that $a^{-1}b^{-1}$ is an inverse of ab.
Sketch. If $a^{-1} = 0$, then $1 = aa^{-1} = 0$. Since $a^{-1}a = 1$, a is the inverse of a^{-1}. By Proposition 1.1.17, $ab \neq 0$, and $(ab)\left(a^{-1}b^{-1}\right) = \left(aa^{-1}\right)\left(bb^{-1}\right) = 1$.

If $a^{-1} = 0$, then $1 = aa^{-1} = a0 = 0$ by Proposition 1.1.15, contradicting $1 \neq 0$. So $a^{-1} \neq 0$, and a^{-1} has a multiplicative inverse. Since $a^{-1}a = aa^{-1} = 1$, the element a is a multiplicative inverse of a^{-1}, so $\left(a^{-1}\right)^{-1} = a$ by Theorem 1.1.9.

By Proposition 1.1.17, $ab \neq 0$. Reordering the product,

$$(ab)\left(a^{-1}b^{-1}\right) = \left(aa^{-1}\right)\left(bb^{-1}\right) = 1 \cdot 1 = 1,$$

so $a^{-1}b^{-1}$ is a multiplicative inverse of ab, and $(ab)^{-1} = a^{-1}b^{-1}$ by Theorem 1.1.9.

Solution to Exercise 1.1.38

Uses. Notation 1.1.10, Exercise 1.1.37, Definition 1.1.1 (Field)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Rewrite every fraction x/y as xy^{-1}, and note that $bd \neq 0$ with $(bd)^{-1} = b^{-1}d^{-1}$ (Exercise 1.1.37).
Hint 2. Then expand with distributivity and reorder products freely. The factors bb^{-1} and dd^{-1} collapse to 1.
Sketch. With $(bd)^{-1} = b^{-1}d^{-1}$, the right-hand sides expand to $ab^{-1}(dd^{-1}) + cd^{-1}(bb^{-1})$ and $(ab^{-1})(cd^{-1})$. For the last claim, multiply $a/b = c/d$ by bd, and $ad = bc$ by $b^{-1}d^{-1}$.

By Exercise 1.1.37, $bd \neq 0$ and $(bd)^{-1} = b^{-1}d^{-1}$. Using distributivity and reordering products,

$$\frac{ad+bc}{bd} = (ad+bc)b^{-1}d^{-1} = ab^{-1}\big(dd^{-1}\big) + cd^{-1}\big(bb^{-1}\big) = ab^{-1} + cd^{-1} = \frac{a}{b} + \frac{c}{d}.$$

Similarly,

$$\frac{a}{b} \cdot \frac{c}{d} = ab^{-1}cd^{-1} = (ac)\big(b^{-1}d^{-1}\big) = (ac)(bd)^{-1} = \frac{ac}{bd}.$$

For the last claim, suppose $ab^{-1} = cd^{-1}$. Multiplying both sides by bd gives $ad\big(bb^{-1}\big) = bc\big(dd^{-1}\big)$, that is, $ad = bc$. Conversely, if $ad = bc$, multiplying both sides by $b^{-1}d^{-1}$ gives $ab^{-1}\big(dd^{-1}\big) = cd^{-1}\big(bb^{-1}\big)$, that is, $a/b = c/d$.

Solution to Exercise 1.1.39

Uses. Proposition 1.1.15 (Multiplying by zero)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Compute $a \cdot 1$ for an arbitrary $a \in F$, in two ways.
Hint 2. You will want $a0 = 0$. Check that the proof of Proposition 1.1.15 never used $1 \neq 0$.
Sketch. $a = a \cdot 1 = a \cdot 0 = 0$ for every a, where $a0 = 0$ holds because the proof of Proposition 1.1.15 used only distributivity, the additive axioms, and cancellation.

The set F is nonempty because it contains 0. The proof of Proposition 1.1.15 used distributivity, the additive identity axiom, and additive cancellation, which in turn rests only on the additive axioms; none of these involve $1 \neq 0$. So $a0 = 0$ for every $a \in F$. Now for every $a \in F$,

$$a = a \cdot 1 = a \cdot 0 = 0,$$

using the multiplicative identity axiom and $1 = 0$. Hence $F = \{0\}$.

Solution to Exercise 1.1.40

Uses. Proposition 1.1.16 (Sign rules), Proposition 1.1.17 (No zero divisors)
Hint 1. Move everything to one side and try to factor $a^2 - b^2$.
Hint 2. $(a-b)(a+b) = a^2 - b^2$ in every field. Now use Proposition 1.1.17.
Sketch. True: $(a-b)(a+b) = a^2 - b^2 = 0$, so $a - b = 0$ or $a + b = 0$.

The claim is true. Expanding with distributivity and Proposition 1.1.16,

$$(a-b)(a+b) = a^2 + ab - ba - b^2 = a^2 - b^2,$$

since $ab - ba = ab - ab = 0$. If $a^2 = b^2$, then $a^2 - b^2 = 0$, so $(a-b)(a+b) = 0$. By Proposition 1.1.17, $a - b = 0$ or $a + b = 0$, that is, $a = b$ or $a = -b$.

In $\mathbb{F}_2$ we have $-b = b$, so the conclusion just says $a = b$. That is consistent: $0^2 = 0$ and $1^2 = 1$ are different.

Solution to Exercise 1.1.41

Uses. Definition 1.1.19 (Integers modulo n)
Hint 1. Try small primes one at a time. Each $\mathbb{F}_p$ has only p elements to test.
Hint 2. $p = 2$ and $p = 5$ work. Try $p = 3$.
Sketch. False: in $\mathbb{F}_3$, $x^2 + 1$ takes the values $1, 2, 2$ at $x = 0, 1, 2$.

The claim is false. In $\mathbb{F}_3$,

$$0^2 + 1 = 1, \quad 1^2 + 1 = 2, \quad 2^2 + 1 = 5 \bmod 3 = 2,$$

so $x^2 + 1 = 0$ has no solution. (It does for some primes: $x = 1$ in $\mathbb{F}_2$, and $x = 2$ or $x = 3$ in $\mathbb{F}_5$.)

Solution to Exercise 1.1.42

> *Uses.* Proposition 1.1.17 (No zero divisors), Example 1.1.4 (The field with two elements)
> *Hint 1.* The factorization and the use of Proposition 1.1.17 are correct. Test the conclusion in the smallest field.
> *Hint 2.* In $\mathbb{F}_2$, what is -1?
> *Sketch.* The argument shows that the solutions are exactly 1 and -1, but these coincide when $1 + 1 = 0$. In $\mathbb{F}_2$ there is only one solution.

The argument correctly shows that the solutions of $x^2 = 1$ are exactly $x = 1$ and $x = -1$. The error is the last sentence: 1 and -1 need not be different. In $\mathbb{F}_2$, $1 + 1 = 0$, so $-1 = 1$, and $x^2 = 1$ has exactly one solution. In general $1 = -1$ if and only if $1 + 1 = 0$, so the correct statement is that $x^2 = 1$ has exactly two solutions when $1 + 1 \neq 0$, and exactly one when $1 + 1 = 0$.

Solution to Exercise 1.1.43

> *Uses.* Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$), Proposition 1.1.15 (Multiplying by zero)
> *Hint 1.* f is a bijection if and only if, for every $y \in F$, the equation $ax + b = y$ has exactly one solution x.
> *Hint 2.* For $a \neq 0$, use Proposition 1.1.13 twice. For $a = 0$, f is constant. Why is a constant function on F not injective?
> *Sketch.* If $a \neq 0$, then $ax + b = y$ if and only if $x = a^{-1}(y - b)$, so f is bijective with $f^{-1}(y) = a^{-1}(y - b)$. If $a = 0$, then $f(0) = f(1) = b$ with $0 \neq 1$.

Suppose $a \neq 0$, and let $y \in F$. By part (a) of Proposition 1.1.13, applied with unknown ax, the equation $ax + b = y$ holds if and only if $ax = y - b$, and by part (b) this holds if and only if $x = a^{-1}(y - b)$. So every y has exactly one preimage, that is, f is a bijection, and $f^{-1}(y) = a^{-1}(y - b)$.

Suppose $a = 0$. Then $f(x) = 0x + b = b$ for every x, by Proposition 1.1.15. In particular $f(0) = f(1)$, and $0 \neq 1$, so f is not injective.

Solution to Exercise 1.1.44

> *Uses.* Proposition 1.1.17 (No zero divisors), Proposition 1.1.30 ($\mathbb{C}$ is a field), Exercise 1.1.40
> *Hint 1.* For (a), factor $z^2 + 1$ using $i^2 = -1$, then use that a field has no zero divisors.
> *Hint 2.* For (b), write $z = a + bi$ with $a, b \in \mathbb{R}$ and compare the real and imaginary parts of $z^2 = i$.
> *Hint 3.* You get $a^2 = b^2$ and $2ab = 1$. So $b = \pm a$, and one sign is impossible.
> *Sketch.* (a) $z^2 + 1 = (z - i)(z + i)$. (b) $z^2 = \left(a^2 - b^2\right) + 2abi$, so $a^2 = b^2$ and $2ab = 1$; this forces $b = a$ and $a = \pm 1/\sqrt{2}$, so $z = \pm(1 + i)/\sqrt{2}$.

(a) Since $i^2 = -1$, we have $(z - i)(z + i) = z^2 - i^2 = z^2 + 1$. So $z^2 = -1$ if and only if $(z - i)(z + i) = 0$, which by Proposition 1.1.17 holds if and only if $z = i$ or $z = -i$. These are different, since $i - (-i) = 2i \neq 0$.

(b) Write $z = a + bi$ with $a, b \in \mathbb{R}$. Then $z^2 = \left(a^2 - b^2\right) + 2abi$, so $z^2 = i$ if and only if $a^2 - b^2 = 0$ and $2ab = 1$. The first equation gives $b = a$ or $b = -a$ (Exercise 1.1.40, in $\mathbb{R}$). If $b = -a$, then $2ab = -2a^2 \leq 0$, which is not 1. So $b = a$ and $2a^2 = 1$, giving $a = \pm 1/\sqrt{2}$. Hence

$$z = \frac{1 + i}{\sqrt{2}} \quad \text{or} \quad z = -\frac{1 + i}{\sqrt{2}},$$

and both work: $\left((1 + i)/\sqrt{2}\right)^2 = \left(1 + 2i + i^2\right)/2 = i$.

Once one square root w of i is known, (a)'s method finds all of them: $z^2 = w^2$ means $(z - w)(z + w) = 0$, so $z = \pm w$.

Solution to Exercise 1.1.45

> *Uses.* Corollary 1.1.14 (Cancellation), Non-example 1.1.23 ($\mathbb{Z}/6\mathbb{Z}$), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
> *Hint 1.* Every step but one is valid for all n. Check what each cited result assumes.
> *Hint 2.* Where did Corollary 1.1.14 come from, and does $\mathbb{Z}/6\mathbb{Z}$ satisfy it? Try $a = 2$.
> *Sketch.* Multiplicative cancellation was proved for fields, using a^{-1}. Applying it to $\mathbb{Z}/n\mathbb{Z}$ assumes the inverse we are trying to produce. It fails in $\mathbb{Z}/6\mathbb{Z}$: $2 \cdot 0 = 2 \cdot 3$.

The error is the appeal to Corollary 1.1.14. That result was proved for fields, and its multiplicative part multiplies by a^{-1}, which is exactly what the argument is trying to show exists. So the argument is circular. Cancellation genuinely fails in $\mathbb{Z}/6\mathbb{Z}$: $2 \cdot 0 = 0 = 2 \cdot 3$ but $0 \neq 3$, so m is not injective there (Nonexample 1.1.23). The rest of the argument is fine, and the counting proof of Proposition 1.1.24 repairs it by proving injectivity from Euclid's lemma, which is where primality of n enters.

Solution to Exercise 1.1.46

> *Uses.* Corollary 1.1.14 (Cancellation), Theorem 1.1.9 (Uniqueness of identities and inverses), Proposition 1.1.15 (Multiplying by zero), Proposition 1.1.17 (No zero divisors), Exercise 1.1.40
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* In (a), determine $1 + 1$ first. Rule out $1 + 1 = 1$ with cancellation, and $1 + 1 = 0$ by asking what $1 + c$ could then be.
> *Hint 2.* The other entries are forced the same way. Products involving 0 or 1 are already known from the axioms and Proposition 1.1.15.
> *Hint 3.* In (b), factor $x^3 - x = x(x-1)(x+1)$ and use Proposition 1.1.17 to show $F \subseteq \{0, 1, -1\}$.
> *Sketch.* (a) Each sum is pinned down by cancellation or by uniqueness of inverses; $c^2 = 1$ because $c^2 = 0$ and $c^2 = c$ are impossible. (b) $x^3 = x$ forces $x \in \{0, 1, -1\}$, so F has two or three elements, and it is $\mathbb{F}_2$ or $\mathbb{F}_3$ up to renaming; both satisfy $x^3 = x$.

(a) Sums involving 0, and products involving 0 or 1, are determined by the identity axioms and Proposition 1.1.15, so we need $1 + 1$, $1 + c$, $c + c$, and $c \cdot c$. Throughout, 0, 1, c are distinct.

$1 + 1$. If $1 + 1 = 1 = 1 + 0$, then $1 = 0$ by cancellation, which is false. Suppose $1 + 1 = 0$. If $1 + c = 1 = 1 + 0$, then $c = 0$; if $1 + c = c = 0 + c$, then $1 = 0$; both are false, so $1 + c = 0$. But then 1 and c are both additive inverses of 1, so $c = 1$ by Theorem 1.1.9, which is false. Hence $1 + 1 = c$.

$1 + c$. As above, $1 + c = 1$ forces $c = 0$ and $1 + c = c$ forces $1 = 0$. So $1 + c = 0$, and in particular $-c = 1$.

$c + c$. If $c + c = c = c + 0$, then $c = 0$. If $c + c = 0$, then $-c = c$, but $-c = 1$, so $c = 1$. Both are false, so $c + c = 1$.

$c \cdot c$. If $c^2 = 0$, then $c = 0$ by Proposition 1.1.17. If $c^2 = c = c \cdot 1$, then $c = 1$ by cancellation (as $c \neq 0$). So $c^2 = 1$.

In $\mathbb{F}_3$, $1 + 1 = 2$, $1 + 2 = 3 \bmod 3 = 0$, $2 + 2 = 4 \bmod 3 = 1$, and $2 \cdot 2 = 4 \bmod 3 = 1$. So the tables of F are those of $\mathbb{F}_3$ with c in place of 2.

(b) Suppose $x^3 = x$ for all $x \in F$. As in Exercise 1.1.40, $(x-1)(x+1) = x^2 - 1$, so $x(x-1)(x+1) = x^3 - x = 0$ for every x. By Proposition 1.1.17 (applied twice), every x equals 0, 1, or -1. So $F \subseteq \{0, 1, -1\}$.

If $1 + 1 = 0$, then $-1 = 1$ and $F = \{0, 1\}$. In a field with two elements every sum and product is forced as above, with $1 + 1 = 0$ because $1 + 1 = 1$ would give $1 = 0$; so F has the tables of $\mathbb{F}_2$. If $1 + 1 \neq 0$, then $-1 \neq 1$, and $-1 \neq 0$, so $F = \{0, 1, -1\}$ has exactly three elements, and by (a) F has the tables of $\mathbb{F}_3$ (with -1 in place of 2).

Conversely, $x^3 = x$ holds in $\mathbb{F}_2$ ($0^3 = 0$, $1^3 = 1$) and in $\mathbb{F}_3$ ($2^3 = 8 \bmod 3 = 2$). So the fields with $x^3 = x$ for all x are exactly $\mathbb{F}_2$ and $\mathbb{F}_3$, up to renaming the elements.

Solution to Exercise 1.1.47

> *Uses.* Proposition 1.1.15 (Multiplying by zero), Proposition 1.1.16 (Sign rules), Theorem 1.1.9 (Uniqueness of identities and inverses), Exercise 1.1.40
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* Prove the cycle (i) ⇒ (ii) ⇒ (iii) ⇒ (i).
> *Hint 2.* (i) ⇒ (ii): $a + a = (1 + 1)a$. (ii) ⇒ (iii): use Exercise 1.1.40. (iii) ⇒ (i): compare the squares of 1 and -1.
> *Sketch.* If $1 + 1 = 0$, then $a + a = (1 + 1)a = 0$, so $a = -a$. If $a = -a$ always, then $x^2 = y^2$ gives $x = \pm y = y$. If squaring is injective, then $1^2 = (-1)^2$ gives $1 = -1$, so $1 + 1 = 0$.

We prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i).

(i) ⇒ (ii). Let $a \in F$. By distributivity and Proposition 1.1.15, $a + a = 1a + 1a = (1 + 1)a = 0a = 0$. So a is an additive inverse of a, and $a = -a$ by Theorem 1.1.9.

(ii) ⇒ (iii). Suppose $x^2 = y^2$. By Exercise 1.1.40, $x = y$ or $x = -y$, and $-y = y$ by (ii). So $x = y$.

(iii) ⇒ (i). By Proposition 1.1.16, $(-1)^2 = 1 \cdot 1 = 1^2$. Injectivity gives $-1 = 1$, so $1 + 1 = 1 + (-1) = 0$.

Solution to Exercise 1.1.48

Uses. Definition 1.1.28 (Characteristic), Proposition 1.1.17 (No zero divisors)
Hint 1. Suppose $\operatorname{char} F = m > 0$ and $m = rs$ with $1 < r, s < m$. What can you say about $(r \cdot 1)(s \cdot 1)$?
Hint 2. Show that $(r \cdot 1)(s \cdot 1) = (rs) \cdot 1$ using distributivity, then use Proposition 1.1.17 and the minimality of m.
Sketch. $(r \cdot 1)(s \cdot 1) = (rs) \cdot 1$ by distributivity. If the characteristic m factored as rs with $1 < r, s < m$, then $(r \cdot 1)(s \cdot 1) = m \cdot 1 = 0$ would force $r \cdot 1 = 0$ or $s \cdot 1 = 0$, contradicting the minimality of m.

First, for positive integers r and s, distributivity gives

$$(r \cdot 1)(s \cdot 1) = (r \cdot 1)(1 + \cdots + 1) = (r \cdot 1) + \cdots + (r \cdot 1) = (rs) \cdot 1,$$

where there are s terms in the middle expressions, and the last sum consists of rs ones.

Suppose $\operatorname{char} F = m > 0$. Then $m \neq 1$, since $1 \cdot 1 = 1 \neq 0$. Suppose for contradiction that m is not prime, so $m = rs$ with $1 < r, s < m$. Then $(r \cdot 1)(s \cdot 1) = m \cdot 1 = 0$, so by Proposition 1.1.17, $r \cdot 1 = 0$ or $s \cdot 1 = 0$. Either way some positive integer smaller than m kills 1, contradicting the choice of m as the least such integer. Hence m is prime.

Solution to Exercise 1.1.49

Uses. Definition 1.1.32 (Subfield), $\sqrt{2}$ is irrational
Hint 1. Check the conditions of Definition 1.1.32 one at a time. All but one are immediate.
Hint 2. For the inverse of $a + b\sqrt{2} \neq 0$, imitate the complex case: multiply by $a - b\sqrt{2}$.
Hint 3. Why is $a^2 - 2b^2 \neq 0$? Use that $\sqrt{2}$ is irrational.
Sketch. Closure under $+$, $\cdot$, and negation is a direct computation. For $a + b\sqrt{2} \neq 0$, $\left(a + b\sqrt{2}\right)\left(a - b\sqrt{2}\right) = a^2 - 2b^2$, which is a nonzero rational because $\sqrt{2} \notin \mathbb{Q}$; dividing $a - b\sqrt{2}$ by it gives the inverse.

Write $K = \mathbb{Q}(\sqrt{2})$. Clearly $0 = 0 + 0\sqrt{2}$ and $1 = 1 + 0\sqrt{2}$ lie in K. For $a, b, c, d \in \mathbb{Q}$,

$$\begin{aligned}
\left(a + b\sqrt{2}\right) + \left(c + d\sqrt{2}\right) &= (a + c) + (b + d)\sqrt{2}, \\
\left(a + b\sqrt{2}\right)\left(c + d\sqrt{2}\right) &= (ac + 2bd) + (ad + bc)\sqrt{2}, \\
-\left(a + b\sqrt{2}\right) &= (-a) + (-b)\sqrt{2},
\end{aligned}$$

and these lie in K because $\mathbb{Q}$ is closed under its operations.

Now let $z = a + b\sqrt{2} \neq 0$. We claim $a^2 - 2b^2 \neq 0$. If $b = 0$, then $a = z \neq 0$, so $a^2 - 2b^2 = a^2 \neq 0$. If $b \neq 0$ and $a^2 = 2b^2$, then $(a/b)^2 = 2$, so $\sqrt{2} = |a/b|$ would be rational, which is false. Hence $q = a^2 - 2b^2$ is a nonzero rational, and

$$\left(a + b\sqrt{2}\right)\left(\frac{a}{q} - \frac{b}{q}\sqrt{2}\right) = \frac{a^2 - 2b^2}{q} = 1.$$

So $z^{-1} = a/q - (b/q)\sqrt{2} \in K$. Thus K is a subfield of $\mathbb{R}$, and by Proposition 1.1.33 it is a field.

Solution to Exercise 1.1.50

Uses. Proposition 1.1.17 (No zero divisors), Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$), Exercise 1.1.40
Hint 1. Complete the square. Multiplying the equation by $4a$ is allowed and reversible, because $4a \neq 0$.
Hint 2. Check that $4a\left(ax^2 + bx + c\right) = (2ax + b)^2 - \left(b^2 - 4ac\right)$.
Hint 3. For (b), note that $4 = 0$ in $\mathbb{F}_2$, so $b^2 - 4ac = b^2$ is always a square. Try $x^2 + x + 1$.
Sketch. Since $4a \neq 0$, the equation is equivalent to $(2ax + b)^2 = b^2 - 4ac$. A solution x gives $d = 2ax + b$; conversely d gives $x = (d - b)/(2a)$, and Exercise 1.1.40 shows every solution has this form for d or $-d$. In $\mathbb{F}_2$, $x^2 + x + 1$ has discriminant $1 = 1^2$ but no root.

(a) Since $2 \neq 0$ and $a \neq 0$, Proposition 1.1.17 gives $2a \neq 0$ and $4a = 2 \cdot 2 \cdot a \neq 0$. Expanding,

$$(2ax + b)^2 - \big(b^2 - 4ac\big) = 4a^2x^2 + 4abx + 4ac = 4a\big(ax^2 + bx + c\big).$$

Because $4a \neq 0$, a product $4ay$ is 0 exactly when $y = 0$. So x is a solution if and only if $(2ax + b)^2 = b^2 - 4ac$.

If x is a solution, then $d = 2ax + b$ satisfies $d^2 = b^2 - 4ac$. Conversely, suppose $d^2 = b^2 - 4ac$. If x is a solution, then $(2ax + b)^2 = d^2$, so $2ax + b = d$ or $2ax + b = -d$ (Exercise 1.1.40), and solving for x (Proposition 1.1.13) gives $x = (-b + d)/(2a)$ or $x = (-b - d)/(2a)$. Both of these are solutions, since for them $(2ax + b)^2 = (\pm d)^2 = d^2 = b^2 - 4ac$. So the solutions are exactly $x = (-b \pm d)/(2a)$.

(b) Take $a = b = c = 1$. In $\mathbb{F}_2$, $4 = 0$, so $b^2 - 4ac = 1 = 1^2$ is a square. But $0^2 + 0 + 1 = 1$ and $1^2 + 1 + 1 = 1$, so $x^2 + x + 1$ has no root in $\mathbb{F}_2$. The proof of (a) breaks at the first step, which multiplies by $4a = 0$.

Solution to Exercise 1.1.51

> *Uses.* Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.17 (No zero divisors), Corollary 1.1.14 (Cancellation)
> *Hint 1.* For (a), show that the $p - 1$ products are nonzero and pairwise distinct. You may now use that $\mathbb{F}_p$ is a field.
> *Hint 2.* For (b), multiply all $p - 1$ products together, in two ways.
> *Hint 3.* The product $P = 1 \cdot 2 \cdots (p - 1)$ is nonzero, so it can be cancelled.
> *Sketch.* The products ax for $x \neq 0$ are nonzero and distinct, so they list the $p - 1$ nonzero elements. Multiplying them all gives $a^{p-1}P = P$ with $P = 1 \cdot 2 \cdots (p - 1) \neq 0$; cancel P.

By Proposition 1.1.24, $\mathbb{F}_p$ is a field.

(a) Each product ax with $x \in \{1, \ldots, p - 1\}$ is nonzero by Proposition 1.1.17. If $ax = ay$, then $x = y$ by Corollary 1.1.14, which applies because $a \neq 0$ in a field. So the products are $p - 1$ distinct elements of the $(p - 1)$-element set $\{1, \ldots, p - 1\}$, hence all of it.

(b) Let $P = 1 \cdot 2 \cdots (p - 1)$, which is nonzero by Proposition 1.1.17. By (a), the product of $a \cdot 1, \ldots, a \cdot (p - 1)$ is the product of $1, \ldots, p - 1$ in another order, which is P. Reordering factors, the same product is $a^{p-1}P$. So $a^{p-1}P = P = 1 \cdot P$, and cancelling P (Corollary 1.1.14) gives $a^{p-1} = 1$.

Contrast this use of cancellation with Exercise 1.1.45: here we already know that $\mathbb{F}_p$ is a field.

Solution to Exercise 1.1.52

> *Uses.* Exercise 1.1.40, Proposition 1.1.17 (No zero divisors), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
> *Hint 1.* 0 is a square. For the nonzero squares, ask when two nonzero elements have the same square.
> *Hint 2.* By Exercise 1.1.40, $x^2 = y^2$ if and only if $y = \pm x$. When p is odd, $x \neq -x$ for $x \neq 0$.
> *Hint 3.* So squaring is two-to-one on the $p - 1$ nonzero elements.
> *Sketch.* The nonzero elements split into $(p - 1)/2$ pairs $\{x, -x\}$ with $x \neq -x$. Elements of a pair have the same square, different pairs have different squares, and squares of nonzero elements are nonzero. With 0, that makes $(p + 1)/2$ squares.

There are $(p + 1)/2$ squares.

First, $x \neq -x$ for $x \neq 0$: otherwise $(1 + 1)x = x + x = 0$, so $1 + 1 = 0$ by Proposition 1.1.17, but $1 + 1 = 2 \neq 0$ in $\mathbb{F}_p$ because p is odd. So the $p - 1$ nonzero elements split into $(p - 1)/2$ disjoint pairs $\{x, -x\}$.

By Exercise 1.1.40, two nonzero elements have the same square if and only if they lie in the same pair. Also $x^2 \neq 0$ for $x \neq 0$ by Proposition 1.1.17. So the nonzero squares are in bijection with the pairs, and there are $(p - 1)/2$ of them. Adding the square $0 = 0^2$ gives $(p + 1)/2$.

For example, the squares in $\mathbb{F}_7$ are 0, 1, $2 = 3^2$, and $4 = 2^2$, which is $(7 + 1)/2 = 4$ elements.

Solution to Exercise 1.1.53

> *Uses.* Definition 1.1.32 (Subfield), $\sqrt[3]{2}$ is irrational
> *Hint 1.* K contains 0 and 1 and is closed under addition and negation. Try products.
> *Hint 2.* Is $\left(\sqrt[3]{2}\right)^2 = \sqrt[3]{4}$ in K? Suppose $\sqrt[3]{4} = a + b\sqrt[3]{2}$ and multiply both sides by $\sqrt[3]{2}$.
> *Hint 3.* You get $2 = ab + \left(a + b^2\right)\sqrt[3]{2}$. Use the irrationality of $\sqrt[3]{2}$, twice.

Sketch. False: K is not closed under multiplication, because $\sqrt[3]{4} \notin K$. Writing $\sqrt[3]{4} = a + b\sqrt[3]{2}$ and multiplying by $\sqrt[3]{2}$ gives $2 = ab + \left(a + b^2\right)\sqrt[3]{2}$; irrationality of $\sqrt[3]{2}$ forces $a = -b^2$ and then $b^3 = -2$, which is impossible for rational b.

The claim is false: K is not closed under multiplication. Write $t = \sqrt[3]{2}$, so $t \in K$ and $t^3 = 2$. We will show that $t^2 \notin K$.

First, t is irrational. If $t = r/s$ with integers r, s, $s > 0$, and $\gcd(r, s) = 1$, then $r^3 = 2s^3$, so r is even. Writing $r = 2k$ gives $8k^3 = 2s^3$, so $s^3 = 4k^3$ and s is even too, a contradiction.

Suppose $t^2 = a + bt$ with $a, b \in \mathbb{Q}$. Multiplying by t,

$$2 = t^3 = at + bt^2 = at + b(a + bt) = ab + \left(a + b^2\right)t.$$

If $a + b^2 \neq 0$, then $t = (2 - ab)/\left(a + b^2\right)$ is rational, which is false. So $a = -b^2$, and then $2 = ab = -b^3$, so $(-b)^3 = 2$ with $-b \in \mathbb{Q}$, again contradicting the irrationality of t. Hence $t^2 \notin K$, although $t \in K$, so K is not a subfield of $\mathbb{R}$.

Solution to Exercise 1.1.54

Uses. Proposition 1.1.30 ($\mathbb{C}$ is a field), Proposition 1.1.17 (No zero divisors), Definition 1.1.20 (Commutative ring)
Hint 1. Show that every axiom except multiplicative inverses holds for every c; for associativity, expand both sides. So the question is when nonzero elements are invertible.
Hint 2. For $c < 0$ imitate $\mathbb{C}$: multiply (a, b) by $(a, -b)$.
Hint 3. For $c \geq 0$, write $c = t^2$ with $t \geq 0$, and look for zero divisors among $(t, 1)$ and $(-t, 1)$.
Sketch. For every c the structure is a commutative ring with identity $(1, 0)$. If $c < 0$, then $(a, b)(a, -b) = \left(a^2 - cb^2, 0\right)$ with $a^2 - cb^2 > 0$ for $(a, b) \neq (0, 0)$, which gives inverses. If $c = t^2 \geq 0$, then $(t, 1)(-t, 1) = (0, 0)$, so there are zero divisors. Answer: exactly when $c < 0$.

The answer is: exactly when $c < 0$.

The other axioms hold for every c. Addition is componentwise, so its axioms hold as in $\mathbb{R}^2$. The multiplication is commutative, because the formula is symmetric under swapping (a, b) and (a', b'). The element $(1, 0)$ is an identity, $(a, b)(1, 0) = (a, b)$, and $(1, 0) \neq (0, 0)$. Distributivity holds because each coordinate of the product is a sum of terms each involving exactly one coordinate of the second factor, so it is additive in the second factor. For associativity, with $(a, b), (a', b'), (a'', b'')$,

$$((a, b)(a', b'))(a'', b'') = ((aa' + cbb')a'' + c(ab' + ba')b'', (aa' + cbb')b'' + (ab' + ba')a''),$$
$$(a, b)((a', b')(a'', b'')) = (a(a'a'' + cb'b'') + cb(a'b'' + b'a''), a(a'b'' + b'a'') + b(a'a'' + cb'b'')),$$

and both first coordinates equal $aa'a'' + c(ab'b'' + ba'b'' + bb'a'')$, while both second coordinates equal $aa'b'' + ab'a'' + ba'a'' + cbb'b''$. So for every c we have a commutative ring, and it is a field exactly when every nonzero element has a multiplicative inverse.

If $c < 0$. For $(a, b) \neq (0, 0)$, $q = a^2 - cb^2 = a^2 + |c|b^2 > 0$. Since $(a, b)(a, -b) = \left(a^2 - cb^2, -ab + ba\right) = (q, 0)$, the element $(a/q, -b/q)$ is an inverse of (a, b). So the structure is a field.

If $c \geq 0$. Let $t = \sqrt{c}$. Then

$$(t, 1)(-t, 1) = \left(-t^2 + c, t - t\right) = (0, 0),$$

while $(t, 1)$ and $(-t, 1)$ are nonzero. By Proposition 1.1.17 the structure is not a field.

Solution to Exercise 1.1.55

Uses. Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.15 (Multiplying by zero), the division algorithm: $x \bmod n = y \bmod n$ if and only if $n \mid x - y$
Hint 1. Without distributivity, $+$ and $\cdot$ are unrelated. Keep the addition of $\mathbb{Z}/3\mathbb{Z}$ and look for a separate multiplication on $\{0, 1, 2\}$.
Hint 2. The axioms about multiplication alone ask for a commutative, associative operation with identity 1 in which every nonzero element has an inverse. The addition of $\mathbb{Z}/3\mathbb{Z}$ has all these properties, but its identity is 0. Shift it so that the identity becomes 1.

Hint 3. Try $a * b = (a + b - 1) \bmod 3$.
Sketch. On $\{0, 1, 2\}$ take the addition of $\mathbb{Z}/3\mathbb{Z}$ and the multiplication $a * b = (a + b - 1) \bmod 3$. This multiplication is commutative and associative with identity 1, and every element has an inverse, but $0 * 0 = 2$.

Let $F = \{0, 1, 2\}$ with the addition $+_3$ of $\mathbb{Z}/3\mathbb{Z}$, and define

$$a * b = (a + b - 1) \bmod 3.$$

The additive axioms hold by Lemma 1.1.21. For the multiplication we use that $x \bmod 3 = y \bmod 3$ whenever $3 \mid x - y$. The operation $*$ is commutative since $a + b - 1 = b + a - 1$. It is associative: $(a + b - 1) \bmod 3$ differs from $a + b - 1$ by a multiple of 3, so

$$(a * b) * c = ((a + b - 1) \bmod 3 + c - 1) \bmod 3 = (a + b + c - 2) \bmod 3,$$

and in the same way $a * (b * c) = (a + b + c - 2) \bmod 3$. The element 1 is an identity, since $a * 1 = a \bmod 3 = a$, and $1 \neq 0$. Every element has an inverse: $(2 - a) \bmod 3$ differs from $2 - a$ by a multiple of 3, so $a * ((2 - a) \bmod 3) = (a + 2 - a - 1) \bmod 3 = 1$. So every field axiom except distributivity holds. However,

$$0 * 0 = (-1) \bmod 3 = 2 \neq 0.$$

By Proposition 1.1.15, distributivity must therefore fail. Indeed $0 * (0 + 0) = 2$, while $0 * 0 + 0 * 0 = 2 +_3 2 = 1$.

Solution to Exercise 1.1.56

Uses. Definition 1.1.1 (Field), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Definition 1.1.28 (Characteristic)
Hint 1. Do not check all 64 triples. Identify the elements with pairs of elements of $\mathbb{F}_2$: $0 = (0, 0)$, $1 = (1, 0)$, $\alpha = (0, 1)$, $\beta = (1, 1)$. What does addition become?
Hint 2. The nonzero elements are 1, α, $\alpha^2 = \beta$, and $\alpha^3 = 1$. How does multiplication of nonzero elements act on exponents?
Hint 3. For distributivity, write multiplication by α as a map on pairs and check that it respects addition. Multiplication by β is that map applied twice.
Sketch. As pairs over $\mathbb{F}_2$, addition is coordinatewise, which gives the additive axioms. Nonzero elements are $\alpha^0, \alpha^1, \alpha^2$ with $\alpha^i \alpha^j = \alpha^{i +_3 j}$, so multiplication of nonzero elements is addition in $\mathbb{Z}/3\mathbb{Z}$, which gives the multiplicative axioms. Multiplication by α is $(u, v) \mapsto (v, u + v)$, which respects addition, and multiplication by β is its square, which gives distributivity. Since $1 + 1 = 0$, the characteristic is 2, whereas $\mathbb{Z}/4\mathbb{Z}$ is not even a field.

(a) We check the axioms in groups.

Addition. Identify $0, 1, \alpha, \beta$ with the pairs $(0, 0), (1, 0), (0, 1), (1, 1)$ of elements of $\mathbb{F}_2$. Under this identification the given addition is coordinatewise addition in $\mathbb{F}_2$: each pair added to itself gives $(0, 0)$, and the sum of two different nonzero pairs is the third, for example $(1, 0) + (0, 1) = (1, 1)$. Coordinatewise addition is commutative and associative with identity $(0, 0)$, because addition in $\mathbb{F}_2$ is, and every element is its own additive inverse.

Multiplication. If some factor is 0, both sides of the commutative and associative laws are 0. The nonzero elements are $\alpha^0 = 1$, $\alpha^1 = \alpha$, and $\alpha^2 = \beta$, and $\alpha^3 = \alpha\beta = 1$. Checking the table, $\alpha^i \alpha^j = \alpha^{i +_3 j}$ for all $i, j \in \{0, 1, 2\}$; for example $\beta\beta = \alpha^{2 +_3 2} = \alpha^1 = \alpha$. So multiplying nonzero elements amounts to adding exponents in $\mathbb{Z}/3\mathbb{Z}$, and multiplication of nonzero elements is commutative and associative because addition in $\mathbb{Z}/3\mathbb{Z}$ is (Lemma 1.1.21). The identity is $\alpha^0 = 1 \neq 0$, and α^i has inverse α^{3-i} for $i = 1, 2$.

Distributivity. We show $x(y + z) = xy + xz$. For $x = 0$ and $x = 1$ this is clear. Multiplication by α sends $0 \mapsto 0$, $1 \mapsto \alpha$, $\alpha \mapsto \beta$, $\beta \mapsto 1$, which in pairs is

$$(u, v) \mapsto (v, u + v).$$

This map respects addition: $(u + u', v + v')$ goes to $(v + v', u + u' + v + v') = (v, u + v) + (v', u' + v')$. So $\alpha(y + z) = \alpha y + \alpha z$. Since $\beta = \alpha\alpha$, associativity gives $\beta y = \alpha(\alpha y)$, and therefore

$$\beta(y+z) = \alpha(\alpha y + \alpha z) = \alpha(\alpha y) + \alpha(\alpha z) = \beta y + \beta z.$$

So $\mathbb{F}_4$ is a field.

(b) $1 + 1 = 0$ and $1 \neq 0$, so $\operatorname{char} \mathbb{F}_4 = 2$. Renaming the elements of $\mathbb{Z}/4\mathbb{Z}$ cannot change whether it satisfies the field axioms, and $\mathbb{Z}/4\mathbb{Z}$ is not a field by Proposition 1.1.24 (in it $2 \cdot 2 = 0$). So $\mathbb{F}_4$ is not a renamed $\mathbb{Z}/4\mathbb{Z}$. Another way to see it: $x + x = 0$ for every $x \in \mathbb{F}_4$, while $1 + 1 = 2 \neq 0$ in $\mathbb{Z}/4\mathbb{Z}$.

Solution to Exercise 1.1.57

> *Uses.* Proposition 1.1.13 (Solving $a + x = b$ and $ax = b$), Proposition 1.1.17 (No zero divisors)
> *Hint 1.* Call the sum S. Multiplying every element of F by a fixed nonzero c permutes F. What does that say about cS?
> *Hint 2.* You get $cS = S$. To conclude $S = 0$ you need such a c with $c \neq 1$. When does one exist?
> *Sketch.* If $|F| > 2$, pick $c \notin \{0, 1\}$. Since $x \mapsto cx$ is a bijection of F, $cS = S$, so $(c-1)S = 0$ with $c - 1 \neq 0$, and $S = 0$. If $F = \{0, 1\}$, then $S = 1$.

Let S be the sum of all elements of F. If F has exactly two elements, then $F = \{0, 1\}$ and $S = 1 \neq 0$, which is the exception.

Suppose F has more than two elements, and pick $c \in F$ with $c \neq 0$ and $c \neq 1$. By part (b) of Proposition 1.1.13, for each $y \in F$ the equation $cx = y$ has exactly one solution, so $x \mapsto cx$ is a bijection $F \to F$. Hence the elements cx, for $x \in F$, are the elements of F in another order, and their sum is S. By distributivity their sum is also cS. So $cS = S$, that is, $(c-1)S = 0$. Since $c - 1 \neq 0$, Proposition 1.1.17 gives $S = 0$.

Solution to Exercise 1.1.58

> *Uses.* Definition 1.1.1 (Field)
> *Techniques.* Proving from axioms (Technique 1.1.7)
> *Hint 1.* Expand $(1 + 1)(a + b)$ in two different ways using distributivity.
> *Hint 2.* You get $a + a + b + b = a + b + a + b$. To finish you want to cancel a on the left and b on the right, but our proof of cancellation used commutativity of addition.
> *Hint 3.* Re-prove what you need from the remaining axioms. First show that if $a + b = 0$ then $b + a = 0$, and that $0 + a = a$. Then both cancellation laws follow.
> *Sketch.* Without commutativity of $+$, the right inverse b of a is also a left inverse, and 0 is also a left identity; this gives cancellation on both sides. Distributivity on each side (multiplication is still commutative) expands $(1 + 1)(a + b)$ as both $a + a + b + b$ and $a + b + a + b$. Cancelling a on the left and b on the right gives $a + b = b + a$.

We may still use commutativity of multiplication, so both distributive laws hold and $1x = x1 = x$. Associativity of $+$ lets us drop parentheses in sums (the remark in Section 1.1.2 does not need commutativity for this).

Inverses work on both sides. Let $a + b = 0$, and choose c with $b + c = 0$. Then

$$b + a = (b + a) + 0 = (b + a) + (b + c) = b + (a + b) + c = b + 0 + c = b + c = 0.$$

0 *is a left identity.* With $a + b = 0$ as before, $0 + a = (a + b) + a = a + (b + a) = a + 0 = a$.

Cancellation on both sides. Suppose $x + y = x + z$, and let $x + w = 0$, so $w + x = 0$ as well. Adding w on the left, $y = 0 + y = (w + x) + y = w + (x + y) = w + (x + z) = (w + x) + z = 0 + z = z$. Similarly, if $y + x = z + x$, adding w on the right gives $y = y + (x + w) = (y + x) + w = (z + x) + w = z$.

The main computation. By distributivity, applied in two different ways,

$$(1 + 1)(a + b) = (1 + 1)a + (1 + 1)b = a + a + b + b,$$
$$(1 + 1)(a + b) = 1(a + b) + 1(a + b) = a + b + a + b.$$

So $a + (a + b) + b = a + (b + a) + b$. Cancelling a on the left and b on the right gives $a + b = b + a$, as desired.

Section 1.2

Solution to Exercise 1.2.8

Uses. Definition 1.2.4 (The space F^n), Definition 1.2.7 (Matrices), Proposition 1.1.13 (Solving $a+x=b$ and $ax=b$), Proposition 1.1.30 ($\mathbb{C}$ is a field)
Hint 1. The operations are coordinatewise, so each equation is a separate equation in F for each coordinate or entry.
Hint 2. Solve each one with Proposition 1.1.13: subtract, then multiply by the inverse of the coefficient. In $\mathbb{F}_3$, $2^{-1}=2$; in $\mathbb{F}_5$, $3^{-1}=2$; in $\mathbb{C}$, $(1+i)^{-1}=(1-i)/2$.
Sketch. (a) $2\mathbf{x}=(1,1,1)$, so $\mathbf{x}=(2,2,2)$. (b) $3X=B-A=\begin{pmatrix}4&4\\3&1\end{pmatrix}$, so $X=\begin{pmatrix}3&3\\1&2\end{pmatrix}$. (c) $(1+i)\mathbf{v}=(-1,2i)$, so $\mathbf{v}=((-1+i)/2,1+i)$.

In each part, both sides are equal exactly when they agree in every coordinate (or entry), so we solve one equation in F per coordinate, using Proposition 1.1.13.

(a) In $\mathbb{F}_3$, the coordinates of $(2,0,1)$ minus those of $(1,2,0)$ are $2-1=1$, $0-2=-2=1$, and $1-0=1$, so $2\mathbf{x}=(1,1,1)$. Since $2\cdot 2=4=1$ in $\mathbb{F}_3$, $2^{-1}=2$, and $\mathbf{x}=(2,2,2)$. Check: $2(2,2,2)+(1,2,0)=(1,1,1)+(1,2,0)=(2,0,1)$ in $\mathbb{F}_3$.

(b) Entrywise in $\mathbb{F}_5$,

$$3X=\begin{pmatrix}0-1&1-2\\1-3&0-4\end{pmatrix}=\begin{pmatrix}4&4\\3&1\end{pmatrix}.$$

Since $3\cdot 2=6=1$ in $\mathbb{F}_5$, $3^{-1}=2$, so

$$X=2\begin{pmatrix}4&4\\3&1\end{pmatrix}=\begin{pmatrix}8&8\\6&2\end{pmatrix}=\begin{pmatrix}3&3\\1&2\end{pmatrix}.$$

Check: $3X=\begin{pmatrix}9&9\\3&6\end{pmatrix}=\begin{pmatrix}4&4\\3&1\end{pmatrix}$, and adding A gives $\begin{pmatrix}5&6\\6&5\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}=B$ in $\mathbb{F}_5$.

(c) We need $(1+i)v_1=-1$ and $(1+i)v_2=2i$. By Proposition 1.1.30, $(1+i)^{-1}=(1-i)/2$, so

$$v_1=\frac{-(1-i)}{2}=\frac{-1+i}{2},\quad v_2=\frac{2i(1-i)}{2}=i-i^2=1+i.$$

Check: $(1+i)(-1+i)/2=\left(-1+i-i+i^2\right)/2=-1$, and $(1+i)(1+i)=2i$.

Solution to Exercise 1.2.24

Uses. Proposition 1.2.22 (Transport of structure), Example 1.2.23 (Repairing the line)
Hint 1. The zero vector is defined by (A3) relative to the addition. Which addition is the argument assuming?
Hint 2. To refute the claim, transport the structure of $\mathbb{R}$ along a bijection $\mathbb{R}\to(1,2)$. The function $g(t)=t/(1+|t|)$ is a bijection $\mathbb{R}\to(-1,1)$.
Sketch. The zero vector depends on the operations; 0 is the zero of $\mathbb{R}$ for its usual addition, which is not an operation on $(1,2)$ anyway. Transporting along the bijection $\varphi(t)=(3+g(t))/2$, $g(t)=t/(1+|t|)$, makes $(1,2)$ a real vector space with zero vector $\varphi(0)=3/2$.

The error is the assertion that the zero vector must be 0. The zero vector is whatever element satisfies (A3) for the addition in question, and 0 plays that role only for the usual addition of $\mathbb{R}$, which is not even an operation on $(1,2)$. A different addition on $(1,2)$ can have a different zero vector.

The conclusion is false. Let $g(t)=t/(1+|t|)$. For $t\geq 0$, $g(t)=t/(1+t)\in[0,1)$, and for $s\in[0,1)$ the equation $s=t/(1+t)$ has the unique solution $t=s/(1-s)\geq 0$. For $t<0$, $g(t)=t/(1-t)\in(-1,0)$, and for $s\in(-1,0)$ the equation $s=t/(1-t)$ has the unique solution $t=s/(1+s)<0$. So g is a bijection $\mathbb{R}\to(-1,1)$ with inverse $s\mapsto s/(1-|s|)$. Then $\varphi(t)=(3+g(t))/2$ is a bijection $\mathbb{R}\to(1,2)$, and by Proposition 1.2.22, transporting the structure of $\mathbb{R}$ along φ makes $(1,2)$ a real vector space. Its zero vector is $\varphi(0)=3/2$.

Solution to Exercise 1.2.37

Uses. Proposition 1.2.30 (Additive arithmetic), Corollary 1.2.36 (Cancelling scalars and vectors), Proposition 1.2.32 (The scalar zero), Definition 1.2.1 (Vector space)
Hint 1. First isolate $a\mathbf{x}$ with Proposition 1.2.30(a), as you would isolate ax in a field.
Hint 2. For existence, try $\mathbf{x}=a^{-1}(\mathbf{w}-\mathbf{u})$. For uniqueness, two solutions $\mathbf{x},\mathbf{x}'$ have $a\mathbf{x}=a\mathbf{x}'$.

Sketch. The equation holds if and only if $a\mathbf{x} = \mathbf{w} - \mathbf{u}$. If $a \neq 0$, then $\mathbf{x} = a^{-1}(\mathbf{w} - \mathbf{u})$ works, and it is the only solution by Corollary 1.2.36(a). If $a = 0$, then $a\mathbf{x} = \mathbf{0}$, so there is a solution exactly when $\mathbf{u} = \mathbf{w}$, and then every $\mathbf{x}$ is one.

By Proposition 1.2.30(a), applied with the vector $a\mathbf{x}$ as the unknown, $a\mathbf{x} + \mathbf{u} = \mathbf{w}$ holds if and only if $a\mathbf{x} = \mathbf{w} - \mathbf{u}$ (using (A1) to write $a\mathbf{x} + \mathbf{u} = \mathbf{u} + a\mathbf{x}$).

(a) Let $\mathbf{x} = a^{-1}(\mathbf{w} - \mathbf{u})$. By (A6), $a\mathbf{x} = \left(aa^{-1}\right)(\mathbf{w} - \mathbf{u}) = 1(\mathbf{w} - \mathbf{u}) = \mathbf{w} - \mathbf{u}$, so $\mathbf{x}$ is a solution. If $\mathbf{x}'$ is another, then $a\mathbf{x}' = \mathbf{w} - \mathbf{u} = a\mathbf{x}$, and $\mathbf{x}' = \mathbf{x}$ by Corollary 1.2.36(a). So $a^{-1}(\mathbf{w} - \mathbf{u})$ is the only solution.

(b) If $a = 0$, then $a\mathbf{x} = \mathbf{0}$ for every $\mathbf{x}$ by Proposition 1.2.32, so the equation reads $\mathbf{0} + \mathbf{u} = \mathbf{w}$, that is, $\mathbf{u} = \mathbf{w}$. So if $\mathbf{u} = \mathbf{w}$, every $\mathbf{x} \in V$ is a solution, and otherwise there is none.

Solution to Exercise 1.2.41

Uses. Proposition 1.2.18 (Degree and evaluation versus the operations), Definition 1.2.16 (Degree)
Hint 1. For (a), find two polynomials whose leading terms cancel.
Hint 2. (b) is the contrapositive of part of Proposition 1.2.18(a), plus a look at one coefficient. For (c), check closure.
Sketch. (a) False: $\left(x^2 + 1\right) + \left(-x^2\right) = 1$. (b) True: if $\deg p \neq \deg q$, then $\deg(p + q) = \max(\deg p, \deg q) \geq \deg p$; so $\deg p = \deg q = n$, and the n-th coefficient $a_n + b_n$ of $p + q$ must be 0. (c) False: $x^2 + \left(x - x^2\right) = x$ has degree 1.

(a) False. Take $p = x^2 + 1$ and $q = -x^2$, both of degree 2. Then $p + q = 1$ has degree $0 < 2$.

(b) True. If $\deg p \neq \deg q$, then $\deg(p + q) = \max(\deg p, \deg q) \geq \deg p$ by Proposition 1.2.18(a). So $\deg(p + q) < \deg p$ forces $\deg p = \deg q$; call it n. Write $p = (a_k)$ and $q = (b_k)$. If $a_n + b_n \neq 0$, then the n-th coefficient of $p + q$ is nonzero and $\deg(p + q) \geq n = \deg p$. So $a_n + b_n = 0$.

(c) False. The polynomials x^2 and $x - x^2$ both have degree exactly 2, but their sum x has degree 1. So addition is not an operation on this set. (The zero polynomial, of degree $-\infty$, is not in the set either.)

Solution to Exercise 1.2.42

Uses. Corollary 1.2.36 (Cancelling scalars and vectors)
Hint 1. Pick a nonzero vector $\mathbf{v}$ and look at all of its multiples.
Hint 2. By Corollary 1.2.36(b), the map $F \to V$, $a \mapsto a\mathbf{v}$, is injective.
Sketch. For $\mathbf{v} \neq \mathbf{0}$, $a \mapsto a\mathbf{v}$ is an injective map $F \to V$, so V has at least $|F|$ elements.

Since $V \neq \{\mathbf{0}\}$, there is some $\mathbf{v} \in V$ with $\mathbf{v} \neq \mathbf{0}$. Define $m : F \to V$ by $m(a) = a\mathbf{v}$. If $m(a) = m(b)$, then $a\mathbf{v} = b\mathbf{v}$, so $a = b$ by Corollary 1.2.36(b). So m is injective, and its image is a subset of V with as many elements as F. Hence V has at least as many elements as F. If F is infinite, so is V; and if V is finite, then F, which injects into V, is finite.

Solution to Exercise 1.2.43

Uses. Proposition 1.2.27 (Uniqueness of the zero vector), Proposition 1.2.30 (Additive arithmetic), Remark 1.2.2 (Reading the axioms)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Read Proposition 1.2.27 with its quantifiers. What property does it say only $\mathbf{0}$ has?
Hint 2. For (b), compare $\mathbf{u} + \mathbf{v} = \mathbf{u}$ with $\mathbf{u} + \mathbf{0} = \mathbf{u}$. For (c), the argument in (b) must use (A4), so drop it. Try max on $[0, \infty)$.
Sketch. (a) Proposition 1.2.27 is about vectors $\mathbf{z}$ with $\mathbf{w} + \mathbf{z} = \mathbf{w}$ for *every* $\mathbf{w}$; the hypothesis gives this for one $\mathbf{u}$. (b) $\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{0}$, so $\mathbf{v} = \mathbf{0}$ by cancellation. (c) On $[0, \infty)$ with $s + t = \max(s, t)$, the identity is 0, and $\max(5, 3) = 5$.

(a) Proposition 1.2.27 says that only one vector $\mathbf{z}$ satisfies $\mathbf{w} + \mathbf{z} = \mathbf{w}$ *for every* $\mathbf{w} \in V$. It says nothing about vectors $\mathbf{z}$ that satisfy $\mathbf{u} + \mathbf{z} = \mathbf{u}$ for one particular $\mathbf{u}$, and that is all the hypothesis gives for $\mathbf{v}$. This is the quantifier distinction of Remark 1.2.2. The claim is true, but this argument does not prove it.

(b) By (A3), $\mathbf{u} + \mathbf{v} = \mathbf{u} = \mathbf{u} + \mathbf{0}$. Cancelling $\mathbf{u}$ (Proposition 1.2.30(b)) gives $\mathbf{v} = \mathbf{0}$.

(c) Let $S = [0, \infty)$ with $s + t = \max(s, t)$. This is commutative and associative, and 0 satisfies (A3), since $\max(s, 0) = s$ for $s \geq 0$. But $\max(5, 3) = 5$, so $5 + 3 = 5$ although $3 \neq 0$. The proof in (b) breaks down because cancellation relied on (A4), which fails here: $\max(5, t) \geq 5 > 0$ for every t.

Solution to Exercise 1.2.44

Uses. Corollary 1.2.36 (Cancelling scalars and vectors), Proposition 1.2.30 (Additive arithmetic), Proposition 1.2.28 (Uniqueness of additive inverses), Proposition 1.2.34 (Negatives and the scalar -1), Example 1.2.6 (Small fields)
Hint 1. For (a), a two-element space is $\{\mathbf{0}, \mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$. How many multiples does $\mathbf{v}$ have?
Hint 2. For (b), take a vector space V of the given size, and a vector outside the multiples of some nonzero $\mathbf{v}$. Where can its sum with $\mathbf{v}$ lie?
Hint 3. Over $\mathbb{F}_2$, $-\mathbf{v} = (-1)\mathbf{v} = \mathbf{v}$; over $\mathbb{F}_3$, $-\mathbf{v} = 2\mathbf{v}$. Rule out each possible value of the sum with cancellation.
Sketch. (a) If $V = \{\mathbf{0}, \mathbf{v}\}$, then $a \mapsto a\mathbf{v}$ injects F into V, so F has exactly two elements; conversely such an F is a two-element vector space over itself. (b) In a three-element $\mathbb{F}_2$-space $\{\mathbf{0}, \mathbf{u}, \mathbf{w}\}$, the sum $\mathbf{u} + \mathbf{w}$ can be none of $\mathbf{0}$, $\mathbf{u}$, $\mathbf{w}$. In a four-element $\mathbb{F}_3$-space, $\mathbf{0}, \mathbf{v}, 2\mathbf{v}$ are distinct; for the fourth element $\mathbf{w}$, the sum $\mathbf{w} + \mathbf{v}$ can be none of the four.

(a) Exactly the fields with two elements. Suppose V is a vector space over F with exactly two elements, so $V = \{\mathbf{0}, \mathbf{v}\}$ with $\mathbf{v} \neq \mathbf{0}$. By Corollary 1.2.36(b), $a \mapsto a\mathbf{v}$ is an injective map $F \to V$, so F has at most two elements. A field has at least two elements, $0 \neq 1$, so F has exactly two. Conversely, if F has exactly two elements, then $F = F^1$ is a vector space over itself (Example 1.2.6) with two elements.

(b) *Over* $\mathbb{F}_2$. Suppose $V = \{\mathbf{0}, \mathbf{u}, \mathbf{w}\}$ with three distinct elements. Consider $\mathbf{u} + \mathbf{w} \in V$. If $\mathbf{u} + \mathbf{w} = \mathbf{u} = \mathbf{u} + \mathbf{0}$, then $\mathbf{w} = \mathbf{0}$ by cancellation (Proposition 1.2.30(b)). If $\mathbf{u} + \mathbf{w} = \mathbf{w}$, then likewise $\mathbf{u} = \mathbf{0}$. If $\mathbf{u} + \mathbf{w} = \mathbf{0}$, then $\mathbf{w} = -\mathbf{u} = (-1)\mathbf{u}$ by Proposition 1.2.28 and Proposition 1.2.34, and $-1 = 1$ in $\mathbb{F}_2$, so $\mathbf{w} = 1\mathbf{u} = \mathbf{u}$. Each case contradicts distinctness.

Over $\mathbb{F}_3$. Suppose V has exactly four elements, and pick $\mathbf{v} \neq \mathbf{0}$ in V. By Corollary 1.2.36(b), the vectors $0\mathbf{v} = \mathbf{0}$, $\mathbf{v}$, and $2\mathbf{v}$ are distinct; let $\mathbf{w}$ be the fourth element of V. Consider $\mathbf{w} + \mathbf{v}$. If it equals $\mathbf{w} = \mathbf{w} + \mathbf{0}$, then $\mathbf{v} = \mathbf{0}$. If it equals $\mathbf{0}$, then $\mathbf{w} = -\mathbf{v} = (-1)\mathbf{v} = 2\mathbf{v}$, as $-1 = 2$ in $\mathbb{F}_3$. If it equals $\mathbf{v} = \mathbf{0} + \mathbf{v}$, then $\mathbf{w} = \mathbf{0}$. If it equals $2\mathbf{v} = \mathbf{v} + \mathbf{v}$, then $\mathbf{w} = \mathbf{v}$. Each case is a contradiction.

Solution to Exercise 1.2.45

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.32 (The scalar zero), Proposition 1.2.35 (Zero products), Definition 1.1.28 (Characteristic)
Hint 1. First show $n * \mathbf{v} = (n * 1)\mathbf{v}$ by induction on n, using (A5) and (A8).
Hint 2. Then (i) $\Rightarrow$ (ii) is Proposition 1.2.32, (ii) $\Rightarrow$ (iii) uses $V \neq \{\mathbf{0}\}$, and (iii) $\Rightarrow$ (i) is Proposition 1.2.35.
Sketch. By (A5), (A8), and induction, $n * \mathbf{v} = (n * 1)\mathbf{v}$. If $n * 1 = 0$ this is $\mathbf{0}$ for all $\mathbf{v}$; some nonzero $\mathbf{v}$ exists; and $(n * 1)\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$ forces $n * 1 = 0$. So the n killing all of V are exactly those with $n * 1 = 0$.

First, $n * \mathbf{v} = (n * 1)\mathbf{v}$ for every $n \geq 1$, by induction on n. For $n = 1$ this is (A5). If it holds for n, then

$$(n + 1) * \mathbf{v} = n * \mathbf{v} + \mathbf{v} = (n * 1)\mathbf{v} + 1\mathbf{v} = (n * 1 + 1)\mathbf{v} = ((n + 1) * 1)\mathbf{v},$$

by the inductive hypothesis, (A5), and (A8).

(i) $\Rightarrow$ (ii). If $n * 1 = 0$, then $n * \mathbf{v} = 0\mathbf{v} = \mathbf{0}$ for every $\mathbf{v}$, by Proposition 1.2.32.

(ii) $\Rightarrow$ (iii). Since $V \neq \{\mathbf{0}\}$, some $\mathbf{v} \neq \mathbf{0}$ exists, and $n * \mathbf{v} = \mathbf{0}$ by (ii).

(iii) $\Rightarrow$ (i). If $\mathbf{v} \neq \mathbf{0}$ and $(n * 1)\mathbf{v} = n * \mathbf{v} = \mathbf{0}$, then $n * 1 = 0$ by Proposition 1.2.35.

By the equivalence of (i) and (ii), the positive integers n with $n * \mathbf{v} = \mathbf{0}$ for all $\mathbf{v}$ are exactly those with $n * 1 = 0$. So the least such n is $\operatorname{char} F$, and there is none exactly when $\operatorname{char} F = 0$.

Solution to Exercise 1.2.46

Uses. Exercise 1.2.45, Example 1.2.39 (When does $\mathbf{v} = -\mathbf{v}$ force $\mathbf{v} = \mathbf{0}$?), Proposition 1.2.27 (Uniqueness of the zero vector)
Hint 1. Whatever the scalar multiplication is, Exercise 1.2.45 constrains the addition. What does it say about $p \cdot \mathbf{v}$ over $\mathbb{F}_p$?
Hint 2. For (b), split into the cases $1 + 1 = 0$ and $1 + 1 \neq 0$ in F, and look at the elements 1 and 2 of $\mathbb{Z}/4\mathbb{Z}$.
Sketch. (a) No: over $\mathbb{F}_p$, $p * \mathbf{v} = \mathbf{0}$ for every $\mathbf{v}$, but $p * 1 = p \neq 0$ in $\mathbb{Z}$. (b) No: if $1 + 1 = 0$ in F, then $\mathbf{v} + \mathbf{v} = \mathbf{0}$ for all $\mathbf{v}$, but $1 + 1 = 2 \neq 0$ in $\mathbb{Z}/4\mathbb{Z}$; if $1 + 1 \neq 0$, then $\mathbf{v} + \mathbf{v} = \mathbf{0}$ forces $\mathbf{v} = \mathbf{0}$, but $2 + 2 = 0$ with $2 \neq 0$.

(a) No. Suppose some scalar multiplication made $\mathbb{Z}$ a vector space over $\mathbb{F}_p$. Its zero vector is the identity element 0 of the addition, by Proposition 1.2.27. In $\mathbb{F}_p$ we have $p * 1 = 0$, so by Exercise 1.2.45, adding any vector to itself p times gives the zero vector. But adding $1 \in \mathbb{Z}$ to itself p times gives $p \neq 0$.

(b) No. Suppose F and a scalar multiplication make $\mathbb{Z}/4\mathbb{Z}$ an F-vector space; again its zero vector is 0. If $1+1=0$ in F, then $\mathbf{v}+\mathbf{v}=\mathbf{0}$ for every $\mathbf{v}$ (Example 1.2.39), but $1+1=2\neq 0$ in $\mathbb{Z}/4\mathbb{Z}$. If $1+1\neq 0$ in F, then $\mathbf{v}+\mathbf{v}=\mathbf{0}$ forces $\mathbf{v}=\mathbf{0}$ (Example 1.2.39), but $2+2=0$ in $\mathbb{Z}/4\mathbb{Z}$ although $2\neq 0$. Both cases are impossible.

Solution to Exercise 1.2.47

Uses. Definition 1.2.1 (Vector space)
Hint 1. Suppose such a $\odot$ exists, and let $r=i\odot 1\in\mathbb{R}$. Compute $i\odot(i\odot 1)$ in two ways.
Hint 2. One way uses (A6): $i\odot(i\odot 1)=(i^2)\odot 1$. For the other, write $r=r\odot 1$ and use (A6) and commutativity in $\mathbb{C}$ to move i past r.
Sketch. No. With $r=i\odot 1$, (A6) gives $i\odot r=i\odot(r\odot 1)=(ri)\odot 1=r\odot(i\odot 1)=r^2$, and also $i\odot r=i\odot(i\odot 1)=(-1)\odot 1=-1$. So $r^2=-1$ with $r\in\mathbb{R}$, which is impossible.

No. Suppose such a $\odot$ exists, and let $r=i\odot 1$, a real number. Since r is real, $r=r\cdot 1=r\odot 1$. By (A6) and commutativity of multiplication in $\mathbb{C}$,

$$i\odot r=i\odot(r\odot 1)=(ir)\odot 1=(ri)\odot 1=r\odot(i\odot 1)=r\odot r=r^2.$$

On the other hand, by (A6),

$$i\odot r=i\odot(i\odot 1)=(ii)\odot 1=(-1)\odot 1=-1.$$

So $r^2=-1$, which no real number satisfies.

The requirement that real scalars act as usual is essential. Without it, there is a complex vector space structure on $(\mathbb{R},+)$, but its construction uses the axiom of choice (Section 15.1).

Solution to Exercise 1.2.48

Uses. Proposition 1.2.5 (F^n is a vector space), Definition 1.1.29 (Complex numbers)
Hint 1. Addition is unchanged, so only (A5)–(A8) need checking, and each reduces to a property of the function $a\mapsto\overline{a}$, $|a|$, or $\operatorname{Re}a$.
Hint 2. (A6) needs the function to be multiplicative, and (A8) needs it to be additive. Check $\overline{a+b}=\overline{a}+\overline{b}$ and $\overline{ab}=\overline{a}\,\overline{b}$ by writing a and b in coordinates.
Sketch. (a) Yes: conjugation fixes 1, and is additive and multiplicative. (b) No: (A8) fails for $a=1$, $b=-1$, since $|0|\mathbf{v}=\mathbf{0}$ but $|1|\mathbf{v}+|-1|\mathbf{v}=2\mathbf{v}$. (c) No: (A6) fails for $a=b=i$, since $\operatorname{Re}(i^2)=-1$ but $\operatorname{Re}(i)\operatorname{Re}(i)=0$.

Since the addition is that of $\mathbb{C}^2$, (A1)–(A4) hold in all three cases. For $g:\mathbb{C}\to\mathbb{C}$ and $a\odot\mathbf{v}=g(a)\mathbf{v}$, (A7) also holds, as $g(a)(\mathbf{v}+\mathbf{w})=g(a)\mathbf{v}+g(a)\mathbf{w}$ in $\mathbb{C}^2$. The remaining axioms read $g(1)\mathbf{v}=\mathbf{v}$, $g(ab)\mathbf{v}=g(a)g(b)\mathbf{v}$, and $g(a+b)\mathbf{v}=g(a)\mathbf{v}+g(b)\mathbf{v}$.

(a) Yes. First, $\overline{1}=1$, which gives (A5). Write $a=x+yi$ and $b=u+wi$. Then $\overline{a+b}=(x+u)-(y+w)i=\overline{a}+\overline{b}$, which gives (A8). Also $ab=(xu-yw)+(xw+yu)i$, so $\overline{ab}=(xu-yw)-(xw+yu)i$, while

$$\overline{a}\,\overline{b}=(x-yi)(u-wi)=(xu-yw)+(-xw-yu)i.$$

These agree, so $\overline{ab}\mathbf{v}=\overline{a}\big(\overline{b}\mathbf{v}\big)$, which gives (A6).

(b) No. Take $a=1$, $b=-1$, and $\mathbf{v}=(1,0)$. Then $(a+b)\odot\mathbf{v}=|0|\mathbf{v}=\mathbf{0}$, while $a\odot\mathbf{v}+b\odot\mathbf{v}=|1|\mathbf{v}+|-1|\mathbf{v}=(2,0)$. So (A8) fails.

(c) No. Take $a=b=i$ and $\mathbf{v}=(1,0)$. Then $(ab)\odot\mathbf{v}=\operatorname{Re}(-1)\mathbf{v}=(-1,0)$, while $a\odot(b\odot\mathbf{v})=\operatorname{Re}(i)(\operatorname{Re}(i)\mathbf{v})=\mathbf{0}$. So (A6) fails. (Here (A5) and (A8) do hold, since $\operatorname{Re}1=1$ and Re is additive.)

Solution to Exercise 1.2.49

Uses. Proposition 1.2.5 (F^n is a vector space)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. Test small cases first: try $a=b=1$ in (A8) over $\mathbb{R}$, and then over $\mathbb{F}_2$.
Hint 2. All axioms except (A8) hold for every F. (A8) holds if and only if $(a+b)^2=a^2+b^2$ for all $a,b\in F$.
Sketch. (A1)–(A7) always hold: $1^2=1$ and $(ab)^2=a^2b^2$. (A8) in the second coordinate needs $(a+b)^2=a^2+b^2$, that is, $(1+1)ab=0$ for all a,b. This holds for all a,b exactly when $1+1=0$. So the answer is: the fields with $1+1=0$.

Exactly the fields with $1+1=0$.

Over $\mathbb{R}$, (A8) fails: $(1+1)\odot(0,1)=(0,4)$, while $1\odot(0,1)+1\odot(0,1)=(0,2)$. Over $\mathbb{F}_2$, $a^2=a$ for both elements, so $\odot$ is the usual scalar multiplication and all axioms hold. These two small cases suggest that $1+1$ is what matters.

In general, (A1)–(A4) hold because the addition is that of F^2. (A5) holds since $1^2=1$. (A6) holds since $(ab)\odot(x,y)=\big(abx,a^2b^2y\big)=a\odot\big(bx,b^2y\big)$. (A7) holds since $a\odot(x+x',y+y')=\big(ax+ax',a^2y+a^2y'\big)$. For (A8), the first coordinates of $(a+b)\odot(x,y)$ and $a\odot(x,y)+b\odot(x,y)$ agree, and the second coordinates are $(a+b)^2y$ and $\big(a^2+b^2\big)y$. Since $(a+b)^2=a^2+(1+1)ab+b^2$, (A8) holds for all a,b,x,y if and only if $(1+1)aby=0$ for all $a,b,y\in F$. If $1+1=0$, this holds. If $1+1\neq 0$, it fails for $a=b=y=1$. So F^2 with $\odot$ is a vector space exactly when $1+1=0$ in F.

Over $\mathbb{F}_2$ the new structure is the usual one. Over a larger field with $1+1=0$, such as the four-element field of Exercise 1.1.56, $a\mapsto a^2$ is not the identity, so $\odot$ is a different scalar multiplication from the usual one.

Solution to Exercise 1.2.50

Uses. Proposition 1.2.22 (Transport of structure)
Hint 1. The parabola is in bijection with $\mathbb{R}$. Transport the structure of $\mathbb{R}$ (Proposition 1.2.22).
Hint 2. Use $\varphi(t)=\big(t,t^2\big)$, whose inverse is $(x,y)\mapsto x$.
Sketch. Transport along $\varphi(t)=\big(t,t^2\big)$: $\big(s,s^2\big)\oplus\big(t,t^2\big)=\big(s+t,(s+t)^2\big)$ and $a\odot\big(t,t^2\big)=\big(at,a^2t^2\big)$. Then $(1,1)\oplus(2,4)=\varphi(3)=(3,9)$, the zero vector is $(0,0)$, $-(2,4)=(-2,4)$, and $3\odot(-1,1)=(-3,9)$.

Define $\varphi:\mathbb{R}\to P$ by $\varphi(t)=\big(t,t^2\big)$. It is a bijection: each point of P is $\big(x,x^2\big)=\varphi(x)$, and $\varphi(s)=\varphi(t)$ forces $s=t$ by comparing first coordinates. So $\varphi^{-1}(x,y)=x$. By Proposition 1.2.22, P is a real vector space with

$$\big(s,s^2\big)\oplus\big(t,t^2\big)=\varphi(s+t)=\big(s+t,(s+t)^2\big),\quad a\odot\big(t,t^2\big)=\varphi(at)=\big(at,a^2t^2\big).$$

In particular $(1,1)\oplus(2,4)=\varphi(1+2)=(3,9)$. The zero vector is $\varphi(0)=(0,0)$. The inverse of $(2,4)=\varphi(2)$ is $\varphi(-2)=(-2,4)$, since $\varphi(2)\oplus\varphi(-2)=\varphi(0)$. Finally $3\odot(-1,1)=\varphi(3\cdot(-1))=(-3,9)$.

The set P is curved, but that is no obstacle: the structure lives in the operations, and only the operations of $\mathbb{R}^2$ fail to preserve P.

Solution to Exercise 1.2.51

Uses. Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations), Exercise 1.1.51
Hint 1. Write $f=a_0+a_1x+a_2x^2+a_3x^3$ and compute $f(0)$ and $f(1)$ in terms of the coefficients.
Hint 2. Two polynomials have the same function exactly when their difference has the zero function, by Proposition 1.2.18(c). For (b), use Fermat's little theorem (Exercise 1.1.51).
Sketch. (a) $f(0)=a_0$ and $f(1)=a_0+a_1+a_2+a_3$, so the zero function needs $a_0=0$ and an even number of a_1,a_2,a_3 equal to 1: $f\in\{\mathbf{0},x+x^2,x+x^3,x^2+x^3\}$. There are 16 polynomials and 4 functions; f and g agree as functions iff $f-g$ is one of these four. (b) Coefficient 1 in position p, so nonzero; $c^p=c$ for all $c\in\mathbb{F}_p$.

(a) Write $f=a_0+a_1x+a_2x^2+a_3x^3$ with $a_k\in\mathbb{F}_2$. Then $f(0)=a_0$ and $f(1)=a_0+a_1+a_2+a_3$. So the polynomial function of f is zero exactly when $a_0=0$ and $a_1+a_2+a_3=0$ in $\mathbb{F}_2$, that is, when an even number of a_1,a_2,a_3 equal 1. This gives

$$f\in\{\mathbf{0},x+x^2,x+x^3,x^2+x^3\}.$$

$\mathbb{F}_2[x]_{\leq 3}$ has $2^4=16$ elements, one for each choice of (a_0,a_1,a_2,a_3), while there are $2^2=4$ functions $\mathbb{F}_2\to\mathbb{F}_2$. By Proposition 1.2.18(c), f and g have the same polynomial function if and only if $(f-g)(c)=f(c)-g(c)=0$ for all c, that is, if and only if $f-g$ is one of the four polynomials above.

(b) The coefficient of x^p-x in position $p\geq 2$ is 1, so $x^p-x\neq\mathbf{0}$; in fact $\deg(x^p-x)=p$. Its value at $c\in\mathbb{F}_p$ is c^p-c. For $c=0$ this is 0. For $c\neq 0$, Fermat's little theorem (Exercise 1.1.51) gives $c^{p-1}=1$, so $c^p=c$ and $c^p-c=0$.

Solution to Exercise 1.2.52

Uses. Definition 1.2.1 (Vector space), Corollary 1.2.36 (Cancelling scalars and vectors), Proposition 1.2.34 (Negatives and the scalar -1), Proposition 1.2.28 (Uniqueness of additive inverses), Proposition 1.2.27 (Uniqueness of the zero vector), Proposition 1.2.32 (The scalar zero)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. (a) is induction with (A5) and (A8). For (b), check that $q\mathbf{v}$ works using (A6), and get uniqueness from Corollary 1.2.36(a).
Hint 2. For (c), show first that $m \odot \mathbf{v}$, for every integer m, is expressible using only $+$ and negatives. Then use (b) for both structures.
Sketch. (a) $(n+1)\mathbf{w} = n\mathbf{w} + 1\mathbf{w}$. (b) $n(q\mathbf{v}) = (nq)\mathbf{v} = m\mathbf{v}$, and $n\mathbf{w} = n\mathbf{w}'$ with $n \neq 0$ forces $\mathbf{w} = \mathbf{w}'$. (c) By (a), $m \odot \mathbf{v}$ is an m-fold sum for $m \geq 1$, it is $\mathbf{0}$ for $m = 0$, and $-((-m) \odot \mathbf{v})$ for $m < 0$; so it agrees with $m \odot' \mathbf{v}$. Then $q \odot \mathbf{v}$ and $q \odot' \mathbf{v}$ both solve "$\mathbf{w} + \cdots + \mathbf{w} = m \odot \mathbf{v}$" ($n$ terms), which has at most one solution.

(a) By induction on n. For $n = 1$, $1\mathbf{w} = \mathbf{w}$ by (A5). If $n\mathbf{w}$ is the n-fold sum, then $(n+1)\mathbf{w} = n\mathbf{w} + 1\mathbf{w} = n\mathbf{w} + \mathbf{w}$ by (A8) and (A5), which is the $(n+1)$-fold sum.

(b) By (A6), $n(q\mathbf{v}) = (nq)\mathbf{v} = m\mathbf{v}$, so $q\mathbf{v}$ is such a vector. If $n\mathbf{w} = m\mathbf{v} = n\mathbf{w}'$, then $\mathbf{w} = \mathbf{w}'$ by Corollary 1.2.36(a), since $n \neq 0$ in $\mathbb{Q}$.

(c) Let $\odot$ and $\odot'$ be as stated; both have the same addition, hence the same zero vector and the same negatives (Proposition 1.2.27 and Proposition 1.2.28). First let $m \in \mathbb{Z}$. If $m \geq 1$, then by (a) both $m \odot \mathbf{v}$ and $m \odot' \mathbf{v}$ equal the m-fold sum $\mathbf{v} + \cdots + \mathbf{v}$. If $m = 0$, both are $\mathbf{0}$ by Proposition 1.2.32. If $m < 0$, then by Proposition 1.2.34(b), applied to both structures,

$$m \odot \mathbf{v} = -((-m) \odot \mathbf{v}) = -((-m) \odot' \mathbf{v}) = m \odot' \mathbf{v}.$$

So $m \odot \mathbf{v} = m \odot' \mathbf{v}$ for every integer m.

Now let $q = m/n$ with $n \geq 1$, and write $n * \mathbf{w}$ for the n-fold sum $\mathbf{w} + \cdots + \mathbf{w}$, which by (a) equals both $n \odot \mathbf{w}$ and $n \odot' \mathbf{w}$. By (b) for $\odot$, the vector $q \odot \mathbf{v}$ satisfies $n * (q \odot \mathbf{v}) = m \odot \mathbf{v}$, and it is the only vector $\mathbf{w}$ with $n * \mathbf{w} = m \odot \mathbf{v}$. By (b) for $\odot'$, $n * (q \odot' \mathbf{v}) = m \odot' \mathbf{v} = m \odot \mathbf{v}$. So $q \odot' \mathbf{v}$ is also such a vector, and $q \odot' \mathbf{v} = q \odot \mathbf{v}$.

For instance, the usual product is the only scalar multiplication that makes $(\mathbb{R}, +)$ a vector space over $\mathbb{Q}$. Compare Exercise 1.2.47, where the scalars were $\mathbb{C}$.

Solution to Exercise 1.2.53

Uses. Proposition 1.2.22 (Transport of structure), $\tanh : \mathbb{R} \to (-1, 1)$ is a bijection, the addition formula for tanh
Hint 1. For (a), compare $1 + uv$ with $u + v$ and $-(u + v)$; both differences factor.
Hint 2. For (b), look for a bijection $\varphi : \mathbb{R} \to (-1, 1)$ with $\varphi(s + t) = \varphi(s) \oplus \varphi(t)$. Which function has an addition formula of this shape?
Hint 3. Use the addition formula for tanh, and transport the structure of $\mathbb{R}$ along tanh.
Sketch. (a) $1 + uv > 0$, and $(1 + uv) \mp (u + v) = (1 \mp u)(1 \mp v) > 0$. (b) tanh is a bijection $\mathbb{R} \to (-1, 1)$ carrying $+$ to $\oplus$, so transporting along it gives $a \odot v = \tanh(a \operatorname{artanh} v)$, which can be written with powers of $1 \pm v$. Then $2 \odot v = 2v/(1 + v^2)$, $(1/2) \odot v = \left(1 - \sqrt{1 - v^2}\right)/v$ for $v \neq 0$, the zero vector is 0, and the inverse of v is $-v$.

(a) Let $u, v \in (-1, 1)$. Then $|uv| < 1$, so $1 + uv > 0$. Also

$$(1 + uv) - (u + v) = (1 - u)(1 - v) > 0, \quad (1 + uv) + (u + v) = (1 + u)(1 + v) > 0.$$

So $-(1 + uv) < u + v < 1 + uv$, and dividing by $1 + uv > 0$ gives $-1 < u \oplus v < 1$.

(b) The function $\tanh : \mathbb{R} \to (-1, 1)$ is a bijection: it is continuous and strictly increasing, with limits -1 and 1 at $-\infty$ and ∞. Its inverse is $\operatorname{artanh} v = (1/2)\ln((1 + v)/(1 - v))$. The addition formula

$$\tanh(s + t) = \frac{\tanh s + \tanh t}{1 + \tanh s \tanh t}$$

says exactly that $\tanh(s + t) = \tanh(s) \oplus \tanh(t)$. Hence, with $u = \tanh s$ and $v = \tanh t$, we get $u \oplus v = \tanh(\operatorname{artanh} u + \operatorname{artanh} v)$, which is the addition obtained by transporting the structure of $\mathbb{R}$ along tanh. By Proposition 1.2.22, V becomes a real vector space with this addition and

$$a \odot v = \tanh(a \operatorname{artanh} v).$$

Using $\tanh s = \left(e^{2s}-1\right)/\left(e^{2s}+1\right)$ and $e^{2a \operatorname{artanh} v} = ((1+v)/(1-v))^a$,

$$a \odot v = \frac{(1+v)^a - (1-v)^a}{(1+v)^a + (1-v)^a}.$$

For $a = 2$, this is $(4v)/\left(2+2v^2\right) = 2v/\left(1+v^2\right)$, which equals $v \oplus v$, as (A5) and (A8) require. For $a = 1/2$, multiplying the numerator and denominator by $\sqrt{1+v} - \sqrt{1-v}$ gives, for $v \neq 0$,

$$\frac{1}{2} \odot v = \frac{\sqrt{1+v}-\sqrt{1-v}}{\sqrt{1+v}+\sqrt{1-v}} = \frac{2-2\sqrt{1-v^2}}{2v} = \frac{1-\sqrt{1-v^2}}{v},$$

and $(1/2) \odot 0 = 0$. For example, $(1/2) \odot (4/5) = (1-3/5)/(4/5) = 1/2$, and indeed $(1/2) \oplus (1/2) = 1/(5/4) = 4/5$. The zero vector is $\tanh 0 = 0$, and the inverse of $v = \tanh t$ is $\tanh(-t) = -v$; directly, $v \oplus (-v) = 0/\left(1-v^2\right) = 0$.

Solution to Exercise 1.2.54

Uses. Proposition 1.2.32 (The scalar zero), Proposition 1.2.30 (Additive arithmetic), Definition 1.2.1 (Vector space)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. Trace the proof of Proposition 1.2.32 back to the axioms. Which ones does it use?
Hint 2. For (b), you need an addition with an identity but without inverses. On $\{0,1\}$, try $u+v = \max(u,v)$. Then look for a scalar multiplication satisfying (A5)–(A8); a very lazy one works.
Sketch. (a) Proposition 1.2.32 cancels $0\mathbf{v}$, and cancellation (Proposition 1.2.30) was proved with (A4); the argument is circular. (b) On $\{0,1\}$ let $u+v = \max(u,v)$ and $a \odot v = v$ for all a. (A1)–(A3) hold with identity 0, and (A5)–(A8) hold because $\max(v,v) = v$. But $1+w = 1$ for every w, so 1 has no inverse.

(a) The argument uses Proposition 1.2.32, whose proof cancelled $0\mathbf{v}$ from $0\mathbf{v} + 0\mathbf{v} = 0\mathbf{v} + \mathbf{0}$. That cancellation is Proposition 1.2.30(b), which was proved by adding $-\mathbf{u}$, an inverse supplied by (A4). So the argument assumes (A4) in order to prove it. The computation does show something true: in a vector space, $(-1)\mathbf{v}$ *is* the inverse of $\mathbf{v}$ (Proposition 1.2.34).

(b) Let $V = \{0,1\}$, let F be any field, and define

$$u + v = \max(u,v), \quad a \odot v = v \quad (a \in F,\ u,v \in V).$$

(A1) and (A2) hold because max is commutative and associative. (A3) holds with zero vector 0, as $\max(v,0) = v$. (A5): $1 \odot v = v$. (A6): $(ab) \odot v = v = a \odot (b \odot v)$. (A7): $a \odot (u+v) = u+v = (a \odot u) + (a \odot v)$. (A8): $(a+b) \odot v = v = \max(v,v) = (a \odot v) + (b \odot v)$. But (A4) fails for $v = 1$: $1 + w = \max(1,w) = 1 \neq 0$ for both $w \in V$.

Consistently with (a), the conclusion of Proposition 1.2.32 fails here: $0 \odot 1 = 1$, which is not the zero vector.

Solution to Exercise 1.2.55

Uses. Proposition 1.2.27 (Uniqueness of the zero vector), Proposition 1.2.30 (Additive arithmetic), Proposition 1.2.32 (The scalar zero), Proposition 1.2.33 (The zero vector), Proposition 1.2.28 (Uniqueness of additive inverses)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. First list the consequences of the axioms that remain available: everything in Proposition 1.2.27, Proposition 1.2.28, Proposition 1.2.30, Proposition 1.2.32, and Proposition 1.2.33, since none of their proofs used (A5). Check that $a(-\mathbf{u}) = -(a\mathbf{u})$ still holds.
Hint 2. For existence in (a), try $\mathbf{w} = 1\mathbf{v}$ and $\mathbf{n} = \mathbf{v} - 1\mathbf{v}$. Use (A6) to compute $1(1\mathbf{v})$ and $a(1\mathbf{v})$.
Hint 3. For uniqueness, apply the scalar 1 to $\mathbf{v} = \mathbf{w} + \mathbf{n}$.
Sketch. Without (A5), the results through Proposition 1.2.33 still hold, and $a(-\mathbf{u}) = -(a\mathbf{u})$ follows from (A7). Then $1(1\mathbf{v}) = 1\mathbf{v}$ and $a(\mathbf{v} - 1\mathbf{v}) = a\mathbf{v} - (a1)\mathbf{v} = \mathbf{0}$, so $\mathbf{v} = 1\mathbf{v} + (\mathbf{v} - 1\mathbf{v})$ is a decomposition, and applying 1 to any decomposition recovers $\mathbf{w} = 1\mathbf{v}$. W contains $\mathbf{0}$ and is closed under the operations and negatives. If (A5) holds, $\mathbf{n} = 1\mathbf{n} = \mathbf{0}$; if $N = \{\mathbf{0}\}$, every $\mathbf{v} = 1\mathbf{v}$.

What is still available. The proofs of Proposition 1.2.27, Proposition 1.2.28, and Proposition 1.2.30 used only (A1)–(A4). The proofs of Proposition 1.2.32 and Proposition 1.2.33 (first approach) used in addition only (A8) and (A7). So all of these hold in V. We also need $a(-\mathbf{u}) = -(a\mathbf{u})$, whose earlier proof used (A5). Instead, by (A7) and Proposition 1.2.33, $a\mathbf{u} + a(-\mathbf{u}) = a(\mathbf{u} + (-\mathbf{u})) = a\mathbf{0} = \mathbf{0}$, so $a(-\mathbf{u}) = -(a\mathbf{u})$ by Proposition 1.2.28.

(a) *Existence.* Let $\mathbf{w} = 1\mathbf{v}$ and $\mathbf{n} = \mathbf{v} - 1\mathbf{v}$. By (A6), $1\mathbf{w} = 1(1\mathbf{v}) = (1 \cdot 1)\mathbf{v} = 1\mathbf{v} = \mathbf{w}$, so $\mathbf{w} \in W$. For $a \in F$, by (A7), the rule above, and (A6),

$$a\mathbf{n} = a\mathbf{v} + a(-(1\mathbf{v})) = a\mathbf{v} - a(1\mathbf{v}) = a\mathbf{v} - (a \cdot 1)\mathbf{v} = a\mathbf{v} - a\mathbf{v} = \mathbf{0},$$

so $\mathbf{n} \in N$. And $\mathbf{w} + \mathbf{n} = 1\mathbf{v} + (\mathbf{v} - 1\mathbf{v}) = \mathbf{v}$ by Proposition 1.2.30(a).

Uniqueness. Suppose $\mathbf{v} = \mathbf{w} + \mathbf{n}$ with $\mathbf{w} \in W$ and $\mathbf{n} \in N$. By (A7), $1\mathbf{v} = 1\mathbf{w} + 1\mathbf{n} = \mathbf{w} + \mathbf{0} = \mathbf{w}$. So $\mathbf{w} = 1\mathbf{v}$ is determined by $\mathbf{v}$, and then $\mathbf{n} = \mathbf{v} - \mathbf{w}$ is determined by Proposition 1.2.30(a).

(b) We have $\mathbf{0} \in W$, since $1\mathbf{0} = \mathbf{0}$ by Proposition 1.2.33. For $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$: $1(\mathbf{w} + \mathbf{w}') = 1\mathbf{w} + 1\mathbf{w}' = \mathbf{w} + \mathbf{w}'$ by (A7); $1(a\mathbf{w}) = (1a)\mathbf{w} = a\mathbf{w}$ by (A6); and $1(-\mathbf{w}) = -(1\mathbf{w}) = -\mathbf{w}$ by the rule above. So W is closed under addition, scalar multiplication, and negatives, and the operations of V restrict to operations on W. The equations in (A1), (A2), and (A6)–(A8) hold in W because they hold in V; (A3) and (A4) hold because $\mathbf{0} \in W$ and $-\mathbf{w} \in W$; and (A5) holds in W by the definition of W.

(c) If (A5) holds and $\mathbf{n} \in N$, then $\mathbf{n} = 1\mathbf{n} = \mathbf{0}$, so $N = \{\mathbf{0}\}$ (and $\mathbf{0} \in N$ by Proposition 1.2.33). Conversely, if $N = \{\mathbf{0}\}$, then in the decomposition of (a), $\mathbf{v} = 1\mathbf{v} + \mathbf{0} = 1\mathbf{v}$ for every $\mathbf{v}$, which is (A5).

So without (A5), V splits into a genuine vector space W and a part N that every scalar sends to $\mathbf{0}$; (A5) is exactly what removes N. Example (a) of Example 1.2.25 is an instance, with W the x-axis and N the y-axis.

Solution to Exercise 1.2.56

> *Uses.* Definition 1.2.1 (Vector space), Exercise 1.1.58
> *Techniques.* Proving from axioms (Technique 1.1.7)
> *Hint 1.* Expand $(1 + 1)(\mathbf{u} + \mathbf{v})$ in two ways, once with (A8) first and once with (A7) first.
> *Hint 2.* You get $\mathbf{u} + \mathbf{v} + \mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{u} + \mathbf{v} + \mathbf{v}$. To cancel $\mathbf{u}$ on the left and $\mathbf{v}$ on the right, you need two-sided inverses and a two-sided zero, which (A3) and (A4) do not provide directly.
> *Hint 3.* Use scalars. Show $0\mathbf{v} = \mathbf{0}$ using only right-hand inverses, then check that $(-1)\mathbf{v}$ is an inverse on both sides and that $\mathbf{0} + \mathbf{v} = 0\mathbf{v} + 1\mathbf{v} = \mathbf{v}$.
> *Sketch.* Adding a right inverse of $0\mathbf{v}$ to $0\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$ gives $0\mathbf{v} = \mathbf{0}$. Then (A8) gives $\mathbf{v} + (-1)\mathbf{v} = (-1)\mathbf{v} + \mathbf{v} = 0\mathbf{v} = \mathbf{0}$ and $\mathbf{0} + \mathbf{v} = (0 + 1)\mathbf{v} = \mathbf{v}$, so cancellation works on both sides. Expanding $(1 + 1)(\mathbf{u} + \mathbf{v})$ two ways gives $\mathbf{u} + (\mathbf{v} + \mathbf{u}) + \mathbf{v} = \mathbf{u} + (\mathbf{u} + \mathbf{v}) + \mathbf{v}$; cancel.

By (A2) we may drop parentheses in sums. We cannot reorder terms.

Step 1: $0\mathbf{v} = \mathbf{0}$. By (A8), $0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$. By (A4) choose $\mathbf{y}$ with $0\mathbf{v} + \mathbf{y} = \mathbf{0}$, and add it on the right:

$$\mathbf{0} = 0\mathbf{v} + \mathbf{y} = 0\mathbf{v} + 0\mathbf{v} + \mathbf{y} = 0\mathbf{v} + \mathbf{0} = 0\mathbf{v}.$$

Step 2: two-sided inverses and zero. By (A5), (A8), and Step 1,

$$\mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}, \quad (-1)\mathbf{v} + \mathbf{v} = ((-1) + 1)\mathbf{v} = \mathbf{0},$$

and $\mathbf{0} + \mathbf{v} = 0\mathbf{v} + 1\mathbf{v} = (0 + 1)\mathbf{v} = \mathbf{v}$.

Step 3: cancellation on both sides. If $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$, adding $(-1)\mathbf{x}$ on the left and using Step 2 gives $\mathbf{0} + \mathbf{y} = \mathbf{0} + \mathbf{z}$, so $\mathbf{y} = \mathbf{z}$. If $\mathbf{y} + \mathbf{x} = \mathbf{z} + \mathbf{x}$, adding $(-1)\mathbf{x}$ on the right and using Step 2 and (A3) gives $\mathbf{y} = \mathbf{z}$.

Step 4. By (A8) and (A5),

$$(1 + 1)(\mathbf{u} + \mathbf{v}) = (\mathbf{u} + \mathbf{v}) + (\mathbf{u} + \mathbf{v}) = \mathbf{u} + (\mathbf{v} + \mathbf{u}) + \mathbf{v},$$

and by (A7), then (A8) and (A5),

$$(1 + 1)(\mathbf{u} + \mathbf{v}) = (1 + 1)\mathbf{u} + (1 + 1)\mathbf{v} = \mathbf{u} + \mathbf{u} + \mathbf{v} + \mathbf{v} = \mathbf{u} + (\mathbf{u} + \mathbf{v}) + \mathbf{v}.$$

Cancelling $\mathbf{u}$ on the left and $\mathbf{v}$ on the right (Step 3) gives $\mathbf{v}+\mathbf{u}=\mathbf{u}+\mathbf{v}$, as desired.

Compare Exercise 1.1.58, the same statement for fields. There the two-sided inverses had to be built from the additive axioms alone; here the scalar -1 supplies them.

Solution to Exercise 1.2.57

Uses. Definition 1.2.1 (Vector space), between any two distinct real numbers there is a rational number
Hint 1. For (a), (A7) always holds, and each of (A5), (A6), (A8) is equivalent to one of the three conditions; for the forward direction, take $x=1$.
Hint 2. For (b), first find f on $\mathbb{Z}$ and then on $\mathbb{Q}$ using additivity and $f(1)=1$.
Hint 3. Multiplicativity gives $f(a^2)=f(a)^2\geq 0$, so f sends nonnegative numbers to nonnegative numbers. With additivity, f is nondecreasing. Now squeeze $f(a)$ between rationals.
Sketch. (a) (A5), (A6), (A8) at $x=1$ are the three conditions, and conversely the conditions give the axioms for all x. (b) Additivity and $f(1)=1$ give $f(q)=q$ for rational q. Since $f(b)=f\left(\sqrt{b}\right)^2\geq 0$ for $b\geq 0$, f is nondecreasing. If $f(a)<a$, a rational q with $f(a)<q<a$ gives $q=f(q)\leq f(a)$, a contradiction; similarly $f(a)>a$ is impossible.

(a) Since the addition is that of $\mathbb{R}$, (A1)–(A4) hold, and (A7) holds because $f(a)(x+y)=f(a)x+f(a)y$. The other axioms read, for all $a,b,x\in\mathbb{R}$,

$$\text{(A5)}\ \ f(1)x=x,\quad \text{(A6)}\ \ f(ab)x=f(a)f(b)x,\quad \text{(A8)}\ \ f(a+b)x=f(a)x+f(b)x.$$

Taking $x=1$ shows that they imply $f(1)=1$, $f(ab)=f(a)f(b)$, and $f(a+b)=f(a)+f(b)$. Conversely, multiplying these three equations by x gives the three axioms.

(b) Let f satisfy the three conditions.

f fixes $\mathbb{Q}$. From $f(0)=f(0+0)=f(0)+f(0)$ we get $f(0)=0$, and then $f(a)+f(-a)=f(0)=0$ gives $f(-a)=-f(a)$. For an integer $n\geq 1$, additivity and $f(1)=1$ give $f(n)=f(1)+\cdots+f(1)=n$ by induction, and so $f(n)=n$ for all $n\in\mathbb{Z}$. For $q=m/n$ with $n\geq 1$, $nf(q)=f(nq)=f(m)=m$ (again by additivity), so $f(q)=q$.

f is nondecreasing. If $b\geq 0$, then $f(b)=f\left(\sqrt{b}\sqrt{b}\right)=f\left(\sqrt{b}\right)^2\geq 0$. So if $a\leq b$, then $f(b)-f(a)=f(b-a)\geq 0$.

$f(a)=a$. Let $a\in\mathbb{R}$, and suppose $f(a)<a$. Choose a rational q with $f(a)<q<a$. Since $q<a$ and f is nondecreasing, $q=f(q)\leq f(a)$, contradicting $f(a)<q$. If $f(a)>a$, choose a rational q with $a<q<f(a)$; then $f(a)\leq f(q)=q$, again a contradiction. So $f(a)=a$.

Over $\mathbb{C}$ the analogous question has at least two answers, the identity and conjugation (Exercise 1.2.48). The proof above used the order of $\mathbb{R}$, which $\mathbb{C}$ lacks.

Section 1.3

Solution to Exercise 1.3.6

Uses. Definition 1.3.4 (Linear combination), Definition 1.2.4 (The space F^n)
Hint 1. Compute the vector represented by an arbitrary linear combination $a_1(0,0)+a_2(1,-1)+a_3(-\pi,\pi)$, one coordinate at a time.
Hint 2. The second coordinate is always the negative of the first, so a single equation constrains a_1,a_2,a_3. Which coefficient does not appear in it?
Sketch. The combination represents $(a_2-\pi a_3,-(a_2-\pi a_3))$. It represents $(1,-1)$ exactly when $a_2=1+\pi a_3$, and $\mathbf{0}$ exactly when $a_2=\pi a_3$, with a_1 and a_3 free in both cases. Adding the coefficients of each representation of $\mathbf{0}$ to those of one representation of $(1,-1)$ gives all representations of $(1,-1)$.

By the coordinatewise operations, the linear combination with coefficients a_1,a_2,a_3 represents

$$a_1(0,0)+a_2(1,-1)+a_3(-\pi,\pi)=(a_2-\pi a_3,-a_2+\pi a_3).$$

The second coordinate is the negative of the first. So the combination represents $(1,-1)$ exactly when $a_2-\pi a_3=1$, and it represents $\mathbf{0}$ exactly when $a_2-\pi a_3=0$; the coefficient a_1 plays no role. Hence the linear combinations representing $(1,-1)$ are those with coefficients

$$a_1, \quad 1+\pi t, \quad t \quad (a_1, t \in \mathbb{R}),$$

and those representing $\mathbf{0}$ are those with coefficients $a_1, \pi t, t$ for $a_1, t \in \mathbb{R}$. The first family is obtained from the single representation with coefficients $0, 1, 0$ by adding the coefficients of each member of the second: $(0+a_1, 1+\pi t, 0+t)$. Exercise 1.3.40 shows that this is a general phenomenon.

Solution to Exercise 1.3.16

Uses. Definition 1.3.10 (Span), Definition 1.2.4 (The space F^n)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. Compute the vector represented by $a(1,0,-1)+b(0,1,-1)$ and add up its coordinates.
Hint 2. For the other containment, a vector (x,y,z) with $x+y+z=0$ has $z=-x-y$. Which coefficients reproduce its first two coordinates?
Sketch. $a(1,0,-1)+b(0,1,-1)=(a,b,-a-b)$ has coordinate sum 0. Conversely, if $x+y+z=0$ then $(x,y,z)=(x,y,-x-y)=x(1,0,-1)+y(0,1,-1)$.

Let $P=\{(x,y,z)\in\mathbb{R}^3 \mid x+y+z=0\}$. We proceed by double containment.

span $\subseteq P$. Every element of the span is represented by some linear combination $a(1,0,-1)+b(0,1,-1)=(a,b,-a-b)$, whose coordinates add up to $a+b-a-b=0$. So it lies in P.

$P \subseteq$ span. Let $(x,y,z)\in P$. Then $z=-x-y$, and

$$x(1,0,-1)+y(0,1,-1)=(x,y,-x-y)=(x,y,z).$$

So (x,y,z) is in the span, as desired.

Solution to Exercise 1.3.24

Uses. Corollary 1.3.19 (Equal spans), Proposition 1.3.17 (A list lies in its span)
Hint 1. By Corollary 1.3.19, it suffices to show $\mathcal{L}\subseteq\operatorname{span}(\mathcal{M})$ and $\mathcal{M}\subseteq\operatorname{span}(\mathcal{L})$.
Hint 2. Each entry of $\mathcal{L}$ is an entry of $\mathcal{M}$. Where do the entries of $\mathcal{M}$ lie?
Sketch. Every entry of $\mathcal{L}$ is an entry of $\mathcal{M}$, hence lies in $\operatorname{span}(\mathcal{M})$ by Proposition 1.3.17, and symmetrically. Apply Corollary 1.3.19.

Let $\mathbf{v}\in\mathcal{L}$. By assumption $\mathbf{v}\in\mathcal{M}$, so $\mathbf{v}\in\operatorname{span}(\mathcal{M})$ by Proposition 1.3.17. Hence $\mathcal{L}\subseteq\operatorname{span}(\mathcal{M})$, and by the same argument with the roles exchanged, $\mathcal{M}\subseteq\operatorname{span}(\mathcal{L})$. By Corollary 1.3.19, $\operatorname{span}(\mathcal{L})=\operatorname{span}(\mathcal{M})$.

Solution to Exercise 1.3.25

Uses. Corollary 1.3.19 (Equal spans)
Hint 1. Use Corollary 1.3.19. Each $\mathbf{w}_j$ is visibly a linear combination of the $\mathbf{v}_i$.
Hint 2. Recover $\mathbf{v}_j$ from two consecutive partial sums.
Sketch. $\mathbf{w}_j$ is the combination of the $\mathbf{v}_i$ with coefficients 1 in positions $1,\dots,j$ and 0 after. Conversely $\mathbf{v}_1=\mathbf{w}_1$ and $\mathbf{v}_j=\mathbf{w}_j-\mathbf{w}_{j-1}$ for $j\geq 2$.

Write $\mathcal{V}=\mathbf{v}_1,\dots,\mathbf{v}_k$ and $\mathcal{W}=\mathbf{w}_1,\dots,\mathbf{w}_k$. Each $\mathbf{w}_j$ is represented by the linear combination of $\mathcal{V}$ with coefficient 1 on $\mathbf{v}_1,\dots,\mathbf{v}_j$ and 0 on the remaining entries, so $\mathcal{W}\subseteq\operatorname{span}(\mathcal{V})$. Conversely, $\mathbf{v}_1=\mathbf{w}_1$, and for $j\geq 2$,

$$\mathbf{w}_j-\mathbf{w}_{j-1}=\left(\mathbf{v}_1+\cdots+\mathbf{v}_{j-1}+\mathbf{v}_j\right)-\left(\mathbf{v}_1+\cdots+\mathbf{v}_{j-1}\right)=\mathbf{v}_j$$

by Proposition 1.2.30(a). So each $\mathbf{v}_j$ is a linear combination of $\mathcal{W}$, with coefficient 1 on $\mathbf{w}_j$, -1 on $\mathbf{w}_{j-1}$, and 0 elsewhere, and $\mathcal{V}\subseteq\operatorname{span}(\mathcal{W})$. By Corollary 1.3.19, the spans are equal.

Solution to Exercise 1.3.26

Uses. Definition 1.3.10 (Span), Definition 1.2.4 (The space F^n)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For (a), compute the vector represented by $a(1,2,0)+b(0,1,1)$ and compare its middle coordinate with the outer ones.
Hint 2. For the other containment in (a), the coefficients can be read off from the first and third coordinates. Then (b) is a matter of checking the equation $y=2x+z$.

Sketch. (a) $a(1,2,0)+b(0,1,1)=(a,2a+b,b)$ satisfies $y=2x+z$; conversely, if $y=2x+z$, then $(x,y,z)=x(1,2,0)+z(0,1,1)$. (b) $(1,1,1)$ fails the equation since $2+1\neq 1$; $(2,5,1)$ satisfies it, and $(2,5,1)=2(1,2,0)+(0,1,1)$.

(a) Let $P=\{(x,y,z)\in\mathbb{R}^3 \mid y=2x+z\}$. We proceed by double containment. An element of $\operatorname{span}(\mathcal{L})$ is represented by some linear combination

$$a(1,2,0)+b(0,1,1)=(a,2a+b,b),$$

and its middle coordinate $2a+b$ is twice the first plus the third. So $\operatorname{span}(\mathcal{L})\subseteq P$. Conversely, let $(x,y,z)\in P$. Then

$$x(1,2,0)+z(0,1,1)=(x,2x+z,z)=(x,y,z),$$

so $(x,y,z)\in\operatorname{span}(\mathcal{L})$. Hence $P\subseteq\operatorname{span}(\mathcal{L})$.

(b) For $(1,1,1)$ we have $2\cdot 1+1=3\neq 1$, so $(1,1,1)\notin\operatorname{span}(\mathcal{L})$ by (a). For $(2,5,1)$ we have $2\cdot 2+1=5$, so $(2,5,1)\in\operatorname{span}(\mathcal{L})$, and the proof of (a) gives the representation $(2,5,1)=2(1,2,0)+1(0,1,1)$.

Solution to Exercise 1.3.27

Uses. Definition 1.3.10 (Span), Corollary 1.3.20 (Operations that preserve the span), Definition 1.3.4 (Linear combination)
Hint 1. The combination $a(1,1,0)+b(0,1,2)$ represents $(a,a+b,2b)$. Run through the nine choices of $a,b\in\mathbb{F}_3$, reducing modulo 3.
Hint 2. For (b), $(1,2,2)$ is the sum of the other two entries. Rewrite an arbitrary combination of $\mathcal{M}$ as a combination of $\mathcal{L}$ and use that the nine vectors in (a) are distinct.
Sketch. (a) The span consists of the vectors $(a,a+b,2b)$: $(0,0,0),(0,1,2),(0,2,1),(1,1,0),(1,2,2),(1,0,1),(2,2,0),(2,0,2),(2,1,1)$. It contains $(1,0,1)$ ($a=1$, $b=2$) but not $(1,1,1)$. (b) $(1,2,2)\in\operatorname{span}(\mathcal{L})$. There are 27 combinations; a,b,c represents $(a+c)(1,1,0)+(b+c)(0,1,2)$, which is $(1,0,1)$ exactly when $a+c=1$ and $b+c=2$, giving 3 combinations, one for each c.

(a) The combination with coefficients a,b represents $a(1,1,0)+b(0,1,2)=(a,a+b,2b)$. Running through $a,b\in\mathbb{F}_3$ and reducing modulo 3:

$$\begin{aligned} a=0:&\quad (0,0,0),(0,1,2),(0,2,1),\\ a=1:&\quad (1,1,0),(1,2,2),(1,0,1),\\ a=2:&\quad (2,2,0),(2,0,2),(2,1,1),\end{aligned}$$

with $b=0,1,2$ in each row. These nine vectors are distinct, because the first coordinate determines a and the third determines b (as $2b=2b'$ forces $b=b'$). The vector $(1,0,1)$ appears, with $a=1$ and $b=2$. The vector $(1,1,1)$ does not: $a=1$ and $2b=1$ force $b=2$, and then $a+b=0\neq 1$.

(b) In $\mathbb{F}_3^3$, $(1,1,0)+(0,1,2)=(1,2,2)$, so the third entry of $\mathcal{M}$ lies in $\operatorname{span}(\mathcal{L})$, and $\operatorname{span}(\mathcal{M})=\operatorname{span}(\mathcal{L})$ by Corollary 1.3.20(c). A linear combination of $\mathcal{M}$ is determined by its three coefficients, so there are $3^3=27$ of them. By Lemma 1.3.9, the one with coefficients a,b,c represents

$$a(1,1,0)+b(0,1,2)+c((1,1,0)+(0,1,2))=(a+c)(1,1,0)+(b+c)(0,1,2).$$

By (a), the combinations of $\mathcal{L}$ represent distinct vectors for distinct coefficients, and $(1,0,1)$ is represented by the coefficients $1,2$. So a,b,c represents $(1,0,1)$ exactly when $a+c=1$ and $b+c=2$. For each $c\in\mathbb{F}_3$ there is exactly one such a and b, so exactly 3 linear combinations of $\mathcal{M}$ represent $(1,0,1)$: those with coefficients $1,2,0$, then $0,1,1$, and $2,0,2$. The same count holds for each of the nine vectors of the span, and $9\cdot 3=27$.

Solution to Exercise 1.3.28

Uses. Proposition 1.3.17 (A list lies in its span), Proposition 1.3.18 (Comparing spans), Corollary 1.3.19 (Equal spans), Definition 1.3.10 (Span), Proposition 1.2.33 (The zero vector)
Hint 1. For (a), if some entry is nonzero, what does Proposition 1.3.17 say? If all entries are $\mathbf{0}$, compare with the empty list.
Hint 2. For (b), the forward direction gives $\mathbf{w}=c\mathbf{v}$ for some c. Rule out $c=0$ separately in the cases $\mathbf{v}=\mathbf{0}$ and $\mathbf{v}\neq\mathbf{0}$.

Sketch. (a) $\operatorname{span}(\mathcal{L}) = \{\mathbf{0}\}$ exactly when every entry is $\mathbf{0}$, the empty list included; these are also exactly the lists that span $\{\mathbf{0}\}$. (b) If $\mathbf{w} = c\mathbf{v}$ with $c \neq 0$, then $\mathbf{v} = c^{-1}\mathbf{w}$, so the spans agree. Conversely $\mathbf{w} \in \operatorname{span}(\mathbf{v})$ gives $\mathbf{w} = c\mathbf{v}$; if $c = 0$ then $\mathbf{w} = \mathbf{0}$, so $\mathbf{v} \in \operatorname{span}(\mathbf{0}) = \{\mathbf{0}\}$ and $\mathbf{w} = 1\mathbf{v}$ anyway.

(a) If some entry $\mathbf{v}_j$ of $\mathcal{L}$ is nonzero, then $\mathbf{v}_j \in \operatorname{span}(\mathcal{L})$ by Proposition 1.3.17, so $\operatorname{span}(\mathcal{L}) \neq \{\mathbf{0}\}$. If every entry of $\mathcal{L}$ is $\mathbf{0}$ (this includes the empty list), then $\mathcal{L} \subseteq \{\mathbf{0}\} = \operatorname{span}(())$, so $\operatorname{span}(\mathcal{L}) \subseteq \{\mathbf{0}\}$ by Proposition 1.3.18, and $\mathbf{0} \in \operatorname{span}(\mathcal{L})$ by Proposition 1.3.17. So $\operatorname{span}(\mathcal{L}) = \{\mathbf{0}\}$ exactly when all entries of $\mathcal{L}$ are $\mathbf{0}$. A list in the zero space has all its entries equal to $\mathbf{0}$, so by what we just showed, every list in $\{\mathbf{0}\}$ spans it: $()$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{0}$, and so on.

(b) Suppose $\mathbf{w} = c\mathbf{v}$ with $c \neq 0$. Then $\mathbf{w} \in \operatorname{span}(\mathbf{v})$, and $\mathbf{v} = c^{-1}(c\mathbf{v}) = c^{-1}\mathbf{w} \in \operatorname{span}(\mathbf{w})$ by (A6) and (A5). So the spans are equal by Corollary 1.3.19.

Conversely, suppose $\operatorname{span}(\mathbf{v}) = \operatorname{span}(\mathbf{w})$. Then $\mathbf{w} \in \operatorname{span}(\mathbf{v})$ by Proposition 1.3.17, so $\mathbf{w} = c\mathbf{v}$ for some $c \in F$. If $c \neq 0$ we are done. If $c = 0$, then $\mathbf{w} = \mathbf{0}$, so $\operatorname{span}(\mathbf{v}) = \operatorname{span}(\mathbf{0}) = \{\mathbf{0}\}$, and therefore $\mathbf{v} = \mathbf{0}$ by (a). Then $\mathbf{w} = \mathbf{0} = 1\mathbf{v}$, with $1 \neq 0$.

Solution to Exercise 1.3.29

Uses. Definition 1.3.4 (Linear combination), Example 1.3.5 (One vector, many representations)
Hint 1. What does the equation $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = b_1\mathbf{v}_1 + b_2\mathbf{v}_2$ assert: that two expressions are the same, or that two vectors are equal?
Hint 2. For a counterexample, take $\mathbf{v}_1 = \mathbf{v}_2$.
Sketch. The equation says that the two linear combinations represent the same vector, not that they are the same linear combination. With $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v} \neq \mathbf{0}$, $1\mathbf{v}_1 + 0\mathbf{v}_2 = 0\mathbf{v}_1 + 1\mathbf{v}_2$. For $(1,0), (0,1)$, the two sides are (a_1, a_2) and (b_1, b_2).

(a) The hypothesis is an equation between vectors: it says that the two linear combinations *represent* the same vector. The argument reads it as saying that the two linear combinations are the same, which is what the claim asserts. So the argument assumes its conclusion.

(b) Let $\mathbf{v} \neq \mathbf{0}$ (for instance $\mathbf{v} = (1,0)$ in $\mathbb{R}^2$), and take $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v}$. Then $1\mathbf{v}_1 + 0\mathbf{v}_2 = \mathbf{v} = 0\mathbf{v}_1 + 1\mathbf{v}_2$, but $1 \neq 0$. For $\mathbf{v}_1 = (1,0)$ and $\mathbf{v}_2 = (0,1)$, the two sides represent (a_1, a_2) and (b_1, b_2), and these columns are equal only if $a_1 = b_1$ and $a_2 = b_2$. The lists for which the claim holds are the linearly independent ones (Section 1.5).

Solution to Exercise 1.3.30

Uses. Definition 1.3.10 (Span), Example 1.3.11 (Small spans)
Techniques. Double containment (Technique 1.3.14), Try small cases and small fields (Technique 1.2.38)
Hint 1. The argument proves a containment. Which one, and is it the one with content?
Hint 2. For (b), solve $a + b = x$ and $a - b = y$ for a and b. Which field element do you divide by?
Sketch. (a) It shows only span $\subseteq \mathbb{R}^2$, which holds for any list in $\mathbb{R}^2$; the containment $\mathbb{R}^2 \subseteq$ span is missing. (b) $(x,y) = \frac{x+y}{2}(1,1) + \frac{x-y}{2}(1,-1)$. (c) Over $\mathbb{F}_2$, $(1,-1) = (1,1)$ and the span is $\{\mathbf{0}, (1,1)\} \neq \mathbb{F}_2^2$. Over $\mathbb{F}_3$, $2^{-1} = 2$, so the formula of (b) works and the span is $\mathbb{F}_3^2$.

(a) The argument proves $\operatorname{span}((1,1),(1,-1)) \subseteq \mathbb{R}^2$, which holds for every list in $\mathbb{R}^2$ because the operations of $\mathbb{R}^2$ produce elements of $\mathbb{R}^2$. The containment $\mathbb{R}^2 \subseteq \operatorname{span}((1,1),(1,-1))$ is never addressed, and it is the one with content.

(b) It remains to show $\mathbb{R}^2 \subseteq \operatorname{span}((1,1),(1,-1))$. Let $(x,y) \in \mathbb{R}^2$, and set $a = (x+y)/2$ and $b = (x-y)/2$. Then

$$a(1,1) + b(1,-1) = (a+b, a-b) = (x,y),$$

so (x,y) lies in the span.

(c) In $\mathbb{F}_2^2$, $-1 = 1$, so the list is $(1,1),(1,1)$, and its span is $\operatorname{span}((1,1)) = \{\mathbf{0}, (1,1)\}$ by Corollary 1.3.20(c) and Example 1.3.11. This is not $\mathbb{F}_2^2$; for instance, $(1,0)$ is missing. The proof of (b) breaks down because it divides by $2 = 0$. In $\mathbb{F}_3$, $2 \neq 0$ and $2^{-1} = 2$, so with $a = 2(x+y)$ and $b = 2(x-y)$,

$$a(1,1) + b(1,-1) = (4x, 4y) = (x,y),$$

since $4 = 1$ in $\mathbb{F}_3$. So the list spans $\mathbb{F}_3^2$.

Solution to Exercise 1.3.31

Uses. Definition 1.3.10 (Span), Definition 1.2.9 (Function spaces)
Hint 1. For (a), use a double-angle formula. For (b), compare $f(t)$ with $f(-t)$ for $f = a + b\cos$.
Hint 2. For (c), write $\sin^2 = 1 - \cos^2$. If $\sin^2 = a + b\cos$, evaluate at three values of t where $\cos t$ is 0, 1, and -1.
Sketch. (a) $\cos^2 t = \frac{1}{2} + \frac{1}{2}\cos 2t$. (b) Every $a + b\cos$ is even, while $\sin(\pi/2) = 1 \neq -1 = \sin(-\pi/2)$. (c) No: $\sin^2 = a + b\cos$ at $t = \pi/2, 0, \pi$ gives $1 = a$, $0 = a + b$, $0 = a - b$, which is inconsistent.

(a) By the double-angle formula $\cos 2t = 2\cos^2 t - 1$, we have $\cos^2 t = \frac{1}{2} \cdot 1 + \frac{1}{2}\cos 2t$ for all t. So $t \mapsto \cos^2 t$ is represented by the linear combination $\frac{1}{2} \cdot 1 + \frac{1}{2}c_2$.

(b) Every element of $\operatorname{span}(1, \cos)$ has the form $f = a + b\cos$ with $a, b \in \mathbb{R}$, and $f(-t) = a + b\cos(-t) = a + b\cos t = f(t)$ for all t. But $\sin(\pi/2) = 1$ and $\sin(-\pi/2) = -1$, so sin is not of this form.

(c) No. Suppose $\sin^2 t = a + b\cos t$ for all $t \in \mathbb{R}$. At $t = \pi/2$ this gives $1 = a$. At $t = 0$ it gives $0 = a + b$, and at $t = \pi$ it gives $0 = a - b$. Adding the last two equations gives $a = 0$, contradicting $a = 1$.

Solution to Exercise 1.3.32

Uses. Definition 1.2.17 (Evaluation), Definition 1.2.16 (Degree), Definition 1.3.10 (Span), Notation 1.2.15
Techniques. Double containment (Technique 1.3.14)
Hint 1. Write $p = a_0 + a_1x + a_2x^2 + a_3x^3$. The condition $p(1) = 0$ expresses a_0 in terms of a_1, a_2, a_3.
Hint 2. Substitute $a_0 = -a_1 - a_2 - a_3$ and group the terms by a_1, a_2, and a_3.
Sketch. The list $x - 1, x^2 - 1, x^3 - 1$. A combination $a_1(x-1) + a_2(x^2-1) + a_3(x^3-1)$ has degree at most 3 and vanishes at 1. Conversely, if $p(1) = a_0 + a_1 + a_2 + a_3 = 0$, then $p = a_1(x-1) + a_2(x^2-1) + a_3(x^3-1)$.

We show that $\operatorname{span}(x - 1, x^2 - 1, x^3 - 1) = W$, by double containment.

$\operatorname{span} \subseteq W$. A linear combination of the list represents

$$a_1(x-1) + a_2(x^2-1) + a_3(x^3-1) = -(a_1 + a_2 + a_3) + a_1x + a_2x^2 + a_3x^3,$$

which has no nonzero coefficients beyond position 3, so it lies in $\mathbb{R}[x]_{\leq 3}$. Its value at 1 is $-(a_1 + a_2 + a_3) + a_1 + a_2 + a_3 = 0$ (Definition 1.2.17). So it lies in W.

$W \subseteq \operatorname{span}$. Let $p = a_0 + a_1x + a_2x^2 + a_3x^3 \in W$. Then $p(1) = a_0 + a_1 + a_2 + a_3 = 0$, so $a_0 = -(a_1 + a_2 + a_3)$, and by the computation above,

$$p = -(a_1 + a_2 + a_3) + a_1x + a_2x^2 + a_3x^3 = a_1(x-1) + a_2(x^2-1) + a_3(x^3-1).$$

So p lies in the span.

Solution to Exercise 1.3.33

Uses. Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans), Definition 1.2.16 (Degree), induction on n
Techniques. Double containment (Technique 1.3.14)
Hint 1. One containment is a degree count, via Proposition 1.3.18 and $F[x]_{\leq n} = \operatorname{span}(1, x, \ldots, x^n)$. For the other, it suffices to show that each x^k with $k \leq n$ lies in $\operatorname{span}(p_0, \ldots, p_n)$.
Hint 2. Induct on n. If p_n has leading coefficient $c \neq 0$, then $p_n - cx^n$ has degree less than n, and $x^n = c^{-1}(p_n - (p_n - cx^n))$.
Sketch. (a) Each p_k lies in $F[x]_{\leq n} = \operatorname{span}(1, \ldots, x^n)$, giving $\subseteq$. For $\supseteq$, induct on n: p_0 is a nonzero constant, so $1 \in \operatorname{span}(p_0)$. If $\operatorname{span}(p_0, \ldots, p_{n-1}) = F[x]_{\leq n-1}$ and $p_n = cx^n + r$ with $c \neq 0$ and $\deg r < n$, then $r \in \operatorname{span}(p_0, \ldots, p_{n-1})$, so $x^n = c^{-1}(p_n - r) \in \operatorname{span}(p_0, \ldots, p_n)$, and the lower powers are there by induction. (b) $x^2 = (x^2 - 2x + 1) + 2(x - 1) + 1$.

(a) Write $\mathcal{P}_n = p_0, \ldots, p_n$. We prove $\operatorname{span}(\mathcal{P}_n) = F[x]_{\leq n}$ by induction on n, and in each case by double containment.

$\operatorname{span}(\mathcal{P}_n) \subseteq F[x]_{\leq n}$. Each p_k has degree $k \leq n$, so $\mathcal{P}_n \subseteq F[x]_{\leq n} = \operatorname{span}(1, x, \ldots, x^n)$ (Example 1.3.11). By Proposition 1.3.18, $\operatorname{span}(\mathcal{P}_n) \subseteq \operatorname{span}(1, x, \ldots, x^n) = F[x]_{\leq n}$.

$F[x]_{\leq n} \subseteq \operatorname{span}(\mathcal{P}_n)$. By Proposition 1.3.18, it suffices to show $x^k \in \operatorname{span}(\mathcal{P}_n)$ for $0 \leq k \leq n$. For $n = 0$, p_0 has degree 0, so $p_0 = c$ is a nonzero constant and $1 = c^{-1}p_0 \in \operatorname{span}(p_0)$. Now let $n \geq 1$, and assume $\operatorname{span}(\mathcal{P}_{n-1}) = F[x]_{\leq n-1}$. Let $c \neq 0$ be the leading coefficient of p_n, and let $r = p_n - cx^n$. The coefficients

of r in positions $k > n$ are 0, and so is its coefficient in position n, so $r \in F[x]_{\le n-1} = \operatorname{span}(\mathcal{P}_{n-1})$, say $r = \sum_{k=0}^{n-1} d_k p_k$. Then

$$x^n = c^{-1}(p_n - r) = c^{-1}p_n - \sum_{k=0}^{n-1}(c^{-1}d_k)p_k \in \operatorname{span}(\mathcal{P}_n),$$

using Lemma 1.3.9. For $k < n$, $x^k \in F[x]_{\le n-1} = \operatorname{span}(\mathcal{P}_{n-1}) \subseteq \operatorname{span}(\mathcal{P}_n)$, by the inductive hypothesis and Corollary 1.3.20(b). This completes the induction.

(b) Following the proof of (a) from the top degree down, $x^2 = (x^2 - 2x + 1) + (2x - 1)$ and $2x - 1 = 2(x - 1) + 1$. So

$$x^2 = 1 \cdot 1 + 2(x-1) + 1 \cdot (x^2 - 2x + 1).$$

Check: $1 + 2x - 2 + x^2 - 2x + 1 = x^2$.

Solution to Exercise 1.3.34

> *Uses.* Corollary 1.3.19 (Equal spans), Proposition 1.3.17 (A list lies in its span), Lemma 1.3.9 (Arithmetic of linear combinations)
> *Hint 1.* Write $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. What does the hypothesis $\mathbf{w} \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$ say about a_k?
> *Hint 2.* Since $a_k \neq 0$, solve for $\mathbf{v}_k$. Then use Corollary 1.3.19.
> *Sketch.* Write $\mathbf{w} = \sum_i a_i\mathbf{v}_i$. If $a_k = 0$, then $\mathbf{w} \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$, so $a_k \neq 0$ and $\mathbf{v}_k = a_k^{-1}\mathbf{w} - \sum_{i<k} a_k^{-1}a_i\mathbf{v}_i$. So each list lies in the span of the other.

Since $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, we can write $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. If $a_k = 0$, then $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1}$ (the term $0\mathbf{v}_k = \mathbf{0}$ contributes nothing), contradicting $\mathbf{w} \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$. So $a_k \neq 0$. (In particular $k \geq 1$.) By Lemma 1.3.9 and Proposition 1.2.34,

$$\mathbf{v}_k = a_k^{-1}(\mathbf{w} - a_1\mathbf{v}_1 - \cdots - a_{k-1}\mathbf{v}_{k-1}) = a_k^{-1}\mathbf{w} + \sum_{i=1}^{k-1}(-a_k^{-1}a_i)\mathbf{v}_i.$$

Let $\mathcal{L}' = \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{w}$. The entries of $\mathcal{L}'$ lie in $\operatorname{span}(\mathcal{L})$: the $\mathbf{v}_i$ by Proposition 1.3.17, and $\mathbf{w}$ by assumption. The entries of $\mathcal{L}$ lie in $\operatorname{span}(\mathcal{L}')$: the $\mathbf{v}_i$ with $i < k$ by Proposition 1.3.17, and $\mathbf{v}_k$ by the display. By Corollary 1.3.19, $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$.

This exchange is the step that drives the proof of Lemma 4.1.2.

Solution to Exercise 1.3.35

> *Uses.* Definition 1.3.10 (Span), Example 1.3.11 (Small spans), Exercise 1.3.34
> *Hint 1.* Test (a)–(d) in $\mathbb{R}^2$ with the vectors $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$, and $\mathbf{e}_1 + \mathbf{e}_2$.
> *Hint 2.* (e) is true. Write $\mathbf{u} = a\mathbf{v} + b\mathbf{w}$ and decide what $\mathbf{u} \notin \operatorname{span}(\mathbf{w})$ says about a.
> *Sketch.* (a) False: $\mathbf{e}_1 + \mathbf{e}_2 \in \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ lies on neither axis. (b) False: $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) \cap \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$ contains $\mathbf{e}_1 + \mathbf{e}_2$, but the lists share no entry, so $\mathcal{N} = ()$. (c) False: $()$ and $\mathbf{0}$. (d) False: $\mathbf{u} = \mathbf{v} = \mathbf{e}_1$, $\mathbf{w} = \mathbf{e}_2$. (e) True: $\mathbf{u} = a\mathbf{v} + b\mathbf{w}$ with $a \neq 0$, so $\mathbf{v} = a^{-1}\mathbf{u} - a^{-1}b\mathbf{w}$.

Write $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ in $\mathbb{R}^2$.

(a) False. Let $\mathcal{L} = \mathbf{e}_1$ and $\mathcal{M} = \mathbf{e}_2$. Then $(1, 1) = \mathbf{e}_1 + \mathbf{e}_2 \in \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$, but $\operatorname{span}(\mathbf{e}_1) = \{(a, 0) \mid a \in \mathbb{R}\}$ and $\operatorname{span}(\mathbf{e}_2) = \{(0, b) \mid b \in \mathbb{R}\}$, and $(1, 1)$ lies in neither.

(b) False. Let $\mathcal{L} = \mathbf{e}_1, \mathbf{e}_2$ and $\mathcal{M} = (1, 1)$. No vector is an entry of both lists, so $\mathcal{N} = ()$ and $\operatorname{span}(\mathcal{N}) = \{\mathbf{0}\}$. But $(1, 1) \in \operatorname{span}(\mathcal{M})$ and $(1, 1) \in \operatorname{span}(\mathcal{L}) = \mathbb{R}^2$, so the intersection contains $(1, 1) \neq \mathbf{0}$.

(c) False. $\operatorname{span}(()) = \{\mathbf{0}\} = \operatorname{span}(\mathbf{0})$ by Example 1.3.11, but the lists have lengths 0 and 1.

(d) False. Let $\mathbf{u} = \mathbf{v} = \mathbf{e}_1$ and $\mathbf{w} = \mathbf{e}_2$. Then $\mathbf{u} = 1\mathbf{v} + 0\mathbf{w}$ and $\mathbf{v} = 1\mathbf{u} + 0\mathbf{w}$, but $\operatorname{span}(\mathbf{u}, \mathbf{v}) = \operatorname{span}(\mathbf{e}_1)$ does not contain $\mathbf{e}_2$.

(e) True. Write $\mathbf{u} = a\mathbf{v} + b\mathbf{w}$. If $a = 0$, then $\mathbf{u} = b\mathbf{w} \in \operatorname{span}(\mathbf{w})$, contrary to assumption. So $a \neq 0$, and

$$\mathbf{v} = a^{-1}(\mathbf{u} - b\mathbf{w}) = a^{-1}\mathbf{u} + \big(-a^{-1}b\big)\mathbf{w} \in \operatorname{span}(\mathbf{u}, \mathbf{w}).$$

This is the case $k = 2$ of Exercise 1.3.34, applied to the list $\mathbf{w}, \mathbf{v}$.

Solution to Exercise 1.3.36

Uses. Definition 1.3.10 (Span), (1.3.1), Notation 1.3.3
Techniques. Double containment (Technique 1.3.14)
Hint 1. A linear combination of the concatenation $\mathcal{L}, \mathcal{M}$ has one coefficient for each entry of $\mathcal{L}$ and one for each entry of $\mathcal{M}$.
Hint 2. By (A2), the sum over the concatenation can be grouped as (the terms from $\mathcal{L}$) plus (the terms from $\mathcal{M}$).
Sketch. The combination of $\mathcal{L}, \mathcal{M}$ with coefficients $a_1, \ldots, a_k, b_1, \ldots, b_m$ represents $\sum a_i\mathbf{v}_i + \sum b_j\mathbf{w}_j$, by grouping the terms. So the vectors it represents are exactly the sums $\mathbf{u} + \mathbf{w}$, both containments included.

Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathcal{M} = \mathbf{w}_1, \ldots, \mathbf{w}_m$, and let S be the set on the right. A linear combination of $\mathcal{L}, \mathcal{M}$ has coefficients $a_1, \ldots, a_k, b_1, \ldots, b_m$, and by (1.3.1) and (A2) the vector it represents is

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + b_1\mathbf{w}_1 + \cdots + b_m\mathbf{w}_m = \left(\sum_{i=1}^{k} a_i\mathbf{v}_i\right) + \left(\sum_{j=1}^{m} b_j\mathbf{w}_j\right).$$

(If $k = 0$ or $m = 0$, the corresponding sum is $\mathbf{0}$, and the identity holds by (A1) and (A3).) We use this identity for both containments.

$\operatorname{span}(\mathcal{L}, \mathcal{M}) \subseteq S$. An element of the left side is represented by some such combination, and the identity writes it as $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} = \sum a_i\mathbf{v}_i \in \operatorname{span}(\mathcal{L})$ and $\mathbf{w} = \sum b_j\mathbf{w}_j \in \operatorname{span}(\mathcal{M})$.

$S \subseteq \operatorname{span}(\mathcal{L}, \mathcal{M})$. Let $\mathbf{u} \in \operatorname{span}(\mathcal{L})$ and $\mathbf{w} \in \operatorname{span}(\mathcal{M})$, say $\mathbf{u} = \sum a_i\mathbf{v}_i$ and $\mathbf{w} = \sum b_j\mathbf{w}_j$. By the identity, the combination of $\mathcal{L}, \mathcal{M}$ with coefficients $a_1, \ldots, a_k, b_1, \ldots, b_m$ represents $\mathbf{u} + \mathbf{w}$, so $\mathbf{u} + \mathbf{w} \in \operatorname{span}(\mathcal{L}, \mathcal{M})$.

In the language of Section 4.4, the span of a concatenation is the sum of the spans (Definition 4.4.1).

Solution to Exercise 1.3.37

Uses. Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans), Corollary 1.3.20 (Operations that preserve the span)
Hint 1. If $\delta = ad - bc \neq 0$, compute $d\mathbf{u} - b\mathbf{v}$ and $a\mathbf{v} - c\mathbf{u}$.
Hint 2. If $\delta = 0$, show that every vector (p, q) in the span satisfies $bp - aq = 0$ and $dp - cq = 0$. Then find a vector of F^2 that violates one of these equations.
Hint 3. For (b), compute δ when $\mathbf{v} = t\mathbf{u}$. Conversely, if $\delta = 0$ and $a \neq 0$, try $t = c/a$.
Sketch. (a) $d\mathbf{u} - b\mathbf{v} = (\delta, 0)$ and $a\mathbf{v} - c\mathbf{u} = (0, \delta)$, so if $\delta \neq 0$ then $\mathbf{e}_1, \mathbf{e}_2$ lie in the span. If $\delta = 0$, every $x\mathbf{u} + y\mathbf{v}$ satisfies $bp - aq = 0$ and $dp - cq = 0$; if $\mathbf{u} = \mathbf{v} = \mathbf{0}$ the span is $\{\mathbf{0}\}$, and otherwise one of $\mathbf{e}_1, \mathbf{e}_2$ violates an equation with a nonzero coefficient. (b) $\mathbf{v} = t\mathbf{u}$ gives $\delta = 0$; conversely, $\delta = 0$ with $a \neq 0$ gives $\mathbf{v} = (c/a)\mathbf{u}$, and with $a = 0$, $b \neq 0$ it gives $c = 0$ and $\mathbf{v} = (d/b)\mathbf{u}$.

Let $\delta = ad - bc$.

(a) Suppose $\delta \neq 0$. Then

$$d\mathbf{u} - b\mathbf{v} = (da - bc, db - bd) = (\delta, 0), \quad a\mathbf{v} - c\mathbf{u} = (ac - ca, ad - cb) = (0, \delta).$$

So $\mathbf{e}_1 = \delta^{-1}(d\mathbf{u} - b\mathbf{v})$ and $\mathbf{e}_2 = \delta^{-1}(a\mathbf{v} - c\mathbf{u})$ lie in $\operatorname{span}(\mathbf{u}, \mathbf{v})$, and by Proposition 1.3.18, $F^2 = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) \subseteq \operatorname{span}(\mathbf{u}, \mathbf{v})$.

Now suppose $\delta = 0$. A vector $(p, q) = x\mathbf{u} + y\mathbf{v} = (xa + yc, xb + yd)$ of the span satisfies

$$bp - aq = y(bc - ad) = 0 \quad \text{and} \quad dp - cq = x(ad - bc) = 0.$$

If $a = b = c = d = 0$, then $\operatorname{span}(\mathbf{u}, \mathbf{v}) = \{\mathbf{0}\} \neq F^2$. Otherwise one of these two equations has a nonzero coefficient. If $a \neq 0$ or $c \neq 0$, the vector $\mathbf{e}_2 = (0, 1)$ gives $b \cdot 0 - a \cdot 1 = -a$ and $d \cdot 0 - c \cdot 1 = -c$, one of which is nonzero, so $\mathbf{e}_2$ is not in the span. If $b \neq 0$ or $d \neq 0$, then similarly $\mathbf{e}_1 = (1, 0)$ gives b and d, and $\mathbf{e}_1$ is not in the span. Either way, $\mathbf{u}, \mathbf{v}$ does not span F^2.

(b) If $\mathbf{v} \in \operatorname{span}(\mathbf{u})$, say $\mathbf{v} = t\mathbf{u} = (ta, tb)$, then $\delta = atb - bta = 0$, so the list does not span by (a). Conversely, suppose the list does not span, so $\delta = 0$ by (a). If $a \neq 0$, then $bc = ad$ gives

$$\frac{c}{a}\mathbf{u} = \left(c, \frac{bc}{a}\right) = (c, d) = \mathbf{v}.$$

If $a = 0$, then $b \neq 0$ because $\mathbf{u} \neq \mathbf{0}$, and $\delta = -bc = 0$ forces $c = 0$; then $\mathbf{v} = (0, d) = (d/b)(0, b) = (d/b)\mathbf{u}$. Either way $\mathbf{v} \in \operatorname{span}(\mathbf{u})$.

The quantity $ad - bc$ is the determinant of the 2×2 matrix with columns $\mathbf{u}$ and $\mathbf{v}$ (Theorem 5.1.13).

Solution to Exercise 1.3.38

Uses. Exercise 1.3.37, Corollary 1.2.36 (Cancelling scalars and vectors), Example 1.3.11 (Small spans), Corollary 1.3.20 (Operations that preserve the span)
Hint 1. For (a), how many elements does $\operatorname{span}(\mathbf{v})$ have?
Hint 2. For (b), use Exercise 1.3.37(b). First show that a spanning list $\mathbf{u}, \mathbf{v}$ has $\mathbf{u} \neq \mathbf{0}$. Then count the choices of $\mathbf{v}$ outside $\operatorname{span}(\mathbf{u})$, which has exactly q elements.
Hint 3. For (c), a list in $\mathbb{F}_2^2$ spans exactly when it has two different nonzero entries. Count the complement.
Sketch. (a) $\operatorname{span}(\mathbf{v}) = \{a\mathbf{v}\}$ has at most $q < q^2$ elements. (b) If $\mathbf{u} = \mathbf{0}$, then $\operatorname{span}(\mathbf{0}, \mathbf{v}) = \operatorname{span}(\mathbf{v}) \neq F^2$. For each of the $q^2 - 1$ nonzero $\mathbf{u}$, the multiples of $\mathbf{u}$ are q distinct vectors, and $\mathbf{v}$ must avoid them: $q^2 - q$ choices. (c) A list of length 3 fails to span exactly when its nonzero entries are all equal: $1 + 3\left(2^3 - 1\right) = 22$ lists. So $64 - 22 = 42$ span.

(a) For $\mathbf{v} \in F^2$, $\operatorname{span}(\mathbf{v}) = \{a\mathbf{v} \mid a \in F\}$ has at most q elements, and F^2 has $q^2 > q$ elements.

(b) If $\mathbf{u} = \mathbf{0}$, then $\operatorname{span}(\mathbf{0}, \mathbf{v}) = \operatorname{span}(\mathbf{v})$ by Corollary 1.3.20(a) and (c), which is not F^2 by (a). So a spanning list $\mathbf{u}, \mathbf{v}$ has $\mathbf{u} \neq \mathbf{0}$, and there are $q^2 - 1$ such $\mathbf{u}$. For each of them, Exercise 1.3.37(b) says that $\mathbf{u}, \mathbf{v}$ spans F^2 exactly when $\mathbf{v} \notin \operatorname{span}(\mathbf{u})$. The vectors $a\mathbf{u}$, $a \in F$, are distinct by Corollary 1.2.36(b), so $\operatorname{span}(\mathbf{u})$ has exactly q elements, and there are $q^2 - q$ choices for $\mathbf{v}$. In total, $\left(q^2 - 1\right)\left(q^2 - q\right)$ lists of length 2 span F^2. For $q = 2$ this is $3 \cdot 2 = 6$.

(c) Let $\mathcal{L}$ be a list of length 3 in $\mathbb{F}_2^2$. If $\mathcal{L}$ has two different nonzero entries $\mathbf{u}, \mathbf{v}$, then $\mathbf{v} \notin \{\mathbf{0}, \mathbf{u}\} = \operatorname{span}(\mathbf{u})$, so $\mathbf{u}, \mathbf{v}$ spans $\mathbb{F}_2^2$ by Exercise 1.3.37(b), and so does $\mathcal{L}$, by Corollary 1.3.20(a) and (b). Otherwise all nonzero entries of $\mathcal{L}$ equal a single vector $\mathbf{w}$ (or there are none), and $\operatorname{span}(\mathcal{L}) \subseteq \operatorname{span}(\mathbf{w}) = \{\mathbf{0}, \mathbf{w}\}$, which is not $\mathbb{F}_2^2$. The lists of the second kind are the zero list $\mathbf{0}, \mathbf{0}, \mathbf{0}$, and, for each of the 3 nonzero vectors $\mathbf{w}$, the $2^3 - 1 = 7$ lists with entries in $\{\mathbf{0}, \mathbf{w}\}$ other than the zero list. That makes $1 + 3 \cdot 7 = 22$ lists. Since there are $4^3 = 64$ lists of length 3, exactly $64 - 22 = 42$ of them span $\mathbb{F}_2^2$.

Solution to Exercise 1.3.39

Uses. Definition 1.3.10 (Span), Definition 1.3.4 (Linear combination)
Hint 1. How many linear combinations does $\mathcal{L}$ have? Each element of the span is represented by at least one of them.
Hint 2. For (b), compare q^k with the number of elements of F^n.
Sketch. (a) A linear combination of $\mathcal{L}$ is determined by its k coefficients, so there are q^k of them, and $\operatorname{span}(\mathcal{L})$ is the set of vectors they represent. (b) F^n has $q^n > q^k$ elements when $k < n$.

(a) A linear combination of $\mathcal{L}$ is determined by its list of coefficients $a_1, \ldots, a_k$, so $\mathcal{L}$ has exactly q^k linear combinations (for $k = 0$, exactly one). The span is the set of vectors represented by them, and each linear combination represents exactly one vector. So $\operatorname{span}(\mathcal{L})$ has at most q^k elements.

(b) F^n has q^n elements. If $\mathcal{L}$ has length $k < n$, then $\operatorname{span}(\mathcal{L})$ has at most $q^k < q^n$ elements by (a), so $\operatorname{span}(\mathcal{L}) \neq F^n$.

The statement of (b) holds over every field, but over an infinite field counting is useless, and the proof needs the theory of dimension (Section 4.1).

Solution to Exercise 1.3.40

Uses. Lemma 1.3.9 (Arithmetic of linear combinations), Exercise 1.3.39, Proposition 1.2.30 (Additive arithmetic)
Hint 1. Fix one representation $b_1, \ldots, b_k$ of $\mathbf{w}$. If $a_1, \ldots, a_k$ represents $\mathbf{0}$, what does $a_1 + b_1, \ldots, a_k + b_k$ represent?
Hint 2. Show that $a \mapsto a + b$ is a bijection from the representations of $\mathbf{0}$ to the representations of $\mathbf{w}$. For (b), sort the q^k linear combinations by the vector they represent.

Sketch. (a) If $\sum b_i\mathbf{v}_i = \mathbf{w}$, then $a \mapsto a + b$ sends representations of $\mathbf{0}$ to representations of $\mathbf{w}$, with inverse $a' \mapsto a' - b$, by Lemma 1.3.9. (b) Each of the q^k linear combinations represents exactly one vector of the span, and each vector of the span is represented $N(\mathbf{0})$ times. (c) A divisor of p^k is a power of p.

(a) Let $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, and fix coefficients $b_1, \ldots, b_k$ with $\sum b_i\mathbf{v}_i = \mathbf{w}$. If $a_1, \ldots, a_k$ represents $\mathbf{0}$, then by Lemma 1.3.9(a),

$$\sum_{i=1}^{k}(a_i + b_i)\mathbf{v}_i = \sum_{i=1}^{k} a_i\mathbf{v}_i + \sum_{i=1}^{k} b_i\mathbf{v}_i = \mathbf{0} + \mathbf{w} = \mathbf{w}.$$

Conversely, if $a'_1, \ldots, a'_k$ represents $\mathbf{w}$, then by Lemma 1.3.9(a) and (b) with $c = -1$, and Proposition 1.2.34,

$$\sum_{i=1}^{k}(a'_i - b_i)\mathbf{v}_i = \sum_{i=1}^{k} a'_i\mathbf{v}_i + (-1)\sum_{i=1}^{k} b_i\mathbf{v}_i = \mathbf{w} - \mathbf{w} = \mathbf{0}.$$

So $a \mapsto a + b$ maps the representations of $\mathbf{0}$ to the representations of $\mathbf{w}$, and $a' \mapsto a' - b$ maps back. These maps are inverse to each other, so both sets have the same number of elements: $N(\mathbf{w}) = N(\mathbf{0})$.

(b) As in Exercise 1.3.39, $\mathcal{L}$ has exactly q^k linear combinations, and each represents exactly one vector, which lies in $\operatorname{span}(\mathcal{L})$. Grouping them by the vector they represent and using (a),

$$q^k = \sum_{\mathbf{w} \in \operatorname{span}(\mathcal{L})} N(\mathbf{w}) = |\operatorname{span}(\mathcal{L})| \cdot N(\mathbf{0}).$$

(c) By (b), $|\operatorname{span}(\mathcal{L})|$ divides p^k. By unique factorization, the positive divisors of p^k are p^r with $0 \le r \le k$.

The bijection in (a) is the same idea as Proposition 3.1.19: all representations of $\mathbf{w}$ are one of them plus a representation of $\mathbf{0}$.

Solution to Exercise 1.3.41

Uses. Corollary 1.3.20 (Operations that preserve the span), Proposition 1.3.18 (Comparing spans), Proposition 1.3.17 (A list lies in its span)
Hint 1. For (a), express $\mathbf{w} - \mathbf{u}$ in terms of the other two differences.
Hint 2. For (b), try the standard vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in $\mathbb{R}^3$ and add up coordinates.
Hint 3. For (c), if $\mathbf{u} \in D$, recover $\mathbf{v}$ and $\mathbf{w}$ from $\mathbf{u}$ and the differences.
Sketch. (a) $\mathbf{w} - \mathbf{u} = -(\mathbf{u} - \mathbf{v}) - (\mathbf{v} - \mathbf{w})$, and each difference lies in $\operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$. (b) False: for $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ in $\mathbb{R}^3$, every difference has coordinate sum 0, hence so does every combination, but $\mathbf{e}_1$ does not. (c) If equality holds, $\mathbf{u} \in D$. If $\mathbf{u} \in D$, then $\mathbf{v} = \mathbf{u} - (\mathbf{u} - \mathbf{v}) \in D$ and $\mathbf{w} = \mathbf{v} - (\mathbf{v} - \mathbf{w}) \in D$, so $\operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w}) \subseteq D$.

(a) Since $\mathbf{w} - \mathbf{u} = -(\mathbf{u} - \mathbf{v}) - (\mathbf{v} - \mathbf{w}) \in D$, Corollary 1.3.20(c) gives $\operatorname{span}(\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u}) = D$. Each difference is a linear combination of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ (for instance $\mathbf{u} - \mathbf{v} = 1\mathbf{u} + (-1)\mathbf{v} + 0\mathbf{w}$), so $D \subseteq \operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ by Proposition 1.3.18.

(b) False. In $\mathbb{R}^3$, take $\mathbf{u} = \mathbf{e}_1, \mathbf{v} = \mathbf{e}_2, \mathbf{w} = \mathbf{e}_3$. Then $\operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathbb{R}^3$. The differences $(1, -1, 0)$ and $(0, 1, -1)$ have coordinate sum 0, so every vector $a(1, -1, 0) + b(0, 1, -1) = (a, b - a, -b)$ of D has coordinate sum 0. Since $\mathbf{e}_1$ has coordinate sum 1, $\mathbf{e}_1 \notin D$.

(c) If $D = \operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$, then $\mathbf{u} \in D$ by Proposition 1.3.17. Conversely, suppose $\mathbf{u} \in D$. Then $\mathbf{v} = \mathbf{u} - (\mathbf{u} - \mathbf{v})$ and $\mathbf{w} = \mathbf{v} - (\mathbf{v} - \mathbf{w})$ lie in $\operatorname{span}(\mathbf{u}, \mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w})$, which equals D by Corollary 1.3.20(a) and (c), since $\mathbf{u} \in D$. So $\operatorname{span}(\mathbf{u}, \mathbf{v}, \mathbf{w}) \subseteq D$ by Proposition 1.3.18, and equality follows from (a).

Solution to Exercise 1.3.42

Uses. Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans), Exercise 1.2.45
Hint 1. Add up all the $\mathbf{u}_i$. Each coordinate of the sum is a sum of $n - 1$ ones.
Hint 2. If $(n-1) * 1 \neq 0$, get $\mathbf{s}$ and then each $\mathbf{e}_i = \mathbf{s} - \mathbf{u}_i$. If $(n-1) * 1 = 0$, compute the sum of the coordinates of an arbitrary combination $\sum a_i\mathbf{u}_i$.
Sketch. $\sum_i \mathbf{u}_i = ((n-1) * 1)\mathbf{s}$. If $m = (n-1) * 1 \neq 0$, then $\mathbf{s} = m^{-1}\sum \mathbf{u}_i$ and $\mathbf{e}_i = \mathbf{s} - \mathbf{u}_i$ lie in the span. If $m = 0$, the vector $\sum a_i\mathbf{u}_i$ has coordinates $S - a_j$ with $S = \sum a_i$, whose sum is $n * S - S = mS = 0$, while $\mathbf{e}_1$ has coordinate sum 1.

Let $m = (n-1) * 1 \in F$. Each coordinate of $\mathbf{u}_1 + \cdots + \mathbf{u}_n$ is a sum of $n-1$ ones and one zero, so $\mathbf{u}_1 + \cdots + \mathbf{u}_n = m\mathbf{s}$.

Suppose $m \neq 0$. Then $\mathbf{s} = m^{-1}(\mathbf{u}_1 + \cdots + \mathbf{u}_n) \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_n)$, and so $\mathbf{e}_i = \mathbf{s} - \mathbf{u}_i$ lies in the span for each i. By Proposition 1.3.18, $F^n = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_n) \subseteq \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_n)$.

Suppose $m = 0$. An element of the span is $\mathbf{x} = \sum_i a_i \mathbf{u}_i$ for some $a_1, \ldots, a_n \in F$; let $S = a_1 + \cdots + a_n$. The j-th coordinate of $\mathbf{x}$ is the sum of the a_i with $i \neq j$, which is $S - a_j$. So the sum of the coordinates of $\mathbf{x}$ is

$$\sum_{j=1}^{n} (S - a_j) = (n * 1)S - S = ((n-1) * 1)S = mS = 0.$$

The coordinates of $\mathbf{e}_1$ sum to $1 \neq 0$, so $\mathbf{e}_1$ is not in the span, and the list does not span F^n.

For example, over $\mathbb{R}$ the list always spans, while over $\mathbb{F}_2$ it spans exactly when n is even, and over $\mathbb{F}_3$ exactly when $n-1$ is not a multiple of 3.

Solution to Exercise 1.3.43

Uses. Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans), Corollary 1.3.20 (Operations that preserve the span)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. For (a), combine $\mathbf{e}_i + \mathbf{e}_j$, $\mathbf{e}_i + \mathbf{e}_k$, and $\mathbf{e}_j + \mathbf{e}_k$ to isolate a multiple of $\mathbf{e}_i$.
Hint 2. For (b), the sum of the coordinates of a sum of vectors is the sum of their coordinate sums. For the other containment, pair up the coordinates equal to 1.
Hint 3. For (c), the sum of two vectors of $\mathcal{W}_3$ that share two positions lies in $\mathcal{W}_2$. Then handle vectors with an odd number of ones by adding $(1, 1, 1, 0, \ldots, 0)$.
Sketch. (a) $(\mathbf{e}_i + \mathbf{e}_j) + (\mathbf{e}_i + \mathbf{e}_k) - (\mathbf{e}_j + \mathbf{e}_k) = (1+1)\mathbf{e}_i$. (b) Each entry of $\mathcal{W}_2$ has coordinate sum 0, so every combination does; a vector with $2r$ ones is a sum of r entries of $\mathcal{W}_2$. (c) For $n \geq 4$, $(\mathbf{e}_i + \mathbf{e}_k + \mathbf{e}_l) + (\mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l) = \mathbf{e}_i + \mathbf{e}_j$, so the span contains $\operatorname{span}(\mathcal{W}_2)$; a vector with coordinate sum 1 is $(\mathbf{x} + \mathbf{t}) + \mathbf{t}$ with $\mathbf{t} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. For $n = 3$, $\mathcal{W}_3 = (1, 1, 1)$ and the span is $\{\mathbf{0}, (1, 1, 1)\}$.

The entries of $\mathcal{W}_2$ are the vectors $\mathbf{e}_i + \mathbf{e}_j$ with $i < j$, and those of $\mathcal{W}_3$ are $\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k$ with $i < j < k$. By Corollary 1.3.20(a), the order in which they are listed does not affect the spans.

(a) Fix i, and choose j and k so that i, j, k are distinct, which is possible since $n \geq 3$. Then

$$(\mathbf{e}_i + \mathbf{e}_j) + (\mathbf{e}_i + \mathbf{e}_k) - (\mathbf{e}_j + \mathbf{e}_k) = (1+1)\mathbf{e}_i.$$

Since $1 + 1 \neq 0$, $\mathbf{e}_i = (1+1)^{-1}(1+1)\mathbf{e}_i \in \operatorname{span}(\mathcal{W}_2)$. By Proposition 1.3.18, $F^n = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_n) \subseteq \operatorname{span}(\mathcal{W}_2)$.

(b) For $\mathbf{x} \in \mathbb{F}_2^n$, write $\sigma(\mathbf{x}) = x_1 + \cdots + x_n \in \mathbb{F}_2$. Then $\sigma(\mathbf{x} + \mathbf{y}) = \sigma(\mathbf{x}) + \sigma(\mathbf{y})$ and $\sigma(a\mathbf{x}) = a\sigma(\mathbf{x})$, by the coordinatewise operations and the field axioms. Let $E = \{\mathbf{x} \mid \sigma(\mathbf{x}) = 0\}$. We proceed by double containment.

Each entry $\mathbf{e}_i + \mathbf{e}_j$ of $\mathcal{W}_2$ has $\sigma = 1 + 1 = 0$. So for a linear combination $\mathbf{x} = \sum c_{ij}(\mathbf{e}_i + \mathbf{e}_j)$, applying the two rules for σ term by term (by induction on the number of terms) gives $\sigma(\mathbf{x}) = \sum c_{ij} \cdot 0 = 0$. Hence $\operatorname{span}(\mathcal{W}_2) \subseteq E$.

Conversely, let $\mathbf{x} \in E$, and let $i_1 < i_2 < \cdots < i_t$ be the positions where $x_i = 1$. Then $\sigma(\mathbf{x})$ is the sum of t ones, which is 0 in $\mathbb{F}_2$ exactly when t is even, say $t = 2r$. Then

$$\mathbf{x} = (\mathbf{e}_{i_1} + \mathbf{e}_{i_2}) + (\mathbf{e}_{i_3} + \mathbf{e}_{i_4}) + \cdots + (\mathbf{e}_{i_{2r-1}} + \mathbf{e}_{i_{2r}})$$

is a sum of entries of $\mathcal{W}_2$ (and $\mathbf{x} = \mathbf{0}$ if $r = 0$), so $\mathbf{x} \in \operatorname{span}(\mathcal{W}_2)$.

(c) Let $n \geq 4$. For $i \neq j$, choose k and l so that i, j, k, l are distinct. In $\mathbb{F}_2^n$, $\mathbf{e}_k + \mathbf{e}_k = \mathbf{0}$, so

$$(\mathbf{e}_i + \mathbf{e}_k + \mathbf{e}_l) + (\mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l) = \mathbf{e}_i + \mathbf{e}_j.$$

So $\mathcal{W}_2 \subseteq \operatorname{span}(\mathcal{W}_3)$, and by Proposition 1.3.18 and (b), every $\mathbf{x}$ with $\sigma(\mathbf{x}) = 0$ lies in $\operatorname{span}(\mathcal{W}_3)$. Now let $\sigma(\mathbf{x}) = 1$, and let $\mathbf{t} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, an entry of $\mathcal{W}_3$. Then $\sigma(\mathbf{x} + \mathbf{t}) = 1 + 1 = 0$, so $\mathbf{x} + \mathbf{t} \in \operatorname{span}(\mathcal{W}_3)$, and $\mathbf{x} = (\mathbf{x} + \mathbf{t}) + \mathbf{t}$ is a sum of two elements of $\operatorname{span}(\mathcal{W}_3)$, written as combinations of $\mathcal{W}_3$. By Lemma 1.3.9(a), $\mathbf{x} \in \operatorname{span}(\mathcal{W}_3)$. So $\mathcal{W}_3$ spans $\mathbb{F}_2^n$.

For $n = 3$, the only vector with three coordinates equal to 1 is $(1, 1, 1)$, so $\operatorname{span}(\mathcal{W}_3) = \operatorname{span}((1, 1, 1)) = \{\mathbf{0}, (1, 1, 1)\}$ by Example 1.3.11. The pairing trick needs a fourth position.

Solution to Exercise 1.3.44

Uses. Definition 1.3.10 (Span), Example 1.2.20 (Real and rational scalars)
Hint 1. Suppose $\sqrt{3} = a + b\sqrt{2}$ with $a, b \in \mathbb{Q}$, and square both sides.
Hint 2. The squared equation is $3 = a^2 + 2b^2 + 2ab\sqrt{2}$. If $ab \neq 0$, solve for $\sqrt{2}$. Otherwise treat $a = 0$ and $b = 0$ separately.
Sketch. If $\sqrt{3} = a + b\sqrt{2}$ with a, b rational, then $3 = a^2 + 2b^2 + 2ab\sqrt{2}$. If $ab \neq 0$, $\sqrt{2}$ is rational. If $b = 0$, $\sqrt{3} = a$ is rational. If $a = 0$, then $\sqrt{6} = \sqrt{2}\sqrt{3} = 2b$ is rational. Each case is impossible.

Over $\mathbb{Q}$, $\operatorname{span}\left(1, \sqrt{2}\right) = \left\{a + b\sqrt{2} \;\middle|\; a, b \in \mathbb{Q}\right\}$. Suppose that $\sqrt{3} = a + b\sqrt{2}$ with $a, b \in \mathbb{Q}$. Squaring gives

$$3 = a^2 + 2b^2 + 2ab\sqrt{2}.$$

If $ab \neq 0$, then $\sqrt{2} = \left(3 - a^2 - 2b^2\right)/(2ab)$ is rational, a contradiction. So $a = 0$ or $b = 0$. If $b = 0$, then $\sqrt{3} = a$ is rational, a contradiction. If $a = 0$, then $\sqrt{3} = b\sqrt{2}$, and multiplying by $\sqrt{2}$ gives $\sqrt{6} = 2b$, which is rational, a contradiction. Hence $\sqrt{3} \notin \operatorname{span}\left(1, \sqrt{2}\right)$.

Over $\mathbb{R}$ instead of $\mathbb{Q}$, $\operatorname{span}(1) = \mathbb{R}$ contains every real number. The span depends on the field of scalars, a theme of Section 1.5.

Solution to Exercise 1.3.45

Uses. Definition 1.3.10 (Span), Example 1.2.20 (Real and rational scalars), $\mathbb{Q}^k$ is countable, and $\mathbb{R}$ is uncountable
Hint 1. How many linear combinations with rational coefficients does a list of length k have?
Hint 2. The span is the image of the countable set $\mathbb{Q}^k$ under the map that sends coefficients to the vector they represent.
Sketch. The span of $v_1, \ldots, v_k$ over $\mathbb{Q}$ is the image of $\mathbb{Q}^k$ under $(a_1, \ldots, a_k) \mapsto \sum a_i v_i$. A finite product of countable sets is countable, and so is the image of a countable set, while $\mathbb{R}$ is uncountable.

Let $v_1, \ldots, v_k \in \mathbb{R}$. (Vectors here are real numbers, so we do not write them in bold.) The map $\mathbb{Q}^k \to \mathbb{R}$ sending $(a_1, \ldots, a_k)$ to $a_1 v_1 + \cdots + a_k v_k$ has image $\operatorname{span}(v_1, \ldots, v_k)$, by Definition 1.3.10. The set $\mathbb{Q}^k$ is countable, being a finite product of countable sets (for $k = 0$ it has one element), and the image of a countable set under any map is countable. So $\operatorname{span}(v_1, \ldots, v_k)$ is countable. Since $\mathbb{R}$ is uncountable, $\operatorname{span}(v_1, \ldots, v_k) \neq \mathbb{R}$.

The same argument shows that span over $\mathbb{Q}$ of any countable set is countable, so every spanning set of $\mathbb{R}$ over $\mathbb{Q}$ is uncountable (Aside (Spans of infinite sets)).

Solution to Exercise 1.3.46

Uses. Definition 1.3.10 (Span), Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. For (a), pressing lamps in any order adds a sum of $\mathbf{t}_i$'s, and pressing a lamp twice does nothing. Over $\mathbb{F}_2$, what are the linear combinations of $\mathbf{t}_1, \ldots, \mathbf{t}_n$?
Hint 2. Work out $n = 1, 2, 3, 4, 5$ by hand. For $n = 2$ and $n = 5$, look for a set S of positions such that every $\mathbf{t}_i$ has an even number of ones in S.
Hint 3. For $n = 3m + 2$, take $S = \{i \mid 3 \nmid i\}$. For the other n, show $\mathbf{t}_{3j+1} + \mathbf{t}_{3j+2} = \mathbf{e}_{3j} + \mathbf{e}_{3j+3}$, walk along the row from the left end to get $\mathbf{e}_3, \mathbf{e}_6, \ldots$, then come back.
Sketch. (a) The reachable configurations from $\mathbf{x}$ are $\mathbf{x} + \sum_{i \in S} \mathbf{t}_i$, and over $\mathbb{F}_2$ these sums are exactly the linear combinations of the $\mathbf{t}_i$; $\mathbf{x} + \mathbf{y} = \mathbf{0}$ means $\mathbf{x} = \mathbf{y}$. (b) If $n = 3m + 2$, the parity of the lamps in positions not divisible by 3 is unchanged by every press, since each $\mathbf{t}_i$ meets those positions in 0 or 2 places, while $\mathbf{e}_1$ has parity 1. If $n = 3m$ or $3m + 1$, then $\mathbf{e}_3, \mathbf{e}_6, \ldots, \mathbf{e}_{3m}$ are in the span, $\mathbf{t}_n$ supplies whichever of $\mathbf{e}_{n-1}, \mathbf{e}_n$ is missing, and going back with $\mathbf{e}_{i-1} = \mathbf{t}_i + \mathbf{e}_i + \mathbf{e}_{i+1}$ gives all $\mathbf{e}_i$.

(a) Pressing lamps one after another adds the corresponding $\mathbf{t}_i$'s, and since addition in $\mathbb{F}_2^n$ is commutative and $\mathbf{t}_i + \mathbf{t}_i = \mathbf{0}$, the result of any sequence of presses from $\mathbf{x}$ is $\mathbf{x} + \sum_{i \in S} \mathbf{t}_i$, where S is the set of lamps pressed an odd number of times. Every subset S arises. Since the only scalars in $\mathbb{F}_2$ are 0 and 1, the linear

combinations of $\mathbf{t}_1, \ldots, \mathbf{t}_n$ represent exactly the sums $\sum_{i \in S} \mathbf{t}_i$. Finally, $\mathbf{x} + \mathbf{y} = \mathbf{0}$ in $\mathbb{F}_2^n$ if and only if $\mathbf{x} = -\mathbf{y} = \mathbf{y}$. So $\mathbf{x}$ can be switched off if and only if $\mathbf{x} \in \operatorname{span}(\mathbf{t}_1, \ldots, \mathbf{t}_n)$, and every configuration can be switched off if and only if the list spans $\mathbb{F}_2^n$.

(b) For $n = 1$, $\mathbf{t}_1 = \mathbf{e}_1$ spans $\mathbb{F}_2^1$. From now on let $n \geq 2$, so $\mathbf{t}_1 = \mathbf{e}_1 + \mathbf{e}_2$ and $\mathbf{t}_n = \mathbf{e}_{n-1} + \mathbf{e}_n$.

Case $n = 3m + 2$. For $\mathbf{x} \in \mathbb{F}_2^n$, let $f(\mathbf{x}) = \sum_{i \in S} x_i$, where S is the set of $i \in \{1, \ldots, n\}$ not divisible by 3. Then $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$. We claim $f(\mathbf{t}_i) = 0$ for every i. For $2 \leq i \leq n - 1$, the positions $i - 1, i, i + 1$ are three consecutive integers, exactly one of which is divisible by 3, so $f(\mathbf{t}_i) = 1 + 1 = 0$. For $i = 1$, positions 1 and 2 both lie in S. For $i = n = 3m + 2$, positions $3m + 1$ and $3m + 2$ both lie in S. So f vanishes on each $\mathbf{t}_i$, hence on every sum of them, hence on the span. Since $f(\mathbf{e}_1) = 1$, $\mathbf{e}_1$ is not in the span.

Case $n = 3m$ or $n = 3m + 1$, with $m \geq 1$. First, $\mathbf{t}_1 + \mathbf{t}_2 = (\mathbf{e}_1 + \mathbf{e}_2) + (\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = \mathbf{e}_3$. Next, if $3j + 3 \leq n$, then $3j + 2 \leq n - 1$, so lamp $3j + 2$ has a right neighbor, and

$$\mathbf{t}_{3j+1} + \mathbf{t}_{3j+2} = \left(\mathbf{e}_{3j} + \mathbf{e}_{3j+1} + \mathbf{e}_{3j+2}\right) + \left(\mathbf{e}_{3j+1} + \mathbf{e}_{3j+2} + \mathbf{e}_{3j+3}\right) = \mathbf{e}_{3j} + \mathbf{e}_{3j+3}.$$

So $\mathbf{e}_{3j+3} = \mathbf{t}_{3j+1} + \mathbf{t}_{3j+2} + \mathbf{e}_{3j}$, and by induction on j, the vectors $\mathbf{e}_3, \mathbf{e}_6, \ldots, \mathbf{e}_{3m}$ lie in the span. If $n = 3m$, this gives $\mathbf{e}_n$, and then $\mathbf{e}_{n-1} = \mathbf{t}_n + \mathbf{e}_n$. If $n = 3m + 1$, it gives $\mathbf{e}_{n-1}$, and then $\mathbf{e}_n = \mathbf{t}_n + \mathbf{e}_{n-1}$. In both cases $\mathbf{e}_{n-1}$ and $\mathbf{e}_n$ lie in the span, and for $i = n - 1, n - 2, \ldots, 2$ in turn, $\mathbf{e}_{i-1} = \mathbf{t}_i + \mathbf{e}_i + \mathbf{e}_{i+1}$ does too. So every $\mathbf{e}_i$ lies in the span, and the list spans $\mathbb{F}_2^n$ by Proposition 1.3.18.

Since every $n \geq 1$ has exactly one of the forms $3m$, $3m + 1$, $3m + 2$, this proves (b). Small cases are how one finds S: for $n = 2$ the two presses are equal, $\mathbf{t}_1 = \mathbf{t}_2 = (1, 1)$, and the invariant is $x_1 + x_2$; for $n = 5$ it is $x_1 + x_2 + x_4 + x_5$.

Section 1.4

Solution to Exercise 1.4.6

> *Uses.* Theorem 1.4.3 (Subspace criterion)
> *Hint 1.* Compare with (II) and (III) of Theorem 1.4.3. The vector $a\mathbf{w} + \mathbf{w}'$ is a linear combination of length 2.
> *Hint 2.* For the converse, special values of a and $\mathbf{w}'$ recover the two closure conditions of (III).
> *Sketch.* If W is a subspace, then $\mathbf{0} \in W$ by (III), and $a\mathbf{w} + 1\mathbf{w}' \in W$ by (II). Conversely, $a = 1$ gives $\mathbf{w} + \mathbf{w}' \in W$, and $\mathbf{w}' = \mathbf{0}$ gives $a\mathbf{w} = a\mathbf{w} + \mathbf{0} \in W$, so (III) holds.

Suppose W is a subspace. Then $\mathbf{0} \in W$ by (III) of Theorem 1.4.3, and for $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$, the vector $a\mathbf{w} + \mathbf{w}' = a\mathbf{w} + 1\mathbf{w}'$ is a linear combination of elements of W, so it lies in W by (II).

Conversely, suppose the condition holds, and let $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$. Taking the scalar 1 gives $\mathbf{w} + \mathbf{w}' = 1\mathbf{w} + \mathbf{w}' \in W$, by (A5). Taking $\mathbf{0}$ as the second vector, which is allowed since $\mathbf{0} \in W$, gives $a\mathbf{w} = a\mathbf{w} + \mathbf{0} \in W$, by (A3). So W satisfies (III), and it is a subspace by Theorem 1.4.3.

Solution to Exercise 1.4.15

> *Uses.* Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 1.4.3 (Subspace criterion), Example 1.3.11 (Small spans)
> *Hint 1.* One direction is Proposition 1.4.14(b).
> *Hint 2.* For the other, apply the condition to three kinds of list: the empty list, a list $\mathbf{w}, \mathbf{w}'$ of length 2, and a list $\mathbf{w}$ of length 1.
> *Sketch.* ($\Rightarrow$) is Proposition 1.4.14(b). ($\Leftarrow$) $\operatorname{span}(()) = \{\mathbf{0}\}$ gives $\mathbf{0} \in W$; $\mathbf{w} + \mathbf{w}' \in \operatorname{span}(\mathbf{w}, \mathbf{w}')$ and $a\mathbf{w} \in \operatorname{span}(\mathbf{w})$ give the closure conditions, so (III) of Theorem 1.4.3 holds.

If W is a subspace and $\mathcal{L}$ is a list in W, then $\operatorname{span}(\mathcal{L}) \subseteq W$ by Proposition 1.4.14(b).

Conversely, suppose $\operatorname{span}(\mathcal{L}) \subseteq W$ for every list $\mathcal{L}$ in W. For the empty list, $\operatorname{span}(()) = \{\mathbf{0}\}$ (Definition 1.3.10), so $\mathbf{0} \in W$. Let $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$. Then $\mathbf{w} + \mathbf{w}' = 1\mathbf{w} + 1\mathbf{w}' \in \operatorname{span}(\mathbf{w}, \mathbf{w}') \subseteq W$ and $a\mathbf{w} \in \operatorname{span}(\mathbf{w}) \subseteq W$. So W satisfies (III) of Theorem 1.4.3, and it is a subspace.

In this form the criterion needs no separate nonemptiness condition: the empty list supplies $\mathbf{0}$ (Remark 1.4.5).

Solution to Exercise 1.4.21

> *Uses.* Proposition 1.4.20 (Unions of two subspaces), Example 1.4.7 (Trivial subspaces)
> *Hint 1.* V is a subspace of itself, so $U \cup W$ is a subspace.
> *Hint 2.* Apply Proposition 1.4.20. If $U \subseteq W$, what is $U \cup W$?
> *Sketch.* $U \cup W = V$ is a subspace, so $U \subseteq W$ or $W \subseteq U$ by Proposition 1.4.20. In the first case $V = U \cup W = W$, in the second $V = U$.

Since $U \cup W = V$ is a subspace of V, Proposition 1.4.20 gives $U \subseteq W$ or $W \subseteq U$. If $U \subseteq W$, then $V = U \cup W = W$. If $W \subseteq U$, then $V = U \cup W = U$.

So a vector space is never the union of two proper subspaces. Three can suffice: $\mathbb{F}_2^2$ is the union of its three lines $\{\mathbf{0}, (1,0)\}$, $\{\mathbf{0}, (0,1)\}$, $\{\mathbf{0}, (1,1)\}$ (Example 1.4.12). Exercise 1.4.42 and Exercise 1.4.43 show how this depends on the size of the field.

Solution to Exercise 1.4.26

> *Uses.* Theorem 1.4.3 (Subspace criterion), Example 1.4.8 (One equation in F^n), Proposition 1.4.16 (Intersections of subspaces)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* Check whether $\mathbf{0}$ lies in the set first. When it does, test closure on simple vectors, such as standard vectors and their negatives.
> *Hint 2.* For (e), over $\mathbb{R}$ a sum of two squares vanishes only when both squares do. Over $\mathbb{C}$, look for a nonzero solution of $z^2 + w^2 = 0$.
> *Sketch.* (a) Yes, an equation as in Example 1.4.8. (b) No, $\mathbf{0}$ is missing. (c) No: $\mathbf{e}_1 + \mathbf{e}_2$. (d) No: $(-1)\mathbf{e}_2$. (e) Over $\mathbb{R}$ the set is $\{x_1 = x_2 = 0\}$, a subspace. Over $\mathbb{C}$ it contains $(1, i, 0)$ and $(1, -i, 0)$ but not their sum $(2, 0, 0)$.

(a) Yes. It is the solution set of $x_1 + 0x_2 - 2x_3 = 0$, a subspace by Example 1.4.8.

(b) No. The coordinates of $\mathbf{0}$ add up to $0 \neq 1$.

(c) No. The vectors $(1,0,0)$ and $(0,1,0)$ lie in the set, but their sum $(1,1,0)$ does not, since $1 \cdot 1 \neq 0$.

(d) No. The vector $(0,1,0)$ lies in the set, since $0 \leq 1$, but $(-1)(0,1,0) = (0,-1,0)$ does not, since $0 > -1$.

(e) For real numbers, $x_1^2 \geq 0$ and $x_2^2 \geq 0$, so $x_1^2 + x_2^2 = 0$ holds exactly when $x_1 = x_2 = 0$. So the subset of $\mathbb{R}^3$ is $\{\mathbf{x} \mid x_1 = 0\} \cap \{\mathbf{x} \mid x_2 = 0\}$, an intersection of two subspaces of the form in Example 1.4.8, hence a subspace by Proposition 1.4.16. In $\mathbb{C}^3$, the vectors $(1, i, 0)$ and $(1, -i, 0)$ satisfy the equation, since $1 + i^2 = 1 + (-i)^2 = 0$, but their sum $(2,0,0)$ gives $4 + 0 \neq 0$. So the subset of $\mathbb{C}^3$ is not a subspace. The same equation defines a subspace over one field and not over the other.

Solution to Exercise 1.4.27

> *Uses.* Theorem 1.4.3 (Subspace criterion), Definition 1.2.9 (Function spaces), Definition 1.2.13 (Polynomials), Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 1.4.20 (Unions of two subspaces)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* The sets defined by linear conditions, such as an equation between values or coefficients, are subspaces. For the others, look for constants, simple polynomials, or diagonal matrices that break closure.
> *Hint 2.* For (b), try the constant function 1 and its double. For (d), try $x - 1$ and $x - 2$. For (f), try the two diagonal matrices with a single 1.
> *Sketch.* (a) Yes: the condition survives sums and multiples. (b) No: the constant 1 is in it, the constant 2 is not. (c) Yes: $|f + g| \leq M + N$ and $|af| \leq |a|M$. (d) No: $(x-1) + (x-2) = 2x - 3$ is nonzero at 1 and 2. (e) Yes: coefficientwise operations. (f) No: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = I$.

(a) Yes. The zero function satisfies $0 = 0$. If $f(0) = f(1)$ and $g(0) = g(1)$, then $(f+g)(0) = f(1) + g(1) = (f+g)(1)$ and $(af)(0) = af(1) = (af)(1)$.

(b) No. The constant function 1 lies in the set, since $1 = 1^2$, but $2 \cdot 1$ is the constant function 2, and $2 \neq 2^2$.

(c) Yes. The zero function is bounded by $M = 0$. If $|f(t)| \le M$ and $|g(t)| \le N$ for all t, then $|f(t) + g(t)| \le M + N$ and $|af(t)| \le |a|M$ for all t, by the triangle inequality for real numbers.

(d) No. The polynomials $x - 1$ and $x - 2$ lie in the set, but their sum $p = 2x - 3$ has $p(1)p(2) = (-1) \cdot 1 = -1$. In fact, since $\mathbb{R}$ has no zero divisors, the set is the union of the subspaces $\{p(1) = 0\}$ and $\{p(2) = 0\}$ of Example 1.4.9, and neither contains the other ($x - 1$ is in the first only, $x - 2$ in the second only), so Proposition 1.4.20 also shows that it is not a subspace.

(e) Yes. The operations of $\mathbb{R}[x]$ are coefficientwise, so if $a_k = b_k = 0$ for every odd k, then $a_k + b_k = 0$ and $ca_k = 0$ for every odd k; and $\mathbf{0}$ has all coefficients 0.

(f) No, over every field. The matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ satisfy the equation, since $1 \cdot 0 = 0 \cdot 0$, but their sum $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gives $1 \cdot 1 \neq 0 \cdot 0$.

Solution to Exercise 1.4.28

Uses. Theorem 1.4.3 (Subspace criterion), Proposition 1.2.19 (Restriction of scalars)
Hint 1. The conditions $\mathbf{0} \in W$ and closure under addition do not depend on the field. Only closure under scalar multiplication does, and over $\mathbb{R}$ there are fewer scalars to test.
Hint 2. For (b), test multiplication by i. For D over $\mathbb{R}$, check that $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{rz} = r\overline{z}$ for real r.
Sketch. (a) A set closed under multiplication by all complex scalars is closed under multiplication by the real ones. (b) $\mathbb{R}^2$ is a real subspace but not a complex one: $i(1, 0) = (i, 0)$. D is a real subspace, since conjugation is additive and commutes with real scalars, but not a complex one: $(1, 1) \in D$ and $i(1, 1) = (i, i) \notin D$.

(a) Let W be a subspace of $\mathbb{C}^2$ over $\mathbb{C}$. Then $\mathbf{0} \in W$, W is closed under addition, and $c\mathbf{w} \in W$ for all $c \in \mathbb{C}$ and $\mathbf{w} \in W$, in particular for all real c. The real vector space $\mathbb{C}^2$ has the same addition, and its scalar multiplication is the complex one restricted to real scalars. So W satisfies (III) of Theorem 1.4.3 over $\mathbb{R}$.

(b) $\mathbb{R}^2$. It contains $\mathbf{0} = (0, 0)$, sums of pairs of real numbers are pairs of real numbers, and so are real multiples. So $\mathbb{R}^2$ is a subspace of $\mathbb{C}^2$ over $\mathbb{R}$. It is not a subspace over $\mathbb{C}$: $(1, 0) \in \mathbb{R}^2$ but $i(1, 0) = (i, 0) \notin \mathbb{R}^2$.

D. For $z = a + bi$ and $w = c + di$,

$$\overline{z} + \overline{w} = (a + c) - (b + d)i = \overline{z + w},$$

and for real r, $r\overline{z} = ra - rbi = \overline{rz}$. So D contains $\left(0, \overline{0}\right) = \mathbf{0}$, and

$$(z, \overline{z}) + (w, \overline{w}) = (z + w, \overline{z + w}), \quad r(z, \overline{z}) = (rz, \overline{rz})$$

lie in D. Hence D is a subspace over $\mathbb{R}$. It is not a subspace over $\mathbb{C}$: $(1, 1) = \left(1, \overline{1}\right) \in D$, but $i(1, 1) = (i, i)$, and $\overline{i} = -i \neq i$.

So the converse of (a) fails.

Solution to Exercise 1.4.29

Uses. Definition 1.4.1 (Subspace), Theorem 1.4.3 (Subspace criterion), Remark 1.4.4 (The zero vector of a subspace)
Hint 1. Reread Definition 1.4.1. Under which operations must a subspace be a vector space?
Hint 2. Test P against (III) of Theorem 1.4.3, with the usual operations of $\mathbb{R}$.
Sketch. A subspace must be a vector space under the operations of $\mathbb{R}$ itself, not under new ones. Under the usual operations, $0 \notin P$ and $(-1) \cdot 1 = -1 \notin P$. The claimed zero vector 1 is also a giveaway: a subspace's zero is the zero of $\mathbb{R}$ (Remark 1.4.4).

The argument shows that P is a vector space, but under operations that are not those of $\mathbb{R}$. Definition 1.4.1 requires P to be a vector space under the addition and scalar multiplication of $\mathbb{R}$, restricted to P. Under these, P fails (III) of Theorem 1.4.3: $0 \notin P$, and $1 \in P$ while $(-1) \cdot 1 = -1 \notin P$. So P is not a subspace of $\mathbb{R}$.

The zero vector 1 is a symptom of the error: by Remark 1.4.4, the zero vector of a subspace of $\mathbb{R}$ must be the real number 0.

Solution to Exercise 1.4.30

Uses. Proposition 1.4.20 (Unions of two subspaces), Lemma 1.4.19 (Adding a vector outside a subspace), Non-example 1.4.13 (Three ways to fail)
Hint 1. Two of the three steps are correct. Which one quietly assumes something about where $\mathbf{v}$ and $\mathbf{w}$ lie?
Hint 2. Closure of U under addition says something about two vectors of U. What if $\mathbf{v} \in U$ and $\mathbf{w} \in W$?
Sketch. The last step applies closure of U (or of W) to two vectors that need not both lie in U (or both in W). In the mixed case $\mathbf{v} \in U \setminus W$, $\mathbf{w} \in W \setminus U$, the sum lies in neither (Lemma 1.4.19); the axes of $\mathbb{R}^2$ are a counterexample to the claim.

The first two steps are correct: $U \cup W$ always contains $\mathbf{0}$ and is closed under scalar multiplication. The error is in the last step. Closure of U under addition applies when $\mathbf{v}$ and $\mathbf{w}$ *both* lie in U, and closure of W when both lie in W. The argument ignores the mixed case, $\mathbf{v} \in U$ and $\mathbf{w} \in W$ with $\mathbf{v} \notin W$ and $\mathbf{w} \notin U$, and in that case $\mathbf{v} + \mathbf{w}$ lies in neither subspace, by Lemma 1.4.19. The claim is false: for the two axes of $\mathbb{R}^2$, $(1,0) + (0,1) = (1,1)$ lies on neither axis (Non-example 1.4.13). By Proposition 1.4.20, it holds exactly when one subspace contains the other, which is exactly when the mixed case cannot occur.

Solution to Exercise 1.4.31

Uses. Theorem 1.4.3 (Subspace criterion), Definition 1.1.19 (Integers modulo n), Definition 1.2.1 (Vector space)
Hint 1. In $\mathbb{F}_p$, every scalar is a sum of 1's: $c = 1 + \cdots + 1$ (c terms) for $1 \le c \le p-1$, and 0 is the sum of p ones.
Hint 2. By (A8) and (A5), $(1 + \cdots + 1)\mathbf{w} = \mathbf{w} + \cdots + \mathbf{w}$. For (b), think of the integers inside $\mathbb{Q}$.
Sketch. (a) For $\mathbf{w} \in W$ and $m \ge 1$, $m * \mathbf{w} \in W$ by closure, and $(m*1)\mathbf{w} = m * \mathbf{w}$ by (A8) and (A5). Every $c \in \mathbb{F}_p$ equals $m * 1$ for some $1 \le m \le p$, so $c\mathbf{w} \in W$; with $c = 0$ this gives $\mathbf{0} \in W$. (b) $\mathbb{Z}^2$: closed under sums and negatives, but $\frac{1}{2}(1,0) \notin \mathbb{Z}^2$.

(a) For a positive integer m, write $m * \mathbf{w} = \mathbf{w} + \cdots + \mathbf{w}$ (m terms) and $m * 1 = 1 + \cdots + 1$ in $\mathbb{F}_p$. We claim that $(m*1)\mathbf{w} = m * \mathbf{w}$, by induction on m. For $m = 1$ this is (A5). For $m \ge 2$, by (A8), the inductive hypothesis, and (A5),

$$(m*1)\mathbf{w} = ((m-1)*1 + 1)\mathbf{w} = (m-1) * \mathbf{w} + 1\mathbf{w} = m * \mathbf{w}.$$

Now let $\mathbf{w} \in W$, which exists since $W \neq \emptyset$. By closure under addition and induction on m, $m * \mathbf{w} \in W$ for every $m \ge 1$. Every $c \in \mathbb{F}_p$ is of the form $m * 1$ with $1 \le m \le p$: for $c \in \{1, \ldots, p-1\}$, the sum of c ones in $\mathbb{F}_p$ is c, and the sum of p ones is $p \bmod p = 0$. So $c\mathbf{w} = m * \mathbf{w} \in W$ for every $c \in \mathbb{F}_p$. With $c = 0$ this gives $\mathbf{0} = 0\mathbf{w} \in W$. Hence W satisfies (III) of Theorem 1.4.3.

(b) Let $W = \mathbb{Z}^2 = \{(m,n) \mid m, n \in \mathbb{Z}\}$. It contains $(0,0)$, and sums and negatives of pairs of integers are pairs of integers. But $(1,0) \in W$ and $\frac{1}{2}(1,0) = \left(\frac{1}{2}, 0\right) \notin W$, so W is not a subspace. The proof of (a) fails because $\frac{1}{2}$ is not a sum of 1's in $\mathbb{Q}$.

Solution to Exercise 1.4.32

Uses. Exercise 1.1.40, Proposition 1.4.20 (Unions of two subspaces), Example 1.4.8 (One equation in F^n)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. Factor: $x^2 = y^2$ if and only if $x = y$ or $x = -y$ (Exercise 1.1.40). So S_p is a union of two lines.
Hint 2. When are the two lines $x = y$ and $x = -y$ the same line? If they differ, apply Proposition 1.4.20.
Sketch. By Exercise 1.1.40, $S_p = D \cup D'$ with $D = \{x = y\}$ and $D' = \{x = -y\}$, both subspaces. For $p = 2$, $-1 = 1$, so $S_2 = D$ is a subspace. For odd p, $1 \neq -1$, so $(1,1) \in D \setminus D'$ and $(1,-1) \in D' \setminus D$, and S_p is not a subspace by Proposition 1.4.20; indeed $(1,1) + (1,-1) = (2,0) \notin S_p$. Answer: $p = 2$ only.

By Exercise 1.1.40, $x^2 = y^2$ holds if and only if $x = y$ or $x = -y$. So $S_p = D \cup D'$, where

$$D = \{(x,y) \mid x - y = 0\}, \quad D' = \{(x,y) \mid x + y = 0\}$$

are subspaces of $\mathbb{F}_p^2$ by Example 1.4.8.

If $p = 2$, then $-1 = 1$ in $\mathbb{F}_2$, so $D' = D$ and $S_2 = D$ is a subspace. (It is $\{\mathbf{0}, (1,1)\}$.)

If p is odd, then $1 + 1 = 2 \neq 0$ in $\mathbb{F}_p$, so $1 \neq -1$. Then $(1,1) \in D$ but $(1,1) \notin D'$, since $1 + 1 \neq 0$; and $(1,-1) \in D'$ but $(1,-1) \notin D$, since $1 - (-1) = 2 \neq 0$. So neither of D, D' contains the other, and $S_p = D \cup D'$ is not a subspace by Proposition 1.4.20. Concretely, $(1,1) + (1,-1) = (2,0)$, and $2^2 = 4 \neq 0 = 0^2$ in $\mathbb{F}_p$.

So S_p is a subspace exactly when $p = 2$.

Solution to Exercise 1.4.33

Uses. Theorem 1.4.25 (The subspaces of F^2), Exercise 1.3.28, Corollary 1.2.36 (Cancelling scalars and vectors)
Hint 1. By Theorem 1.4.25, it suffices to count lines. How many elements does a line $\operatorname{span}(\mathbf{v})$ have, and how many of them are nonzero?
Hint 2. Every nonzero vector lies on exactly one line. So the lines split the q^2-1 nonzero vectors into groups of equal size.
Sketch. (a) $\{\mathbf{0}\}$, four lines $\{\mathbf{0},(1,0),(2,0)\}$, $\{\mathbf{0},(0,1),(0,2)\}$, $\{\mathbf{0},(1,1),(2,2)\}$, $\{\mathbf{0},(1,2),(2,1)\}$, and $\mathbb{F}_3^2$. (b) A line has q elements, $q-1$ of them nonzero, and each nonzero $\mathbf{v}$ lies on exactly one line, $\operatorname{span}(\mathbf{v})$. So there are $(q^2-1)/(q-1)=q+1$ lines, plus $\{\mathbf{0}\}$ and F^2.

(b) By Theorem 1.4.25, the subspaces of F^2 are $\{\mathbf{0}\}$, the lines, and F^2, and these are different: $\{\mathbf{0}\}$ has 1 element, a line has q (see below), and F^2 has q^2, with $1<q<q^2$ as $q\geq 2$. So it suffices to count the lines.

Let $\mathbf{v}\neq\mathbf{0}$. By Corollary 1.2.36(b), the vectors $a\mathbf{v}$, $a\in F$, are distinct, so $\operatorname{span}(\mathbf{v})$ has exactly q elements, of which $q-1$ are nonzero. Every nonzero vector $\mathbf{w}$ lies on exactly one line. It lies on $\operatorname{span}(\mathbf{w})$. If it also lies on $\operatorname{span}(\mathbf{v})$, then $\mathbf{w}=c\mathbf{v}$ with $c\neq 0$ (as $\mathbf{w}\neq\mathbf{0}$), and $\operatorname{span}(\mathbf{w})=\operatorname{span}(\mathbf{v})$ by Exercise 1.3.28(b). So the lines partition the q^2-1 nonzero vectors of F^2 into blocks of $q-1$ vectors, and there are

$$\frac{q^2-1}{q-1}=q+1$$

lines. Together with $\{\mathbf{0}\}$ and F^2, this gives $q+3$ subspaces.

(a) For $q=3$ there are 4 lines. Listing the multiples $0\mathbf{v},1\mathbf{v},2\mathbf{v}$ of $\mathbf{v}=(1,0),(0,1),(1,1),(1,2)$ in $\mathbb{F}_3^2$, the six subspaces are

$$\{\mathbf{0}\},\quad \{\mathbf{0},(1,0),(2,0)\},\quad \{\mathbf{0},(0,1),(0,2)\},$$
$$\{\mathbf{0},(1,1),(2,2)\},\quad \{\mathbf{0},(1,2),(2,1)\},\quad \mathbb{F}_3^2.$$

The four lines contain $4\cdot 2=8$ nonzero vectors, which are all the nonzero vectors of $\mathbb{F}_3^2$, as (b) predicts.

Solution to Exercise 1.4.34

Uses. Theorem 1.4.25 (The subspaces of F^2), Exercise 1.1.40, Exercise 1.3.28
Hint 1. By Theorem 1.4.25, only the lines $\operatorname{span}(\mathbf{v})$ need thought. Show that such a line qualifies exactly when $\sigma(\mathbf{v})=c\mathbf{v}$ for some $c\in F$.
Hint 2. Write $\mathbf{v}=(a,b)$. From $b=ca$ and $a=cb$, deduce $c^2=1$, so $c=\pm 1$.
Sketch. $\{\mathbf{0}\}$ and F^2 qualify. A line $\operatorname{span}(\mathbf{v})$ qualifies iff $\sigma(\mathbf{v})\in\operatorname{span}(\mathbf{v})$, since $\sigma(t\mathbf{v})=t\sigma(\mathbf{v})$. With $\mathbf{v}=(a,b)$ and $\sigma(\mathbf{v})=c\mathbf{v}$: $a=c^2a$, $b=c^2b$, so $c^2=1$ and $c=\pm 1$, giving the lines $\operatorname{span}((1,1))$ and $\operatorname{span}((1,-1))$. These are different iff $1+1\neq 0$. So there are four such subspaces if $1+1\neq 0$, and three if $1+1=0$.

Call U *swap-closed* if $\sigma(\mathbf{u})\in U$ for all $\mathbf{u}\in U$. Clearly $\{\mathbf{0}\}$ and F^2 are swap-closed. By Theorem 1.4.25, every other subspace is a line $\operatorname{span}(\mathbf{v})$ with $\mathbf{v}=(a,b)\neq\mathbf{0}$. For $t\in F$, $\sigma(t\mathbf{v})=(tb,ta)=t\sigma(\mathbf{v})$. So if $\sigma(\mathbf{v})\in\operatorname{span}(\mathbf{v})$, then σ maps all of $\operatorname{span}(\mathbf{v})$ into $\operatorname{span}(\mathbf{v})$; and conversely a swap-closed line contains $\sigma(\mathbf{v})$. Hence $\operatorname{span}(\mathbf{v})$ is swap-closed if and only if $\sigma(\mathbf{v})=c\mathbf{v}$ for some $c\in F$, that is, $b=ca$ and $a=cb$.

Suppose $b=ca$ and $a=cb$. Then $a=c^2a$ and $b=c^2b$. Since a or b is nonzero, $c^2=1$, and by Exercise 1.1.40, $c=1$ or $c=-1$. If $c=1$, then $a=b$ and $\mathbf{v}=a(1,1)$. If $c=-1$, then $b=-a$ and $\mathbf{v}=a(1,-1)$. Either way $a\neq 0$, so $\operatorname{span}(\mathbf{v})$ is $\operatorname{span}((1,1))$ or $\operatorname{span}((1,-1))$ by Exercise 1.3.28(b). Conversely, both lines are swap-closed, since $\sigma(1,1)=(1,1)$ and $\sigma(1,-1)=(-1,1)=-(1,-1)$.

If $1+1\neq 0$, the two lines are different, because $(1,-1)=c(1,1)$ would force $c=1$ and $c=-1$. Then the swap-closed subspaces are the four subspaces $\{\mathbf{0}\}$, $\operatorname{span}((1,1))$, $\operatorname{span}((1,-1))$, and F^2. If $1+1=0$, then $-1=1$, the two lines coincide, and there are three: $\{\mathbf{0}\}$, $\operatorname{span}((1,1))$, and F^2. Over $\mathbb{R}$, the two lines are the diagonals, the mirror of σ and the line it reverses. Vectors $\mathbf{v}$ with $\sigma(\mathbf{v})=c\mathbf{v}$ are eigenvectors, the subject of Section 7.2.

Solution to Exercise 1.4.35

Uses. Proposition 1.4.24 (The sum of two subspaces), Example 1.4.10 (Functions)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For $E\cap O$: if f is both even and odd, compare $f(-t)$ computed both ways.

Hint 2. For $E + O$, guess the even part of f by symmetrizing: average $f(t)$ and $f(-t)$. The odd part is what remains.
Sketch. (a) If $f \in E \cap O$, then $f(t) = f(-t) = -f(t)$, so $2f(t) = 0$. Every f is $f_e + f_o$ with $f_{e(t)} = (f(t) + f(-t))/2$ even and $f_{o(t)} = (f(t) - f(-t))/2$ odd. (b) In $\mathbb{F}_2$, $-t = t$ and $-f(t) = f(t)$, so both conditions read $f(t) = f(t)$. Then $E + O$ is still everything, but $E \cap O \neq \{\mathbf{0}\}$: the step $2f(t) = 0 \Rightarrow f(t) = 0$ divides by $2 = 0$.

(a) If $f \in E \cap O$, then for every t, $f(t) = f(-t) = -f(t)$, so $2f(t) = 0$ and $f(t) = 0$. Hence $E \cap O = \{\mathbf{0}\}$.

For the sum, $E + O \subseteq \mathbb{R}^{\mathbb{R}}$ holds trivially. Conversely, let $f \in \mathbb{R}^{\mathbb{R}}$, and define

$$f_e(t) = \frac{f(t) + f(-t)}{2}, \quad f_o(t) = \frac{f(t) - f(-t)}{2}.$$

Then $f_e(-t) = (f(-t) + f(t))/2 = f_e(t)$, so $f_e \in E$; $f_o(-t) = (f(-t) - f(t))/2 = -f_o(t)$, so $f_o \in O$; and $f_e(t) + f_o(t) = f(t)$ for every t. So $f = f_e + f_o \in E + O$, and $E + O = \mathbb{R}^{\mathbb{R}}$.

(b) In $\mathbb{F}_2$, $-a = a$ for every a, since $a + a = (1 + 1)a = 0$. So for every $f : \mathbb{F}_2 \to \mathbb{F}_2$ and every t, $f(-t) = f(t)$ and $-f(t) = f(t)$, and both conditions read $f(t) = f(t)$. Every function is even and odd: $E = O = \mathbb{F}_2^{\mathbb{F}_2}$. So $E + O = \mathbb{F}_2^{\mathbb{F}_2}$ still holds, trivially, but $E \cap O \neq \{\mathbf{0}\}$, and a function no longer has just one decomposition into an even and an odd part. The first argument in (a) fails because it passes from $2f(t) = 0$ to $f(t) = 0$, and the formulas for f_e and f_o in the second involve $1/2$. In $\mathbb{F}_2$, $2 = 0$ has no inverse.

Solution to Exercise 1.4.36

Uses. Proposition 1.4.24 (The sum of two subspaces), Example 1.4.11 (Matrices), Theorem 1.4.3 (Subspace criterion)
Techniques. Double containment (Technique 1.3.14)
Hint 1. N and K are cut out by entrywise equations, as in Example 1.4.11. For $S + N = F^{n\times n}$, build the symmetric part of A from the entries of A on and below the diagonal.
Hint 2. For (b) with $1 + 1 \neq 0$, use $\left(A + A^{\mathsf{T}}\right)/2$ and $\left(A - A^{\mathsf{T}}\right)/2$, written entrywise. If $1 + 1 = 0$, compare the conditions defining K and S.
Sketch. (a) A matrix in $S \cap N$ vanishes on and below the diagonal, and its entries above equal entries below. For A, let B copy A on and below the diagonal and be symmetric; then $A - B \in N$. (b) If $1 + 1 \neq 0$, $a_{ij} = h\big(a_{ij} + a_{ji}\big) + h\big(a_{ij} - a_{ji}\big)$ with $h = (1 + 1)^{-1}$. If $1 + 1 = 0$, then $-a = a$, so $K = S$ and $S + K = S$, which misses the matrix with a single 1 in position $(1, 2)$.

(a) The zero matrix satisfies the conditions defining N and K. If $A, B \in N$ and $c \in F$, then for $i \geq j$, $a_{ij} + b_{ij} = 0$ and $ca_{ij} = 0$, so $A + B, cA \in N$. If $A, B \in K$, then $a_{ij} + b_{ij} = -a_{ji} - b_{ji} = -\big(a_{ji} + b_{ji}\big)$ and $ca_{ij} = -\big(ca_{ji}\big)$, so $A + B, cA \in K$. So N and K are subspaces.

Let $A \in S \cap N$. For $i \geq j$, $a_{ij} = 0$ because $A \in N$. For $i < j$, $a_{ij} = a_{ji}$ because $A \in S$, and $a_{ji} = 0$ because $j > i$. So $A = \mathbf{0}$.

Let $A \in F^{n\times n}$, and define B by $b_{ij} = a_{ij}$ if $i \geq j$ and $b_{ij} = a_{ji}$ if $i < j$. Then B is symmetric: for $i > j$, $b_{ij} = a_{ij}$, and $b_{ji} = a_{ij}$ since $j < i$. Let $C = A - B$. For $i \geq j$, $c_{ij} = a_{ij} - a_{ij} = 0$, so $C \in N$. Hence $A = B + C \in S + N$, and $S + N = F^{n\times n}$.

(b) Suppose $1 + 1 \neq 0$, and let $h = (1 + 1)^{-1}$. For $A \in F^{n\times n}$, define B and C by

$$b_{ij} = h\big(a_{ij} + a_{ji}\big), \quad c_{ij} = h\big(a_{ij} - a_{ji}\big).$$

Then $b_{ji} = b_{ij}$ and $c_{ji} = -c_{ij}$, so $B \in S$ and $C \in K$, and $b_{ij} + c_{ij} = h(1 + 1)a_{ij} = a_{ij}$. So $A = B + C$, and $S + K = F^{n\times n}$.

Suppose $1 + 1 = 0$. Then $-a = a$ for all $a \in F$, so the condition $a_{ij} = -a_{ji}$ is the condition $a_{ij} = a_{ji}$, and $K = S$. Hence $S + K = S$ by Proposition 1.4.24. Since $n \geq 2$, the matrix E with 1 in position $(1, 2)$ and 0 elsewhere is not symmetric ($e_{12} = 1 \neq 0 = e_{21}$), so $S + K \neq F^{n\times n}$.

Solution to Exercise 1.4.37

Uses. Proposition 1.4.24 (The sum of two subspaces), Proposition 1.4.16 (Intersections of subspaces)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Double containment (Technique 1.3.14)
Hint 1. Look for counterexamples among $\{\mathbf{0}\}$, $\mathbb{R}^2$, and lines in $\mathbb{R}^2$. For (b), use three different lines.

Hint 2. For (c), one containment follows from Proposition 1.4.24(b). For the other, write $\mathbf{v} \in U \cap (W + X)$ as $\mathbf{w} + \mathbf{x}$ and show that $\mathbf{x} \in U$.
Sketch. (a) False: $U = \mathbb{R}^2$, $W = \{\mathbf{0}\}$, $X = \mathbb{R}^2$. (b) False: with the lines $U = \operatorname{span}((1,1))$, $W = \operatorname{span}(\mathbf{e}_1)$, $X = \operatorname{span}(\mathbf{e}_2)$, the left side is U and the right side is $\{\mathbf{0}\}$. (c) True. W and $U \cap X$ lie in $U \cap (W + X)$, which gives $\supseteq$. If $\mathbf{v} = \mathbf{w} + \mathbf{x} \in U$, then $\mathbf{x} = \mathbf{v} - \mathbf{w} \in U$ because $\mathbf{w} \in W \subseteq U$, so $\mathbf{v} \in W + (U \cap X)$.

(a) False. In $V = \mathbb{R}^2$, take $U = \mathbb{R}^2$, $W = \{\mathbf{0}\}$, and $X = \mathbb{R}^2$. Then $U + W = \mathbb{R}^2 = U + X$, but $W \neq X$.

(b) False. In $\mathbb{R}^2$, take $U = \operatorname{span}((1,1))$, $W = \operatorname{span}(\mathbf{e}_1)$, and $X = \operatorname{span}(\mathbf{e}_2)$. Then $W + X = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) = \mathbb{R}^2$ (Exercise 1.3.36), so the left side is U. On the other hand, $U \cap W = \{\mathbf{0}\}$: if $(t,t) = (s,0)$ then $t = 0$. Likewise $U \cap X = \{\mathbf{0}\}$. So the right side is $\{\mathbf{0}\} + \{\mathbf{0}\} = \{\mathbf{0}\} \neq U$. (The containment $\supseteq$ always holds: $U \cap W$ lies in U and in $W \subseteq W + X$, likewise $U \cap X$, and Proposition 1.4.24(b) applies.)

(c) True. We proceed by double containment.

$\supseteq$. $W \subseteq U$ by hypothesis and $W \subseteq W + X$ by Proposition 1.4.24(a), so $W \subseteq U \cap (W + X)$. Also $U \cap X \subseteq U$, and $U \cap X \subseteq X \subseteq W + X$, so $U \cap X \subseteq U \cap (W + X)$. The set $U \cap (W + X)$ is a subspace by Proposition 1.4.24(a) and Proposition 1.4.16, so it contains $W + (U \cap X)$ by Proposition 1.4.24(b).

$\subseteq$. Let $\mathbf{v} \in U \cap (W + X)$, and write $\mathbf{v} = \mathbf{w} + \mathbf{x}$ with $\mathbf{w} \in W$ and $\mathbf{x} \in X$. Since $\mathbf{w} \in W \subseteq U$ and $\mathbf{v} \in U$, we get $\mathbf{x} = \mathbf{v} - \mathbf{w} \in U$. So $\mathbf{x} \in U \cap X$, and $\mathbf{v} = \mathbf{w} + \mathbf{x} \in W + (U \cap X)$.

Statement (c) is called the modular law. The counterexample in (b) shows that the hypothesis $W \subseteq U$ cannot simply be dropped.

Solution to Exercise 1.4.38

Uses. Lemma 1.4.19 (Adding a vector outside a subspace), Theorem 1.4.3 (Subspace criterion)
Hint 1. Let X be a subspace containing $V \setminus U$. You need to show that $U \subseteq X$ too. Fix a vector $\mathbf{v}$ outside U.
Hint 2. For $\mathbf{u} \in U$, where does $\mathbf{u} + \mathbf{v}$ lie? Then write $\mathbf{u}$ in terms of $\mathbf{u} + \mathbf{v}$ and $\mathbf{v}$.
Sketch. Let X contain $V \setminus U$, and fix $\mathbf{v} \notin U$. For $\mathbf{u} \in U$, $\mathbf{u} + \mathbf{v} \notin U$ by Lemma 1.4.19, so $\mathbf{u} + \mathbf{v} \in X$; also $\mathbf{v} \in X$, so $\mathbf{u} = (\mathbf{u} + \mathbf{v}) - \mathbf{v} \in X$. Hence $X \supseteq U \cup (V \setminus U) = V$.

Let X be a subspace of V that contains every vector of $V \setminus U$. Since $U \neq V$, we may fix $\mathbf{v} \in V$ with $\mathbf{v} \notin U$; then $\mathbf{v} \in X$. Let $\mathbf{u} \in U$. By Lemma 1.4.19, $\mathbf{u} + \mathbf{v} \notin U$, so $\mathbf{u} + \mathbf{v} \in X$. Hence

$$\mathbf{u} = (\mathbf{u} + \mathbf{v}) + (-1)\mathbf{v} \in X$$

by (II) of Theorem 1.4.3. So $U \subseteq X$, and with $V \setminus U \subseteq X$ this gives $X = V$.

In other words, the vectors outside a proper subspace generate everything: every $\mathbf{u} \in U$ is the difference of two vectors outside U.

Solution to Exercise 1.4.39

Uses. Exercise 1.3.33, Proposition 1.4.14 (The span is the smallest subspace containing a list), Example 1.4.23 ($F[x]$ as a union of a chain), Exercise 1.4.27
Hint 1. Choose $p_k \in U$ of degree k for each k. What is $\operatorname{span}(p_0, \ldots, p_n)$? See Exercise 1.3.33.
Hint 2. Every polynomial lies in some $F[x]_{\leq n}$. For (b), consider the even polynomials.
Sketch. (a) With $\deg p_k = k$, $F[x]_{\leq n} = \operatorname{span}(p_0, \ldots, p_n) \subseteq U$ for every n, and every polynomial lies in some $F[x]_{\leq n}$. (b) The polynomials with $a_k = 0$ for all odd k: a subspace containing x^{2k} for every k, but not x.

(a) For each $k \geq 0$, choose $p_k \in U$ with $\deg p_k = k$. Let $n \geq 0$. By Exercise 1.3.33(a), $\operatorname{span}(p_0, \ldots, p_n) = F[x]_{\leq n}$, and since $p_0, \ldots, p_n \in U$, Proposition 1.4.14(b) gives $F[x]_{\leq n} \subseteq U$. Every polynomial lies in $F[x]_{\leq n}$ for some n: take $n = \deg p$ if $p \neq \mathbf{0}$ and $n = 0$ if $p = \mathbf{0}$. Hence $F[x] = \bigcup_{n \geq 0} F[x]_{\leq n} \subseteq U$.

(b) Let U be the set of polynomials whose coefficients a_k vanish for every odd k. It is a subspace (the argument of Exercise 1.4.27(e) works over every field), it contains x^{2k} for every $k \geq 0$, and it does not contain x. So it is proper. Part (a) shows that it must miss some degree, and it misses all the odd ones.

Solution to Exercise 1.4.40

Uses. Theorem 1.4.3 (Subspace criterion), Definition 1.2.9 (Function spaces)

Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (b), the trouble is addition: two functions with periods that have no common multiple. You need a sum whose lack of a period you can actually prove.
Hint 2. Let g be 1 on $\mathbb{Q}$ and 0 elsewhere, and let h be 1 on $\sqrt{2}\mathbb{Q} = \{\sqrt{2}r \mid r \in \mathbb{Q}\}$ and 0 elsewhere. Find periods of g and h.
Hint 3. At which points does $g + h$ take the value 2? A period would move one such point to another.
Sketch. (a) The condition $f(t+1) = f(t)$ survives sums and multiples. (b) No. $g = 1$ on $\mathbb{Q}$ has period 1, and $h = 1$ on $\sqrt{2}\mathbb{Q}$ has period $\sqrt{2}$. Since $\mathbb{Q} \cap \sqrt{2}\mathbb{Q} = \{0\}$, the function $g + h$ equals 2 at 0 and at most 1 elsewhere. If $p > 0$ were a period, then $(g+h)(p) = (g+h)(0) = 2$, forcing $p = 0$.

(a) The zero function has period 1. If $f(t+1) = f(t)$ and $g(t+1) = g(t)$ for all t, then $(f+g)(t+1) = f(t) + g(t) = (f+g)(t)$ and $(af)(t+1) = af(t) = (af)(t)$.

(b) No: the periodic functions are not closed under addition. Let $g(t) = 1$ if $t \in \mathbb{Q}$ and $g(t) = 0$ otherwise, and let $h(t) = 1$ if $t \in \sqrt{2}\mathbb{Q}$ and $h(t) = 0$ otherwise, where $\sqrt{2}\mathbb{Q} = \{\sqrt{2}r \mid r \in \mathbb{Q}\}$.

The function g has period 1, since $t + 1 \in \mathbb{Q}$ if and only if $t \in \mathbb{Q}$. The function h has period $\sqrt{2}$, since $t + \sqrt{2} = \sqrt{2}r$ with $r \in \mathbb{Q}$ if and only if $t = \sqrt{2}(r-1)$. So g and h are periodic.

Next, $\mathbb{Q} \cap \sqrt{2}\mathbb{Q} = \{0\}$. Indeed, if $r = \sqrt{2}s$ with $r, s \in \mathbb{Q}$ and $s \neq 0$, then $\sqrt{2} = r/s$ would be rational; so $s = 0$ and $r = 0$. Hence $(g+h)(0) = 2$, while $(g+h)(t) \leq 1$ for every $t \neq 0$, because t lies in at most one of $\mathbb{Q}$ and $\sqrt{2}\mathbb{Q}$.

Suppose $g + h$ had a period $p > 0$. Then $(g+h)(p) = (g+h)(0+p) = (g+h)(0) = 2$, so $p = 0$, a contradiction. So $g + h$ is not periodic, although g and h are.

The same happens for many nicer functions, such as $\sin t + \sin\left(\sqrt{2}t\right)$, but proving that this one has no period takes more work. The indicator functions are chosen so that the proof is short.

Solution to Exercise 1.4.41

Uses. Proposition 1.4.14 (The span is the smallest subspace containing a list), Example 1.4.8 (One equation in F^n), Exercise 1.2.45
Hint 1. The four subspaces are symmetric. For the converse, split into two cases: every vector of U has all coordinates equal, or some vector $\mathbf{x} \in U$ has $x_i \neq x_j$.
Hint 2. In the second case, subtract from $\mathbf{x}$ the vector obtained by swapping coordinates i and j. Then move the result around with other permutations to get every $\mathbf{e}_k - \mathbf{e}_l$ into U.
Hint 3. The vectors $\mathbf{e}_k - \mathbf{e}_n$ span Z. If U is bigger than Z, pick $\mathbf{u} \in U$ with nonzero coordinate sum and write any $\mathbf{y}$ as an element of Z plus a multiple of $\mathbf{u}$.
Sketch. If all vectors of U are constant, $U \subseteq \operatorname{span}(\mathbf{s})$, so $U = \{\mathbf{0}\}$ or $\operatorname{span}(\mathbf{s})$. Otherwise some $\mathbf{x} \in U$ has $x_i \neq x_j$; subtracting its swap gives $(x_i - x_j)(\mathbf{e}_i - \mathbf{e}_j) \in U$, and permuting gives every $\mathbf{e}_k - \mathbf{e}_l \in U$, so $Z = \operatorname{span}(\mathbf{e}_1 - \mathbf{e}_n, \ldots, \mathbf{e}_{n-1} - \mathbf{e}_n) \subseteq U$. If some $\mathbf{u} \in U$ has coordinate sum $\sigma \neq 0$, then $\mathbf{y} = (\mathbf{y} - (\tau/\sigma)\mathbf{u}) + (\tau/\sigma)\mathbf{u} \in U$, where τ is the coordinate sum of $\mathbf{y}$. (b) Exactly when $n * 1 = 0$ in F.

For a bijection π of $\{1, \ldots, n\}$ and $\mathbf{x} \in F^n$, write $\pi\mathbf{x} = \left(x_{\pi(1)}, \ldots, x_{\pi(n)}\right)$; so U is symmetric if $\pi\mathbf{x} \in U$ for all $\mathbf{x} \in U$ and all π. The sum of the coordinates of $\mathbf{x}$ is denoted $\Sigma(\mathbf{x})$; it does not change under permutations.

(a) The four subspaces are symmetric: for $\{\mathbf{0}\}$ and F^n there is nothing to check, $\pi(c\mathbf{s}) = c\mathbf{s}$, and $\Sigma(\pi\mathbf{x}) = \Sigma(\mathbf{x})$. Note that Z is a subspace by Example 1.4.8. Now let U be a symmetric subspace.

Case 1: every $\mathbf{x} \in U$ has all coordinates equal. Then $U \subseteq \operatorname{span}(\mathbf{s})$. If $U \neq \{\mathbf{0}\}$, it contains some $c\mathbf{s}$ with $c \neq 0$, hence $\mathbf{s} = c^{-1}(c\mathbf{s})$, hence $\operatorname{span}(\mathbf{s})$ by Proposition 1.4.14(b). So $U = \{\mathbf{0}\}$ or $U = \operatorname{span}(\mathbf{s})$.

Case 2: some $\mathbf{x} \in U$ has $x_i \neq x_j$ for some $i \neq j$. Let $\mathbf{y}$ be $\mathbf{x}$ with coordinates i and j swapped. Then $\mathbf{y} \in U$, and $\mathbf{x} - \mathbf{y}$ has i-th coordinate $x_i - x_j$, j-th coordinate $x_j - x_i$, and all others 0:

$$\mathbf{x} - \mathbf{y} = (x_i - x_j)(\mathbf{e}_i - \mathbf{e}_j) \in U.$$

Multiplying by $(x_i - x_j)^{-1}$ gives $\mathbf{e}_i - \mathbf{e}_j \in U$. Now let $k \neq l$, and choose a bijection π of $\{1, \ldots, n\}$ with $\pi(k) = i$ and $\pi(l) = j$. Then $\pi(\mathbf{e}_i - \mathbf{e}_j)$ has k-th coordinate 1, l-th coordinate -1, and all others 0; that is, $\mathbf{e}_k - \mathbf{e}_l \in U$.

Every $\mathbf{z} \in Z$ satisfies $z_n = -(z_1 + \cdots + z_{n-1})$, so

$$\mathbf{z} = z_1(\mathbf{e}_1 - \mathbf{e}_n) + \cdots + z_{n-1}(\mathbf{e}_{n-1} - \mathbf{e}_n),$$

as both sides have the same first $n-1$ coordinates and the same n-th coordinate. So $Z \subseteq \operatorname{span}(\mathbf{e}_1 - \mathbf{e}_n, \ldots, \mathbf{e}_{n-1} - \mathbf{e}_n) \subseteq U$ by Proposition 1.4.14(b). If $U = Z$, we are done. Otherwise there is $\mathbf{u} \in U$ with $\sigma = \Sigma(\mathbf{u}) \neq 0$. For any $\mathbf{y} \in F^n$, let $\tau = \Sigma(\mathbf{y})$. Then $\Sigma(\mathbf{y} - (\tau/\sigma)\mathbf{u}) = \tau - (\tau/\sigma)\sigma = 0$, so

$$\mathbf{y} = (\mathbf{y} - (\tau/\sigma)\mathbf{u}) + (\tau/\sigma)\mathbf{u} \in Z + U \subseteq U.$$

Hence $U = F^n$.

(b) $\operatorname{span}(\mathbf{s}) \subseteq Z$ if and only if $\mathbf{s} \in Z$, that is, $\Sigma(\mathbf{s}) = n * 1 = 0$ in F (in the notation of Exercise 1.2.45). This happens exactly when $\operatorname{char} F$ divides n; over $\mathbb{F}_2$ with n even, for example. In that case the four subspaces form a chain $\{\mathbf{0}\} \subseteq \operatorname{span}(\mathbf{s}) \subseteq Z \subseteq F^n$. For $F = \mathbb{F}_2$ and $n = 2$, $\operatorname{span}(\mathbf{s}) = Z = \{\mathbf{0}, (1,1)\}$, so there are only three symmetric subspaces.

Solution to Exercise 1.4.42

Uses. Proposition 1.4.20 (Unions of two subspaces), Lemma 1.4.19 (Adding a vector outside a subspace), Example 1.4.12 (The subspaces of $\mathbb{F}_2^2$), Corollary 1.2.36 (Cancelling scalars and vectors)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. For (a), use the lines of Example 1.4.12. For (b), let $U = U_1 \cup U_2 \cup U_3$, and suppose no U_i equals U. First show that U_1 is not contained in $U_2 \cup U_3$.
Hint 2. Pick $\mathbf{u} \in U_1$ outside $U_2 \cup U_3$ and $\mathbf{v} \in U$ outside U_1. Consider the vectors $\mathbf{v} + \lambda\mathbf{u}$ for $\lambda \in F$. Which of the U_i can contain them?
Hint 3. None lies in U_1, and U_2 contains at most one of them: if it contains two, it contains their difference, a nonzero multiple of $\mathbf{u}$.
Sketch. (a) The three lines of $\mathbb{F}_2^2$. (b) Suppose no U_i is all of $U = U_1 \cup U_2 \cup U_3$. If $U_1 \subseteq U_2 \cup U_3$, then $U = U_2 \cup U_3$, and Proposition 1.4.20 gives a contradiction; so pick $\mathbf{u} \in U_1$ outside $U_2 \cup U_3$, and $\mathbf{v} \in U \setminus U_1$. The vectors $\mathbf{v} + \lambda\mathbf{u}$, $\lambda \in F$, are at least three distinct vectors of U, none in U_1 (Lemma 1.4.19), and at most one in each of U_2, U_3. Contradiction.

(a) By Example 1.4.12, the lines $\{\mathbf{0}, (1,0)\}$, $\{\mathbf{0}, (0,1)\}$, $\{\mathbf{0}, (1,1)\}$ are subspaces of $\mathbb{F}_2^2$, and their union is $\mathbb{F}_2^2$. Each contains exactly one nonzero vector, and these are different, so no line contains another.

(b) Let $U = U_1 \cup U_2 \cup U_3$, a subspace by hypothesis. One of the U_i contains the other two if and only if it equals U, so suppose, to the contrary, that $U_i \neq U$ for each i.

First, $U_1 \not\subseteq U_2 \cup U_3$. Otherwise $U = U_2 \cup U_3$ is a subspace, so by Proposition 1.4.20 one of U_2, U_3 contains the other and equals U, contrary to our assumption. So we may pick $\mathbf{u} \in U_1$ with $\mathbf{u} \notin U_2$ and $\mathbf{u} \notin U_3$. In particular $\mathbf{u} \neq \mathbf{0}$. Since $U_1 \neq U$, we may also pick $\mathbf{v} \in U$ with $\mathbf{v} \notin U_1$.

For $\lambda \in F$, the vector $\mathbf{v} + \lambda\mathbf{u}$ lies in the subspace U. It does not lie in U_1, by Lemma 1.4.19, since $\lambda\mathbf{u} \in U_1$ and $\mathbf{v} \notin U_1$. Suppose U_2 contained $\mathbf{v} + \lambda\mathbf{u}$ and $\mathbf{v} + \mu\mathbf{u}$ with $\lambda \neq \mu$. Then it would contain their difference $(\lambda - \mu)\mathbf{u}$, and hence $\mathbf{u} = (\lambda - \mu)^{-1}(\lambda - \mu)\mathbf{u}$, which is false. So U_2 contains at most one of the vectors $\mathbf{v} + \lambda\mathbf{u}$, and likewise U_3. These vectors are distinct for distinct λ, by Corollary 1.2.36(b), since $\mathbf{u} \neq \mathbf{0}$, and F has at least three elements. So at least three distinct vectors of U lie in $U_2 \cup U_3$, which contains at most two of them, and none of them lies in U_1. This contradicts $U = U_1 \cup U_2 \cup U_3$.

Part (a) shows that the hypothesis on F is needed. The proof needs three values of λ to outnumber U_2 and U_3; over $\mathbb{F}_2$ there are only two.

Solution to Exercise 1.4.43

Uses. Lemma 1.4.19 (Adding a vector outside a subspace), Exercise 1.4.33, Corollary 1.2.36 (Cancelling scalars and vectors)
Hint 1. Suppose $V = U_1 \cup \cdots \cup U_n$, and choose such a covering with n as small as possible. Why is U_1 then not contained in $U_2 \cup \cdots \cup U_n$?
Hint 2. Pick $\mathbf{u} \in U_1$ outside the other U_j, and $\mathbf{v} \notin U_1$. Count how many of the vectors $\mathbf{v} + \lambda\mathbf{u}$, $\lambda \in F$, each U_j can contain.
Sketch. (a) Take a counterexample with n minimal. Then $n \geq 2$ and $U_1 \not\subseteq U_2 \cup \cdots \cup U_n$; pick $\mathbf{u}$ in U_1 only, and $\mathbf{v} \notin U_1$. The $|F| \geq n$ distinct vectors $\mathbf{v} + \lambda\mathbf{u}$ avoid U_1, and each U_j with $j \geq 2$ contains at most one of them, so at most $n-1$ are covered. (b) An infinite field has at least n elements for every n. (c) F^2 is the union of its $q+1$ lines.

(a) Suppose the statement fails, and among all coverings $V = U_1 \cup \cdots \cup U_n$ of V by proper subspaces with $n \le |F|$, choose one with n as small as possible. Then $n \ge 2$, since $U_1 \ne V$. If $U_1 \subseteq U_2 \cup \cdots \cup U_n$, then $V = U_2 \cup \cdots \cup U_n$ is a covering by $n-1 \le |F|$ proper subspaces, contradicting minimality. So we may pick $\mathbf{u} \in U_1$ with $\mathbf{u} \notin U_j$ for all $j \ge 2$; in particular $\mathbf{u} \ne \mathbf{0}$. Since $U_1 \ne V$, we may also pick $\mathbf{v} \in V$ with $\mathbf{v} \notin U_1$.

Consider the vectors $\mathbf{v} + \lambda\mathbf{u}$ for $\lambda \in F$. They are distinct, by Corollary 1.2.36(b), since $\mathbf{u} \ne \mathbf{0}$, so there are $|F| \ge n$ of them. None lies in U_1, by Lemma 1.4.19, since $\lambda\mathbf{u} \in U_1$ and $\mathbf{v} \notin U_1$. For $j \ge 2$, U_j contains at most one of them: if $\mathbf{v} + \lambda\mathbf{u}$ and $\mathbf{v} + \mu\mathbf{u}$ lie in U_j with $\lambda \ne \mu$, then so does $(\lambda - \mu)^{-1}((\mathbf{v} + \lambda\mathbf{u}) - (\mathbf{v} + \mu\mathbf{u})) = \mathbf{u}$, which is false. So at most $n-1$ of these at least n vectors lie in $U_1 \cup \cdots \cup U_n = V$, a contradiction.

(b) If F is infinite, it has at least n elements for every n, so (a) applies to every finite collection of proper subspaces.

(c) By Exercise 1.4.33, F^2 has $q+1$ lines, each a proper subspace with $q < q^2$ elements. Every nonzero $\mathbf{v} \in F^2$ lies on the line $\operatorname{span}(\mathbf{v})$, and $\mathbf{0}$ lies on every line, so F^2 is the union of the $q+1$ lines. Since $q+1 > q$, this does not contradict (a), and it shows that the bound $n \le |F|$ in (a) cannot be relaxed to $n \le |F| + 1$.

Solution to Exercise 1.4.44

> *Uses.* Proposition 1.4.14 (The span is the smallest subspace containing a list), Example 1.4.8 (One equation in F^n), Corollary 1.3.19 (Equal spans)
> *Techniques.* Double containment (Technique 1.3.14)
> *Hint 1.* By Proposition 1.4.14, the smallest subspace containing a list is its span.
> *Hint 2.* For (a), both vectors satisfy $x_3 = 0$. For the reverse containment, solve $a + b = x_1$ and $a - b = x_2$. For (b), add and subtract the two polynomials.
> *Sketch.* (a) $\operatorname{span}((1,1,0),(1,-1,0)) = \{x_3 = 0\}$: the plane is a subspace containing both vectors, and $(x_1, x_2, 0) = \frac{x_1+x_2}{2}(1,1,0) + \frac{x_1-x_2}{2}(1,-1,0)$. (b) $\operatorname{span}(1+x^2, 1-x^2) = \operatorname{span}(1, x^2) = \{a + bx^2 \mid a, b \in \mathbb{R}\}$.

By Proposition 1.4.14, the smallest subspace containing the entries of a list is the span of the list.

(a) Let $P = \{\mathbf{x} \in \mathbb{R}^3 \mid x_3 = 0\}$, a subspace by Example 1.4.8. Both vectors lie in P, so $\operatorname{span}((1,1,0),(1,-1,0)) \subseteq P$ by Proposition 1.4.14(b). Conversely, for $(x_1, x_2, 0) \in P$,

$$\frac{x_1 + x_2}{2}(1,1,0) + \frac{x_1 - x_2}{2}(1,-1,0) = (x_1, x_2, 0).$$

So the smallest subspace is the plane $x_3 = 0$.

(b) Since $1 = \frac{1}{2}\big((1+x^2) + (1-x^2)\big)$ and $x^2 = \frac{1}{2}\big((1+x^2) - (1-x^2)\big)$, and conversely $1 \pm x^2$ are combinations of 1 and x^2, Corollary 1.3.19 gives

$$\operatorname{span}(1+x^2, 1-x^2) = \operatorname{span}(1, x^2) = \{a + bx^2 \mid a, b \in \mathbb{R}\}.$$

Section 1.5

Solution to Exercise 1.5.8

> *Uses.* Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.1.17 (No zero divisors)
> *Hint 1.* Expand $a(\mathbf{v} + \mathbf{w}) + b(\mathbf{v} - \mathbf{w})$ as a linear combination of $\mathbf{v}, \mathbf{w}$, and use that $\mathbf{v}, \mathbf{w}$ has only the trivial relation.
> *Hint 2.* You get $a + b = 0$ and $a - b = 0$. Add them. When does $(1+1)a = 0$ force $a = 0$?
> *Sketch.* Expanding, the combination with coefficients a, b of the new list is the combination with coefficients $a + b$, $a - b$ of $\mathbf{v}, \mathbf{w}$. So a relation of the new list gives $a + b = a - b = 0$, hence $(1+1)a = 0$ and $(1+1)b = 0$. If $1 + 1 \ne 0$, then $a = b = 0$. If $1 + 1 = 0$, then $(\mathbf{v} + \mathbf{w}) + (\mathbf{v} - \mathbf{w}) = (1+1)\mathbf{v} = \mathbf{0}$ is a dependence relation.

By Lemma 1.3.9, for $a, b \in F$,

$$a(\mathbf{v} + \mathbf{w}) + b(\mathbf{v} - \mathbf{w}) = (a + b)\mathbf{v} + (a - b)\mathbf{w}.$$

Suppose that $1+1 \neq 0$, and that the left side is $\mathbf{0}$. Since $\mathbf{v}, \mathbf{w}$ is independent, Theorem 1.5.5 gives $a+b=0$ and $a-b=0$. Adding and subtracting these equations gives $(1+1)a=0$ and $(1+1)b=0$, so $a=b=0$ by Proposition 1.1.17. Hence $\mathbf{v}+\mathbf{w}, \mathbf{v}-\mathbf{w}$ is independent.

Suppose that $1+1=0$. Taking $a=b=1$, the right side is $(1+1)\mathbf{v}+0\mathbf{w}=\mathbf{0}$. So $1,1$ is a dependence relation, and the list is dependent. In fact $-1=1$, so $\mathbf{v}-\mathbf{w}=\mathbf{v}+\mathbf{w}$, and the list repeats an entry.

Solution to Exercise 1.5.12

Uses. Lemma 1.5.10 (Linear dependence lemma), Definition 1.3.10 (Span), Corollary 1.3.20 (Operations that preserve the span)
Hint 1. Go through $j=1,\ldots,5$ in order. For $j=2$, compare $\mathbf{v}_2$ with $\mathbf{v}_1$.
Hint 2. Once you know that $\mathbf{v}_2$ is redundant, $\operatorname{span}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)=\operatorname{span}(\mathbf{v}_1,\mathbf{v}_3)$, whose elements are $(a,a+b,b)$. Test $\mathbf{v}_4$ and $\mathbf{v}_5$ against this description.
Sketch. $j=2$, since $\mathbf{v}_2=2\mathbf{v}_1$, and $j=4$, since $\mathbf{v}_4=\mathbf{v}_1+\mathbf{v}_3$, with relations $2\mathbf{v}_1-\mathbf{v}_2=\mathbf{0}$ and $\mathbf{v}_1+\mathbf{v}_3-\mathbf{v}_4=\mathbf{0}$. No other j works: $\mathbf{v}_1\neq\mathbf{0}$, $\mathbf{v}_3$ is not a multiple of $\mathbf{v}_1$, and $\operatorname{span}(\mathbf{v}_1,\ldots,\mathbf{v}_4)=\{(a,a+b,b)\}$ does not contain $(0,0,1)$.

$j=1$: $\mathbf{v}_1\neq\mathbf{0}$, so $\mathbf{v}_1\notin\operatorname{span}(())=\{\mathbf{0}\}$.

$j=2$: $\mathbf{v}_2=2\mathbf{v}_1$. The relation $2\mathbf{v}_1+(-1)\mathbf{v}_2=\mathbf{0}$, with zero coefficients on the other entries, has largest nonzero index 2.

$j=3$: By Corollary 1.3.20(c), $\operatorname{span}(\mathbf{v}_1,\mathbf{v}_2)=\operatorname{span}(\mathbf{v}_1)$, whose elements $(a,a,0)$ have third coordinate 0. So $\mathbf{v}_3=(0,1,1)$ is not in it.

$j=4$: $\mathbf{v}_1+\mathbf{v}_3=(1,2,1)=\mathbf{v}_4$. The relation $\mathbf{v}_1+\mathbf{v}_3+(-1)\mathbf{v}_4=\mathbf{0}$ has largest nonzero index 4.

$j=5$: By Corollary 1.3.20(a) and (c), deleting $\mathbf{v}_2$ and $\mathbf{v}_4$ shows $\operatorname{span}(\mathbf{v}_1,\ldots,\mathbf{v}_4)=\operatorname{span}(\mathbf{v}_1,\mathbf{v}_3)$, which consists of the vectors $a(1,1,0)+b(0,1,1)=(a,a+b,b)$. For $(0,0,1)$ we would need $a=0$, $b=1$, and $a+b=0$, which is impossible. So $\mathbf{v}_5\notin\operatorname{span}(\mathbf{v}_1,\ldots,\mathbf{v}_4)$.

Hence $j=2$ and $j=4$ are the indices that work.

Solution to Exercise 1.5.16

Uses. Proposition 1.5.15 (Short lists, zeros, and repeats), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Which step turns a linear combination of two vectors into a multiple of one of them?
Hint 2. Test the claim on $(1,0),(0,1),(1,1)$ in $\mathbb{R}^2$.
Sketch. A combination of $\mathbf{v}_2$ and $\mathbf{v}_3$ with both coefficients nonzero need not be a multiple of either. The list $(1,0),(0,1),(1,1)$ satisfies the hypothesis but has the relation $(1,0)+(0,1)-(1,1)=\mathbf{0}$.

The error is the step "so $\mathbf{v}_1$ is a scalar multiple of $\mathbf{v}_2$ or of $\mathbf{v}_3$". It holds when $a_2=0$ or $a_3=0$, but a linear combination of $\mathbf{v}_2$ and $\mathbf{v}_3$ with both coefficients nonzero is in general a multiple of neither.

The claim is false. In $\mathbb{R}^2$, no two of $(1,0)$, $(0,1)$, $(1,1)$ are multiples of each other: their coordinates show that $(0,1)$ and $(1,1)$ are not multiples of $(1,0)$, and so on. But $(1,0)+(0,1)+(-1)(1,1)=\mathbf{0}$ is a dependence relation. For two vectors the claim is correct (Proposition 1.5.15(d)), and the argument above is a valid proof in that case.

Solution to Exercise 1.5.22

Uses. Proposition 1.5.15 (Short lists, zeros, and repeats), Theorem 1.5.5 (Independence is decided at zero), Definition 1.1.29 (Complex numbers), Proposition 1.5.21 (Real lists in $\mathbb{C}^m$)
Hint 1. Over $\mathbb{C}$, is the second vector a complex multiple of the first? Compare first coordinates to find the only candidate.
Hint 2. Over $\mathbb{R}$, write $a(1,1+i)+b(1-i,2)=\mathbf{0}$ with a,b real, and take the real and imaginary parts of the first coordinate.
Sketch. $(1-i)(1,1+i)=(1-i,2)$, so the list is dependent over $\mathbb{C}$. Over $\mathbb{R}$, the first coordinate $a+b-bi$ of a relation must vanish, which forces $b=0$ and $a=0$. No contradiction: the vectors are not real.

Over $\mathbb{C}$:

$$(1-i)(1,1+i)=(1-i,(1-i)(1+i))=\left(1-i,1-i^2\right)=(1-i,2).$$

So the second vector is a multiple of the first, and the list is dependent by Proposition 1.5.15(d).

Over $\mathbb{R}$: let $a, b \in \mathbb{R}$ with $a(1, 1+i) + b(1-i, 2) = \mathbf{0}$. The first coordinate is $a + b(1-i) = (a+b) + (-b)i$, which is 0 only if $a + b = 0$ and $-b = 0$, by Definition 1.1.29. So $a = b = 0$, and the list is independent over $\mathbb{R}$.

This does not contradict Proposition 1.5.21, which concerns lists of vectors with real coordinates. Here the coordinates $1 + i$ and $1 - i$ are not real.

Solution to Exercise 1.5.28

Uses. Theorem 1.5.5 (Independence is decided at zero), Lemma 1.5.10 (Linear dependence lemma), Proposition 1.5.15 (Short lists, zeros, and repeats)
Hint 1. Write $a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 = \mathbf{0}$ as three equations in a, b, c, one per coordinate, and eliminate.
Hint 2. In (b), the equations can be solved from the top down. In (c), add up the three vectors.
Sketch. (a) Dependent: $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$, and only $j = 3$ works. (b) Independent: the relation reads $(a, a+b, a+b+c) = \mathbf{0}$. (c) Dependent: $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$, and only $j = 3$ works.

(a) The relation $a(1, 2, 3) + b(4, 5, 6) + c(7, 8, 9) = \mathbf{0}$ reads

$$a + 4b + 7c = 0, \quad 2a + 5b + 8c = 0, \quad 3a + 6b + 9c = 0.$$

Subtracting twice the first equation from the second gives $-3b - 6c = 0$, so $b = -2c$, and then the first gives $a = -4b - 7c = c$. The third equation then holds: $3c - 12c + 9c = 0$. Taking $c = 1$ gives the dependence relation $\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$, so the list is dependent. Here $j = 1$ fails as $\mathbf{v}_1 \neq \mathbf{0}$, and $j = 2$ fails as $(4, 5, 6)$ is not a multiple of $(1, 2, 3)$ (the multiplier would have to be 4 and $5/2$). The relation gives $j = 3$: $\mathbf{v}_3 = -\mathbf{v}_1 + 2\mathbf{v}_2$.

(b) A relation reads

$$a(1, 1, 1) + b(0, 1, 1) + c(0, 0, 1) = (a, a+b, a+b+c) = \mathbf{0},$$

which gives $a = 0$, then $b = 0$, then $c = 0$. So the list is independent.

(c) The three vectors add up to $(1 + 0 - 1, -1 + 1 + 0, 0 - 1 + 1) = \mathbf{0}$, so 1, 1, 1 is a dependence relation. Here $\mathbf{v}_1 \neq \mathbf{0}$, and $\mathbf{v}_2 = (0, 1, -1)$ is not a multiple of $(1, -1, 0)$, since a multiple $t(1, -1, 0)$ with first coordinate 0 has $t = 0$. So only $j = 3$ works, with $\mathbf{v}_3 = -\mathbf{v}_1 - \mathbf{v}_2$.

Solution to Exercise 1.5.29

Uses. Theorem 1.5.5 (Independence is decided at zero), Proposition 1.1.17 (No zero divisors), Proposition 1.5.15 (Short lists, zeros, and repeats)
Hint 1. The relation $a(1, 1, 1) + b(1, 2, 3) + c(1, 4, 9) = \mathbf{0}$ gives one equation per coordinate. The t-th coordinates of the three vectors are $1, t, t^2$.
Hint 2. Eliminate a, then b. You should reach $2c = 0$. Which primes make $2 = 0$?
Sketch. The equations $a + b + c = 0$, $a + 2b + 4c = 0$, $a + 3b + 9c = 0$ reduce to $b + 3c = 0$ and $2c = 0$, using only operations valid in every field. If $2 \neq 0$ (every $p \neq 2$, and $\mathbb{Q}$), the list is independent. In $\mathbb{F}_2^3$ the list is $(1, 1, 1), (1, 0, 1), (1, 0, 1)$, which repeats an entry.

Reading the coordinates of $a(1, 1, 1) + b(1, 2, 3) + c(1, 4, 9) = \mathbf{0}$ one at a time gives

$$a + b + c = 0, \quad a + 2b + 4c = 0, \quad a + 3b + 9c = 0.$$

Subtracting the first equation from the other two gives $b + 3c = 0$ and $2b + 8c = 0$, and subtracting twice the first of these from the second gives $2c = 0$. These steps are valid over every field.

If $2 \neq 0$ in F, then $c = 0$ by Proposition 1.1.17, so $b = -3c = 0$ and $a = -b - c = 0$, and the list is independent by Theorem 1.5.5. This is the case for $F = \mathbb{Q}$ and $F = \mathbb{F}_p$ with p odd.

If $p = 2$, then $2 = 0$ and $4 = 0$ and $3 = 9 = 1$ in $\mathbb{F}_2$, so the list is $(1, 1, 1), (1, 0, 1), (1, 0, 1)$. It repeats an entry, so it is dependent (Proposition 1.5.15(c)).

So the list is independent in $\mathbb{F}_p^3$ exactly for $p \neq 2$, and it is independent in $\mathbb{Q}^3$.

Solution to Exercise 1.5.30

Uses. Lemma 1.5.10 (Linear dependence lemma), Proposition 1.5.14 (Appending a vector), Example 1.5.11 (Which entries are redundant)
Hint 1. Is the largest index with a nonzero coefficient necessarily k? What if $a_k = 0$?
Hint 2. For (b), compare with Proposition 1.5.14.
Sketch. (a) j is the largest index with $a_j \neq 0$, which is k only if $a_k \neq 0$. Counterexample: $(1,0),(2,0),(0,1)$ in $\mathbb{R}^2$. (b) This is the contrapositive of the backward direction of Proposition 1.5.14.

(a) The step "$j = k$" is wrong. The index j is the largest one with $a_j \neq 0$, and nothing prevents $a_k = 0$. The claim itself is false: in $\mathbb{R}^2$, the list $(1,0),(2,0),(0,1)$ is dependent (Example 1.5.11), but $(0,1) \notin \operatorname{span}((1,0),(2,0))$, since every vector of that span has second coordinate 0.

(b) Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ is independent and that $\mathbf{v}_k \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is independent by Proposition 1.5.14. Taking the contrapositive: if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is dependent, then $\mathbf{v}_k \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$. In terms of the argument in (a): if $a_k = 0$, the relation would be a nontrivial relation of the independent list $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$, which is impossible, so here $j = k$ is correct.

Solution to Exercise 1.5.31

Uses. Corollary 1.5.13 (Independence, one entry at a time), Corollary 1.3.20 (Operations that preserve the span)
Hint 1. By Corollary 1.3.20(b), $S_{j-1} \subseteq S_j$ always. When are they equal?
Hint 2. Corollary 1.3.20(c) says $S_j = S_{j-1}$ exactly when $\mathbf{v}_j \in S_{j-1}$. Now use Corollary 1.5.13.
Sketch. $S_{j-1} \subseteq S_j$, with equality if and only if $\mathbf{v}_j \in S_{j-1}$ (Corollary 1.3.20(c)). So the chain is strict exactly when $\mathbf{v}_j \notin S_{j-1}$ for all j, which is independence by Corollary 1.5.13.

For $1 \leq j \leq k$, the list $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ is a sublist of $\mathbf{v}_1, \ldots, \mathbf{v}_j$, so $S_{j-1} \subseteq S_j$ by Corollary 1.3.20(b). By Corollary 1.3.20(c), applied to the list $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ and the vector $\mathbf{v}_j$, we have $S_j = S_{j-1}$ if and only if $\mathbf{v}_j \in S_{j-1}$. So $S_{j-1} \subsetneq S_j$ if and only if $\mathbf{v}_j \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$. The chain is strict at every step if and only if this holds for every j, which by Corollary 1.5.13 is exactly the independence of $\mathcal{L}$.

Solution to Exercise 1.5.32

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.9 (Function spaces), Example 1.5.26 (Trigonometric functions)
Hint 1. For (a), recall a double-angle formula. For (b), evaluate a relation at points where the sines are simple, such as $\pi/2$, $\pi/3$, and $\pi/4$.
Hint 2. For (c), evaluate at $t = 1$ and $t = -1$. Which of these points is available in $\mathbb{R}^{[0,\infty)}$?
Sketch. (a) Dependent: $1 + \cos 2t - 2\cos^2 t = 0$. (b) Independent: evaluating $a \sin t + b \sin 2t + c \sin 3t = 0$ at $\pi/2$, $\pi/3$, $\pi/4$ gives $a = c$, $b = -a$, and $\left(\sqrt{2} - 1\right)a = 0$. (c) Independent in $\mathbb{R}^{\mathbb{R}}$ (evaluate at ± 1); dependent in $\mathbb{R}^{[0,\infty)}$, where $|t| = t$.

(a) Dependent. By the double-angle formula $\cos 2t = 2\cos^2 t - 1$, the relation with coefficients $1, 1, -2$ holds: $1 + \cos 2t - 2\cos^2 t = 0$ for every t.

(b) Independent. Suppose $a \sin t + b \sin 2t + c \sin 3t = 0$ for all t. At $t = \pi/2$: $\sin(\pi/2) = 1$, $\sin \pi = 0$, and $\sin(3\pi/2) = -1$, so $a - c = 0$. At $t = \pi/3$: $\sin(\pi/3) = \sin(2\pi/3) = \sqrt{3}/2$ and $\sin \pi = 0$, so $\left(\sqrt{3}/2\right)(a + b) = 0$ and $b = -a$. At $t = \pi/4$: $\sin(\pi/4) = \sin(3\pi/4) = \sqrt{2}/2$ and $\sin(\pi/2) = 1$, so

$$\frac{\sqrt{2}}{2}a + b + \frac{\sqrt{2}}{2}c = 0.$$

Substituting $b = -a$ and $c = a$ gives $\left(\sqrt{2} - 1\right)a = 0$, so $a = 0$, and then $b = c = 0$.

(c) In $\mathbb{R}^{\mathbb{R}}$, suppose $at + b|t| = 0$ for all t. At $t = 1$, $a + b = 0$; at $t = -1$, $-a + b = 0$. So $a = b = 0$, and the list is independent. In $\mathbb{R}^{[0,\infty)}$, the two functions are equal, since $|t| = t$ for $t \geq 0$, so the list repeats an entry and is dependent. The relation $1 \cdot t + (-1)|t| = 0$ holds on $[0,\infty)$ but not at $t = -1$: whether functions are independent depends on their domain.

Solution to Exercise 1.5.33

Uses. Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Exercise 1.3.34, Corollary 1.3.19 (Equal spans)
Hint 1. If $a_j = 0$, then $\mathbf{w}$ is a combination of the other entries of $\mathcal{L}'$. Write down a relation.

Hint 2. If $a_j \neq 0$, take a relation $\sum_{i\neq j} b_i\mathbf{v}_i + b_j\mathbf{w} = \mathbf{0}$, substitute for $\mathbf{w}$, and collect the coefficients of $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
Sketch. If $a_j = 0$: $\sum_{i\neq j} a_i\mathbf{v}_i - \mathbf{w} = \mathbf{0}$ is a nontrivial relation of $\mathcal{L}'$. If $a_j \neq 0$: a relation of $\mathcal{L}'$ becomes $\sum_{i\neq j}\big(b_i + b_j a_i\big)\mathbf{v}_i + b_j a_j\mathbf{v}_j = \mathbf{0}$, so $b_j a_j = 0$, whence $b_j = 0$, and then all $b_i = 0$. For the spans, solve for $\mathbf{v}_j$.

Suppose $a_j = 0$. Then $\mathbf{w} = \sum_{i\neq j} a_i\mathbf{v}_i$, and the linear combination of $\mathcal{L}'$ with coefficient a_i on $\mathbf{v}_i$ for $i \neq j$ and -1 on $\mathbf{w}$ represents $\mathbf{w} - \mathbf{w} = \mathbf{0}$. It is nontrivial, so $\mathcal{L}'$ is dependent.

Suppose $a_j \neq 0$. Let $\sum_{i\neq j} b_i\mathbf{v}_i + b_j\mathbf{w} = \mathbf{0}$. Substituting $\mathbf{w} = \sum_i a_i\mathbf{v}_i$ and collecting coefficients with Lemma 1.3.9,

$$\sum_{i\neq j}\big(b_i + b_j a_i\big)\mathbf{v}_i + \big(b_j a_j\big)\mathbf{v}_j = \mathbf{0}.$$

Since $\mathcal{L}$ is independent, all these coefficients are 0 (Theorem 1.5.5). From $b_j a_j = 0$ and $a_j \neq 0$ we get $b_j = 0$, and then $b_i = b_i + b_j a_i = 0$ for $i \neq j$. So $\mathcal{L}'$ is independent.

For the spans, $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, and since $a_j \neq 0$,

$$\mathbf{v}_j = a_j^{-1}\mathbf{w} + \sum_{i\neq j}\big(-a_j^{-1}a_i\big)\mathbf{v}_i \in \operatorname{span}(\mathcal{L}').$$

So each list lies in the span of the other, and $\operatorname{span}(\mathcal{L}') = \operatorname{span}(\mathcal{L})$ by Corollary 1.3.19. (Compare Exercise 1.3.34, which made the same exchange using only spans.)

Solution to Exercise 1.5.34

Uses. Proposition 1.5.15 (Short lists, zeros, and repeats), Example 1.5.11 (Which entries are redundant), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Look for counterexamples in $\mathbb{R}^2$ among $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_1 + \mathbf{e}_2$, and repeated vectors.
Hint 2. Exactly one statement is true. For it, write down a relation in which the last entry has coefficient -1.
Sketch. (a) False: $\mathbf{u} = \mathbf{w} = \mathbf{e}_1$, $\mathbf{v} = \mathbf{e}_2$. (b) False: $(1,0), (2,0), (0,1)$. (c) False: $\mathcal{L} = \mathcal{M} = \mathbf{e}_1$. (d) True: $\mathbf{v}_1 + \cdots + \mathbf{v}_k - (\mathbf{v}_1 + \cdots + \mathbf{v}_k) = \mathbf{0}$ has coefficient -1 on the last entry. (e) False: $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$.

Write $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ in $\mathbb{R}^2$.

(a) False. Take $\mathbf{u} = \mathbf{w} = \mathbf{e}_1$ and $\mathbf{v} = \mathbf{e}_2$. Then $\mathbf{u}, \mathbf{v}$ and $\mathbf{v}, \mathbf{w}$ are $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_2, \mathbf{e}_1$, which are independent (Example 1.5.2, and neither is a multiple of the other). But $\mathbf{u}, \mathbf{w} = \mathbf{e}_1, \mathbf{e}_1$ repeats an entry (Proposition 1.5.15(c)).

(b) False. The list $(1,0), (2,0), (0,1)$ is dependent, but $(0,1) \notin \operatorname{span}((1,0), (2,0))$ (Example 1.5.11).

(c) False. Take $\mathcal{L} = \mathcal{M} = \mathbf{e}_1$, which is independent by Proposition 1.5.15(b). The concatenation $\mathbf{e}_1, \mathbf{e}_1$ repeats an entry.

(d) True. The linear combination with coefficient 1 on each $\mathbf{v}_i$ and -1 on the last entry represents $(\mathbf{v}_1 + \cdots + \mathbf{v}_k) - (\mathbf{v}_1 + \cdots + \mathbf{v}_k) = \mathbf{0}$, and it is nontrivial. (Independence of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ was not needed. For $k = 0$ the list is the single vector $\mathbf{0}$, the empty sum.)

(e) False. Take $\mathbf{u} = \mathbf{e}_1$, $\mathbf{v} = \mathbf{e}_2$, and $\mathbf{w} = (1,1)$. Then $\mathbf{w}$ is on neither axis, but $\mathbf{u} + \mathbf{v} - \mathbf{w} = \mathbf{0}$. Avoiding the two lines spanned by $\mathbf{u}$ and $\mathbf{v}$ is much weaker than avoiding $\operatorname{span}(\mathbf{u}, \mathbf{v})$, which is what Proposition 1.5.14 requires.

Solution to Exercise 1.5.35

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 1.1.17 (No zero divisors), Proposition 1.5.15 (Short lists, zeros, and repeats)
Hint 1. Evaluate each p_i at 0, 1, and 2. Each polynomial vanishes at two of these points.
Hint 2. If $ap_1 + bp_2 + cp_3 = \mathbf{0}$, evaluate at 0, 1, 2 using Proposition 1.2.18(c). For the case $1 + 1 = 0$, rewrite the three polynomials with $2 = 0$ and $3 = 1$.
Sketch. $p_1(0) = 2$, $p_2(1) = -1$, $p_3(2) = 2$, and each p_i vanishes at the other two points. A relation evaluated at $0, 1, 2$ gives $2a = 0$, $-b = 0$, $2c = 0$, so if $2 \neq 0$ the list is independent. If $2 = 0$, then $p_1 = x^2 + x = p_3$, a repeated entry.

By Definition 1.2.17,

$$\begin{aligned} p_1(0) &= 2, \quad p_1(1) = 1-3+2 = 0, \quad p_1(2) = 4-6+2 = 0, \\ p_2(0) &= 0, \quad p_2(1) = 1-2 = -1, \qquad\quad p_2(2) = 4-4 = 0, \\ p_3(0) &= 0, \quad p_3(1) = 1-1 = 0, \qquad\quad p_3(2) = 4-2 = 2. \end{aligned}$$

Suppose that $1+1 \neq 0$, and that $ap_1 + bp_2 + cp_3 = \mathbf{0}$. The zero polynomial has value 0 everywhere, so by Proposition 1.2.18(c), evaluating at 0, 1, 2 gives

$$2a = 0, \quad -b = 0, \quad 2c = 0.$$

Since $2 = 1+1 \neq 0$, Proposition 1.1.17 gives $a = c = 0$, and $b = 0$. So the list is independent by Theorem 1.5.5.

Suppose that $1+1 = 0$. Then $2 = 0$, $3 = 1$, and $-1 = 1$ in F, so

$$p_1 = x^2 + x, \quad p_2 = x^2, \quad p_3 = x^2 + x.$$

The list repeats an entry, so it is dependent (Proposition 1.5.15(c)). The evaluation argument breaks down because the values $p_1(0) = p_3(2) = 2$ vanish when $2 = 0$, so the equations at 0 and 2 become $0 = 0$, and only $b = 0$ remains.

Solution to Exercise 1.5.36

Uses. Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.1.17 (No zero divisors)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. Expand $\sum_i a_i \mathbf{w}_i$ in terms of the $\mathbf{v}_i$. Each $\mathbf{v}_i$ occurs in exactly two of the $\mathbf{w}$'s.
Hint 2. Independence of the $\mathbf{v}_i$ gives $a_i + a_{i-1} = 0$ for every i, with $a_0 = a_n$. So the signs alternate around the cycle. What happens when you get back to the start?
Hint 3. Try $n = 2, 3, 4$ first, and over $\mathbb{F}_2$.
Sketch. With $a_0 = a_n$, $\sum a_i \mathbf{w}_i = \sum_i (a_{i-1} + a_i)\mathbf{v}_i$, so a relation means $a_i = -a_{i-1}$ for all i, that is, $a_i = (-1)^{i-1} a_1$ and $\left(1 + (-1)^{n-1}\right) a_1 = 0$. For n even, $a_i = (-1)^{i-1}$ is a nontrivial relation. For n odd, the condition is $(1+1)a_1 = 0$. Answer: independent exactly when n is odd and $1 + 1 \neq 0$.

Let $a_1, \ldots, a_n \in F$, and set $a_0 = a_n$. The vector $\mathbf{v}_i$ occurs in $\mathbf{w}_{i-1}$ and in $\mathbf{w}_i$ (in $\mathbf{w}_n$ and $\mathbf{w}_1$ for $i = 1$), each time with coefficient 1. So by Lemma 1.3.9,

$$\sum_{i=1}^{n} a_i \mathbf{w}_i = \sum_{i=1}^{n} (a_{i-1} + a_i)\mathbf{v}_i.$$

Since the $\mathbf{v}_i$ are independent, $a_1, \ldots, a_n$ is a linear relation of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ if and only if $a_i = -a_{i-1}$ for $2 \leq i \leq n$ and $a_1 = -a_n$. The first $n-1$ conditions say that $a_i = (-1)^{i-1} a_1$ for all i, and then the last one says $a_1 = -(-1)^{n-1} a_1$, that is,

$$\left(1 + (-1)^{n-1}\right) a_1 = 0.$$

If n is even, $(-1)^{n-1} = -1$, so the last condition always holds, and $a_i = (-1)^{i-1}$ is a nontrivial relation. The list is dependent, over every field.

If n is odd, the last condition is $(1+1)a_1 = 0$. If $1 + 1 \neq 0$, it forces $a_1 = 0$ (Proposition 1.1.17), hence every $a_i = 0$, and the list is independent by Theorem 1.5.5. If $1 + 1 = 0$, then $a_1 = 1$ satisfies it, giving the nontrivial relation $a_i = (-1)^{i-1} = 1$ (as $-1 = 1$). Indeed,

$$\mathbf{w}_1 + \cdots + \mathbf{w}_n = (1+1)(\mathbf{v}_1 + \cdots + \mathbf{v}_n) = \mathbf{0}.$$

So $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is independent exactly when n is odd and $1 + 1 \neq 0$ in F. For $n = 2$ the list is $\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_1$, which repeats an entry, as it must.

Solution to Exercise 1.5.37

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.1.29 (Complex numbers), Proposition 1.2.19 (Restriction of scalars)
Hint 1. For real a_j, b_j, rewrite $a_j\mathbf{v}_j + b_j(i\mathbf{v}_j)$ as a single complex multiple of $\mathbf{v}_j$.
Hint 2. Complex relations of $\mathbf{v}_1, \dots, \mathbf{v}_k$ and real relations of the doubled list correspond to each other, via $c_j = a_j + b_j i$.
Sketch. By (A6) and (A8), $a_j\mathbf{v}_j + b_j(i\mathbf{v}_j) = (a_j + b_j i)\mathbf{v}_j$. So real coefficients $a_1, b_1, \dots, a_k, b_k$ form a relation of the doubled list if and only if $a_1 + b_1 i, \dots, a_k + b_k i$ form a complex relation of the original list, and one is trivial exactly when the other is.

For $a_j, b_j \in \mathbb{R}$, (A6) gives $b_j(i\mathbf{v}_j) = (b_j i)\mathbf{v}_j$, and then (A8) gives $a_j\mathbf{v}_j + (b_j i)\mathbf{v}_j = (a_j + b_j i)\mathbf{v}_j$. Hence

$$\sum_{j=1}^{k}\bigl(a_j\mathbf{v}_j + b_j(i\mathbf{v}_j)\bigr) = \sum_{j=1}^{k}(a_j + b_j i)\mathbf{v}_j. \tag{C.1.1}$$

(⇒) Suppose that $\mathbf{v}_1, \dots, \mathbf{v}_k$ is independent over $\mathbb{C}$, and that real numbers $a_1, b_1, \dots, a_k, b_k$ form a relation of the doubled list. By (C.1.1), the complex numbers $a_j + b_j i$ form a relation of $\mathbf{v}_1, \dots, \mathbf{v}_k$, so $a_j + b_j i = 0$, that is, $a_j = b_j = 0$ (Definition 1.1.29).

(⇐) Suppose that the doubled list is independent over $\mathbb{R}$, and that $c_1, \dots, c_k \in \mathbb{C}$ form a relation of $\mathbf{v}_1, \dots, \mathbf{v}_k$. Write $c_j = a_j + b_j i$ with $a_j, b_j \in \mathbb{R}$. By (C.1.1), the real numbers $a_1, b_1, \dots, a_k, b_k$ form a relation of the doubled list, so they are all 0, and every $c_j = 0$.

In the language of Section 4.1: if $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis of V over $\mathbb{C}$, the doubled list is a basis over $\mathbb{R}$, so the real dimension of V is twice its complex dimension.

Solution to Exercise 1.5.38

Uses. Definition 1.3.10 (Span), Definition 1.1.29 (Complex numbers), Definition 1.2.4 (The space F^n), Proposition 1.5.21 (Real lists in $\mathbb{C}^m$)
Hint 1. Every $\mathbf{z} \in \mathbb{C}^m$ can be written as $\mathbf{x} + i\mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, by taking real and imaginary parts coordinatewise.
Hint 2. For (⇐), write a real vector as a complex combination of $\mathcal{L}$ and take real parts coordinate by coordinate, as in the proof of Proposition 1.5.21.
Sketch. (⇒) Write $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, expand $\mathbf{x} = \sum a_j\mathbf{v}_j$ and $\mathbf{y} = \sum b_j\mathbf{v}_j$ over $\mathbb{R}$; then $\mathbf{z} = \sum(a_j + b_j i)\mathbf{v}_j$. (⇐) If $\mathbf{x} \in \mathbb{R}^m$ equals $\sum c_j\mathbf{v}_j$ with $c_j = a_j + b_j i$, then comparing real parts in each coordinate gives $\mathbf{x} = \sum a_j\mathbf{v}_j$.

Write v_{jt} for the t-th coordinate of $\mathbf{v}_j$, a real number.

(⇒) Suppose $\mathcal{L}$ spans $\mathbb{R}^m$, and let $\mathbf{z} \in \mathbb{C}^m$ have coordinates $z_t = x_t + y_t i$ with $x_t, y_t \in \mathbb{R}$. Let $\mathbf{x} = (x_1, \dots, x_m)$ and $\mathbf{y} = (y_1, \dots, y_m)$ in $\mathbb{R}^m$, and write $\mathbf{x} = \sum_j a_j\mathbf{v}_j$ and $\mathbf{y} = \sum_j b_j\mathbf{v}_j$ with $a_j, b_j \in \mathbb{R}$. Then for each t,

$$\sum_{j=1}^{k}(a_j + b_j i)v_{jt} = \sum_{j=1}^{k} a_j v_{jt} + \left(\sum_{j=1}^{k} b_j v_{jt}\right) i = x_t + y_t i = z_t,$$

so $\mathbf{z} = \sum_j(a_j + b_j i)\mathbf{v}_j \in \operatorname{span}_{\mathbb{C}}(\mathcal{L})$.

(⇐) Suppose $\mathcal{L}$ spans $\mathbb{C}^m$, and let $\mathbf{x} \in \mathbb{R}^m$. Then $\mathbf{x} = \sum_j c_j\mathbf{v}_j$ for some $c_j = a_j + b_j i$ in $\mathbb{C}$. In coordinate t,

$$x_t = \sum_{j=1}^{k} a_j v_{jt} + \left(\sum_{j=1}^{k} b_j v_{jt}\right) i.$$

Both sums are real, and x_t is real, so comparing real parts (Definition 1.1.29) gives $x_t = \sum_j a_j v_{jt}$. Hence $\mathbf{x} = \sum_j a_j\mathbf{v}_j$ with real coefficients, and $\mathbf{x} \in \operatorname{span}_{\mathbb{R}}(\mathcal{L})$.

Together with Proposition 1.5.21, this shows that $\mathcal{L}$ is a basis of $\mathbb{R}^m$ if and only if it is a basis of $\mathbb{C}^m$ (Section 1.6).

Solution to Exercise 1.5.39

Uses. Theorem 1.5.5 (Independence is decided at zero), Example 1.2.20 (Real and rational scalars)
Hint 1. Suppose $a\ln 2 + b\ln 3 + c\ln 5 = 0$ with $a, b, c \in \mathbb{Q}$. Clear denominators to make the coefficients integers.
Hint 2. Exponentiate: $2^A 3^B 5^C = 1$ with integers A, B, C. Move the factors with negative exponents to the other side.

Sketch. Multiply a relation by a common denominator to get $A\ln 2 + B\ln 3 + C\ln 5 = 0$ with $A, B, C \in \mathbb{Z}$, so $2^A 3^B 5^C = 1$. Collecting negative exponents on the right gives two factorizations of one positive integer into primes, and uniqueness forces $A = B = C = 0$.

Suppose $a\ln 2 + b\ln 3 + c\ln 5 = 0$ with $a, b, c \in \mathbb{Q}$. Let $N \geq 1$ be a common denominator, so that $A = Na$, $B = Nb$, $C = Nc$ are integers, and $A\ln 2 + B\ln 3 + C\ln 5 = 0$. Exponentiating, $2^A 3^B 5^C = e^0 = 1$. For an integer n, write $n^+ = \max(n, 0)$ and $n^- = \max(-n, 0)$, so that $n = n^+ - n^-$ and at most one of n^+, n^- is nonzero. Multiplying by $2^{A^-} 3^{B^-} 5^{C^-}$,

$$2^{A^+} 3^{B^+} 5^{C^+} = 2^{A^-} 3^{B^-} 5^{C^-}.$$

Both sides are positive integers. By unique factorization, the exponent of 2 on the two sides agrees, so $A^+ = A^-$; since one of them is 0, both are, and $A = 0$. Likewise $B = C = 0$. Hence $a = b = c = 0$, and the list is independent by Theorem 1.5.5.

The same argument shows that $\ln p_1, \ldots, \ln p_k$ is independent over $\mathbb{Q}$ for any distinct primes $p_1, \ldots, p_k$. Since there are infinitely many primes, $\mathbb{R}$ over $\mathbb{Q}$ contains independent lists of every length, which is another way to see that no finite list spans it (Exercise 1.3.45, Section 4.1).

Solution to Exercise 1.5.40

Uses. Exercise 1.3.44, Example 1.5.19 (The same vectors over two fields), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Group a relation $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0$ as $\alpha + \beta\sqrt{3} = 0$ with $\alpha = a + b\sqrt{2}$ and $\beta = c + d\sqrt{2}$.
Hint 2. If $\beta \neq 0$, then $\sqrt{3} = -\alpha/\beta$. Rationalize the denominator: multiply numerator and denominator by $c - d\sqrt{2}$.
Hint 3. Check that $c^2 - 2d^2 \neq 0$ when $(c, d) \neq (0, 0)$, so that $-\alpha/\beta$ lies in $\operatorname{span}\left(1, \sqrt{2}\right)$.
Sketch. Write the relation as $\left(a + b\sqrt{2}\right) + \left(c + d\sqrt{2}\right)\sqrt{3} = 0$. If $c + d\sqrt{2} \neq 0$, then $c^2 - 2d^2 \neq 0$, and $\sqrt{3} = -\left(a + b\sqrt{2}\right)\left(c - d\sqrt{2}\right)/\left(c^2 - 2d^2\right)$ is of the form $r + s\sqrt{2}$ with $r, s \in \mathbb{Q}$, contradicting Exercise 1.3.44. So $c + d\sqrt{2} = 0$, hence $c = d = 0$, and then $a = b = 0$.

Suppose $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} = 0$ with $a, b, c, d \in \mathbb{Q}$. Since $\sqrt{6} = \sqrt{2}\sqrt{3}$, this reads

$$\left(a + b\sqrt{2}\right) + \left(c + d\sqrt{2}\right)\sqrt{3} = 0.$$

Suppose $c + d\sqrt{2} \neq 0$. Then $(c, d) \neq (0, 0)$, and $c^2 - 2d^2 \neq 0$: otherwise $d \neq 0$ (as $c^2 = 2d^2$ with $d = 0$ forces $c = 0$) and $(c/d)^2 = 2$, making $\sqrt{2}$ or $-\sqrt{2}$ rational. So $c - d\sqrt{2} \neq 0$ as well, since $\left(c + d\sqrt{2}\right)\left(c - d\sqrt{2}\right) = c^2 - 2d^2$. Hence

$$\sqrt{3} = -\frac{a + b\sqrt{2}}{c + d\sqrt{2}} = -\frac{\left(a + b\sqrt{2}\right)\left(c - d\sqrt{2}\right)}{c^2 - 2d^2} = -\frac{ac - 2bd}{c^2 - 2d^2} - \frac{bc - ad}{c^2 - 2d^2}\sqrt{2},$$

using

$$\left(a + b\sqrt{2}\right)\left(c - d\sqrt{2}\right) = (ac - 2bd) + (bc - ad)\sqrt{2}.$$

Both coefficients on the right are rational, so $\sqrt{3} \in \operatorname{span}\left(1, \sqrt{2}\right)$ over $\mathbb{Q}$, contradicting Exercise 1.3.44.

So $c + d\sqrt{2} = 0$. Since $1, \sqrt{2}$ is independent over $\mathbb{Q}$ (Example 1.5.19(b)), $c = d = 0$. The relation becomes $a + b\sqrt{2} = 0$, and for the same reason $a = b = 0$. By Theorem 1.5.5, the list is independent.

Solution to Exercise 1.5.41

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.1.19 (Integers modulo n), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$)
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. Prove the contrapositive of (a). Starting from a nontrivial rational relation, clear denominators to get integer coefficients.
Hint 2. Divide the integer coefficients by their greatest common divisor. Then not all of them are divisible by p, so reducing the relation modulo p gives a nontrivial relation over $\mathbb{F}_p$.

Sketch. (a) A nontrivial relation over $\mathbb{Q}$ can be scaled to have integer coefficients $n_1, \dots, n_k$ with greatest common divisor 1. Reducing each coordinate equation $\sum_j n_j v_{jt} = 0$ modulo p gives a relation over $\mathbb{F}_p$, which is nontrivial because p does not divide every n_j. (b) The list (2) in $\mathbb{Q}^1$ is independent, but its reduction modulo 2 is (0).

(a) We prove the contrapositive. Suppose that $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ with $c_j \in \mathbb{Q}$ not all 0. Multiplying by a common denominator of the c_j gives a relation with integer coefficients, not all 0; dividing these by their greatest common divisor gives a relation with integer coefficients $n_1, \dots, n_k$ whose greatest common divisor is 1. In particular, p does not divide some n_j.

Write v_{jt} for the t-th coordinate of $\mathbf{v}_j$. For each t, $\sum_j n_j v_{jt} = 0$ is an equation between integers. Reducing modulo p is compatible with sums and products (Definition 1.1.19, Lemma 1.1.21), so $\sum_j \overline{n}_j \overline{v}_{jt} = \overline{0}$ in $\mathbb{F}_p$, where bars denote remainders modulo p. That is, $\overline{n}_1, \dots, \overline{n}_k$ is a linear relation of $\overline{\mathbf{v}}_1, \dots, \overline{\mathbf{v}}_k$ in $\mathbb{F}_p^m$. It is nontrivial, since $\overline{n}_j \neq 0$ for the j with $p \nmid n_j$. So the reduced list is dependent.

(b) The list (2) of length one in $\mathbb{Q}^1$ is independent, since $(2) \neq \mathbf{0}$ (Proposition 1.5.15(b)). Its reduction modulo 2 is (0), which is dependent. For a list of length two: $(1,1), (1,-1)$ is independent in $\mathbb{Q}^2$ (neither vector is a multiple of the other), but modulo 2 both vectors reduce to $(1,1)$.

Part (a) gives a quick test: to show that an integer list is independent over $\mathbb{Q}$, it suffices to check independence over a small field such as $\mathbb{F}_2$, where there are only finitely many relations to try.

Solution to Exercise 1.5.42

Uses. Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.19 (Restriction of scalars), Example 1.2.6 (Small fields), Theorem 1.6.10 (Every spanning list contains a basis), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Hint 1. One direction holds because a relation over K is a relation over E. For the other, take a relation $\sum_j c_j\mathbf{v}_j = \mathbf{0}$ with $c_j \in E$.
Hint 2. The coefficients $c_1, \dots, c_k$ are vectors of the K-vector space E. By Theorem 1.6.10, some sublist $r_1, \dots, r_s$ of $c_1, \dots, c_k$ is K-independent and has the same K-span.
Hint 3. Write $c_j = \sum_l q_{jl} r_l$ with $q_{jl} \in K$, substitute into each coordinate of the relation, and use K-independence of the r_l, then of $\mathcal{L}$.
Sketch. (Independence over E implies independence over K.) A K-relation is an E-relation. (Independence over K implies independence over E.) Given an E-relation $\sum c_j\mathbf{v}_j = \mathbf{0}$, choose a K-independent sublist $r_1, \dots, r_s$ of $c_1, \dots, c_k$ with the same K-span, and write $c_j = \sum_l q_{jl} r_l$ with $q_{jl} \in K$. In coordinate t, $0 = \sum_l \big(\sum_j q_{jl} v_{jt}\big) r_l$ with coefficients in K, so $\sum_j q_{jl} v_{jt} = 0$ for all l, t: each $q_{1l}, \dots, q_{kl}$ is a K-relation of $\mathcal{L}$, hence trivial. So every $c_j = 0$.

The operations of K^m are those of E^m restricted to vectors and scalars from K. Write $v_{jt} \in K$ for the t-th coordinate of $\mathbf{v}_j$.

Independence over E implies independence over K. A relation $a_1, \dots, a_k \in K$ of $\mathcal{L}$ in K^m is also a relation in E^m, so it is trivial.

Independence over K implies independence over E. Suppose $\mathcal{L}$ is independent over K, and let $c_1, \dots, c_k \in E$ with $\sum_j c_j\mathbf{v}_j = \mathbf{0}$ in E^m. The field E is a vector space over itself (Example 1.2.6), hence over K (Proposition 1.2.19), with scalar multiplication the product in E. Let W be the K-span of $c_1, \dots, c_k$ in E. It is a K-vector space (Proposition 1.4.14), and $c_1, \dots, c_k$ spans it, so by Theorem 1.6.10 some sublist $r_1, \dots, r_s$ of $c_1, \dots, c_k$ is a basis of W over K. In particular, $r_1, \dots, r_s$ is K-independent, and each c_j lies in its K-span: $c_j = \sum_{l=1}^{s} q_{jl} r_l$ with $q_{jl} \in K$.

Fix a coordinate t. The t-th coordinate of the relation is, after expanding and regrouping in the field E,

$$0 = \sum_{j=1}^{k} c_j v_{jt} = \sum_{j=1}^{k} \sum_{l=1}^{s} q_{jl} r_l v_{jt} = \sum_{l=1}^{s} \left(\sum_{j=1}^{k} q_{jl} v_{jt} \right) r_l.$$

Each coefficient $\sum_j q_{jl} v_{jt}$ lies in K. Since $r_1, \dots, r_s$ is K-independent, all these coefficients are 0. As this holds for every t, we get $\sum_j q_{jl}\mathbf{v}_j = \mathbf{0}$ in K^m for each l. Since $\mathcal{L}$ is K-independent, $q_{jl} = 0$ for all j and l. Therefore $c_j = \sum_l q_{jl} r_l = 0$ for every j, and $\mathcal{L}$ is independent over E by Theorem 1.5.5.

For $K = \mathbb{R}$ and $E = \mathbb{C}$, this is Proposition 1.5.21. There the list $1, i$ played the role of $r_1, \dots, r_s$, chosen once and for all; here the list depends on the relation, which is what makes the argument work when E has no finite spanning list over K (Remark 1.5.23).

Solution to Exercise 1.5.43

Uses. Definition 1.5.1 (Linear independence), Definition 1.3.10 (Span), Exercise 1.3.39, Proposition 1.5.14 (Appending a vector), Corollary 1.5.17 (Sublists), induction on k
Hint 1. In (a), consider the map sending a list of coefficients $a_1, \dots, a_k$ to the vector it represents. It maps F^k onto $\operatorname{span}(\mathcal{L})$. What does independence say about it?
Hint 2. For (b), build an independent list one entry at a time with Proposition 1.5.14. Given an independent list of length j, how many vectors lie outside its span?
Sketch. (a) The map $F^k \to \operatorname{span}(\mathcal{L})$, coefficients $\mapsto$ represented vector, is onto by definition, and injective exactly when $\mathcal{L}$ is independent; a map from a set with q^k elements onto a set is injective iff the target has q^k elements. (b) By Proposition 1.5.14, an independent list of length $j+1$ is an independent list of length j followed by a vector outside its span, which has q^j elements by (a): $q^n - q^j$ choices. (c) $7 \cdot 6 \cdot 4 = 168$ and 0.

(a) Let $\mathcal{L} = \mathbf{v}_1, \dots, \mathbf{v}_k$, and let $\varphi : F^k \to \operatorname{span}(\mathcal{L})$ send $(a_1, \dots, a_k)$ to $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$. By Definition 1.3.10, φ is surjective, and by Definition 1.5.1, $\mathcal{L}$ is independent exactly when every vector of $\operatorname{span}(\mathcal{L})$ has one preimage, that is, when φ is injective. The set F^k has q^k elements. A surjection from a set with q^k elements onto a finite set is injective if and only if the target also has q^k elements. So $\mathcal{L}$ is independent if and only if $|\operatorname{span}(\mathcal{L})| = q^k$.

(b) Let N_k be the number of independent lists of length k in F^n; we show $N_k = \prod_{i=0}^{k-1}\left(q^n - q^i\right)$ by induction on k. For $k = 0$, the empty list is independent, so $N_0 = 1$, the empty product. Let $k \geq 0$. By Corollary 1.5.17 and Proposition 1.5.14, an independent list of length $k+1$ is the same thing as an independent list $\mathcal{L}$ of length k followed by a vector $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$. For each of the N_k choices of $\mathcal{L}$, the span has q^k elements by (a), so there are $q^n - q^k$ choices of $\mathbf{w}$. Hence $N_{k+1} = N_k\left(q^n - q^k\right)$, which completes the induction. As soon as $k > n$, the product contains the factor $q^n - q^n = 0$.

(c) For $q = 2$, $n = 3$, $k = 3$: $(8-1)(8-2)(8-4) = 7 \cdot 6 \cdot 4 = 168$. For $k = 4$ the product contains the factor $8 - 2^3 = 0$, so there are none: no list of length 4 in $\mathbb{F}_2^3$ is independent. Over a finite field this counting argument already proves that an independent list in F^n has length at most n, a fact that Section 4.1 proves over every field (Theorem 4.1.10).

Solution to Exercise 1.5.44

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.9 (Function spaces), Proposition 1.1.17 (No zero divisors)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Suppose the list is dependent, and choose a dependence relation $c_1, \dots, c_k$ with as few nonzero coefficients as possible. A relation says $\sum_i c_i\lambda_i^m = 0$ for every $m \geq 0$.
Hint 2. Shift: the same identity holds with $m+1$ in place of m. Subtract λ_{i_0} times the original identity, where $c_{i_0} \neq 0$.
Hint 3. The new relation has coefficients $c_i\left(\lambda_i - \lambda_{i_0}\right)$. It has fewer nonzero coefficients, so it is trivial. What does that say about c_i for $i \neq i_0$?
Sketch. Take a dependence relation $\sum c_i\lambda_i^m = 0$ (all m) with the fewest nonzero coefficients, and i_0 with $c_{i_0} \neq 0$. Then $\sum_i c_i\left(\lambda_i - \lambda_{i_0}\right)\lambda_i^m = 0$ for all m, a relation with fewer nonzero coefficients (the i_0 one vanishes), hence trivial. Since $\lambda_i \neq \lambda_{i_0}$, $c_i = 0$ for $i \neq i_0$, and then the $m = 0$ term gives $c_{i_0} = 0$, a contradiction.

The m-th term of $\mathbf{s}_\lambda$ is λ^m (with $\lambda^0 = 1$). So scalars $c_1, \dots, c_k$ form a linear relation of the list exactly when

$$\sum_{i=1}^{k} c_i\lambda_i^m = 0 \quad \text{for every } m \geq 0.$$

Suppose the list is dependent. Among its dependence relations, choose one, $c_1, \dots, c_k$, with the fewest nonzero coefficients, and fix i_0 with $c_{i_0} \neq 0$. Applying the identity with $m+1$ in place of m, and subtracting λ_{i_0} times the identity for m,

$$\sum_{i=1}^{k} c_i\big(\lambda_i - \lambda_{i_0}\big)\lambda_i^m = \sum_{i=1}^{k} c_i\lambda_i^{m+1} - \lambda_{i_0}\sum_{i=1}^{k} c_i\lambda_i^m = 0 \quad \text{for every } m \geq 0.$$

So $d_i = c_i\big(\lambda_i - \lambda_{i_0}\big)$ is a linear relation. Every i with $d_i \neq 0$ has $c_i \neq 0$, and $d_{i_0} = 0$ while $c_{i_0} \neq 0$. So d has fewer nonzero coefficients than c, and by the choice of c it is not a dependence relation: it is trivial. For $i \neq i_0$, $\lambda_i - \lambda_{i_0} \neq 0$, so $d_i = 0$ forces $c_i = 0$ (Proposition 1.1.17). The relation c thus reads $c_{i_0}\lambda_{i_0}^m = 0$ for all m, and at $m = 0$ this gives $c_{i_0} = 0$, a contradiction. So the list is independent.

Shifting a sequence multiplies $\mathbf{s}_\lambda$ by λ, just as differentiation multiplies $t \mapsto e^{\lambda t}$ by λ; the argument is the first proof of Example 1.5.27, done once for all k. A variant of the argument, by induction on k instead of a minimal relation, shows that the truncated vectors $\big(1, \lambda_i, \lambda_i^2, \ldots, \lambda_i^{k-1}\big) \in F^k$, $1 \leq i \leq k$, are independent. This is the Vandermonde determinant in disguise (Theorem 5.5.21).

Solution to Exercise 1.5.45

Uses. Theorem 1.5.5 (Independence is decided at zero), $e^{-st} \to 0$ as $t \to \infty$, for $s > 0$
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Follow the second proof of Example 1.5.27. In a dependence relation, which term dominates as $t \to \infty$?
Hint 2. Among the i with $c_i \neq 0$, take the one with the largest λ_i, divide by $e^{\lambda_i t}$, and let $t \to \infty$.
Sketch. Suppose $\sum c_i e^{\lambda_i t} = 0$ for all t with some $c_i \neq 0$, and let $\mu = \lambda_{i_0}$ be the largest λ_i with $c_i \neq 0$. Dividing by $e^{\mu t}$ gives $c_{i_0} + \sum c_i e^{(\lambda_i - \mu)t} = 0$, the sum over the other i with $c_i \neq 0$, all of which have $\lambda_i - \mu < 0$. Letting $t \to \infty$ gives $c_{i_0} = 0$, a contradiction.

Suppose, to the contrary, that $\sum_{i=1}^{k} c_i e^{\lambda_i t} = 0$ for all $t \in \mathbb{R}$, with not all $c_i = 0$. Let $I = \{i \mid c_i \neq 0\}$, and choose $i_0 \in I$ with $\mu = \lambda_{i_0}$ as large as possible; since the λ_i are distinct, $\lambda_i < \mu$ for all $i \in I$ other than i_0. Dropping the terms with $c_i = 0$ and multiplying by $e^{-\mu t}$,

$$c_{i_0} + \sum_{i \in I, i \neq i_0} c_i e^{(\lambda_i - \mu)t} = 0 \quad \text{for all } t.$$

Each exponent $\lambda_i - \mu$ in the sum is negative, so each term tends to 0 as $t \to \infty$, and the left side tends to c_{i_0}. Since it is 0 for every t, $c_{i_0} = 0$, contradicting $i_0 \in I$. So the functions are independent.

In the language of Aside (Independent sets), the set of all functions $t \mapsto e^{\lambda t}$ with $\lambda \in \mathbb{R}$ is linearly independent, and it is uncountable. The first proof of Example 1.5.27, organized as in Exercise 1.5.44 with d/dt in place of the shift, also works.

Solution to Exercise 1.5.46

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.4 (The space F^n), Example 1.1.4 (The field with two elements)
Hint 1. For $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$, let $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n \in \mathbb{F}_2$. What is $\mathbf{x}_C \cdot \mathbf{x}_D$ in terms of the clubs C and D?
Hint 2. Take a relation $\sum_j a_j \mathbf{x}_{C_j} = \mathbf{0}$ and take the dot product of both sides with $\mathbf{x}_{C_i}$. Which terms survive?
Sketch. $\mathbf{x}_C \cdot \mathbf{x}_D$ is the parity of $|C \cap D|$, so $\mathbf{x}_{C_i} \cdot \mathbf{x}_{C_j}$ is 1 if $i = j$ and 0 otherwise. The dot product satisfies $\big(\sum_j a_j \mathbf{x}_j\big) \cdot \mathbf{y} = \sum_j a_j(\mathbf{x}_j \cdot \mathbf{y})$. Dotting a relation with $\mathbf{x}_{C_i}$ leaves $a_i = 0$.

For $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$, define $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n \in \mathbb{F}_2$. By the coordinatewise operations and distributivity in $\mathbb{F}_2$, $(\mathbf{x} + \mathbf{x}') \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{y} + \mathbf{x}' \cdot \mathbf{y}$ and $(a\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y})$, so by induction on the number of terms,

$$\left(\sum_{j=1}^{k} a_j \mathbf{x}_j\right) \cdot \mathbf{y} = \sum_{j=1}^{k} a_j(\mathbf{x}_j \cdot \mathbf{y}),$$

and $\mathbf{0} \cdot \mathbf{y} = 0$. For clubs C and D, the product $(\mathbf{x}_C)_t(\mathbf{x}_D)_t$ is 1 if $t \in C \cap D$ and 0 otherwise, so $\mathbf{x}_C \cdot \mathbf{x}_D$ is the sum of $|C \cap D|$ ones in $\mathbb{F}_2$, which is 1 if $|C \cap D|$ is odd and 0 if it is even (Example 1.1.4). By hypothesis, writing $\mathbf{x}_j = \mathbf{x}_{C_j}$,

$$\mathbf{x}_i \cdot \mathbf{x}_i = 1, \quad \mathbf{x}_j \cdot \mathbf{x}_i = 0 \quad (j \neq i).$$

Now suppose $\sum_j a_j \mathbf{x}_j = \mathbf{0}$, and fix i. Taking the dot product of both sides with $\mathbf{x}_i$,

$$0 = \mathbf{0} \cdot \mathbf{x}_i = \sum_{j=1}^{k} a_j(\mathbf{x}_j \cdot \mathbf{x}_i) = a_i.$$

So every $a_i = 0$, and the list is independent by Theorem 1.5.5.

Taking a dot product with $\mathbf{x}_i$ eliminated every term of the relation but one, the idea of Technique 4.1.25. Once we know that an independent list in $\mathbb{F}_2^n$ has length at most n (Exercise 1.5.43, Theorem 4.1.10), it follows that the town has at most n clubs.

Solution to Exercise 1.5.47

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.9 (Function spaces)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. For a function g, a point s, and $h > 0$, consider $\Delta_h g(s) = g(s+h) - 2g(s) + g(s-h)$. What is it for $g(t) = c + dt$?
Hint 2. If h is smaller than the distance from a_i to every other a_l, compute $\Delta_h f_l(a_i)$ for $l \neq i$ and for $l = i$.
Sketch. The operation $\Delta_h g(s) = g(s+h) - 2g(s) + g(s-h)$ respects linear combinations, kills affine functions, and kills $|t - a_l|$ at $s = a_i$ when $l \neq i$ and h is small (the function is affine near a_i). But $\Delta_h f_i(a_i) = 2h$. Applying it to a relation at $s = a_i$ gives $2hb_i = 0$. Once all $b_i = 0$, evaluating $c + dt = 0$ at $t = 0, 1$ gives $c = d = 0$.

For a function $g : \mathbb{R} \to \mathbb{R}$, a point s, and $h > 0$, let $\Delta_h g(s) = g(s+h) - 2g(s) + g(s-h)$. If $g = \sum_r c_r g_r$, then $\Delta_h g(s) = \sum_r c_r \Delta_h g_r(s)$, since the pointwise operations let us expand and regroup.

Affine functions. For $g(t) = c + dt$,

$$\Delta_h g(s) = (c + ds + dh) - 2(c + ds) + (c + ds - dh) = 0.$$

The functions f_l near a_i. Fix i, and let $h > 0$ be smaller than $|a_l - a_i|$ for every $l \neq i$. For $l \neq i$, the interval $[a_i - h, a_i + h]$ lies entirely on one side of a_l, so on it $f_l(t) = \varepsilon(t - a_l)$ for a fixed sign $\varepsilon \in \{1, -1\}$. This is an affine function of t, so $\Delta_h f_l(a_i) = 0$ by the previous computation. For $l = i$, $\Delta_h f_i(a_i) = |h| - 2 \cdot 0 + |-h| = 2h$.

Now suppose that $c + dt + \sum_{l=1}^{n} b_l f_l(t) = 0$ for all t. Applying Δ_h at $s = a_i$, with h as above, every term but one vanishes, and we get $2hb_i = 0$, so $b_i = 0$. This holds for every i, so $c + dt = 0$ for all t. At $t = 0$ this gives $c = 0$, and at $t = 1$ it gives $d = 0$. By Theorem 1.5.5, the list is independent.

The operation Δ_h at a_i measures the kink of a function at a_i; only f_i has a kink there. This is the move of Technique 4.1.25 with a "second difference" as the operator.

Solution to Exercise 1.5.48

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.4 (The space F^n), a sum of nonnegative real numbers is 0 only if every term is 0
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, let $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$. What is $\mathbf{x}_A \cdot \mathbf{x}_B$? Given a relation $\sum_i c_i \mathbf{x}_{A_i} = \mathbf{0}$, take the dot product of the left side with itself.
Hint 2. You should get $0 = \sum_i c_i^2 |A_i| + \lambda \sum_{i \neq j} c_i c_j$. Rewrite $\sum_{i \neq j} c_i c_j$ using $\left(\sum_i c_i\right)^2$, and check that every resulting term is nonnegative.
Hint 3. Each $|A_i| \geq \lambda$. Can two of the sets have exactly λ elements?
Sketch. With $\mathbf{x}_i = \mathbf{x}_{A_i}$, $\mathbf{x}_i \cdot \mathbf{x}_j = |A_i \cap A_j|$. Dotting a relation $\mathbf{u} = \sum c_i \mathbf{x}_i = \mathbf{0}$ with itself gives $0 = \sum_i c_i^2(|A_i| - \lambda) + \lambda\left(\sum_i c_i\right)^2$, a sum of nonnegative terms, so each vanishes: $\sum_i c_i = 0$, and $c_i = 0$ unless $|A_i| = \lambda$. Two sets with λ elements would both equal their intersection, so at most one c_i survives, and $\sum_i c_i = 0$ kills it.

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, let $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$. By the coordinatewise operations and the arithmetic of $\mathbb{R}$, $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$, and the dot product respects linear combinations in each argument: expanding and regrouping finite sums,

$$\left(\sum_i c_i \mathbf{x}_i\right) \cdot \left(\sum_j c_j \mathbf{x}_j\right) = \sum_i \sum_j c_i c_j (\mathbf{x}_i \cdot \mathbf{x}_j).$$

Write $\mathbf{x}_i = \mathbf{x}_{A_i}$. The product of the t-th coordinates of $\mathbf{x}_i$ and $\mathbf{x}_j$ is 1 if $t \in A_i \cap A_j$ and 0 otherwise, so $\mathbf{x}_i \cdot \mathbf{x}_j = |A_i \cap A_j|$. This is λ for $i \neq j$, and $|A_i|$ for $i = j$.

Suppose $\mathbf{u} = \sum_i c_i \mathbf{x}_i = \mathbf{0}$ with $c_i \in \mathbb{R}$. Then $\mathbf{u} \cdot \mathbf{u} = 0$, and by the display,

$$0 = \sum_i c_i^2 |A_i| + \lambda \sum_{i \neq j} c_i c_j = \sum_i c_i^2 (|A_i| - \lambda) + \lambda \left(\sum_i c_i \right)^2,$$

using $\left(\sum_i c_i\right)^2 = \sum_i c_i^2 + \sum_{i \neq j} c_i c_j$. Since $m \geq 2$, each A_i meets some other A_j in λ elements, so $|A_i| \geq \lambda$. So every term on the right is a nonnegative real number, and since they add up to 0, each of them is 0. As $\lambda \geq 1$, this gives $\sum_i c_i = 0$, and $c_i = 0$ for every i with $|A_i| > \lambda$.

At most one i has $|A_i| = \lambda$. Indeed, if $|A_i| = |A_j| = \lambda$ with $i \neq j$, then $A_i \cap A_j$ is a subset of A_i with as many elements as A_i, so $A_i \cap A_j = A_i$; likewise $A_i \cap A_j = A_j$. Then $A_i = A_j$, contradicting distinctness. So all c_i vanish except possibly one, and that one equals $\sum_i c_i = 0$. Hence the list is independent by Theorem 1.5.5.

The hypotheses are needed: with $\lambda = 0$, the sets $\emptyset, \{1\}$ give the dependent list $\mathbf{0}, \mathbf{e}_1$. Once we know that an independent list in $\mathbb{R}^n$ has at most n entries (Theorem 4.1.10), it follows that $m \leq n$. This is Fisher's inequality from combinatorics, and the dot product trick is the one of Exercise 1.5.46, with positivity over $\mathbb{R}$ in place of parity over $\mathbb{F}_2$.

Section 1.6

Solution to Exercise 1.6.6

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Exercise 1.3.32, Proposition 1.5.24 (Distinct degrees)
Hint 1. By Proposition 1.6.2, check spanning and independence separately. One of them is already done in Section 1.3.
Hint 2. The three polynomials are nonzero and have different degrees.
Sketch. The list spans W by Exercise 1.3.32. Its entries have degrees $1, 2, 3$, so it is independent by Proposition 1.5.24.

Each entry has degree at most 3 and value $1 - 1 = 0$ at 1, so the list lies in W, and it spans W by Exercise 1.3.32. Its entries are nonzero polynomials of the distinct degrees 1, 2, and 3, so it is linearly independent by Proposition 1.5.24. By Proposition 1.6.2, it is a basis of W.

Solution to Exercise 1.6.12

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Proposition 1.5.15 (Short lists, zeros, and repeats), Exercise 1.3.37, Corollary 1.5.17 (Sublists)
Hint 1. Sort the sublists by length. A sublist of length 0 or 1 cannot span $\mathbb{R}^2$. Why?
Hint 2. Every sublist of length 3 or 4 has a dependence relation; find one in each case. For length 2, use Proposition 1.5.15(d) and Exercise 1.3.37.
Sketch. Five sublists: all pairs except $(1,1), (2,2)$. Length 0 and 1: the span is $\{\mathbf{0}\}$ or a set of multiples of one vector, which misses $\mathbf{e}_1$ or $\mathbf{e}_2$. Length 3 and 4: each contains a dependent sublist, since $(2,2) = 2(1,1)$ and $(1,1) = (1,0) + (0,1)$ and $(2,2) = 2(1,0) + 2(0,1)$. Length 2: $(1,1), (2,2)$ is dependent, and the other five pairs $(a,b), (c,d)$ have $ad - bc \neq 0$.

Length 0 or 1. The span of () is $\{\mathbf{0}\}$. The span of a single vector $\mathbf{v}$ consists of its multiples. If both $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ were multiples of $\mathbf{v} = (a,b)$, say $\mathbf{e}_1 = c\mathbf{v}$ and $\mathbf{e}_2 = d\mathbf{v}$, then $ca = 1$ and $cb = 0$ would give $c \neq 0$ and $b = 0$, while $db = 1$ would give $b \neq 0$. So no sublist of length 0 or 1 spans $\mathbb{R}^2$.

Length 3 or 4. The four sublists of length 3 have the dependence relations

$$\begin{aligned} &2(1,1) - (2,2) = \mathbf{0} && \text{for } (1,0), (1,1), (2,2) \text{ and for } (1,1), (2,2), (0,1), \\ &(1,0) - (1,1) + (0,1) = \mathbf{0} && \text{for } (1,0), (1,1), (0,1), \\ &2(1,0) - (2,2) + 2(0,1) = \mathbf{0} && \text{for } (1,0), (2,2), (0,1). \end{aligned}$$

The whole list contains the dependent sublist $(1,1), (2,2)$, so it is dependent by Corollary 1.5.17. None of these is a basis.

Length 2. The pair $(1,1), (2,2)$ is dependent. For each of the other five pairs $(a,b), (c,d)$, namely

$$(1,0), (1,1) \quad (1,0), (2,2) \quad (1,0), (0,1) \quad (1,1), (0,1) \quad (2,2), (0,1),$$

the number $ad - bc$ is $1, 2, 1, 1, 2$ respectively, which is nonzero. So each of them spans $\mathbb{R}^2$ by Exercise 1.3.37(a). Neither vector of the pair is a multiple of the other: a multiple of $(1, 0)$ has second coordinate 0, a multiple of $(0, 1)$ has first coordinate 0, and a multiple of $(1, 1)$ or $(2, 2)$ has equal coordinates. So each pair is independent by Proposition 1.5.15(d). By Proposition 1.6.2, these five sublists are exactly the bases of $\mathbb{R}^2$ contained in $\mathcal{L}$.

Solution to Exercise 1.6.18

Uses. Exercise 1.3.33, Proposition 1.5.24 (Distinct degrees), Proposition 1.6.2 (Bases are independent spanning lists), (1.6.1), Proposition 1.2.18 (Degree and evaluation versus the operations)
Hint 1. For (a), the entries have degrees $0, 1, 2$. Use Exercise 1.3.33 and Proposition 1.5.24.
Hint 2. For (b), by (1.6.1) it suffices to expand the combination with the proposed coefficients and check that it gives $a + bx + cx^2$.
Hint 3. For (c), evaluate $p = \alpha + \beta(x - 1) + \gamma\big(x^2 - 2x + 1\big)$ at 1.
Sketch. (a) Degrees $0, 1, 2$: spanning by Exercise 1.3.33, independence by Proposition 1.5.24. (b) Expanding $(a + b + c) + (b + 2c)(x - 1) + c\big(x^2 - 2x + 1\big)$ gives $a + bx + cx^2$. (c) Evaluation respects linear combinations, and the second and third entries of $\mathcal{T}$ vanish at 1.

(a) The entries of $\mathcal{T}$ have degrees 0, 1, and 2. So $\mathcal{T}$ spans $F[x]_{\leq 2}$ by Exercise 1.3.33(a), and it is independent by Proposition 1.5.24. By Proposition 1.6.2, it is a basis.

(b) Expanding,

$$\begin{aligned}(a + b + c) \cdot 1 + (b + 2c)(x - 1) + c\big(x^2 - 2x + 1\big) &= (a + b + c - b - 2c + c) + (b + 2c - 2c)x + cx^2 \\ &= a + bx + cx^2.\end{aligned}$$

By (1.6.1), $\big[a + bx + cx^2\big]_{\mathcal{T}} = (a + b + c, b + 2c, c)$.

(c) Write $[p]_{\mathcal{T}} = (\alpha, \beta, \gamma)$, so $p = \alpha + \beta(x - 1) + \gamma\big(x^2 - 2x + 1\big)$. By Proposition 1.2.18(c), evaluation at 1 respects sums and scalar multiples, so

$$p(1) = \alpha + \beta(1 - 1) + \gamma(1 - 2 + 1) = \alpha.$$

(In (b), this is the identity $p(1) = a + b + c$.) Similarly, over $\mathbb{R}$ the second coordinate $b + 2c$ is the derivative of p at 1. These are the first terms of the Taylor expansion of p around 1.

Solution to Exercise 1.6.22

Uses. Definition 1.6.14 (Coordinates), (1.6.1)
Hint 1. Which vector of V has $\mathcal{B}$-coordinates $\mathbf{e}_j$?
Hint 2. Compute $\big[\mathbf{v}_j\big]_{\mathcal{C}}$ using the hypothesis, and read it in reverse.
Sketch. $\big[\mathbf{v}_j\big]_{\mathcal{C}} = \big[\mathbf{v}_j\big]_{\mathcal{B}} = \mathbf{e}_j$, so $\mathbf{v}_j = \mathbf{w}_j$ by (1.6.1).

Fix j. Since $\mathbf{v}_j$ is represented by the combination of $\mathcal{B}$ with coefficient 1 on $\mathbf{v}_j$ and 0 elsewhere, $\big[\mathbf{v}_j\big]_{\mathcal{B}} = \mathbf{e}_j$. By hypothesis $\big[\mathbf{v}_j\big]_{\mathcal{C}} = \mathbf{e}_j$ as well, which by Definition 1.6.14 means $\mathbf{v}_j = 0\mathbf{w}_1 + \cdots + 1\mathbf{w}_j + \cdots + 0\mathbf{w}_n = \mathbf{w}_j$. This holds for every j, and the lists have the same length, so $\mathcal{B} = \mathcal{C}$.

Solution to Exercise 1.6.27

Uses. Definition 1.6.1 (Basis), Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.1.17 (No zero divisors), Example 1.4.11 (Matrices)
Hint 1. For (a), solve $a(1, 2, 0) + b(0, 1, 2) + c(2, 0, 1) = (x, y, z)$ by expressing a in terms of c, then b, and substituting into the last equation. Which element of F do you need to divide by?
Hint 2. For (b), look for a way to write $2 + x + x^2$ using the other two entries. For (c), which matrices can a combination of these three be?
Sketch. (a) The equations give $9c = 4x - 2y + z$, and cyclically for a and b. If $3 \neq 0$ in F, there is exactly one solution, so the list is a basis. If $3 = 0$, the three vectors add up to $(3, 3, 3) = \mathbf{0}$. (b) No: $2 + x + x^2 = 2\big(1 + x^2\big) + \big(x - x^2\big)$. (c) No: the combinations are the symmetric matrices, and U_{12} is not symmetric. The list is a basis of the symmetric matrices, though.

(a) For $(x, y, z) \in F^3$, the equation $a(1, 2, 0) + b(0, 1, 2) + c(2, 0, 1) = (x, y, z)$ reads

$$a + 2c = x, \quad 2a + b = y, \quad 2b + c = z.$$

Any solution has $a = x - 2c$, then $b = y - 2a = y - 2x + 4c$, and then the third equation becomes $2y - 4x + 8c + c = z$, that is,

$$9c = 4x - 2y + z.$$

Suppose $3 \neq 0$ in F. Then $9 = 3 \cdot 3 \neq 0$ by Proposition 1.1.17, so $c = (4x - 2y + z)/9$, and then

$$a = \frac{x + 4y - 2z}{9}, \quad b = \frac{-2x + y + 4z}{9}.$$

So there is at most one solution. These values are a solution: for instance $a + 2c = (x + 4y - 2z + 8x - 4y + 2z)/9 = x$, and the other two equations check the same way, as the formulas are cyclic shifts of each other. So every vector of F^3 has exactly one representation, and the list is a basis (Definition 1.6.1).

Suppose $3 = 0$ in F, as in $\mathbb{F}_3$. Then $(1, 2, 0) + (0, 1, 2) + (2, 0, 1) = (3, 3, 3) = \mathbf{0}$ is a dependence relation, so the list is not a basis.

(b) Not a basis: $2\left(1 + x^2\right) + \left(x - x^2\right) = 2 + x + x^2$, so $2\left(1 + x^2\right) + \left(x - x^2\right) - \left(2 + x + x^2\right) = \mathbf{0}$ is a dependence relation.

(c) Not a basis. Every linear combination

$$aU_{11} + b(U_{12} + U_{21}) + dU_{22} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

has equal off-diagonal entries, so U_{12} is not in the span, and the list does not span $\mathbb{R}^{2\times 2}$. (The display also shows that each symmetric matrix $\left(\begin{smallmatrix} a & b \\ b & d \end{smallmatrix}\right)$ is represented by exactly one combination, so the list is a basis of the subspace S of symmetric matrices of Example 1.4.11.)

Solution to Exercise 1.6.28

Uses. Definition 1.6.1 (Basis), Definition 1.6.14 (Coordinates), (1.6.1)
Hint 1. Solve for the coefficients directly. In (a), the combination $a(1, 1, 1) + b(0, 1, 1) + c(0, 0, 1)$ is $(a, a + b, a + b + c)$.
Hint 2. In (b), write the combination with coefficients $\alpha, \beta, \gamma, \delta$ as a single matrix and solve from the bottom right entry up.
Sketch. (a) $(a, a + b, a + b + c) = (x, y, z)$ has the unique solution $(x, y - x, z - y)$, which is $[(x, y, z)]_{\mathcal{B}}$; and $\mathbf{v} = (1, 1 + 2, 1 + 2 + 3) = (1, 3, 6)$. (b) The combination is $\left(\begin{smallmatrix} \alpha+\beta+\gamma+\delta & \beta+\gamma+\delta \\ \gamma+\delta & \delta \end{smallmatrix}\right)$, so $[A]_{\mathcal{C}} = (p - q, q - r, r - s, s)$.

(a) The combination with coefficients a, b, c represents $(a, a + b, a + b + c)$. This equals (x, y, z) if and only if $a = x$, $b = y - x$, and $c = z - y$. So each vector has exactly one representation, $\mathcal{B}$ is a basis, and

$$[(x, y, z)]_{\mathcal{B}} = (x, y - x, z - y).$$

By (1.6.1), the vector with coordinates $(1, 2, 3)$ is $1(1, 1, 1) + 2(0, 1, 1) + 3(0, 0, 1) = (1, 3, 6)$.

(b) The combination with coefficients $\alpha, \beta, \gamma, \delta$ represents

$$\begin{pmatrix} \alpha + \beta + \gamma + \delta & \beta + \gamma + \delta \\ \gamma + \delta & \delta \end{pmatrix}.$$

This equals $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ if and only if $\delta = s$, $\gamma = r - s$, $\beta = q - r$, and $\alpha = p - q$. So every $A \in F^{2\times 2}$ has exactly one representation, $\mathcal{C}$ is a basis, and $[A]_{\mathcal{C}} = (p - q, q - r, r - s, s)$.

Solution to Exercise 1.6.29

Uses. Proposition 1.6.20 (Coordinates translate questions about V into F^n), Definition 1.6.1 (Basis)
Hint 1. By Proposition 1.6.20, the list is a basis exactly when $(1, 0, 1), (0, 1, 1), (1, 1, c)$ is a basis of $\mathbb{R}^3$.
Hint 2. Solve $a(1, 0, 1) + b(0, 1, 1) + d(1, 1, c) = (x, y, z)$. You will need to divide by $c - 2$.

Sketch. The coordinate vectors are $(1,0,1),(0,1,1),(1,1,c)$. The equations $a+d=x$, $b+d=y$, $a+b+cd=z$ give $(c-2)d=z-x-y$, which has a unique solution for every right side exactly when $c\neq 2$. For $c=2$, $1+x+2x^2=\left(1+x^2\right)+\left(x+x^2\right)$.

In the standard basis $1, x, x^2$ the coordinate vectors are $(1,0,1)$, $(0,1,1)$, and $(1,1,c)$. By Proposition 1.6.20(c), the polynomials form a basis of $\mathbb{R}[x]_{\leq 2}$ if and only if these columns form a basis of $\mathbb{R}^3$. The equation $a(1,0,1)+b(0,1,1)+d(1,1,c)=(x,y,z)$ reads

$$a+d=x,\quad b+d=y,\quad a+b+cd=z.$$

Any solution has $a=x-d$ and $b=y-d$, and then the last equation becomes $(c-2)d=z-x-y$.

If $c\neq 2$, this forces $d=(z-x-y)/(c-2)$, and then a and b are determined; conversely, these values solve all three equations. So every column has exactly one representation, and the list is a basis.

If $c=2$, the columns satisfy $(1,0,1)+(0,1,1)-(1,1,2)=\mathbf{0}$, so they are dependent, and so are the polynomials: $1+x+2x^2=\left(1+x^2\right)+\left(x+x^2\right)$. So the list is a basis exactly when $c\neq 2$.

Solution to Exercise 1.6.30

Uses. Lemma 1.5.10 (Linear dependence lemma), Example 1.5.11 (Which entries are redundant), Theorem 1.6.10 (Every spanning list contains a basis)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. The inductive hypothesis is about *spanning* lists of length k. Is $\mathbf{v}_1,\ldots,\mathbf{v}_k$ one?
Hint 2. Test the step on the dependent spanning list $(1,0),(2,0),(0,1)$ of $\mathbb{R}^2$.
Sketch. The inductive hypothesis applies only to lists that span V, and deleting the last entry of a dependent spanning list can destroy spanning: $(1,0),(2,0),(0,1)$ spans $\mathbb{R}^2$, but $(1,0),(2,0)$ does not. The repair deletes the entry $\mathbf{v}_j$ given by Lemma 1.5.10, which keeps the span equal to V.

(a) The inductive hypothesis is a statement about lists of length k that span V. The argument applies it to $\mathbf{v}_1,\ldots,\mathbf{v}_k$ without checking that this list spans V, and it need not. In $\mathbb{R}^2$, the list $(1,0),(2,0),(0,1)$ spans $\mathbb{R}^2$ and is dependent (Example 1.5.11), but deleting its last entry leaves $(1,0),(2,0)$, whose span is the first axis. The claim itself is true (Theorem 1.6.10); only the choice of the deleted entry is wrong.

(b) Since $\mathcal{L}$ is dependent, Lemma 1.5.10 gives an index j such that deleting $\mathbf{v}_j$ from $\mathcal{L}$ does not change the span. The resulting list has length k and spans $\operatorname{span}(\mathcal{L})=V$, so the inductive hypothesis applies to it, and the rest of the argument goes through. (In the example, the lemma deletes $(2,0)$.)

Solution to Exercise 1.6.31

Uses. Definition 1.6.14 (Coordinates), (1.6.1), Non-example 1.6.4 (One property without the other)
Hint 1. How many linear combinations of $\mathcal{L}$ represent $\mathbf{0}$? What does "the coefficients" refer to?
Hint 2. For (b), consider $\Phi(a,b,d)=a(1,0)+b(0,1)+d(1,1)$. Then $\Phi(c(\mathbf{v}))=\mathbf{v}$. Which two elements of $\mathbb{R}^3$ does Φ send to $\mathbf{0}$?
Sketch. (a) $\mathcal{L}$ is dependent, so a vector has many representations, and "the" coefficients do not name one column; the proof's step "so the coordinates of $\mathbf{x}+\mathbf{y}$ are a_i+b_i" needs the backward direction of (1.6.1), which uses uniqueness. (b) $\Phi\circ c$ is the identity, so c is injective. Both $(0,0,0)$ and $(1,1,-1)$ represent $\mathbf{0}$, but $c(\mathbf{0})$ is only one of them, and the other is $c(\mathbf{v})$ for no $\mathbf{v}$.

(a) The list $\mathcal{L}$ is dependent (Non-example 1.6.4), so a vector of $\mathbb{R}^2$ is represented by many linear combinations of $\mathcal{L}$; for instance $(1,1)$ has coefficients $(1,1,0)$ and $(0,0,1)$. So "the coefficients" does not specify a column, and the student has not defined a function. In the proof of Proposition 1.6.19(a), the step that fails is the last one: from "$\mathbf{x}+\mathbf{y}=\sum(a_i+b_i)\mathbf{v}_i$" it concludes that (a_i+b_i) *is* the coordinate vector of $\mathbf{x}+\mathbf{y}$. That is the backward direction of (1.6.1), which holds because a basis allows only one representation. For $\mathcal{L}$ the column (a_i+b_i) is one representation among many, and it need not be the one the student's rule picks for $\mathbf{x}+\mathbf{y}$.

(b) Let $\Phi:\mathbb{R}^3\to\mathbb{R}^2$, $\Phi(a,b,d)=a(1,0)+b(0,1)+d(1,1)$. The hypothesis on c says $\Phi(c(\mathbf{v}))=\mathbf{v}$ for every $\mathbf{v}\in\mathbb{R}^2$. If $c(\mathbf{v})=c(\mathbf{v}')$, then $\mathbf{v}=\Phi(c(\mathbf{v}))=\Phi(c(\mathbf{v}'))=\mathbf{v}'$, so c is injective.

Now $\Phi(0,0,0) = \mathbf{0}$ and $\Phi(1,1,-1) = (1,0) + (0,1) - (1,1) = \mathbf{0}$. Let $\mathbf{a}$ be whichever of these two columns differs from $c(\mathbf{0})$. If $\mathbf{a} = c(\mathbf{v})$ for some $\mathbf{v}$, then $\mathbf{v} = \Phi(c(\mathbf{v})) = \Phi(\mathbf{a}) = \mathbf{0}$, so $\mathbf{a} = c(\mathbf{0})$, contradicting the choice of $\mathbf{a}$. So $\mathbf{a}$ is not a value of c, and c is not surjective. By contrast, the coordinate map of a basis is a bijection (Proposition 1.6.16).

Solution to Exercise 1.6.32

Uses. Definition 1.6.1 (Basis), Definition 1.6.8 (Standard bases), Proposition 1.1.17 (No zero divisors), Theorem 1.4.3 (Subspace criterion)
Hint 1. A symmetric matrix is determined by its entries on and above the diagonal. For each position (i,j) with $i < j$, use $U_{ij} + U_{ji}$.
Hint 2. For (b), what does $a_{ii} = -a_{ii}$ say when $1 + 1 \neq 0$? For (c), what is -1 in $\mathbb{F}_2$?
Sketch. (a) U_{ii} $(1 \leq i \leq n)$ and $U_{ij} + U_{ji}$ $(i < j)$: a combination with coefficients c has entries c_{ii} on the diagonal and c_{ij} in both positions (i,j) and (j,i), so each symmetric A is represented exactly by $c =$ its entries on and above the diagonal. Count $n + n(n-1)/2$. (b) $(1+1)a_{ii} = 0$ forces a zero diagonal; use $U_{ij} - U_{ji}$, $i < j$. (c) In $\mathbb{F}_2$, $-a = a$, so $K = S$, with the basis from (a).

(a) Consider the list consisting of the U_{ii} for $1 \leq i \leq n$ and the $U_{ij} + U_{ji}$ for $1 \leq i < j \leq n$, in any fixed order. Each entry is symmetric. A linear combination with coefficient c_{ii} on U_{ii} and c_{ij} on $U_{ij} + U_{ji}$ is the matrix with diagonal entries c_{ii} and with c_{ij} in both positions (i,j) and (j,i) for $i < j$. So it equals a given $A \in S$ if and only if $c_{ij} = a_{ij}$ for all $i \leq j$; these values work because $a_{ji} = a_{ij}$. Every $A \in S$ is represented by exactly one combination, and the list is a basis of S (Definition 1.6.1). Its length is $n + n(n-1)/2 = n(n+1)/2$.

(b) The set K is a subspace by Theorem 1.4.3: the zero matrix satisfies $0 = -0$, and the conditions $a_{ji} = -a_{ij}$ are preserved by entrywise sums and scalar multiples. For $A \in K$, the condition with $j = i$ gives $a_{ii} = -a_{ii}$, that is, $(1+1)a_{ii} = 0$, so $a_{ii} = 0$ by Proposition 1.1.17. Consider the matrices $U_{ij} - U_{ji}$ for $i < j$, which lie in K. A combination with coefficients c_{ij} has zero diagonal, c_{ij} in position (i,j), and $-c_{ij}$ in position (j,i), for $i < j$. It equals $A \in K$ if and only if $c_{ij} = a_{ij}$ for all $i < j$: the diagonal entries of A are 0, and $a_{ji} = -a_{ij}$. So this list of length $n(n-1)/2$ is a basis of K.

(c) In $\mathbb{F}_2$, $-a = a$ for every a, so the condition $a_{ji} = -a_{ij}$ is the condition $a_{ji} = a_{ij}$, and $K = S$. A basis is the one from (a), of length $n(n+1)/2$. The argument of (b) breaks down because $(1+1)a_{ii} = 0$ holds for every a_{ii}; for instance $U_{11} \in K$.

Solution to Exercise 1.6.33

Uses. Definition 1.6.1 (Basis), (1.6.1), Definition 1.6.8 (Standard bases), Theorem 1.4.3 (Subspace criterion)
Hint 1. A matrix in Z is determined by its top-left 2×2 block. Express the entries in row 3 and column 3 in terms of that block.
Hint 2. Compute every entry of $\sum_{k,l \leq 2} c_{kl} Z_{kl}$. Its top-left block is (c_{kl}).
Sketch. Each Z_{kl} lies in Z. The combination $\sum c_{kl} Z_{kl}$ has top-left block (c_{kl}), so at most one combination represents a given A, with $c_{kl} = a_{kl}$. That one works: its entries in column 3, row 3, and the corner are $-(c_{k1} + c_{k2})$, $-(c_{1l} + c_{2l})$, and $\sum c_{kl}$, and for $A \in Z$ the zero row and column sums give exactly these values.

First, Z is a subspace of $F^{3\times 3}$: the zero matrix lies in it, and if every row and column of A and of B sums to 0, the same holds for $A + B$ and cA, since sums of entries respect the entrywise operations (Theorem 1.4.3). Each Z_{kl} lies in Z: its nonzero entries are 1 in positions (k,l) and $(3,3)$ and -1 in positions $(k,3)$ and $(3,l)$, so every row and every column contains either no nonzero entry or one 1 and one -1.

Let $M = \sum_{k,l \in \{1,2\}} c_{kl} Z_{kl}$. Its entry in position (k,l) with $k,l \leq 2$ is c_{kl}, its entry $(k,3)$ is $-(c_{k1} + c_{k2})$, its entry $(3,l)$ is $-(c_{1l} + c_{2l})$, and its entry $(3,3)$ is $c_{11} + c_{12} + c_{21} + c_{22}$.

Now let $A \in Z$. If $M = A$, comparing the top-left blocks gives $c_{kl} = a_{kl}$, so at most one combination represents A. Conversely, take $c_{kl} = a_{kl}$. Since rows 1 and 2 of A sum to 0, $a_{k3} = -(a_{k1} + a_{k2})$, which is the $(k,3)$ entry of M. Since columns 1 and 2 sum to 0, $a_{3l} = -(a_{1l} + a_{2l})$, the $(3,l)$ entry of M. Since row 3 sums to 0,

$$a_{33} = -(a_{31} + a_{32}) = (a_{11} + a_{21}) + (a_{12} + a_{22}),$$

the $(3,3)$ entry of M. So $M = A$. Every $A \in Z$ is represented by exactly one linear combination, the list is a basis of Z (Definition 1.6.1), and by (1.6.1) the coordinates of A are $(a_{11}, a_{12}, a_{21}, a_{22})$.

Solution to Exercise 1.6.34

> *Uses.* Exercise 1.3.33, Proposition 1.5.24 (Distinct degrees), Example 1.4.9 (Polynomials), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 1.3.18 (Comparing spans), Example 1.3.11 (Small spans)
> *Hint 1.* For (a), see Exercise 1.3.33 and Proposition 1.5.24. For (b), which subspace contains every p_k?
> *Hint 2.* For (c), start with x^n and add x^n to each lower power of x.
> *Sketch.* (a) Spanning by Exercise 1.3.33, independence by Proposition 1.5.24. (b) All p_k lie in the subspace $F[x]_{\leq n-1}$, so their span does too, and x^n is missing. (c) $x^n, x^n+1, x^n+x, \ldots, x^n+x^{n-1}$: each $x^k = \left(x^n + x^k\right) - x^n$ is in the span, and in a relation the coefficient of x^k for $k < n$ is the coefficient of $x^n + x^k$.

(a) The list spans $F[x]_{\leq n}$ by Exercise 1.3.33(a). Its entries are nonzero (they have degrees $k \geq 0$) and have distinct degrees, so it is independent by Proposition 1.5.24. So it is a basis.

(b) Each p_k has degree at most n and not equal to n, so $p_k \in F[x]_{\leq n-1}$, a subspace (Example 1.4.9). By Proposition 1.4.14(b), $\operatorname{span}(p_0, \ldots, p_n) \subseteq F[x]_{\leq n-1}$, which does not contain x^n. So the list does not span $F[x]_{\leq n}$.

(c) Let $q_n = x^n$ and $q_k = x^n + x^k$ for $0 \leq k < n$, and consider $q_n, q_0, q_1, \ldots, q_{n-1}$. Each entry has degree n, since $k < n$.

Spanning. $x^n = q_n$ and $x^k = q_k - q_n$ for $k < n$, so $1, x, \ldots, x^n$ lie in the span, and the span is $F[x]_{\leq n}$ by Proposition 1.3.18 and Example 1.3.11.

Independence. Suppose that $cq_n + \sum_{k=0}^{n-1} c_k q_k = \mathbf{0}$. The left side is

$$(c + c_0 + \cdots + c_{n-1})x^n + c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}.$$

Its coefficient of x^k for $k < n$ is c_k, so every $c_k = 0$, and then its coefficient of x^n is $c = 0$. So the list is a basis. Degrees alone therefore do not decide whether a list is a basis: (a) is a sufficient condition, not a necessary one.

Solution to Exercise 1.6.35

> *Uses.* Proposition 1.6.2 (Bases are independent spanning lists), Corollary 1.5.13 (Independence, one entry at a time), Proposition 1.5.18 (Operations that preserve independence), Definition 1.5.1 (Linear independence), Example 1.6.11 (Sifting a list of polynomials)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* Test (a) with the standard basis of $\mathbb{R}^2$ and a line that is not an axis. For (b), the first entry of $\mathcal{C}$ comes after all of $\mathcal{B}$.
> *Hint 2.* For (d), what does "at most one" say about vectors outside the span? For (e), look at Example 1.6.11.
> *Sketch.* (a) False: $U = \operatorname{span}((1,1))$ in $\mathbb{R}^2$ contains no sublist of $\mathbf{e}_1, \mathbf{e}_2$ that spans it. (b) True: $\mathcal{C} \neq ()$ and its first entry lies in $\operatorname{span}(\mathcal{B}) = V$, so Corollary 1.5.13 makes $\mathcal{B}, \mathcal{C}$ dependent. (c) True: scaling by nonzero scalars preserves span and independence (Proposition 1.5.18(b)). (d) False: "at most one" is independence; $\mathbf{e}_1$ in $\mathbb{R}^2$. (e) False: $(1,0), (0,1), (1,1)$ contains three bases of $\mathbb{R}^2$.

(a) False. In $\mathbb{R}^2$, let $\mathcal{B} = \mathbf{e}_1, \mathbf{e}_2$ and $U = \operatorname{span}((1,1))$. The sublists of $\mathcal{B}$ are $()$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathcal{B}$. The last three contain $\mathbf{e}_1$ or $\mathbf{e}_2$, which are not in U (they are not multiples of $(1,1)$), so they are not lists in U at all; and $()$ spans $\{\mathbf{0}\} \neq U$.

(b) True. Since $V \neq \{\mathbf{0}\}$, neither $\mathcal{B}$ nor $\mathcal{C}$ is the empty list (it spans only $\{\mathbf{0}\}$). Let $\mathbf{w}$ be the first entry of $\mathcal{C}$. In $\mathcal{B}, \mathcal{C}$ the entries before $\mathbf{w}$ are those of $\mathcal{B}$, and $\mathbf{w} \in V = \operatorname{span}(\mathcal{B})$. So $\mathcal{B}, \mathcal{C}$ is dependent by Corollary 1.5.13, and it is not a basis.

(c) True. Replacing $\mathbf{v}_j$ by $c_j \mathbf{v}_j$ with $c_j \neq 0$ preserves both the span and independence (Proposition 1.5.18(b)). Doing this for $j = 1, \ldots, n$ in turn produces a list that spans V and is independent, which is a basis by Proposition 1.6.2.

(d) False. Every vector outside $\operatorname{span}(\mathcal{L})$ is represented by no combination, so the hypothesis says exactly that every vector of $\operatorname{span}(\mathcal{L})$ is represented by at most one combination, which is independence (Definition 1.5.1). Independent lists need not span: in $\mathbb{R}^2$, the list $\mathbf{e}_1$ satisfies the hypothesis and is not a basis.

(e) False. In $\mathbb{R}^2$, the spanning list $(1,0),(0,1),(1,1)$ has the three sublists $(1,0),(0,1)$ and $(1,0),(1,1)$ and $(0,1),(1,1)$, each of which is a basis (check spanning and independence as in Example 1.6.3(a)). See also Example 1.6.11.

Solution to Exercise 1.6.36

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Corollary 1.3.19 (Equal spans), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.6.16 (The coordinate map is a bijection), Exercise 1.3.34
Hint 1. If $a_j = 0$, then $\mathbf{w}$ is a combination of the other entries of $\mathcal{B}'$. If $a_j \neq 0$, solve $\mathbf{w} = \sum a_i \mathbf{v}_i$ for $\mathbf{v}_j$.
Hint 2. For independence when $a_j \neq 0$, substitute $\mathbf{w} = \sum_i a_i \mathbf{v}_i$ into a relation of $\mathcal{B}'$ and use the independence of $\mathcal{B}$.
Hint 3. For (b), which column is $[\mathbf{0}]_{\mathcal{B}}$?
Sketch. If $a_j = 0$, then $\mathbf{w} = \sum_{i\neq j} a_i \mathbf{v}_i$, which gives a relation of $\mathcal{B}'$ with coefficient -1 on $\mathbf{w}$. If $a_j \neq 0$, then $\mathbf{v}_j = a_j^{-1}\big(\mathbf{w} - \sum_{i\neq j} a_i \mathbf{v}_i\big)$, so $\mathcal{B}'$ spans V; and a relation $\sum_{i\neq j} c_i \mathbf{v}_i + c\mathbf{w} = \mathbf{0}$ becomes $\sum_{i\neq j}(c_i + ca_i)\mathbf{v}_i + ca_j \mathbf{v}_j = \mathbf{0}$, so $ca_j = 0$, $c = 0$, and then all $c_i = 0$. (b) $[\mathbf{w}]_{\mathcal{B}} \neq \mathbf{0}$ since the coordinate map is injective, so some $a_j \neq 0$.

(a) Write $\mathbf{w} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$.

Suppose $a_j = 0$. Then $\mathbf{w} = \sum_{i\neq j} a_i \mathbf{v}_i$, and the combination of $\mathcal{B}'$ with coefficients a_i on $\mathbf{v}_i$ $(i \neq j)$ and -1 on $\mathbf{w}$ represents $\mathbf{0}$. It is nontrivial, so $\mathcal{B}'$ is dependent and not a basis.

Suppose $a_j \neq 0$. *Spanning.* Solving for $\mathbf{v}_j$,

$$\mathbf{v}_j = a_j^{-1}\mathbf{w} + \sum_{i\neq j}\big(-a_j^{-1}a_i\big)\mathbf{v}_i \in \operatorname{span}(\mathcal{B}').$$

The other entries of $\mathcal{B}$ are entries of $\mathcal{B}'$, so $\mathcal{B} \subseteq \operatorname{span}(\mathcal{B}')$, and also $\mathcal{B}' \subseteq V = \operatorname{span}(\mathcal{B})$. By Corollary 1.3.19, $\operatorname{span}(\mathcal{B}') = \operatorname{span}(\mathcal{B}) = V$. (This is also Exercise 1.3.34, after moving $\mathbf{v}_j$ to the end; its hypothesis that $\mathbf{w}$ is not in the span of the other entries holds because $a_j \neq 0$ and coordinates are unique.) *Independence.* Suppose $\sum_{i\neq j} c_i \mathbf{v}_i + c\mathbf{w} = \mathbf{0}$. Substituting for $\mathbf{w}$ and collecting terms,

$$\sum_{i\neq j}(c_i + ca_i)\mathbf{v}_i + ca_j\mathbf{v}_j = \mathbf{0}.$$

Since $\mathcal{B}$ is independent, $ca_j = 0$, so $c = 0$ because $a_j \neq 0$, and then $c_i = c_i + ca_i = 0$ for all $i \neq j$. So $\mathcal{B}'$ is independent, and it is a basis by Proposition 1.6.2.

(b) Since $[\mathbf{0}]_{\mathcal{B}} = \mathbf{0}$ and the coordinate map is injective (Proposition 1.6.16), $\mathbf{w} \neq \mathbf{0}$ implies $[\mathbf{w}]_{\mathcal{B}} \neq \mathbf{0}$, so $a_j \neq 0$ for some j. By (a), replacing $\mathbf{v}_j$ with $\mathbf{w}$ gives a basis. Repeating such replacements is the idea behind the exchange lemma of Section 4.1 (Lemma 4.1.2).

Solution to Exercise 1.6.37

Uses. Definition 1.6.1 (Basis), (1.6.1), Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 1.6.16 (The coordinate map is a bijection)
Hint 1. If $f = [\cdot]_{\mathcal{C}}$ with $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_n$, then $f(\mathbf{c}_j) = \mathbf{e}_j$. So there is only one candidate for $\mathcal{C}$.
Hint 2. Let $\mathbf{c}_j = f^{-1}(\mathbf{e}_j)$. To show $\mathcal{C}$ is a basis, write $f(\mathbf{v}) = (a_1, \ldots, a_n) = \sum a_i \mathbf{e}_i$ and apply f^{-1}, using that f respects linear combinations.
Sketch. Uniqueness: $[\mathbf{c}_j]_{\mathcal{C}} = \mathbf{e}_j$ forces $\mathbf{c}_j = f^{-1}(\mathbf{e}_j)$. Existence: with this $\mathcal{C}$, $f(\sum b_i \mathbf{c}_i) = \sum b_i \mathbf{e}_i = (b_1, \ldots, b_n)$. Since f is injective, $\sum b_i \mathbf{c}_i = \mathbf{v}$ holds exactly when $(b_1, \ldots, b_n) = f(\mathbf{v})$; so each $\mathbf{v}$ has exactly one representation, with coefficients $f(\mathbf{v})$.

Uniqueness. Suppose $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_n$ is a basis with $f(\mathbf{v}) = [\mathbf{v}]_{\mathcal{C}}$ for all $\mathbf{v}$. Then $f(\mathbf{c}_j) = [\mathbf{c}_j]_{\mathcal{C}} = \mathbf{e}_j$, so $\mathbf{c}_j = f^{-1}(\mathbf{e}_j)$ for each j. So $\mathcal{C}$ is determined by f.

Existence. Let $\mathbf{c}_j = f^{-1}(\mathbf{e}_j)$ and $\mathcal{C} = \mathbf{c}_1, \ldots, \mathbf{c}_n$. For all $b_1, \ldots, b_n \in F$, since f respects linear combinations,

$$f(b_1\mathbf{c}_1 + \cdots + b_n\mathbf{c}_n) = b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n = (b_1, \ldots, b_n).$$

Let $\mathbf{v} \in V$. Since f is injective, $\sum_i b_i \mathbf{c}_i = \mathbf{v}$ holds if and only if $f\left(\sum_i b_i \mathbf{c}_i\right) = f(\mathbf{v})$, that is, if and only if $(b_1, \ldots, b_n) = f(\mathbf{v})$. So exactly one linear combination of $\mathcal{C}$ represents $\mathbf{v}$, namely the one with coefficients $f(\mathbf{v})$. Hence $\mathcal{C}$ is a basis (Definition 1.6.1) and $[\mathbf{v}]_{\mathcal{C}} = f(\mathbf{v})$ (Definition 1.6.14).

Conversely, every coordinate map is a bijection that respects linear combinations, by Proposition 1.6.16 and Proposition 1.6.19(c). In the language of Section 4.5, the coordinate maps $V \to F^n$ are the isomorphisms $V \to F^n$.

Solution to Exercise 1.6.38

Uses. Exercise 1.5.43, Proposition 1.6.2 (Bases are independent spanning lists), Proposition 1.6.7 (Standard bases)
Hint 1. By Exercise 1.5.43(a), an independent list of length k spans a set of exactly q^k elements. How many elements does F^n have?
Hint 2. For (b), (a) turns the count of bases into the count of independent lists of length n, which Exercise 1.5.43(b) computes.
Sketch. (a) If $\mathcal{L}$ is independent of length k, then $|\operatorname{span}(\mathcal{L})| = q^k$, while $|F^n| = q^n$. So an independent list spans F^n exactly when $q^k = q^n$, that is, $k = n$. (b) $(q^n - 1)(q^n - q)\cdots\left(q^n - q^{n-1}\right)$. For $\mathbb{F}_2^2$ this is $3 \cdot 2 = 6$: the ordered pairs of distinct vectors among $(1,0), (0,1), (1,1)$.

(a) F^n has q^n elements. Let $\mathcal{L}$ be an independent list of length k. By Exercise 1.5.43(a), $\operatorname{span}(\mathcal{L})$ has q^k elements. Since $\operatorname{span}(\mathcal{L}) \subseteq F^n$, it equals F^n if and only if $q^k = q^n$, which holds if and only if $k = n$, because $q \geq 2$. So an independent list spans F^n exactly when its length is n. By Proposition 1.6.2, the bases of F^n are the independent lists of length n.

(b) By (a) and Exercise 1.5.43(b), F^n has exactly

$$(q^n - 1)(q^n - q)\left(q^n - q^2\right)\cdots\left(q^n - q^{n-1}\right)$$

bases. For $q = 2$ and $n = 2$ this is $(4-1)(4-2) = 6$. An independent list of length 2 in $\mathbb{F}_2^2$ consists of a nonzero vector and a vector outside its span; the span of a nonzero $\mathbf{v}$ is $\{\mathbf{0}, \mathbf{v}\}$. So the bases are the ordered pairs of distinct vectors from $(1,0), (0,1), (1,1)$:

$$\begin{array}{lll} (1,0),(0,1) & (0,1),(1,0) & (1,0),(1,1) \\ (1,1),(1,0) & (0,1),(1,1) & (1,1),(0,1). \end{array}$$

Part (a) shows that over a finite field every basis of F^n has length n. Section 4.1 proves this over every field (Theorem 4.1.1), where counting elements is not available.

Solution to Exercise 1.6.39

Uses. Theorem 1.4.3 (Subspace criterion), Definition 1.6.1 (Basis), (1.6.1), Proposition 1.6.2 (Bases are independent spanning lists), induction on k
Hint 1. For (a), check the subspace criterion one condition at a time. For the second claim, prove by induction on k that $s_k = s_{k+1} = 0$.
Hint 2. For (b), given $\mathbf{s} \in V$, the sequence $\mathbf{s} - s_0\mathbf{g} - s_1\mathbf{f}$ lies in V. What are its first two terms?
Hint 3. For (c), $\varphi^{k+2} = \varphi^k\varphi^2 = \varphi^k(\varphi + 1)$. To represent $\mathbf{s}$, choose α, β with $\alpha + \beta = s_0$ and $\alpha\varphi + \beta\psi = s_1$, and argue as in (b).
Sketch. (a) The recurrence is preserved by sums and multiples, and the zero sequence satisfies it; if $s_0 = s_1 = 0$, induction gives $s_k = 0$ for all k. (b) $\mathbf{s} - s_0\mathbf{g} - s_1\mathbf{f} \in V$ starts $0, 0$, so it is $\mathbf{0}$; and a combination $c\mathbf{g} + d\mathbf{f}$ starts c, d, so the coefficients are forced. (c) $\mathbf{a}, \mathbf{b} \in V$ because $\varphi^2 = \varphi + 1$ and $\psi^2 = \psi + 1$. Since $\varphi \neq \psi$, the system $\alpha + \beta = s_0$, $\alpha\varphi + \beta\psi = s_1$ has exactly one solution, and by (a) the coefficients represent $\mathbf{s}$. For $\mathbf{f}$: $\alpha = 1/\sqrt{5}$, $\beta = -1/\sqrt{5}$.

(a) The zero sequence satisfies $0 = 0 + 0$. If $\mathbf{s}, \mathbf{t} \in V$ and $c \in \mathbb{R}$, then $s_{k+2} + t_{k+2} = (s_{k+1} + t_{k+1}) + (s_k + t_k)$ and $cs_{k+2} = cs_{k+1} + cs_k$ for all k, since the operations of $\mathbb{R}^X$ are termwise. So V is a subspace by Theorem 1.4.3. Now let $\mathbf{s} \in V$ with $s_0 = s_1 = 0$. We show by induction on k that $s_k = s_{k+1} = 0$. For $k = 0$ this is the hypothesis. If $s_k = s_{k+1} = 0$, then $s_{k+2} = s_{k+1} + s_k = 0$, so $s_{k+1} = s_{k+2} = 0$. Hence $s_k = 0$ for all k, and $\mathbf{s} = \mathbf{0}$.

(b) (The sequences $\mathbf{g}$ and $\mathbf{f}$ exist and are unique: choose the first two terms and define the rest by the recurrence.) Let $\mathbf{s} \in V$ and $c, d \in \mathbb{R}$. The combination $c\mathbf{g} + d\mathbf{f}$ has first two terms c and d. So if it represents

$\mathbf{s}$, then $c = s_0$ and $d = s_1$. Conversely, $\mathbf{t} = \mathbf{s} - s_0\mathbf{g} - s_1\mathbf{f}$ lies in V, since V is a subspace, and $t_0 = s_0 - s_0 = 0$, $t_1 = s_1 - s_1 = 0$. By (a), $\mathbf{t} = \mathbf{0}$, that is, $\mathbf{s} = s_0\mathbf{g} + s_1\mathbf{f}$. So $\mathbf{s}$ is represented by exactly one combination, $\mathbf{g}, \mathbf{f}$ is a basis of V, and $[\mathbf{s}]_{(\mathbf{g},\mathbf{f})} = (s_0, s_1)$.

(c) Since $\varphi^2 = \varphi + 1$, we get $\varphi^{k+2} = \varphi^k\varphi^2 = \varphi^{k+1} + \varphi^k$ for all k, so $\mathbf{a} \in V$; likewise $\mathbf{b} \in V$. Let $\mathbf{s} \in V$. A combination $\alpha\mathbf{a} + \beta\mathbf{b}$ has first two terms $\alpha + \beta$ and $\alpha\varphi + \beta\psi$. The system

$$\alpha + \beta = s_0, \quad \alpha\varphi + \beta\psi = s_1$$

has exactly one solution, because $\varphi - \psi = \sqrt{5} \neq 0$: subtracting ψ times the first equation from the second gives $\alpha = (s_1 - \psi s_0)/\sqrt{5}$, and then $\beta = s_0 - \alpha$. For these α, β, the sequence $\mathbf{s} - \alpha\mathbf{a} - \beta\mathbf{b}$ lies in V and starts with $0, 0$, so it is $\mathbf{0}$ by (a). For any other α, β, the combination differs from $\mathbf{s}$ in one of the first two terms. So $\mathbf{s}$ has exactly one representation, and $\mathbf{a}, \mathbf{b}$ is a basis of V.

For $\mathbf{f}$, $s_0 = 0$ and $s_1 = 1$ give $\alpha = 1/\sqrt{5}$ and $\beta = -1/\sqrt{5}$. So $[\mathbf{f}]_{(\mathbf{a},\mathbf{b})} = \left(1/\sqrt{5}, -1/\sqrt{5}\right)$, which means $\mathbf{f} = (\mathbf{a} - \mathbf{b})/\sqrt{5}$. Comparing k-th terms, $f_k = \left(\varphi^k - \psi^k\right)/\sqrt{5}$.

The same space has two useful bases. In $\mathbf{g}, \mathbf{f}$ the coordinates are the initial values, and in $\mathbf{a}, \mathbf{b}$ the recurrence acts on each basis vector by a single multiplication, which is what makes the closed formula possible. Choosing such a basis is the idea of diagonalization (Section 7.4).

Solution to Exercise 1.6.40

Uses. Exercise 1.6.33, Theorem 1.4.3 (Subspace criterion), Definition 1.6.1 (Basis), (1.6.1), Proposition 1.3.18 (Comparing spans), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Hint 1. For (a), if $A \in S$ has common sum s, what are the row and column sums of $A - sI$?
Hint 2. For uniqueness in (a), compute the first row sum of $c_0 I + \sum c_{kl} Z_{kl}$; then use the uniqueness in Exercise 1.6.33.
Hint 3. For (c), it is enough to write I and the four Z_{kl} as combinations of permutation matrices. Each Z_{kl} is a difference of two of them. For the relation, add up the three permutation matrices that place the 1 of row 1 in different columns "cyclically".
Sketch. (a) Line sums are preserved by sums and scalar multiples, so S is a subspace. For $A \in S$ with sum s, $A - sI \in Z$, so $A = sI + \sum (A - sI)_{kl} Z_{kl}$ by Exercise 1.6.33. Conversely, the first row sum of $c_0 I + \sum c_{kl} Z_{kl}$ is c_0, which forces $c_0 = s$, and then the c_{kl} are the $\mathcal{Z}$-coordinates of $A - sI$. (b) $s(J) = 3$, so $[J] = (3, -2, 1, 1, -2)$; $J \in Z$ iff $3 = 0$ in F. (c) With T_k the permutation matrix fixing position k and C, C' the two cyclic ones: $Z_{11} = I - T_2$, $Z_{22} = I - T_1$, $Z_{12} = T_3 - C'$, $Z_{21} = T_3 - C$. Relation: $I + C + C' = J = T_1 + T_2 + T_3$.

(a) *Subspace.* The zero matrix lies in S with $s = 0$. If $A, B \in S$ and $c \in F$, then every row and column of $A + B$ sums to $s(A) + s(B)$, and every row and column of cA sums to $cs(A)$, because sums of entries respect the entrywise operations. So S is a subspace by Theorem 1.4.3. The matrix I lies in S with $s(I) = 1$, and each Z_{kl} lies in $Z \subseteq S$.

Existence. Let $A \in S$ with $s = s(A)$. Every row and column of $A - sI$ sums to $s - s = 0$, so $A - sI \in Z$. By Exercise 1.6.33,

$$A - sI = \sum_{k,l \in \{1,2\}} (A - sI)_{kl} Z_{kl},$$

where $(A - sI)_{kl}$ is $a_{kl} - s$ for $k = l$ and a_{kl} for $k \neq l$. So A is represented with the coefficients claimed.

Uniqueness. Suppose $A = c_0 I + \sum_{k,l} c_{kl} Z_{kl}$. Every row of a matrix in Z sums to 0, so the first row of the right side sums to c_0, and hence $c_0 = s$. Then $\sum_{k,l} c_{kl} Z_{kl} = A - sI$, and the uniqueness in Exercise 1.6.33 gives $c_{kl} = (A - sI)_{kl}$. So every $A \in S$ has exactly one representation, $I, \mathcal{Z}$ is a basis of S, and the coordinates are as stated ((1.6.1)).

(b) Every row and column of J sums to $3 = 1 + 1 + 1$, so $J \in S$ with $s(J) = 3$, and by (a)

$$[J]_{(I,\mathcal{Z})} = (3, \quad 1 - 3, \quad 1, \quad 1, \quad 1 - 3) = (3, -2, 1, 1, -2).$$

The matrix J lies in Z exactly when its line sums are 0, that is, when $3 = 0$ in F, as in $\mathbb{F}_3$. In that case its first coordinate is 0, and its coordinates are $(0, 1, 1, 1, 1)$.

(c) Each permutation matrix lies in S, with every line sum equal to 1. There are six of them: I; the two "cyclic" ones

$$C = U_{12} + U_{23} + U_{31}, \quad C' = U_{13} + U_{21} + U_{32};$$

and the three that keep exactly one 1 of I in place,

$$T_1 = U_{11} + U_{23} + U_{32}, \quad T_2 = U_{13} + U_{22} + U_{31}, \quad T_3 = U_{12} + U_{21} + U_{33}.$$

Comparing entries,

$$Z_{11} = I - T_2, \quad Z_{22} = I - T_1, \quad Z_{12} = T_3 - C', \quad Z_{21} = T_3 - C.$$

For instance, $I - T_2 = (U_{11} + U_{22} + U_{33}) - (U_{13} + U_{22} + U_{31}) = U_{11} - U_{13} - U_{31} + U_{33} = Z_{11}$. So the basis $I, \mathcal{Z}$ of S lies in the span of the permutation matrices, and $S \subseteq$ that span by Proposition 1.3.18. Conversely, the permutation matrices lie in the subspace S, so their span is contained in S (Proposition 1.4.14). Hence they span S.

The matrices I, C, C' have their 1's in disjoint positions covering all nine, and so do T_1, T_2, T_3. Hence

$$I + C + C' = J = T_1 + T_2 + T_3,$$

and $I + C + C' - T_1 - T_2 - T_3 = 0$ is a dependence relation, with coefficient 1 on I, over every field. So six matrices span S, which has a basis of length 5, and they are dependent; Section 4.1 shows that this had to happen.

Solution to Exercise 1.6.41

Uses. Proposition 1.2.19 (Restriction of scalars), Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.1 (Vector space), Lemma 1.3.9 (Arithmetic of linear combinations)
Hint 1. For spanning, write $\mathbf{v} = \sum_j a_j \mathbf{v}_j$ with $a_j \in E$, and then write each a_j as a K-combination of the α_i.
Hint 2. For independence, group a K-relation $\sum_{i,j} b_{ij}\alpha_i \mathbf{v}_j = \mathbf{0}$ as $\sum_j \big(\sum_i b_{ij}\alpha_i\big)\mathbf{v}_j = \mathbf{0}$. Use independence over E, then over K.
Sketch. Spanning: $\mathbf{v} = \sum_j a_j \mathbf{v}_j$ with $a_j = \sum_i b_{ij}\alpha_i$, $b_{ij} \in K$, so $\mathbf{v} = \sum_{i,j} b_{ij}(\alpha_i \mathbf{v}_j)$ by (A6) and (A8). Independence: a K-relation regroups as $\sum_j c_j \mathbf{v}_j = \mathbf{0}$ with $c_j = \sum_i b_{ij}\alpha_i \in E$; independence of the $\mathbf{v}_j$ over E gives each $c_j = 0$, and independence of the α_i over K gives each $b_{ij} = 0$. (b) $(1, 0), (i, 0), (0, 1), (0, i)$, and coordinates $(1, 2, 3, -1)$.

Scalar multiplication of K on V is that of E, restricted to K, and the same holds on E, where it is multiplication in E. So the vector space axioms of V over E apply whenever the scalars lie in K or E.

(a) *Spanning.* Let $\mathbf{v} \in V$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ spans V over E, $\mathbf{v} = \sum_j a_j \mathbf{v}_j$ with $a_j \in E$. Since $\alpha_1, \ldots, \alpha_m$ spans E over K, each $a_j = \sum_i b_{ij}\alpha_i$ with $b_{ij} \in K$. By (A8) and (A6) in the E-vector space V,

$$a_j \mathbf{v}_j = \left(\sum_{i=1}^{m} b_{ij}\alpha_i\right)\mathbf{v}_j = \sum_{i=1}^{m} b_{ij}\big(\alpha_i \mathbf{v}_j\big),$$

so $\mathbf{v} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ij}\big(\alpha_i \mathbf{v}_j\big)$ is a K-linear combination of the list.

Independence. Suppose that $\sum_{j=1}^{n}\sum_{i=1}^{m} b_{ij}\big(\alpha_i \mathbf{v}_j\big) = \mathbf{0}$ with all $b_{ij} \in K$. Regrouping with the same axioms, $\sum_j c_j \mathbf{v}_j = \mathbf{0}$, where $c_j = \sum_i b_{ij}\alpha_i \in E$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is independent over E, every $c_j = 0$ (Theorem 1.5.5). So for each j, $\sum_i b_{ij}\alpha_i = 0$ is a K-relation of $\alpha_1, \ldots, \alpha_m$, and independence over K gives $b_{ij} = 0$ for all i. So the list is independent over K, and it is a basis by Proposition 1.6.2.

(b) Take $K = \mathbb{R}$, $E = \mathbb{C}$ with basis $1, i$ (Example 1.6.5), and $V = \mathbb{C}^2$ with the standard basis $\mathbf{e}_1, \mathbf{e}_2$ over $\mathbb{C}$. The basis of (a) is

$$(1, 0), \quad (i, 0), \quad (0, 1), \quad (0, i).$$

Since $(1+2i, 3-i) = 1(1,0) + 2(i,0) + 3(0,1) + (-1)(0,i)$ with real coefficients, its coordinates are $(1,2,3,-1)$ by (1.6.1). In general, the coordinates of a vector of $\mathbb{C}^n$ in this basis over $\mathbb{R}$ list the real and imaginary parts of its entries.

Solution to Exercise 1.6.42

Uses. Exercise 1.5.43, Proposition 1.6.2 (Bases are independent spanning lists), Definition 1.2.9 (Function spaces), Proposition 1.3.18 (Comparing spans)
Hint 1. Count: V has 2^{2^n} elements, and the list has length 2^n. By Exercise 1.5.43(a), it is enough to show that the list spans V.
Hint 2. For $\mathbf{a} \in X$, let $\delta_{\mathbf{a}}$ be 1 at $\mathbf{a}$ and 0 elsewhere. Every f is a combination of the $\delta_{\mathbf{a}}$. Show that $\delta_{\mathbf{a}}(\mathbf{y}) = \prod_{i=1}^{n}(y_i + a_i + 1)$.
Hint 3. Expand that product of n binomials. Each term is a constant times $\prod_{i\in S} y_i$ for some S.
Sketch. Every $f = \sum_{\mathbf{a}} f(\mathbf{a})\delta_{\mathbf{a}}$. In $\mathbb{F}_2$, $y_i + a_i + 1 = 1$ exactly when $y_i = a_i$, so $\delta_{\mathbf{a}}(\mathbf{y}) = \prod_i((a_i+1)+y_i)$, and expanding gives $\sum_S\big(\prod_{i\notin S}(a_i+1)\big)m_S(\mathbf{y})$. So the m_S span V. The span of a list of length 2^n then has 2^{2^n} elements, so by Exercise 1.5.43(a) the list is independent, hence a basis. The coordinates in this basis give the unique expression.

Write $N = 2^n$ for the number of subsets S, and fix a listing $m_{S_1}, \dots, m_{S_N}$. The space V has 2^N elements, since a function $X \to \mathbb{F}_2$ chooses one of two values at each of the N points of X.

Spanning. For $\mathbf{a} \in X$, let $\delta_{\mathbf{a}} \in V$ be 1 at $\mathbf{a}$ and 0 elsewhere. Every $f \in V$ satisfies $f = \sum_{\mathbf{a}\in X} f(\mathbf{a})\delta_{\mathbf{a}}$, since both sides take the value $f(\mathbf{y})$ at each $\mathbf{y}$ (only the term with $\mathbf{a} = \mathbf{y}$ contributes). So by Proposition 1.3.18 it suffices to show that each $\delta_{\mathbf{a}}$ lies in the span of the m_S.

In $\mathbb{F}_2$, the element $y_i + a_i + 1$ is 1 if $y_i = a_i$ and 0 if $y_i \neq a_i$ (as $a_i + a_i = 0$ and $1 + 1 = 0$). A product of elements of $\mathbb{F}_2$ is 1 exactly when every factor is 1, so

$$\delta_{\mathbf{a}}(\mathbf{y}) = \prod_{i=1}^{n}((a_i + 1) + y_i) \quad \text{for all } \mathbf{y} \in X.$$

Expanding a product of n sums $(u_i + y_i)$ by distributivity (induction on n) gives the sum, over all subsets $S \subseteq \{1, \dots, n\}$, of the products that take y_i from the factors with $i \in S$ and u_i from the others:

$$\prod_{i=1}^{n}(u_i + y_i) = \sum_{S\subseteq\{1,\dots,n\}}\left(\prod_{i\notin S} u_i\right)\left(\prod_{i\in S} y_i\right).$$

With $u_i = a_i + 1$, this says that $\delta_{\mathbf{a}} = \sum_S c_S m_S$ with $c_S = \prod_{i\notin S}(a_i + 1) \in \mathbb{F}_2$, an identity of functions because it holds at every $\mathbf{y}$. So the list spans V.

Independence. By Exercise 1.5.43(a), a list of length N over $\mathbb{F}_2$ is independent if and only if its span has exactly 2^N elements. Our list has length N, and its span is V, which has 2^N elements. So the list is independent, and it is a basis by Proposition 1.6.2.

The deduction. For $f \in V$, Definition 1.6.1 gives exactly one list of coefficients $c_S \in \mathbb{F}_2$ with $f = \sum_S c_S m_S$. Over $\mathbb{F}_2$ a coefficient is 0 or 1, so this says that f is the sum of the m_S with $c_S = 1$, and the set of these S is uniquely determined. For instance, for $n = 2$, the function that is 1 exactly at $(0,0)$ is $\delta_{(0,0)} = (1 + y_1)(1 + y_2) = 1 + y_1 + y_2 + y_1y_2$. This unique expression is the *algebraic normal form* of a Boolean function.

Section 1.7

Solution to Exercise 1.7.5

Uses. Corollary 1.3.19 (Equal spans), Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 1.4.3 (Subspace criterion), Proposition 1.2.18 (Degree and evaluation versus the operations), Definition 1.2.17 (Evaluation)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. In (a) both sides are spans, so Corollary 1.3.19 reduces the equality to six memberships. In (b) the left side is a subspace.

Hint 2. For (a), write $\mathbf{v}_1$ and $\mathbf{v}_2$ in terms of the new list. For (b), $\operatorname{span}(1, x^2 - x) \subseteq W$ follows from Proposition 1.4.14(b); for the other containment, write $p = a + bx + cx^2$ and see what $p(0) = p(1)$ says about b and c.
Sketch. (a) The new entries are visibly combinations of the old ones, and $\mathbf{v}_3 = \mathbf{v}_3$, $\mathbf{v}_2 = (\mathbf{v}_2 - \mathbf{v}_3) + \mathbf{v}_3$, $\mathbf{v}_1 = (\mathbf{v}_1 - \mathbf{v}_2) + (\mathbf{v}_2 - \mathbf{v}_3) + \mathbf{v}_3$. (b) The left side W is a subspace containing 1 and $x^2 - x$, so it contains their span. Conversely $p = a + bx + cx^2 \in W$ means $b + c = 0$, so $p = a \cdot 1 + c(x^2 - x)$.

(a) Both sides are spans, so by Corollary 1.3.19 it suffices to check that each list lies in the span of the other. Each of $\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \mathbf{v}_3$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, with coefficients $1, -1, 0$, then $0, 1, -1$, then $0, 0, 1$. In the other direction,

$$\begin{aligned} \mathbf{v}_3 &= \mathbf{v}_3, \\ \mathbf{v}_2 &= (\mathbf{v}_2 - \mathbf{v}_3) + \mathbf{v}_3, \\ \mathbf{v}_1 &= (\mathbf{v}_1 - \mathbf{v}_2) + (\mathbf{v}_2 - \mathbf{v}_3) + \mathbf{v}_3. \end{aligned}$$

Neither containment needs a generic element; the constructions are the last two lines, and there are only finitely many of them.

(b) Let W be the left side. First, W is a subspace of $F[x]_{\leq 2}$: the zero polynomial has $\mathbf{0}(0) = \mathbf{0}(1) = 0$, and if $p(0) = p(1)$ and $q(0) = q(1)$, then $(p + q)(0) = p(0) + q(0) = p(1) + q(1) = (p + q)(1)$ and $(ap)(0) = ap(0) = ap(1) = (ap)(1)$ by Proposition 1.2.18(c). So Theorem 1.4.3 applies.

$\operatorname{span}(1, x^2 - x) \subseteq W$. The constant 1 takes the value 1 at 0 and at 1, and $x^2 - x$ takes the value 0 at both. So both entries lie in the subspace W, and their span does by Proposition 1.4.14(b). This half is a finite check.

$W \subseteq \operatorname{span}(1, x^2 - x)$. Let $p = a + bx + cx^2 \in W$. By Definition 1.2.17, $p(0) = a$ and $p(1) = a + b + c$, so $p(0) = p(1)$ gives $b + c = 0$, that is, $b = -c$. Hence

$$p = a + (-c)x + cx^2 = a \cdot 1 + c(x^2 - x) \in \operatorname{span}(1, x^2 - x).$$

This half is the construction: it produces coefficients for a generic element of W. The argument works over every field.

Solution to Exercise 1.7.6

Uses. Proposition 1.2.30 (Additive arithmetic), Proposition 1.2.34 (Negatives and the scalar -1), Definition 1.2.1 (Vector space), Exercise 1.2.37, Proposition 1.2.32 (The scalar zero)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Try small cases and small fields (Technique 1.2.38)
Hint 1. Rewrite the equation so that all occurrences of $\mathbf{m}$ are on one side. You should get $\mathbf{m} + \mathbf{m} = \mathbf{u} + \mathbf{w}$.
Hint 2. By (A5) and (A8), $\mathbf{m} + \mathbf{m} = (1 + 1)\mathbf{m}$. When can you divide by $1 + 1$?
Sketch. Adding $\mathbf{m} + \mathbf{u}$ to both sides shows that the equation is equivalent to $(1 + 1)\mathbf{m} = \mathbf{u} + \mathbf{w}$. If $1 + 1 \neq 0$, this has exactly one solution, $\mathbf{m} = (1 + 1)^{-1}(\mathbf{u} + \mathbf{w})$. If $1 + 1 = 0$, the left side is always $\mathbf{0}$, so there is a solution exactly when $\mathbf{u} + \mathbf{w} = \mathbf{0}$, that is, $\mathbf{w} = -\mathbf{u} = \mathbf{u}$, and then every $\mathbf{m}$ works.

We first rewrite the equation. If $\mathbf{m} - \mathbf{u} = \mathbf{w} - \mathbf{m}$, then adding $\mathbf{m} + \mathbf{u}$ to both sides and regrouping with (A1)–(A4) gives $\mathbf{m} + \mathbf{m} = \mathbf{w} + \mathbf{u}$. Conversely, if $\mathbf{m} + \mathbf{m} = \mathbf{w} + \mathbf{u}$, then adding $-\mathbf{m} - \mathbf{u}$ to both sides gives back the original equation. So the equation is equivalent to $\mathbf{m} + \mathbf{m} = \mathbf{u} + \mathbf{w}$, and by (A5) and (A8), $\mathbf{m} + \mathbf{m} = 1\mathbf{m} + 1\mathbf{m} = (1 + 1)\mathbf{m}$. So it is equivalent to

$$(1 + 1)\mathbf{m} = \mathbf{u} + \mathbf{w}.$$

(a) If $1 + 1 \neq 0$, this equation has exactly one solution, $\mathbf{m} = (1 + 1)^{-1}(\mathbf{u} + \mathbf{w})$, by Exercise 1.2.37(a), applied with $a = 1 + 1$, with $\mathbf{0}$ in place of its $\mathbf{u}$, and with $\mathbf{u} + \mathbf{w}$ in place of its $\mathbf{w}$. Over $\mathbb{R}$ this is the midpoint $(\mathbf{u} + \mathbf{w})/2$. Uniqueness here is the two-candidate argument in disguise: two solutions $\mathbf{m}, \mathbf{m}'$ satisfy $(1 + 1)\mathbf{m} = (1 + 1)\mathbf{m}'$, and the nonzero scalar $1 + 1$ cancels.

(b) Over $\mathbb{F}_2$, $1 + 1 = 0$, so the left side is $0\mathbf{m} = \mathbf{0}$ for every $\mathbf{m}$ (Proposition 1.2.32). The equation becomes $\mathbf{0} = \mathbf{u} + \mathbf{w}$, which holds for all $\mathbf{m}$ or for none. It holds exactly when $\mathbf{w} = -\mathbf{u}$ (Proposition 1.2.30(a)), and $-\mathbf{u} = (-1)\mathbf{u} = 1\mathbf{u} = \mathbf{u}$ by Proposition 1.2.34, since $-1 = 1$ in $\mathbb{F}_2$. So there is a solution only if $\mathbf{w} = \mathbf{u}$, and then

every $\mathbf{m}$ is one. If $V \neq \{\mathbf{0}\}$, choosing $\mathbf{u} \neq \mathbf{w}$ gives no solution and choosing $\mathbf{u} = \mathbf{w}$ gives more than one, so the hypothesis $1 + 1 \neq 0$ in (a) cannot be dropped.

Solution to Exercise 1.7.7

Uses. Definition 1.5.3 (Linear relation), Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.32 (The scalar zero), Proposition 1.5.15 (Short lists, zeros, and repeats)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), if $c_j = 0$, what does the relation say about the list with $\mathbf{v}_j$ deleted?
Hint 2. For (b), the two relations are two candidates; compare them after making one coefficient agree. Subtract (d_1/c_1) times the first relation from the second, and apply (a) to the result.
Sketch. (a) If $c_j = 0$, dropping the j-th term gives a nontrivial relation of the sublist without $\mathbf{v}_j$, which is independent. (b) With $\lambda = d_1/c_1$, the coefficients $d_i - \lambda c_i$ form a linear relation whose first coefficient is 0; by (a) it is not a dependence relation, so it is trivial. (c) $(1,0), (0,1), (1,1)$, whose relations are $\lambda, \lambda, -\lambda$.

(a) Let $c_1, \dots, c_n$ be a dependence relation, and suppose $c_j = 0$ for some j. The term $c_j\mathbf{v}_j = 0\mathbf{v}_j = \mathbf{0}$ can be dropped, so the coefficients c_i, $i \neq j$, form a linear relation of the sublist obtained by deleting $\mathbf{v}_j$. Not all of them are 0, since $c_j = 0$ and the relation is nontrivial. So that sublist, of length $n - 1$, is dependent by Theorem 1.5.5, contrary to the hypothesis. Hence all $c_i \neq 0$.

(b) Let c and d be dependence relations. By (a), $c_1 \neq 0$, so we may set $\lambda = d_1/c_1$. By Lemma 1.3.9,

$$\sum_{i=1}^{n}(d_i - \lambda c_i)\mathbf{v}_i = \sum_{i=1}^{n} d_i\mathbf{v}_i - \lambda \sum_{i=1}^{n} c_i\mathbf{v}_i = \mathbf{0} - \lambda\mathbf{0} = \mathbf{0},$$

so the scalars $e_i = d_i - \lambda c_i$ form a linear relation. Its first coefficient is $e_1 = d_1 - (d_1/c_1)c_1 = 0$. By (a), a dependence relation has no zero coefficient, so e is not a dependence relation: it is the trivial relation. Hence $d_i = \lambda c_i$ for every i.

(c) Take $(1,0), (0,1), (1,1)$. It is dependent, since $(1,0) + (0,1) - (1,1) = \mathbf{0}$, and no entry is a multiple of another, so each sublist of length 2 is independent by Proposition 1.5.15(d). By (b), every linear relation is a multiple of $1, 1, -1$, and conversely each $\lambda, \lambda, -\lambda$ is a relation, by Lemma 1.3.9(b). So the relations are $\lambda, \lambda, -\lambda$ for $\lambda \in \mathbb{R}$ (with $\lambda = 0$ giving the trivial relation).

So for such a list the relations form a single line through the trivial relation. This is the uniqueness-by-subtraction pattern: two candidates, normalized so that one coordinate agrees, differ by a solution of a problem that has only the zero solution.

Solution to Exercise 1.7.8

Uses. Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.34 (Negatives and the scalar -1)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For ($\Leftarrow$), take a nontrivial relation and solve it for one entry. Which choice of entry leaves only *later* entries on the other side?
Hint 2. Let i be the smallest index with $a_i \neq 0$. For ($\Rightarrow$), turn $\mathbf{v}_j = \sum_{l>j} b_l\mathbf{v}_l$ into a relation.
Sketch. ($\Leftarrow$) In a nontrivial relation, let i be the smallest index with $a_i \neq 0$. The earlier terms vanish, so $\mathbf{v}_i = \sum_{l>i}(-a_l/a_i)\mathbf{v}_l \in \operatorname{span}(\mathbf{v}_{i+1}, \dots, \mathbf{v}_k)$, a contradiction. ($\Rightarrow$) If $\mathbf{v}_j = \sum_{l>j} b_l\mathbf{v}_l$, the coefficients $0, \dots, 0, -1, b_{j+1}, \dots, b_k$ form a nontrivial relation.

($\Leftarrow$) Suppose the condition holds, and suppose for contradiction that $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ with not all $a_l = 0$. Let i be the *smallest* index with $a_i \neq 0$. Then $a_l = 0$ for $l < i$, and those terms vanish, so

$$a_i\mathbf{v}_i + \sum_{l=i+1}^{k} a_l\mathbf{v}_l = \mathbf{0}, \quad \text{hence} \quad \mathbf{v}_i = \sum_{l=i+1}^{k}\left(-\frac{a_l}{a_i}\right)\mathbf{v}_l,$$

by the computation in the proof of Lemma 1.5.10. So $\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_{i+1}, \dots, \mathbf{v}_k)$; if $i = k$, the sum is empty and this says $\mathbf{v}_k = \mathbf{0} \in \operatorname{span}(())$. Either way the condition fails for $j = i$, a contradiction. By Theorem 1.5.5 the list is independent.

(⇒) We prove the contrapositive. Suppose $\mathbf{v}_j = b_{j+1}\mathbf{v}_{j+1} + \cdots + b_k\mathbf{v}_k$ for some j. The linear combination with coefficients 0 on $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$, -1 on $\mathbf{v}_j$, and b_l on $\mathbf{v}_l$ for $l > j$ represents $-\mathbf{v}_j + \mathbf{v}_j = \mathbf{0}$, by Proposition 1.2.34 and Lemma 1.3.9. Its coefficient on $\mathbf{v}_j$ is $-1 \neq 0$, so the list is dependent.

The proof of Lemma 1.5.10 chose the largest index to get earlier entries on the other side. Choosing the smallest gets later ones. Which extreme to choose is decided by which side the conclusion wants.

Solution to Exercise 1.7.9

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.2.9 (Function spaces), Proposition 1.1.17 (No zero divisors)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. In a nontrivial relation $\sum a_l f_l = \mathbf{0}$, choose the smallest index i with $a_i \neq 0$, and evaluate the relation at t_i.
Hint 2. For (b), list the functions as $1, r_0, \ldots, r_{n-1}$ and choose the points $-1, 1, 2, \ldots, n$.
Sketch. (a) Evaluating at t_i kills the terms with $l < i$ (their coefficients vanish) and with $l > i$ (as $f_l(t_i) = 0$), leaving $a_i f_i(t_i) = 0$, so $a_i = 0$. (b) With $g_0 = 1$, $g_{i+1} = r_i$ and points $s_0 = -1$, $s_{i+1} = i + 1$: $g_m(s_m) = 1$, and $r_i(s_j) = 0$ for $j \leq i$ because $s_j \leq i$.

(a) Suppose $a_1 f_1 + \cdots + a_n f_n = \mathbf{0}$, the zero function, and suppose for contradiction that some $a_l \neq 0$. Let i be the smallest index with $a_i \neq 0$. A relation among functions holds at every point (Definition 1.2.9), so at t_i,

$$\sum_{l=1}^{n} a_l f_l(t_i) = 0.$$

For $l < i$ we have $a_l = 0$, and for $l > i$ we have $f_l(t_i) = 0$ by hypothesis (as $i < l$). So the sum is $a_i f_i(t_i)$, and $a_i f_i(t_i) = 0$ with $f_i(t_i) \neq 0$ forces $a_i = 0$ (Proposition 1.1.17), a contradiction. By Theorem 1.5.5 the list is independent.

(b) Write $g_0 = 1$ and $g_{i+1} = r_i$ for $0 \leq i \leq n-1$, and choose the points $s_0 = -1$ and $s_{i+1} = i + 1$. We check the hypotheses of (a) for the list $g_0, \ldots, g_n$ and the points $s_0, \ldots, s_n$. First, $g_0(s_0) = 1 \neq 0$ and $g_{i+1}(s_{i+1}) = r_i(i + 1) = \max(0, 1) = 1 \neq 0$. Next, let $j < i + 1$, that is, $j \leq i$. Then $s_j = -1$ or $s_j = j$, so $s_j \leq i$ and $s_j - i \leq 0$, which gives $g_{i+1}(s_j) = r_i(s_j) = \max(0, s_j - i) = 0$. (For g_0 there is nothing to check, since no index is smaller than 0.) By (a), with the indices shifted by one, $1, r_0, \ldots, r_{n-1}$ is independent.

The pattern is a triangular one: each function is invisible at the points before its own, so evaluating at the points in order reads off one coefficient at a time.

Solution to Exercise 1.7.10

Uses. Exercise 1.3.28
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Run the inductive step with $k = 1$. Which vectors are in the two shorter lists?
Hint 2. For the corrected statement, start the induction at $k = 2$ and check that the step from k to $k + 1$ is valid when $k \geq 2$.
Sketch. For $k = 1$ the shorter lists are $\mathbf{v}_1$ and $\mathbf{v}_2$, which share no entry, so "$\mathbf{v}_2$ lies on both lines" is false; the step from 1 to 2 fails. The claim is false for $(1, 0), (0, 1)$. With base case $k = 2$ (the hypothesis), the step is valid for $k \geq 2$, because then the two shorter lists share $\mathbf{v}_2$, and a nonzero $\mathbf{v}_2 \in \operatorname{span}(\mathbf{u})$ has $\operatorname{span}(\mathbf{v}_2) = \operatorname{span}(\mathbf{u})$.

The error. The inductive step uses the vector $\mathbf{v}_2$ as a common entry of the lists $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{v}_2, \ldots, \mathbf{v}_{k+1}$. When $k = 1$ these lists are $\mathbf{v}_1$ and $\mathbf{v}_2$, which have no entry in common, and $\mathbf{v}_2$ is not in the first. So the step from $k = 1$ to $k = 2$ is invalid. The base case and the steps from k to $k + 1$ for $k \geq 2$ are fine, but without the step from 1 to 2 the induction proves nothing beyond $k = 1$.

The claim is false. In $\mathbb{R}^2$, suppose $(1, 0) = a\mathbf{u}$ and $(0, 1) = b\mathbf{u}$. Both vectors are nonzero, so $a, b \neq 0$, and $(0, 1) = b\mathbf{u} = (b/a)(1, 0)$, whose second coordinate is 0. This is a contradiction, so $(1, 0)$ and $(0, 1)$ lie on no common line.

The corrected statement. Suppose every two of the nonzero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lie on a common line; we show by induction on $k \geq 2$ that all of them do (for $k = 1$ there is nothing to prove). For $k = 2$ this is the hypothesis. Suppose it holds for lists of length $k \geq 2$ satisfying the hypothesis, and let $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ satisfy

it. Both shorter lists satisfy it as well, so the argument above gives $\mathbf{u}$ and $\mathbf{u}'$. Since $k \geq 2$, the vector $\mathbf{v}_2$ is an entry of both shorter lists. It lies in $\operatorname{span}(\mathbf{u})$, say $\mathbf{v}_2 = c\mathbf{u}$, and $c \neq 0$ because $\mathbf{v}_2 \neq \mathbf{0}$. So $\operatorname{span}(\mathbf{v}_2) = \operatorname{span}(\mathbf{u})$ by Exercise 1.3.28(b); likewise $\operatorname{span}(\mathbf{v}_2) = \operatorname{span}(\mathbf{u}')$. Hence $\mathbf{v}_{k+1} \in \operatorname{span}(\mathbf{u}') = \operatorname{span}(\mathbf{u})$, and all the vectors lie on the line $\operatorname{span}(\mathbf{u})$.

When an inductive step combines two shorter lists through a shared entry, check that the shared entry exists for the smallest k at which the step is used.

Solution to Exercise 1.7.11

> *Uses.* Definition 1.3.4 (Linear combination), Lemma 1.3.9 (Arithmetic of linear combinations), Theorem 1.5.5 (Independence is decided at zero), Theorem 1.6.10 (Every spanning list contains a basis), Corollary 1.5.13 (Independence, one entry at a time)
> *Techniques.* Induction on the length of a list (Technique 1.6.9), Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* Over $\mathbb{F}_2$, a subset I is the same thing as a list of coefficients in $\{0,1\}$, that is, a linear combination of $\mathcal{L}$. So $Z(\mathcal{L})$ counts the linear relations of $\mathcal{L}$.
> *Hint 2.* Split the zero-sum subsets of $\{1,\dots,k+1\}$ according to whether they contain $k+1$. Those that do correspond to linear combinations of $\mathcal{L}$ representing $\mathbf{w}$. Fix one such representation and add relations to it.
> *Hint 3.* For (b), induct on k. Sifting reads from left to right, so sifting $\mathcal{L}, \mathbf{w}$ deletes what sifting $\mathcal{L}$ deletes, plus $\mathbf{w}$ exactly when $\mathbf{w} \in \operatorname{span}(\mathcal{L})$.
> *Sketch.* Subsets of $\{1,\dots,k\}$ are linear combinations of $\mathcal{L}$ over $\mathbb{F}_2$. A zero-sum subset of $\{1,\dots,k+1\}$ either avoids $k+1$ ($Z(\mathcal{L})$ of them) or is $J \cup \{k+1\}$ with $\sum_J \mathbf{v}_j = -\mathbf{w} = \mathbf{w}$. The representations of $\mathbf{w}$ number 0 if $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and otherwise as many as the relations, since $c \mapsto a + c$ is a bijection from relations to representations for any fixed representation a. For (b), induct on k with $Z(()) = 1 = 2^0$.

Over $\mathbb{F}_2$ the only scalars are 0 and 1, so a linear combination of $\mathcal{L}$ is a list of coefficients in $\{0,1\}$, and it represents $\sum_{i \in I} \mathbf{v}_i$, where I is the set of positions with coefficient 1. This is a bijection between subsets of $\{1,\dots,k\}$ and linear combinations of $\mathcal{L}$. In particular, $Z(\mathcal{L})$ is the number of linear relations of $\mathcal{L}$, and for $\mathbf{w} \in V$ the number $N(\mathbf{w})$ of linear combinations of $\mathcal{L}$ representing $\mathbf{w}$ is the number of I with $\sum_{i \in I} \mathbf{v}_i = \mathbf{w}$.

(a) Split the subsets $I' \subseteq \{1,\dots,k+1\}$ with $\sum_{i \in I'} \mathbf{v}_i = \mathbf{0}$ (writing $\mathbf{v}_{k+1} = \mathbf{w}$) into those that do not contain $k+1$ and those that do. The first kind are the zero-sum subsets of $\{1,\dots,k\}$, and there are $Z(\mathcal{L})$ of them. The second kind are the sets $J \cup \{k+1\}$ with $J \subseteq \{1,\dots,k\}$ and $\sum_{j \in J} \mathbf{v}_j + \mathbf{w} = \mathbf{0}$, that is, $\sum_{j \in J} \mathbf{v}_j = -\mathbf{w} = \mathbf{w}$ (over $\mathbb{F}_2$, $-\mathbf{w} = (-1)\mathbf{w} = \mathbf{w}$). There are $N(\mathbf{w})$ of them.

If $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, then $N(\mathbf{w}) = 0$. If $\mathbf{w} \in \operatorname{span}(\mathcal{L})$, fix one representation $a_1, \dots, a_k$ of $\mathbf{w}$. For a linear relation c, the coefficients $a + c$ represent $\mathbf{w} + \mathbf{0} = \mathbf{w}$ (Lemma 1.3.9). Conversely, for a representation b of $\mathbf{w}$, the coefficients $b - a$ represent $\mathbf{w} - \mathbf{w} = \mathbf{0}$. These two maps are inverse to each other, so $N(\mathbf{w})$ equals the number of relations, $Z(\mathcal{L})$. Hence $Z(\mathcal{L}, \mathbf{w}) = Z(\mathcal{L}) + N(\mathbf{w})$ is $Z(\mathcal{L})$ or $2Z(\mathcal{L})$, as claimed.

(b) For a list $\mathcal{M}$, write $d(\mathcal{M})$ for the number of entries that sifting deletes from it. We show $Z(\mathcal{L}) = 2^{d(\mathcal{L})}$ by induction on k. For $k = 0$, the only subset is $\emptyset$, whose sum is $\mathbf{0}$, so $Z(()) = 1 = 2^0$, and sifting deletes nothing. Suppose the claim holds for lists of length k, and consider $\mathcal{L}, \mathbf{w}$. Sifting decides about each entry by looking only at the entries before it, so on the first k entries, sifting $\mathcal{L}, \mathbf{w}$ makes the same decisions as sifting $\mathcal{L}$. It deletes $\mathbf{w}$ exactly when $\mathbf{w} \in \operatorname{span}(\mathcal{L})$. So $d(\mathcal{L}, \mathbf{w}) = d(\mathcal{L})$ if $\mathbf{w} \notin \operatorname{span}(\mathcal{L})$, and $d(\mathcal{L}) + 1$ otherwise, and by (a) and the inductive hypothesis, $Z(\mathcal{L}, \mathbf{w}) = 2^{d(\mathcal{L}, \mathbf{w})}$ in both cases.

Finally, $Z(\mathcal{L}) = 1$ means that the trivial relation is the only one, which is independence (Theorem 1.5.5). Equivalently $d = 0$: sifting deletes nothing exactly when no entry lies in the span of the earlier ones, which is independence by Corollary 1.5.13.

Solution to Exercise 1.7.12

> *Uses.* Theorem 1.5.5 (Independence is decided at zero), Proposition 1.5.15 (Short lists, zeros, and repeats), Lemma 1.3.9 (Arithmetic of linear combinations), Definition 1.3.10 (Span), Proposition 1.3.17 (A list lies in its span)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* Take V as small as possible and both lists of length one.

Hint 2. In your counterexample, the two spans share a nonzero vector. For (c), a vector in both spans gives a relation of $\mathcal{L}, \mathcal{M}$, and a relation of $\mathcal{L}, \mathcal{M}$ gives a vector in both spans.
Sketch. (a) $\mathcal{L} = \mathcal{M} = \mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$: each is independent, but $\mathbf{v}, \mathbf{v}$ is not. (b) The spans share $\mathbf{v}$; conjecture that a common nonzero vector is the only obstruction. (c) If $\mathbf{x} = \sum a_i \mathbf{u}_i = \sum b_j \mathbf{w}_j$, then $a, -b$ is a relation of $\mathcal{L}, \mathcal{M}$, trivial by independence, so $\mathbf{x} = \mathbf{0}$. Conversely a relation $\sum a_i \mathbf{u}_i + \sum b_j \mathbf{w}_j = \mathbf{0}$ gives $\sum a_i \mathbf{u}_i = \sum(-b_j)\mathbf{w}_j$ in both spans, hence $\mathbf{0}$, and independence of each list kills the coefficients.

(a) Let V be any nonzero vector space, $\mathbf{v} \neq \mathbf{0}$, and $\mathcal{L} = \mathcal{M} = \mathbf{v}$. Each list is independent by Proposition 1.5.15(b), but $\mathcal{L}, \mathcal{M} = \mathbf{v}, \mathbf{v}$ has a repeated entry and is dependent by Proposition 1.5.15(c). No counterexample is smaller: lists of length 0 add nothing, and in the zero space every nonempty list is dependent. (The same counterexample settles Exercise 1.5.34(c).)

(b) The two spans are both span($\mathbf{v}$), and they share the nonzero vector $\mathbf{v}$. Any counterexample must be like this: a relation among the concatenation that involves both lists produces a vector that both lists can represent. This suggests that concatenating independent lists can fail only when their spans share a nonzero vector, which is (c).

(c) Write $\mathcal{L} = \mathbf{u}_1, ..., \mathbf{u}_k$ and $\mathcal{M} = \mathbf{w}_1, ..., \mathbf{w}_m$.

(⇒) Suppose $\mathcal{L}, \mathcal{M}$ is independent, and let $\mathbf{x} \in \text{span}(\mathcal{L}) \cap \text{span}(\mathcal{M})$, say $\mathbf{x} = \sum_i a_i \mathbf{u}_i = \sum_j b_j \mathbf{w}_j$. By Lemma 1.3.9, the coefficients $a_1, ..., a_k, -b_1, ..., -b_m$ represent $\mathbf{x} - \mathbf{x} = \mathbf{0}$, so they form a linear relation of $\mathcal{L}, \mathcal{M}$. It is trivial, so every $a_i = 0$ and $\mathbf{x} = \mathbf{0}$. Since $\mathbf{0}$ lies in both spans (Proposition 1.3.17), the intersection is $\{\mathbf{0}\}$.

(⇐) Suppose $\text{span}(\mathcal{L}) \cap \text{span}(\mathcal{M}) = \{\mathbf{0}\}$, and let $\sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{w}_j = \mathbf{0}$. Then

$$\mathbf{x} = \sum_{i=1}^{k} a_i \mathbf{u}_i = \sum_{j=1}^{m} (-b_j)\mathbf{w}_j$$

lies in both spans, so $\mathbf{x} = \mathbf{0}$. Since $\mathcal{L}$ is independent, $\sum_i a_i \mathbf{u}_i = \mathbf{0}$ forces all $a_i = 0$, and since $\mathcal{M}$ is independent, $\sum_j (-b_j)\mathbf{w}_j = \mathbf{0}$ forces all $b_j = 0$. So the relation is trivial, and $\mathcal{L}, \mathcal{M}$ is independent by Theorem 1.5.5.

For two subspaces, the condition $\text{span}(\mathcal{L}) \cap \text{span}(\mathcal{M}) = \{\mathbf{0}\}$ characterizes direct sums (Section 4.4).

Solution to Exercise 1.7.13

Uses. Theorem 1.4.3 (Subspace criterion), Definition 1.2.1 (Vector space)
Techniques. Proving a disjunction (Technique 1.7.4), Try small cases and small fields (Technique 1.2.38)
Hint 1. Assume that $\mathbf{v} + \mathbf{w} \in U$, and show that $\mathbf{v} - \mathbf{w} \notin U$.
Hint 2. If both were in U, so would be their sum. What is it? For (b), try the smallest possible V and U.
Sketch. If $\mathbf{v} + \mathbf{w} \in U$ and also $\mathbf{v} - \mathbf{w} \in U$, then their sum $(1+1)\mathbf{v}$ lies in U, and so does $(1+1)^{-1}(1+1)\mathbf{v} = \mathbf{v}$, a contradiction. Over $\mathbb{F}_2$: $V = \mathbb{F}_2, U = \{0\}, \mathbf{v} = \mathbf{w} = 1$, where $\mathbf{v} + \mathbf{w} = \mathbf{v} - \mathbf{w} = 0 \in U$.

(a) Assume that $\mathbf{v} + \mathbf{w} \in U$; we show that $\mathbf{v} - \mathbf{w} \notin U$. Suppose instead that $\mathbf{v} - \mathbf{w} \in U$. Since U is closed under addition (Theorem 1.4.3), it contains

$$(\mathbf{v} + \mathbf{w}) + (\mathbf{v} - \mathbf{w}) = \mathbf{v} + \mathbf{v} = (1+1)\mathbf{v},$$

using (A1)–(A4) to cancel $\mathbf{w}$, then (A5) and (A8). Since $1 + 1 \neq 0$, closure under scalar multiplication gives $(1+1)^{-1}((1+1)\mathbf{v}) = \mathbf{v} \in U$ by (A6) and (A5), contradicting $\mathbf{v} \notin U$. So $\mathbf{v} - \mathbf{w} \notin U$, and the disjunction holds.

(b) Take $V = \mathbb{F}_2$, regarded as a vector space over itself, $U = \{0\}$, and $\mathbf{v} = \mathbf{w} = 1$. Then $\mathbf{v} \notin U$, but $\mathbf{v} + \mathbf{w} = 1 + 1 = 0 \in U$ and $\mathbf{v} - \mathbf{w} = 0 \in U$. The proof of (a) breaks exactly at the division by $1 + 1$.

Solution to Exercise 1.7.14

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.27 (Uniqueness of the zero vector), Proposition 1.2.32 (The scalar zero), Example 1.2.39 (When does $\mathbf{v} = -\mathbf{v}$ force $\mathbf{v} = \mathbf{0}$?)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. The scalar multiplication is unknown, so look for a consequence of the axioms that involves only the addition. Example 1.2.39 computed $\mathbf{v} + \mathbf{v} = (1+1)\mathbf{v}$.

Hint 2. Split into the cases $1+1=0$ and $1+1\neq 0$ in F. In the second case, consider the vector $(1+1)^{-1}\odot 1$, where $1\in\mathbb{Z}$.
Sketch. Write $e=1+1\in F$. In every vector space, $\mathbf{v}+\mathbf{v}=e\mathbf{v}$. If $e=0$, then $\mathbf{v}+\mathbf{v}=\mathbf{0}$ for all $\mathbf{v}$, but $1+1=2\neq 0$ in $\mathbb{Z}$. If $e\neq 0$, then $\mathbf{x}=e^{-1}\odot 1$ satisfies $\mathbf{x}+\mathbf{x}=e\odot\left(e^{-1}\odot 1\right)=1$, but no integer x has $x+x=1$.

Suppose, for contradiction, that some field F and some scalar multiplication $\odot$ make $\mathbb{Z}$, with its usual addition, a vector space over F. Its zero vector is the integer 0, the unique element with $n+0=n$ for all n. Write $e=1+1\in F$. For every integer n, by (A5) and (A8),

$$n+n=1\odot n+1\odot n=(1+1)\odot n=e\odot n.$$

Case $e=0$. Then $n+n=0\odot n=0$ for every n, by Proposition 1.2.32. But $1+1=2\neq 0$ in $\mathbb{Z}$.

Case $e\neq 0$. Let $x=e^{-1}\odot 1\in\mathbb{Z}$. By the display, (A6), and (A5),

$$x+x=e\odot\left(e^{-1}\odot 1\right)=\left(ee^{-1}\right)\odot 1=1\odot 1=1.$$

But $x+x=2x$ is even for every integer x, so it is not 1.

Both cases are impossible, so no such F and $\odot$ exist. The argument never examined a particular scalar multiplication. It used a consequence of the axioms, $\mathbf{v}+\mathbf{v}=(1+1)\mathbf{v}$, that every candidate must satisfy and that mentions the scalar multiplication only through the field, and it showed that the addition of $\mathbb{Z}$ is incompatible with both possible values of $1+1$.

Solution to Exercise 1.7.15

Uses. Theorem 1.4.25 (The subspaces of F^2), Definition 1.3.10 (Span)
Techniques. Try small cases and small fields (Technique 1.2.38), Finding counterexamples and minimal examples (Technique 1.4.18), Double containment (Technique 1.3.14)
Hint 1. Test the smallest subspace first.
Hint 2. By Theorem 1.4.25, the other subspaces are F^2 and the lines $\operatorname{span}((p,q))$ with $(p,q)\neq\mathbf{0}$. For a line, the equation $qx-py=0$ is a natural guess; prove the equality by double containment, splitting into the cases $p\neq 0$ and $p=0$.
Sketch. False for $\{\mathbf{0}\}$: if $(a,b)=\mathbf{0}$ the solution set is F^2, and otherwise it contains $(-b,a)\neq\mathbf{0}$. True for every other subspace: F^2 is the solution set of $0x+0y=0$, and $\operatorname{span}((p,q))$ is that of $qx-py=0$. Every multiple of (p,q) satisfies the equation, and a solution (x,y) is $(x/p)(p,q)$ if $p\neq 0$, or $(y/q)(p,q)$ if $p=0$.

The claim is false, and the smallest subspace is the counterexample. Let $a,b\in F$ and let S be the solution set of $ax+by=0$. If $a=b=0$, then $S=F^2$. Otherwise $(-b,a)\neq\mathbf{0}$ and $a(-b)+ba=0$, so S contains a nonzero vector. Either way $S\neq\{\mathbf{0}\}$, so $\{\mathbf{0}\}$ is not the solution set of a single equation. (It is the solution set of the two equations $x=0$ and $y=0$.)

Every other subspace is. By Theorem 1.4.25, the subspaces other than $\{\mathbf{0}\}$ are F^2, which is the solution set of $0x+0y=0$, and the lines $L=\operatorname{span}((p,q))$ with $(p,q)\neq\mathbf{0}$. We show that $L=S$ for the equation $qx-py=0$, by double containment.

$L\subseteq S$. A vector of L is (cp,cq) for some $c\in F$, and $q(cp)-p(cq)=0$.

$S\subseteq L$. Let (x,y) satisfy $qx=py$. If $p\neq 0$, then $y=qx/p$, and $(x,y)=(x/p)(p,q)\in L$. If $p=0$, then $q\neq 0$, and $qx=0$ forces $x=0$; then $(x,y)=(0,y)=(y/q)(0,q)=(y/q)(p,q)\in L$.

Problems for Chapter 1

Solution to Problem 1.1

Uses. Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.17 (No zero divisors), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.13 (Solving $a+x=b$ and $ax=b$), Bézout's identity and the Euclidean algorithm
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. Only multiplicative inverses are in question (Lemma 1.1.21). For a composite $n=rs$, look at $r\cdot s$.
Hint 2. For a prime p and $0<a<p$, Bézout gives $ax+py=1$. For (b), run the Euclidean algorithm on 101 and 17, and then substitute back.

Sketch. (a) $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime. If $n = rs$ with $1 < r, s < n$, then $rs = 0$ is a product of nonzero elements; if $n = p$ is prime and $0 < a < p$, then $\gcd(a, p) = 1$, and Bézout's $ax + py = 1$ makes $x \bmod p$ an inverse of a. (b) $101 = 5 \cdot 17 + 16$ and $17 = 16 + 1$, so $1 = 6 \cdot 17 - 101$, and $17^{-1} = 6$. Then $17x = 2$, so $x = 12$.

(a) *Statement.* For an integer $n \geq 2$, $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime (Proposition 1.1.24).

Proof. By Lemma 1.1.21, $\mathbb{Z}/n\mathbb{Z}$ satisfies every field axiom except possibly the existence of multiplicative inverses, so it is a field exactly when every nonzero element has an inverse.

($\Rightarrow$) We prove the contrapositive. If n is not prime, write $n = rs$ with $1 < r, s < n$. Then r and s are nonzero elements of $\mathbb{Z}/n\mathbb{Z}$ with $r \cdot_n s = n \bmod n = 0$. A field has no zero divisors (Proposition 1.1.17), so $\mathbb{Z}/n\mathbb{Z}$ is not a field.

($\Leftarrow$) Let $n = p$ be prime and $a \in \{1, \ldots, p-1\}$. The only positive divisors of p are 1 and p, and $p \nmid a$, so $\gcd(a, p) = 1$. By Bézout's identity, $ax + py = 1$ for some integers x, y. Let $b = x \bmod p$. By Lemma 1.1.21(a),

$$a \cdot_p b = (ax) \bmod p = (1 - py) \bmod p = 1,$$

so b is an inverse of a.

(b) The Bézout proof computes inverses: run the Euclidean algorithm and substitute back. Here

$$101 = 5 \cdot 17 + 16, \quad 17 = 1 \cdot 16 + 1,$$

so

$$1 = 17 - 16 = 17 - (101 - 5 \cdot 17) = 6 \cdot 17 - 101.$$

Hence $17 \cdot 6 = 102 = 101 + 1$, and $17^{-1} = 6$ in $\mathbb{F}_{101}$. By Proposition 1.1.13, $17x + 3 = 5$ holds exactly when $17x = 2$, that is, $x = 6 \cdot 2 = 12$. Check: $17 \cdot 12 + 3 = 207 = 2 \cdot 101 + 5$.

The counting proof of Proposition 1.1.24 shows that 17 has an inverse without producing it; using it to find 17^{-1} would mean trying the candidates one by one.

Solution to Problem 1.2

Uses. Definition 1.2.1 (Vector space), Proposition 1.2.30 (Additive arithmetic), Proposition 1.2.28 (Uniqueness of additive inverses), Proposition 1.2.33 (The zero vector), Proposition 1.2.32 (The scalar zero), Example 1.2.25 (Changing one operation on $\mathbb{R}^2$)
Techniques. Proving from axioms (Technique 1.1.7)
Hint 1. For (a), it suffices to show $a(-\mathbf{w}) = -(a\mathbf{w})$. Compute $a\mathbf{w} + a(-\mathbf{w})$ with (A7). You will need $a\mathbf{0} = \mathbf{0}$: which of the two proofs of Proposition 1.2.33 avoids (A8)?
Hint 2. For (b), show $(-b)\mathbf{v} = -(b\mathbf{v})$ by computing $b\mathbf{v} + (-b)\mathbf{v}$ with (A8). For (c), look at Example 1.2.25.
Sketch. (a) The first proof of Proposition 1.2.33 uses (A7), (A3), and cancellation, so $a\mathbf{0} = \mathbf{0}$ without (A8). Then $a\mathbf{w} + a(-\mathbf{w}) = a(\mathbf{w} + (-\mathbf{w})) = a\mathbf{0} = \mathbf{0}$, so $a(-\mathbf{w}) = -(a\mathbf{w})$, and (A7) gives the identity. (b) $b\mathbf{v} + (-b)\mathbf{v} = (b + (-b))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$ by (A8) and Proposition 1.2.32, which itself uses (A8); then (A8) again. (c) On $\mathbb{R}^2$ with $a \odot (x, y) = (ax, y)$: $(2 - 1) \odot (0, 1) = (0, 1)$, but $2 \odot (0, 1) - 1 \odot (0, 1) = \mathbf{0}$.

(a) We use only (A1)–(A7) and facts proved from them. Cancellation (Proposition 1.2.30(b)) and uniqueness of negatives (Proposition 1.2.28) use only (A1)–(A4). The first proof of Proposition 1.2.33 uses (A7), (A3), and cancellation, so $a\mathbf{0} = \mathbf{0}$ holds without (A8). Now

$$\begin{aligned} a\mathbf{w} + a(-\mathbf{w}) &= a(\mathbf{w} + (-\mathbf{w})) && \text{(A7)} \\ &= a\mathbf{0} = \mathbf{0} && \text{(A4, then } a\mathbf{0} = \mathbf{0}\text{)}, \end{aligned}$$

so $a(-\mathbf{w})$ is an additive inverse of $a\mathbf{w}$, and $a(-\mathbf{w}) = -(a\mathbf{w})$ by Proposition 1.2.28. Hence, by (A7),

$$a(\mathbf{v} - \mathbf{w}) = a(\mathbf{v} + (-\mathbf{w})) = a\mathbf{v} + a(-\mathbf{w}) = a\mathbf{v} + (-(a\mathbf{w})) = a\mathbf{v} - a\mathbf{w}.$$

(b) First,

$$b\mathbf{v} + (-b)\mathbf{v} = (b + (-b))\mathbf{v} \quad \text{(A8)}$$

$$= 0\mathbf{v} = \mathbf{0} \qquad \text{(@prop:vs-zero-mul)},$$

so $(-b)\mathbf{v} = -(b\mathbf{v})$ by Proposition 1.2.28. Then by (A8),

$$(a-b)\mathbf{v} = (a+(-b))\mathbf{v} = a\mathbf{v} + (-b)\mathbf{v} = a\mathbf{v} - b\mathbf{v}.$$

The proof uses (A8) twice directly, and once more through Proposition 1.2.32, whose proof starts from $(0+0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$. Besides (A8) it uses (A1)–(A4) through Proposition 1.2.28 and cancellation.

(c) By Example 1.2.25(b), $\mathbb{R}^2$ with its usual addition and $a \odot (x, y) = (ax, y)$ satisfies every axiom except (A8). There,

$$(2-1) \odot (0,1) = 1 \odot (0,1) = (0,1),$$

while $2 \odot (0,1) - 1 \odot (0,1) = (0,1) - (0,1) = (0,0)$. So $(a-b) \odot \mathbf{v} \neq a \odot \mathbf{v} - b \odot \mathbf{v}$ for $a = 2$, $b = 1$, $\mathbf{v} = (0,1)$. Every step of (b) that invokes (A8) fails here, and so do the derived facts $0 \odot \mathbf{v} = \mathbf{0}$ and $(-b) \odot \mathbf{v} = -(b \odot \mathbf{v})$: for instance $0 \odot (0,1) = (0,1)$, and $(-1) \odot (0,1) = (0,1) \neq (0,-1)$. Part (a), whose proof avoided (A8), still holds in this structure, as it must.

Solution to Problem 1.3

> *Uses.* Definition 1.5.1 (Linear independence), Definition 1.5.3 (Linear relation), Theorem 1.5.5 (Independence is decided at zero), Lemma 1.3.9 (Arithmetic of linear combinations), Proposition 1.2.34 (Negatives and the scalar -1), Proposition 1.5.15 (Short lists, zeros, and repeats)
> *Techniques.* Contrapositive (Technique 1.5.4), Uniqueness via two candidates (Technique 1.1.8)
> *Hint 1.* For (a), one direction is the definition applied to $\mathbf{0}$. For the other, two different representations of one vector can be subtracted.
> *Hint 2.* For (b), fix the representation a of $\mathbf{w}$. If c is a linear relation, what does $a + c$ represent?
> *Hint 3.* For (c), compare $(1,0)$ and $(1,0),(2,0)$ in $\mathbb{R}^2$, with $\mathbf{w} = (0,1)$.
> *Sketch.* (a) Independence (every vector of the span has exactly one representation) applied to $\mathbf{0}$ gives the test at zero. Conversely, two different representations of $\mathbf{w}$ subtract to a nontrivial relation. (b) If c is a relation, then $a + c$ represents $\mathbf{w}$, so $a + c = a$ and $c = 0$. (c) Neither: $\mathbf{w} = (0,1)$ is outside the spans of the independent list $(1,0)$ and of the dependent list $(1,0),(2,0)$.

(a) *Statement.* $\mathcal{L}$ is linearly independent if every vector of $\operatorname{span}(\mathcal{L})$ is represented by exactly one linear combination of $\mathcal{L}$ (Definition 1.5.1). The test at zero: $\mathcal{L}$ is independent if and only if $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ implies $a_1 = \cdots = a_k = 0$.

Proof. If $\mathcal{L}$ is independent, then $\mathbf{0} \in \operatorname{span}(\mathcal{L})$ is represented by the trivial relation, and by nothing else, so every relation is trivial. Conversely, we prove the contrapositive: suppose $\mathcal{L}$ is dependent, so that some $\mathbf{w}$ has two different representations, with coefficients a_i and b_i. By Lemma 1.3.9 and $-\mathbf{x} = (-1)\mathbf{x}$ (Proposition 1.2.34),

$$\sum_{i=1}^{k}(a_i - b_i)\mathbf{v}_i = \sum_{i=1}^{k} a_i\mathbf{v}_i - \sum_{i=1}^{k} b_i\mathbf{v}_i = \mathbf{w} - \mathbf{w} = \mathbf{0},$$

and $a_j - b_j \neq 0$ for some j. So the test at zero fails.

(b) Let $a_1, \ldots, a_k$ be the unique representation of $\mathbf{w}$, and let $c_1, \ldots, c_k$ be any linear relation of $\mathcal{L}$. By Lemma 1.3.9, the coefficients $a_i + c_i$ represent $\mathbf{w} + \mathbf{0} = \mathbf{w}$. By uniqueness, $a_i + c_i = a_i$ for every i, so $c_i = 0$. Every relation is trivial, and $\mathcal{L}$ is independent by (a).

So uniqueness at a single vector of the span forces uniqueness everywhere. This is the reason the test can be run at $\mathbf{0}$: any vector of the span would do, and $\mathbf{0}$ is the one whose representation we always know.

(c) Neither. In $\mathbb{R}^2$ let $\mathbf{w} = (0,1)$. The list $(1,0)$ is independent (Proposition 1.5.15(b)), the list $(1,0),(2,0)$ is dependent (Proposition 1.5.15(d)), and $\mathbf{w}$ lies in neither span, since both spans consist of vectors with second coordinate 0. A vector outside the span says something about spanning, not about independence.

Solution to Problem 1.4

Uses. Lemma 1.5.10 (Linear dependence lemma), Theorem 1.5.5 (Independence is decided at zero), Corollary 1.3.20 (Operations that preserve the span)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Start from a nontrivial relation. Solving it for one entry requires that entry's coefficient to be nonzero; the choice of *which* entry decides what is on the other side.
Hint 2. For the deletion statement, Corollary 1.3.20(c) says that appending a vector already in the span does not change the span.
Hint 3. In (c), the spans involved are the first axis, the second axis, and $\mathbb{R}^2$.
Sketch. (a) With j largest, the terms after j vanish, so $\mathbf{v}_j = \sum_{l<j}(-a_l/a_j)\mathbf{v}_l$. If $\mathbf{v}_j$ lies in the span of the earlier entries, it lies in the span of the list without it, so appending it back changes nothing (Corollary 1.3.20(c)), and the result is a reordering of $\mathcal{L}$. (b) With i smallest, $\mathbf{v}_i = \sum_{l>i}(-a_l/a_i)\mathbf{v}_l$, a combination of later entries. (c) Only $j = 3$, and only $i = 2$.

(a) *Statement.* If $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$ is linearly dependent, then $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ for some j, and for every such j, deleting $\mathbf{v}_j$ from $\mathcal{L}$ does not change the span (Lemma 1.5.10).

Proof. By Theorem 1.5.5 there is a relation $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}$ with some $a_l \neq 0$. Let j be the largest index with $a_j \neq 0$. The terms with index above j are $\mathbf{0}$, so $a_1\mathbf{v}_1 + \cdots + a_j\mathbf{v}_j = \mathbf{0}$, and since $a_j \neq 0$,

$$\mathbf{v}_j = \sum_{l=1}^{j-1}\left(-\frac{a_l}{a_j}\right)\mathbf{v}_l \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}),$$

where for $j = 1$ the sum is empty and the statement is $\mathbf{v}_1 = \mathbf{0}$. Now let j be any index with $\mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$, and let $\mathcal{L}'$ be $\mathcal{L}$ without $\mathbf{v}_j$. Then $\mathbf{v}_j \in \operatorname{span}(\mathcal{L}')$, since $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$ is a sublist of $\mathcal{L}'$ (Corollary 1.3.20(b)). By Corollary 1.3.20(c), $\operatorname{span}(\mathcal{L}', \mathbf{v}_j) = \operatorname{span}(\mathcal{L}')$, and $\mathcal{L}', \mathbf{v}_j$ is a reordering of $\mathcal{L}$, so its span is $\operatorname{span}(\mathcal{L})$ by Corollary 1.3.20(a).

(b) With i the smallest index with $a_i \neq 0$, the terms before i vanish, so $a_i\mathbf{v}_i + a_{i+1}\mathbf{v}_{i+1} + \cdots + a_k\mathbf{v}_k = \mathbf{0}$, and

$$\mathbf{v}_i = \sum_{l=i+1}^{k}\left(-\frac{a_l}{a_i}\right)\mathbf{v}_l \in \operatorname{span}(\mathbf{v}_{i+1}, \ldots, \mathbf{v}_k).$$

So a dependent list also has an entry in the span of the entries *after* it. The extremal choice decides the side: the largest index leaves earlier entries, the smallest leaves later ones.

(c) Write $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 = (1, 0), (0, 1), (0, 2)$.

Earlier entries. $\mathbf{v}_1 \neq \mathbf{0}$, so $j = 1$ fails. $\operatorname{span}(\mathbf{v}_1)$ is the first axis, which does not contain $\mathbf{v}_2$, so $j = 2$ fails. $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2) = \mathbb{R}^2$ contains $\mathbf{v}_3$, so $j = 3$ works.

Later entries. $\operatorname{span}(\mathbf{v}_2, \mathbf{v}_3)$ is the second axis, which does not contain $\mathbf{v}_1$, so $i = 1$ fails. $\mathbf{v}_2 = (1/2)\mathbf{v}_3$, so $i = 2$ works. $\operatorname{span}(()) = \{\mathbf{0}\}$ and $\mathbf{v}_3 \neq \mathbf{0}$, so $i = 3$ fails.

The two versions of the lemma pick out different redundant entries, here $\mathbf{v}_3$ and $\mathbf{v}_2$. Deleting either one leaves a basis of $\mathbb{R}^2$.

Solution to Problem 1.5

Uses. Theorem 1.4.25 (The subspaces of F^2), Exercise 1.3.28, Definition 1.6.1 (Basis), (1.6.1), Exercise 1.3.37, Proposition 1.5.15 (Short lists, zeros, and repeats), Proposition 1.6.2 (Bases are independent spanning lists)
Hint 1. A nonzero vector (a, b) with $a \neq 0$ lies on the same line as $(1, b/a)$. Inverses in $\mathbb{F}_5$: $2^{-1} = 3$, $3^{-1} = 2$, $4^{-1} = 4$.
Hint 2. For (b), solve $a(1, 2) + b(1, 3) = (x, y)$ for arbitrary (x, y). For (c), a list $\mathbf{u}, \mathbf{v}$ in $\mathbb{F}_5^2$ is a basis exactly when $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \notin \operatorname{span}(\mathbf{u})$.
Sketch. (a) The six lines are $\operatorname{span}((1, c))$ for $c \in \mathbb{F}_5$ and $\operatorname{span}((0, 1))$; $(2, 3) = 2(1, 4)$. (b) The equations $a + b = x$, $2a + 3b = y$ force $b = y - 2x$ and $a = 3x - y$, which work; so $[(x, y)]_{\mathcal{B}} = (3x - y, y - 2x)$ and $[(2, 3)]_{\mathcal{B}} = (3, 4)$. (c) 4 choices of first entry, then $25 - 5 = 20$ choices outside its line: 80.

(a) By Theorem 1.4.25, the lines are the sets $\operatorname{span}(\mathbf{v})$ with $\mathbf{v} \neq \mathbf{0}$, and $\operatorname{span}(\mathbf{v}) = \operatorname{span}(c\mathbf{v})$ for $c \neq 0$ (Exercise 1.3.28(b)). If $\mathbf{v} = (a, b)$ with $a \neq 0$, then $\operatorname{span}(\mathbf{v}) = \operatorname{span}((1, b/a))$; if $a = 0$, then $b \neq 0$ and $\operatorname{span}(\mathbf{v}) =$

$\operatorname{span}((0,1))$. The lines $\operatorname{span}((1,c))$, $c \in \mathbb{F}_5$, are different from each other and from $\operatorname{span}((0,1))$, since the only multiple of $(1,c)$ with first coordinate 1 is $(1,c)$ itself, and $\operatorname{span}((0,1))$ contains no vector with first coordinate 1. So the six lines are

$$\operatorname{span}((1,0)), \quad \operatorname{span}((1,1)), \quad \operatorname{span}((1,2)), \quad \operatorname{span}((1,3)), \quad \operatorname{span}((1,4)), \quad \operatorname{span}((0,1)).$$

For $(2,3)$: $3/2 = 3 \cdot 2^{-1} = 3 \cdot 3 = 9 = 4$, so $(2,3) = 2(1,4)$ lies on $\operatorname{span}((1,4))$.

(b) For $(x,y) \in \mathbb{F}_5^2$, the equation $a(1,2) + b(1,3) = (x,y)$ reads $a + b = x$ and $2a + 3b = y$. Subtracting twice the first equation from the second gives $b = y - 2x$, and then $a = x - b = 3x - y$. Conversely these values satisfy both equations: $a + b = x$, and $2a + 3b = 6x - 2y + 3y - 6x = y$. So every (x,y) has exactly one representation, $\mathcal{B}$ is a basis (Definition 1.6.1), and by (1.6.1),

$$[(x,y)]_{\mathcal{B}} = (3x - y, y - 2x).$$

For $(2,3)$: $3 \cdot 2 - 3 = 3$ and $3 - 4 = -1 = 4$, so $[(2,3)]_{\mathcal{B}} = (3,4)$. Check: $3(1,2) + 4(1,3) = (7,18) = (2,3)$ in $\mathbb{F}_5^2$.

(c) A list $\mathbf{u}, \mathbf{v}$ in $\mathbb{F}_5^2$ with $\mathbf{u} \neq \mathbf{0}$ spans $\mathbb{F}_5^2$ exactly when $\mathbf{v} \notin \operatorname{span}(\mathbf{u})$ (Exercise 1.3.37(b)), and then it is also independent by Proposition 1.5.14, since the list $\mathbf{u}$ is independent (Proposition 1.5.15(b)). So it is a basis exactly when $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \notin \operatorname{span}(\mathbf{u})$ (Proposition 1.6.2). The first entry is one of the 4 nonzero vectors of $\operatorname{span}((1,2))$, and its span is that line, which has 5 elements. So there are $25 - 5 = 20$ choices for the second entry, and $4 \cdot 20 = 80$ bases in all.

Solution to Problem 1.6

> *Uses.* Definition 1.2.17 (Evaluation), Proposition 1.5.24 (Distinct degrees), Proposition 1.6.2 (Bases are independent spanning lists), Proposition 1.4.24 (The sum of two subspaces), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 1.6.7 (Standard bases)
> *Hint 1.* Write $p = a + bx + cx^2 + dx^3$. Then $p(1) = a + b + c + d$ and $p(-1) = a - b + c - d$. Add and subtract.
> *Hint 2.* For $U + W$, find $1, x, x^2, x^3$ in $U + W$. The vectors $x - 1 \in U$ and $x + 1 \in W$ are a good start. Over $\mathbb{F}_2$, what is -1?
> *Sketch.* (a) $p(1) = p(-1) = 0$ if and only if $a + c = 0$ and $b + d = 0$ (this uses $1 + 1 \neq 0$), so $U \cap W = \operatorname{span}\big(1 - x^2, x - x^3\big)$, a basis by distinct degrees. $U + W$ contains $(x + 1) - (x - 1) = 2$, hence 1, then x, $x^2 = \big(x^2 - 1\big) + 1$, $x^3 = \big(x^3 - 1\big) + 1$. (b) Over $\mathbb{F}_2$, $-1 = 1$, so $U = W = U \cap W = U + W = \operatorname{span}\big(1 + x, 1 + x^2, 1 + x^3\big)$, a proper subspace.

Write $p = a + bx + cx^2 + dx^3$. By Definition 1.2.17, $p(1) = a + b + c + d$ and $p(-1) = a - b + c - d$.

(a) If $a + c = 0$ and $b + d = 0$, then $p(1) = (a + c) + (b + d) = 0$ and $p(-1) = (a + c) - (b + d) = 0$. Conversely, if $p(1) = p(-1) = 0$, then adding and subtracting gives $(1 + 1)(a + c) = 0$ and $(1 + 1)(b + d) = 0$, so $a + c = 0$ and $b + d = 0$ because $1 + 1 \neq 0$ (Proposition 1.1.17). Hence $p \in U \cap W$ exactly when $c = -a$ and $d = -b$, that is,

$$p = a\big(1 - x^2\big) + b\big(x - x^3\big).$$

So $U \cap W = \operatorname{span}\big(1 - x^2, x - x^3\big)$. The two polynomials are nonzero of degrees 2 and 3, so they are independent (Proposition 1.5.24), and they form a basis of $U \cap W$ (Proposition 1.6.2).

For $U + W$: the polynomials $x - 1$, $x^2 - 1$, $x^3 - 1$ vanish at 1 and lie in U, and $x + 1$ vanishes at -1 and lies in W. The sum $U + W$ is a subspace containing U and W (Proposition 1.4.24), so it contains

$$(x + 1) - (x - 1) = 1 + 1,$$

hence $1 = (1 + 1)^{-1}(1 + 1)$, and then $x = (x - 1) + 1$, $x^2 = \big(x^2 - 1\big) + 1$, and $x^3 = \big(x^3 - 1\big) + 1$. These span $F[x]_{\leq 3}$ (Proposition 1.6.7), so $U + W = F[x]_{\leq 3}$ by Proposition 1.4.14(b).

(b) Over $\mathbb{F}_2$, $-1 = 1$, so $p(-1) = p(1)$ for every p, and $U = W$. Hence $U \cap W = U + W = U$. Since $-(b + c + d) = b + c + d$ over $\mathbb{F}_2$, $p \in U$ exactly when $a = b + c + d$, that is,

$$p = b(1 + x) + c\big(1 + x^2\big) + d\big(1 + x^3\big).$$

So $U = \operatorname{span}\big(1+x, 1+x^2, 1+x^3\big)$, and this list is a basis of U by distinct degrees. The subspace $U + W = U$ is proper, since $1 \notin U$. Over $\mathbb{F}_2$, the two conditions are one condition, and the computation in (a) broke exactly where it divided by $1+1$.

Solution to Problem 1.7

> *Uses.* Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), (1.6.1), the addition and double-angle formulas
> *Hint 1.* For independence, evaluate a relation $a + b\cos t + c\sin t = 0$ at $t = 0$, π, and $\pi/2$.
> *Hint 2.* For (b), use $\cos(s+u) = \cos s\cos u - \sin s\sin u$, $\cos t = 1 - 2\sin^2(t/2)$, and $\sin t = 2\sin(t/2)\cos(t/2)$. For (c), the same three evaluations force the coefficients; test a fourth point.
> *Sketch.* (a) The list spans V by definition; evaluating at 0, π, $\pi/2$ gives $a+b = a-b = a+c = 0$. (b) $\big(0, 1/2, -\sqrt{3}/2\big)$, $(1/2, -1/2, 0)$, and $(1, 0, 1)$. (c) No: if $\cos^2 t = a + b\cos t + c\sin t$, the points 0, π, $\pi/2$ force $a = 1$, $b = 0$, $c = -1$, and $t = -\pi/2$ gives $0 = 2$.

(a) The list spans V by definition. Suppose $a + b\cos t + c\sin t = 0$ for all t. At $t = 0$ this gives $a + b = 0$, at $t = \pi$ it gives $a - b = 0$, so $a = b = 0$, and at $t = \pi/2$ it gives $a + c = 0$, so $c = 0$. The list is independent (Theorem 1.5.5), hence a basis of V (Proposition 1.6.2).

(b) By (1.6.1), it suffices to write each function as a combination of $1, \cos, \sin$, valid for every t.

- $\cos(t+\pi/3) = \cos(\pi/3)\cos t - \sin(\pi/3)\sin t = (1/2)\cos t - \big(\sqrt{3}/2\big)\sin t$, with coordinates $\big(0, 1/2, -\sqrt{3}/2\big)$.
- $\cos t = 1 - 2\sin^2(t/2)$, so $\sin^2(t/2) = 1/2 - (1/2)\cos t$, with coordinates $(1/2, -1/2, 0)$.
- $(\cos(t/2) + \sin(t/2))^2 = \cos^2(t/2) + \sin^2(t/2) + 2\sin(t/2)\cos(t/2) = 1 + \sin t$, with coordinates $(1, 0, 1)$.

(c) No. Suppose $\cos^2 t = a + b\cos t + c\sin t$ for all t. At $t = 0$ and $t = \pi$ we get $1 = a + b$ and $1 = a - b$, so $a = 1$ and $b = 0$. At $t = \pi/2$ we get $0 = a + c$, so $c = -1$. So the only candidate is $1 - \sin t$. But at $t = -\pi/2$, $\cos^2 t = 0$ while $1 - \sin t = 2$. Hence $t \mapsto \cos^2 t$ is not in V. Three evaluations determined the only possible coefficients, and a fourth refuted them.

Solution to Problem 1.8

> *Uses.* Definition 1.1.32 (Subfield), Proposition 1.2.19 (Restriction of scalars), Theorem 1.4.3 (Subspace criterion), Definition 1.2.4 (The space F^n), Definition 1.1.29 (Complex numbers)
> *Hint 1.* Closure of K^n under scalars from K comes from closure of K under products. For the failure over F, pick $c \in F$ outside K.
> *Hint 2.* For (b), take a nonzero $\mathbf{v} \in U$ and a coordinate $v_i \neq 0$. Which multiples of $\mathbf{v}$ lie in K^n?
> *Hint 3.* For (c), check the subspace criterion directly. In $\mathbb{C}^2$, compare $\operatorname{span}((1,i))$ and $\operatorname{span}((1,1))$.
> *Sketch.* (a) K^n contains $\mathbf{0}$ and is closed under sums and K-multiples coordinatewise; but $c\mathbf{e}_1 \notin K^n$ for $c \in F \setminus K$. (b) If $\mathbf{v} \in U$ is nonzero with $v_i \neq 0$, then $\big(cv_i^{-1}\big)\mathbf{v} \in U \subseteq K^n$ has i-th coordinate c, so every $c \in F$ lies in K. (c) Sums and K-multiples of vectors in $U \cap K^n$ stay in both sets. $\operatorname{span}((1,i)) \cap \mathbb{R}^2 = \{\mathbf{0}\}$, while $\operatorname{span}((1,1)) \cap \mathbb{R}^2$ contains $(1,1)$.

(a) K^n contains $\mathbf{0}$, since $0 \in K$. For $\mathbf{x}, \mathbf{y} \in K^n$ and $a \in K$, the coordinates $x_i + y_i$ and ax_i lie in K, because K is closed under the operations of F (Definition 1.1.32). So K^n is a subspace of the K-vector space F^n by Theorem 1.4.3. Over F it is not closed under scalar multiplication: choose $c \in F$ with $c \notin K$; then $\mathbf{e}_1 \in K^n$, but $c\mathbf{e}_1 = (c, 0, \ldots, 0) \notin K^n$.

(b) $\{\mathbf{0}\}$ is such a subspace. Let U be a subspace of the F-vector space F^n with $U \subseteq K^n$, and suppose for contradiction that U contains some $\mathbf{v} \neq \mathbf{0}$. Choose i with $v_i \neq 0$. For any $c \in F$, the vector $\big(cv_i^{-1}\big)\mathbf{v}$ lies in U, hence in K^n, and its i-th coordinate is $cv_i^{-1}v_i = c$. So $c \in K$ for every $c \in F$, that is, $F = K$, a contradiction.

(c) $U \cap K^n$ contains $\mathbf{0}$. Let $\mathbf{x}, \mathbf{y} \in U \cap K^n$ and $a \in K$. Then $\mathbf{x} + \mathbf{y} \in U$ and $a\mathbf{x} \in U$, because U is closed under sums and under multiples by scalars of F, and $a \in K \subseteq F$. Also $\mathbf{x} + \mathbf{y} \in K^n$ and $a\mathbf{x} \in K^n$, by the computation in (a). So $U \cap K^n$ is a subspace of the K-vector space K^n by Theorem 1.4.3.

Now let $F = \mathbb{C}$, $K = \mathbb{R}$, $n = 2$. For $U = \operatorname{span}((1,i))$, a vector $c(1,i) = (c, ci)$ lies in $\mathbb{R}^2$ only if $c \in \mathbb{R}$ and $ci \in \mathbb{R}$. Writing $ci = 0 + ci$ (Definition 1.1.29), the second condition forces $c = 0$. So $U \cap \mathbb{R}^2 = \{\mathbf{0}\}$, although U is a line in $\mathbb{C}^2$. For $U = \operatorname{span}((1,1))$, the intersection $U \cap \mathbb{R}^2$ contains $(1,1)$; in fact it is $\{(c,c) \mid c \in \mathbb{R}\}$, the real line through $(1,1)$.

By (b), $U \cap \mathbb{R}^2$ is never all of U unless $U = \{\mathbf{0}\}$.

Solution to Problem 1.9

> *Uses.* Proposition 1.2.22 (Transport of structure), Proposition 1.2.19 (Restriction of scalars), Proposition 1.5.15 (Short lists, zeros, and repeats), Theorem 1.5.5 (Independence is decided at zero), unique factorization of positive integers into primes
> *Hint 1.* The zero vector satisfies $x \oplus \mathbf{0} = x$, that is, $x \cdot \mathbf{0} = x$. A linear combination $a \odot 2 \oplus b \odot 3$ is the real number $2^a 3^b$.
> *Hint 2.* Over $\mathbb{Q}$, a relation reads $2^a 3^b = 1$ with a, b rational. Clear denominators by raising to a suitable integer power, and use unique factorization.
> *Sketch.* (a) The zero vector is 1, so $1, 2$ is dependent. $4 = 2 \odot 2$ and $3 = (\log_2 3) \odot 2$, so $2, 4$ and $2, 3$ are dependent over $\mathbb{R}$. (b) If $2^{m/N} 3^{n/N} = 1$ with integers m, n and $N \geq 1$, then $2^m 3^n = 1$, so $m = n = 0$ by unique factorization. $2 \oplus 3 \oplus (-1) \odot 6 = 1$ and $2 \odot 2 \oplus (-1) \odot 4 = 1$. (c) $5 = 2^a 3^b$ with rational a, b would give $5^N = 2^m 3^n$ after raising to the power N.

(a) The zero vector is 1, since $x \oplus 1 = x \cdot 1 = x$ for all $x \in P$. So $1, 2$ is dependent: it contains the zero vector (Proposition 1.5.15(c)). Next, $2 \odot 2 = 2^2 = 4$, so 4 is a scalar multiple of 2, and $2, 4$ is dependent (Proposition 1.5.15(d)). Finally, $(\log_2 3) \odot 2 = 2^{\log_2 3} = 3$, so $2, 3$ is dependent as well. In fact any two vectors x, y of P are dependent over $\mathbb{R}$: if $x = 1$, the list contains the zero vector, and if $x \neq 1$, then $y = (\ln y / \ln x) \odot x$ is a multiple of x.

(b) A linear combination over $\mathbb{Q}$ represents $a \odot 2 \oplus b \odot 3 = 2^a 3^b$. Suppose $2^a 3^b = 1$ with $a, b \in \mathbb{Q}$. Write $a = m/N$ and $b = n/N$ with integers m, n and $N \geq 1$. Raising to the power N gives $2^m 3^n = 1$, that is,

$$2^{m^+} 3^{n^+} = 2^{m^-} 3^{n^-},$$

where $m^+ = \max(m, 0)$ and $m^- = \max(-m, 0)$, and similarly for n. Both sides are positive integers, so by unique factorization the exponents of 2 agree and the exponents of 3 agree: $m^+ = m^-$ and $n^+ = n^-$. One of m^+, m^- is always 0, so $m = 0$, and likewise $n = 0$. Hence $a = b = 0$, and $2, 3$ is independent over $\mathbb{Q}$ by Theorem 1.5.5.

On the other hand, $1 \odot 2 \oplus 1 \odot 3 \oplus (-1) \odot 6 = 2 \cdot 3 \cdot 6^{-1} = 1$, the zero vector, so $2, 3, 6$ is dependent over $\mathbb{Q}$. And $2 \odot 2 \oplus (-1) \odot 4 = 4 \cdot 4^{-1} = 1$, so $2, 4$ is dependent over $\mathbb{Q}$ too.

(c) Suppose $5 = 2^a 3^b$ with $a = m/N$, $b = n/N$ as before. Raising to the power N gives $5^N = 2^m 3^n$, so $5^N 2^{m^-} 3^{n^-} = 2^{m^+} 3^{n^+}$. The prime 5 divides the left side and not the right side, contradicting unique factorization. So $5 \notin \operatorname{span}(2, 3)$ over $\mathbb{Q}$.

In this structure 1 is the zero vector, so "a list containing 1" is the analogue of a list containing $\mathbf{0}$. Under the bijection $\ln : P \to \mathbb{R}$ the problem becomes a statement about $\ln 2$, $\ln 3$, and $\ln 5$ over $\mathbb{Q}$, which is Exercise 1.5.39.

Solution to Problem 1.10

> *Uses.* Proposition 1.5.15 (Short lists, zeros, and repeats), Proposition 1.5.14 (Appending a vector), Corollary 1.5.17 (Sublists), Proposition 1.4.24 (The sum of two subspaces), Proposition 1.4.14 (The span is the smallest subspace containing a list), Example 1.6.5 ($\mathbb{C}$ over $\mathbb{R}$ and over $\mathbb{C}$), Exercise 1.5.8
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* For (a), use Proposition 1.5.15(d): if one of $\mathbf{u}, \mathbf{w}$ were a multiple of the other, where would it lie?
> *Hint 2.* Test (b) over $\mathbb{F}_2$. For (c), treat independent and dependent $\mathcal{L}$ separately. For (d), (e), and (f), try $V = \mathbb{C}$, a list containing $\mathbf{0}$, and the subspace $U + W$.
> *Sketch.* (a) True: $\mathbf{u} = c\mathbf{w}$ would put $\mathbf{u}$ in W, and $\mathbf{w} = c\mathbf{u}$ would put $\mathbf{w}$ in U. (b) False over $\mathbb{F}_2$: the axes of $\mathbb{F}_2^2$ with $\mathbf{u} = \mathbf{e}_1$, $\mathbf{w} = \mathbf{e}_2$ give $\mathbf{u} + \mathbf{w} = \mathbf{u} - \mathbf{w}$. (c) True: the set is $\operatorname{span}(\mathcal{L})$ if $\mathcal{L}$ is independent, and V otherwise. (d) False: 1 spans $\mathbb{C}$ over $\mathbb{C}$, not over $\mathbb{R}$. (e) False: $\mathcal{L} = \mathbf{0}$. (f) True: $U + W$ is a subspace containing the entries.

(a) True. By Proposition 1.5.15(d), if $\mathbf{u}, \mathbf{w}$ were dependent, one would be a multiple of the other. If $\mathbf{u} = c\mathbf{w}$, then $\mathbf{u} \in W$, since W is closed under scalar multiplication; if $\mathbf{w} = c\mathbf{u}$, then $\mathbf{w} \in U$. Both contradict the hypotheses.

(b) False. Over $\mathbb{F}_2$, let $V = \mathbb{F}_2^2$, $U = \operatorname{span}(\mathbf{e}_1)$, $W = \operatorname{span}(\mathbf{e}_2)$, $\mathbf{u} = \mathbf{e}_1$, $\mathbf{w} = \mathbf{e}_2$. The hypotheses hold, but $\mathbf{u} - \mathbf{w} = \mathbf{u} + \mathbf{w}$ because $-1 = 1$, so the list repeats an entry and is dependent (Proposition 1.5.15(c)). (If $1 + 1 \neq 0$, the statement is true, by (a) and Exercise 1.5.8.)

(c) True. If $\mathcal{L}$ is independent, then by Proposition 1.5.14, $\mathcal{L}, \mathbf{v}$ is dependent exactly when $\mathbf{v} \in \operatorname{span}(\mathcal{L})$, so the set is $\operatorname{span}(\mathcal{L})$, a subspace (Proposition 1.4.14). If $\mathcal{L}$ is dependent, then $\mathcal{L}, \mathbf{v}$ has the dependent sublist $\mathcal{L}$ for every $\mathbf{v}$, so it is dependent (Corollary 1.5.17), and the set is V.

(d) False. The list 1 spans $\mathbb{C}$ over $\mathbb{C}$, but over $\mathbb{R}$ its span is $\mathbb{R} \neq \mathbb{C}$ (Example 1.6.5).

(e) False. For $\mathcal{L} = \mathbf{0}$, the list of length one, $\operatorname{span}(\mathcal{L}) = \{\mathbf{0}\}$, so $\operatorname{span}(\mathcal{L}) \cap U = \{\mathbf{0}\}$ for every U, but $\mathcal{L}$ is dependent (Proposition 1.5.15(b)).

(f) True. $U + W$ is a subspace containing U and W (Proposition 1.4.24), so it contains every entry of $\mathcal{L}$, and hence $\operatorname{span}(\mathcal{L})$ by Proposition 1.4.14(b).

Solution to Problem 1.11

Uses. Theorem 1.6.10 (Every spanning list contains a basis), Corollary 1.5.13 (Independence, one entry at a time), Exercise 1.3.36, Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Which entries come before $\mathbf{w}_j$ in the concatenation? Try $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \mathbb{R}^2$ with $\mathcal{M} = \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$.
Hint 2. For (b), the entries of $\mathcal{L}$ come first, and $\mathcal{L}$ is independent. For (c), write $\mathbf{w}_j$ as a combination of the earlier entries and move the $\mathbf{w}$-terms to the left.
Sketch. (a) The entries before $\mathbf{w}_j$ include $\mathbf{w}_1, \ldots, \mathbf{w}_{j-1}$, whose span need not lie in U. With $\mathcal{L} = \mathbf{e}_1$ and $\mathcal{M} = \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$ in $\mathbb{R}^2$, sifting deletes $\mathbf{e}_1 + \mathbf{e}_2 \notin U$. (b) The first k entries form the independent list $\mathcal{L}$, so none lies in the span of earlier ones; the kept entries are independent and span $\operatorname{span}(\mathcal{L}, \mathcal{M}) = U + W$. (c) $\mathbf{w}_j = \sum a_i \mathbf{u}_i + \sum_{l<j} b_l \mathbf{w}_l$ gives $\mathbf{w}_j - \mathbf{y} = \sum a_i \mathbf{u}_i \in U \cap W$ with $\mathbf{y} = \sum_{l<j} b_l \mathbf{w}_l$, and it is nonzero because $\mathcal{M}$ is independent.

(a) The error is "these are the entries of $\mathcal{L}$". The entries before $\mathbf{w}_j$ in $\mathcal{L}, \mathcal{M}$ are $\mathbf{u}_1, \ldots, \mathbf{u}_k$ *and* $\mathbf{w}_1, \ldots, \mathbf{w}_{j-1}$, and the span of these need not lie in U. For a counterexample, take $V = \mathbb{R}^2$, $U = \operatorname{span}(\mathbf{e}_1)$ with basis $\mathcal{L} = \mathbf{e}_1$, and $W = \mathbb{R}^2$ with basis $\mathcal{M} = \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$. Sifting $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$ keeps $\mathbf{e}_1$ (it is nonzero), keeps $\mathbf{e}_2$ (it is not a multiple of $\mathbf{e}_1$), and deletes $\mathbf{e}_1 + \mathbf{e}_2 \in \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$. But $\mathbf{e}_1 + \mathbf{e}_2 \notin U$, so the deleted entry is not in $U \cap W = U$.

(b) For $j \leq k$, the entry $\mathbf{u}_j$ is preceded only by $\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}$. Since $\mathcal{L}$ is independent, $\mathbf{u}_j \notin \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{j-1})$ by Corollary 1.5.13, so $\mathbf{u}_j$ is kept. By Theorem 1.6.10 (third proof), the kept entries form an independent list with the same span as $\mathcal{L}, \mathcal{M}$, which is $\operatorname{span}(\mathcal{L}) + \operatorname{span}(\mathcal{M}) = U + W$ by Exercise 1.3.36. So they form a basis of $U + W$ (Proposition 1.6.2).

(c) If $\mathbf{w}_j$ is deleted, then $\mathbf{w}_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_{j-1})$, say

$$\mathbf{w}_j = \sum_{i=1}^{k} a_i \mathbf{u}_i + \sum_{l=1}^{j-1} b_l \mathbf{w}_l.$$

Let $\mathbf{y} = \sum_{l<j} b_l \mathbf{w}_l$. Then $\mathbf{w}_j - \mathbf{y} = \sum_i a_i \mathbf{u}_i$ lies in U, and it lies in W because $\mathbf{w}_j$ and $\mathbf{y}$ do. It is nonzero: otherwise $\mathbf{w}_j - \sum_{l<j} b_l \mathbf{w}_l = \mathbf{0}$ would be a nontrivial relation of $\mathcal{M}$ (the coefficient of $\mathbf{w}_j$ is 1), contradicting the independence of $\mathcal{M}$.

So each deleted entry produces a nonzero vector of $U \cap W$, though not necessarily the entry itself. Showing that these vectors form a basis of $U \cap W$ gives one route to the formula of Section 4.4 relating the sizes of bases of $U, W, U + W$, and $U \cap W$.

Solution to Problem 1.12

Uses. Theorem 1.5.5 (Independence is decided at zero), Example 1.5.19 (The same vectors over two fields), Definition 1.1.32 (Subfield), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 1.1.17 (No zero divisors)
Hint 1. First, $1, t$ is independent over $\mathbb{Q}$ because $t \notin \mathbb{Q}$. Given $a + bt + ct^2 = 0$, multiply by t (using $t^3 = 2$) and combine the two equations to eliminate t^2.

Hint 2. Eliminating t^2 leaves $\big(2c^2 - ab\big) + \big(ac - b^2\big)t = 0$, so $ab = 2c^2$ and $b^2 = ac$. If $c \neq 0$, compute b^3.
Hint 3. For (c), expand with $t^3 = 2$ and $t^4 = 2t$. For inverses, show that $\alpha \neq 0$ forces $\alpha\beta \neq 0$; if $\alpha\beta = 0$ then $\beta = 0$, and (a) turns this into equations like those in the previous hint.
Sketch. (a) From $a + bt + ct^2 = 0$ and $2c + at + bt^2 = 0$, eliminating t^2 gives $\big(2c^2 - ab\big) + \big(ac - b^2\big)t = 0$, so $ab = 2c^2$ and $b^2 = ac$. Then $b^3 = abc = 2c^3$, so $c = 0$ (else $(b/c)^3 = 2$), then $b = 0$, $a = 0$. (b) Expand with $t^3 = 2$. (c) Expanding, the coefficients of t and t^2 cancel. If $\alpha \neq 0$ and $N = \alpha\beta = 0$, then $\beta = 0$, so $a^2 = 2bc$, $ab = 2c^2$, $b^2 = ac$, which force $a = b = c = 0$ as in (a). So $\alpha^{-1} = \beta/N \in K$.

Throughout, $t^3 = 2$ and $t^4 = 2t$. Since $t \notin \mathbb{Q}$, the list $1, t$ is independent over $\mathbb{Q}$: if $q + rt = 0$ with $r \neq 0$, then $t = -q/r \in \mathbb{Q}$; so $r = 0$, and then $q = 0$ (as in Example 1.5.19(b)).

(a) Suppose $a + bt + ct^2 = 0$ with $a, b, c \in \mathbb{Q}$. Multiplying by t gives $2c + at + bt^2 = 0$. Now c times the second equation minus b times the first eliminates t^2:

$$c\big(2c + at + bt^2\big) - b\big(a + bt + ct^2\big) = \big(2c^2 - ab\big) + \big(ac - b^2\big)t = 0.$$

By independence of $1, t$, we get $ab = 2c^2$ and $b^2 = ac$. Suppose $c \neq 0$. Then $b^3 = b \cdot b^2 = abc = (ab)c = 2c^3$, so $(b/c)^3 = 2$ with $b/c \in \mathbb{Q}$, which is impossible. So $c = 0$. Then $b^2 = ac = 0$ gives $b = 0$, and the original equation gives $a = 0$. By Theorem 1.5.5, $1, t, t^2$ is independent over $\mathbb{Q}$.

(b) For $\alpha = a + bt + ct^2$ and $\alpha' = a' + b't + c't^2$ in K, expanding and using $t^3 = 2$, $t^4 = 2t$,

$$\alpha\alpha' = (aa' + 2bc' + 2cb') + (ab' + ba' + 2cc')t + (ac' + bb' + ca')t^2 \in K.$$

(c) Write $\beta = A + Bt + Ct^2$ with $A = a^2 - 2bc$, $B = 2c^2 - ab$, $C = b^2 - ac$. By the formula in (b), the coefficients of $\alpha\beta$ are

$$\begin{aligned} aA + 2bC + 2cB &= a^3 - 2abc + 2b^3 - 2abc + 4c^3 - 2abc = a^3 + 2b^3 + 4c^3 - 6abc, \\ aB + bA + 2cC &= 2ac^2 - a^2b + a^2b - 2b^2c + 2b^2c - 2ac^2 = 0, \\ aC + bB + cA &= ab^2 - a^2c + 2bc^2 - ab^2 + a^2c - 2bc^2 = 0. \end{aligned}$$

So $\alpha\beta = N$, where $N = a^3 + 2b^3 + 4c^3 - 6abc \in \mathbb{Q}$.

Now K contains 0 and 1, and it is closed under addition and negation because it is a span, hence a subspace of the $\mathbb{Q}$-vector space $\mathbb{R}$ (Proposition 1.4.14). It is closed under multiplication by (b). For inverses, let $\alpha \in K$ be nonzero. We claim $N \neq 0$. If $N = 0$, then $\alpha\beta = 0$ with $\alpha \neq 0$, so $\beta = 0$ (Proposition 1.1.17), and by (a) all its coefficients vanish:

$$a^2 = 2bc, \quad ab = 2c^2, \quad b^2 = ac.$$

The last two are the equations of (a), which force $c = 0$ and $b = 0$, and then $a^2 = 0$ gives $a = 0$. So $\alpha = 0$, a contradiction. Hence $N \neq 0$, and

$$\alpha^{-1} = N^{-1}\beta = N^{-1}A + N^{-1}Bt + N^{-1}Ct^2 \in K.$$

By Definition 1.1.32, K is a subfield of $\mathbb{R}$.

For instance, $(1 + t)^{-1}$: here $a = b = 1$, $c = 0$, so $N = 1 + 2 = 3$ and $\beta = 1 - t + t^2$; indeed $(1 + t)\big(1 - t + t^2\big) = 1 + t^3 = 3$.

Solution to Problem 1.13

Uses. Exercise 1.5.43, Proposition 1.5.14 (Appending a vector), Proposition 1.4.14 (The span is the smallest subspace containing a list), Exercise 1.3.28, Proposition 1.5.15 (Short lists, zeros, and repeats)
Hint 1. For (a), imitate the proof of Theorem 1.4.25: pick vectors of U one at a time, each outside the span of the previous ones. Three independent vectors span q^3 elements. The four kinds have different sizes.
Hint 2. Lines: each nonzero vector lies on exactly one line, and each line has $q - 1$ nonzero vectors.
Hint 3. Planes: count independent lists of length 2 in F^3 in two ways. Each spans exactly one plane, and each plane contains $\big(q^2 - 1\big)\big(q^2 - q\big)$ of them, all spanning it.

Sketch. (a) For $U \neq \{\mathbf{0}\}$, choose $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{w} \in U \setminus \operatorname{span}(\mathbf{v})$ if possible, then $\mathbf{x} \in U \setminus \operatorname{span}(\mathbf{v}, \mathbf{w})$ if possible; the lists stay independent (Proposition 1.5.14), and $\operatorname{span}(\mathbf{v}, \mathbf{w}, \mathbf{x})$ has q^3 elements, so it is F^3. Sizes $1 < q < q^2 < q^3$ separate the kinds. (b) Lines: $(q^3-1)/(q-1) = q^2+q+1$. Planes: $(q^3-1)(q^3-q)/((q^2-1)(q^2-q)) = q^2+q+1$. Total $2q^2+2q+4$, which is 16 for $q = 2$.

(a) The four kinds are subspaces, since spans are (Proposition 1.4.14). Let U be a subspace of F^3 other than $\{\mathbf{0}\}$, and choose $\mathbf{v} \in U$ with $\mathbf{v} \neq \mathbf{0}$; then $\operatorname{span}(\mathbf{v}) \subseteq U$ by Proposition 1.4.14(b). If $U = \operatorname{span}(\mathbf{v})$, it is a line. Otherwise choose $\mathbf{w} \in U$ with $\mathbf{w} \notin \operatorname{span}(\mathbf{v})$. The list $\mathbf{v}$ is independent (Proposition 1.5.15(b)), so $\mathbf{v}, \mathbf{w}$ is independent by Proposition 1.5.14, and $\operatorname{span}(\mathbf{v}, \mathbf{w}) \subseteq U$. If equality holds, U is a plane. Otherwise choose $\mathbf{x} \in U$ with $\mathbf{x} \notin \operatorname{span}(\mathbf{v}, \mathbf{w})$. Then $\mathbf{v}, \mathbf{w}, \mathbf{x}$ is independent by Proposition 1.5.14, so its span has q^3 elements (Exercise 1.5.43(a)). That span is contained in F^3, which also has q^3 elements, so it equals F^3, and $F^3 \subseteq U$. So U is one of the four kinds.

By Exercise 1.5.43(a), a line has q elements and a plane has q^2; also $|\{\mathbf{0}\}| = 1$ and $|F^3| = q^3$. Since $q \geq 2$, the numbers $1 < q < q^2 < q^3$ are different, so no subspace is of two kinds.

(b) *Lines.* Every nonzero vector $\mathbf{v}$ lies on the line $\operatorname{span}(\mathbf{v})$, and on no other: if $\mathbf{v} \in \operatorname{span}(\mathbf{u})$ with $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{v} = c\mathbf{u}$ with $c \neq 0$, so $\operatorname{span}(\mathbf{v}) = \operatorname{span}(\mathbf{u})$ (Exercise 1.3.28(b)). So the lines split the $q^3 - 1$ nonzero vectors of F^3 into disjoint groups, each consisting of the $q - 1$ nonzero vectors of one line. Hence there are

$$\frac{q^3-1}{q-1} = q^2+q+1$$

lines.

Planes. We count the independent lists $\mathbf{v}, \mathbf{w}$ of length 2 in F^3 in two ways. By Exercise 1.5.43(b), there are $(q^3-1)(q^3-q)$ of them. On the other hand, each of them spans exactly one plane, namely $\operatorname{span}(\mathbf{v}, \mathbf{w})$. Fix a plane P. An independent list $\mathbf{v}, \mathbf{w}$ with both entries in P has a span with q^2 elements contained in P, which has q^2 elements, so it spans P. The number of such lists is $(q^2-1)(q^2-q)$: there are $q^2 - 1$ choices of $\mathbf{v} \in P$ with $\mathbf{v} \neq \mathbf{0}$, and then, by Proposition 1.5.14, the choices for $\mathbf{w}$ are the $q^2 - q$ vectors of P outside $\operatorname{span}(\mathbf{v})$. So each plane is spanned by exactly $(q^2-1)(q^2-q)$ independent lists, and the number of planes is

$$\frac{(q^3-1)(q^3-q)}{(q^2-1)(q^2-q)} = \frac{(q^3-1)q(q^2-1)}{(q^2-1)q(q-1)} = \frac{q^3-1}{q-1} = q^2+q+1.$$

All subspaces. $1 + (q^2+q+1) + (q^2+q+1) + 1 = 2q^2+2q+4$. For $q = 2$ this is 16: one zero subspace, seven lines with 2 elements each, seven planes with 4 elements each, and $\mathbb{F}_2^3$ itself.

Over a finite field, counting elements did the work that dimension does in general: a subspace containing an independent list of length 3 is all of F^3 because it is too big to be anything else. That lines and planes come in equal numbers is not an accident; it is explained by duality (Section 13.1).

Solution to Problem 1.14

Uses. Proposition 1.2.30 (Additive arithmetic), Theorem 1.4.3 (Subspace criterion), Proposition 1.4.14 (The span is the smallest subspace containing a list), Example 1.4.8 (One equation in F^n)
Hint 1. Pick $\mathbf{v} \notin U$, and for each scalar a consider the set $a\mathbf{v} + U = \{a\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\}$.
Hint 2. These q sets each have $|U|$ elements. Show that they are pairwise disjoint.
Sketch. (a) For $\mathbf{v} \notin U$, the q sets $a\mathbf{v} + U$ have $|U|$ elements each and are disjoint: $a\mathbf{v} + \mathbf{u} = b\mathbf{v} + \mathbf{u}'$ with $a \neq b$ would give $\mathbf{v} = (a-b)^{-1}(\mathbf{u}' - \mathbf{u}) \in U$. So $q|U| \leq q^n$. (b) $\operatorname{span}(S)$ is a subspace with more than q^{n-1} elements, so it is not proper. (c) The subspace $x_1 = 0$.

(a) Since U is proper, choose $\mathbf{v} \in F^n$ with $\mathbf{v} \notin U$. For $a \in F$, let $a\mathbf{v} + U = \{a\mathbf{v} + \mathbf{u} \mid \mathbf{u} \in U\}$. The map $\mathbf{u} \mapsto a\mathbf{v} + \mathbf{u}$ from U to $a\mathbf{v} + U$ is surjective by definition and injective by cancellation (Proposition 1.2.30(b)), so $a\mathbf{v} + U$ has exactly $|U|$ elements.

The q sets $a\mathbf{v}+U$, $a\in F$, are pairwise disjoint. Suppose $a\mathbf{v}+\mathbf{u}=b\mathbf{v}+\mathbf{u}'$ with $\mathbf{u},\mathbf{u}'\in U$ and $a\neq b$. Then $(a-b)\mathbf{v}=\mathbf{u}'-\mathbf{u}\in U$, and since $a-b\neq 0$, $\mathbf{v}=(a-b)^{-1}(\mathbf{u}'-\mathbf{u})\in U$ by closure under scalar multiplication (Theorem 1.4.3), a contradiction.

So F^n contains q disjoint sets of $|U|$ elements each, and $q|U|\le |F^n|=q^n$. Hence $|U|\le q^{n-1}$.

(b) Let $\mathcal{L}$ be a list whose entries run through the elements of S. By Proposition 1.4.14(a), $\operatorname{span}(\mathcal{L})$ is a subspace of F^n containing every element of S, so it has more than q^{n-1} elements. By (a) it is not a proper subspace, so $\operatorname{span}(\mathcal{L})=F^n$.

(c) Let $S=\{\mathbf{x}\in F^n \mid x_1=0\}$. It is a subspace (Example 1.4.8), so the span of any list running through S is S itself (Proposition 1.4.14), and $S\neq F^n$ since $\mathbf{e}_1\notin S$. Its elements are determined by the free coordinates $x_2,\dots,x_n$, so $|S|=q^{n-1}$.

So a subset of F^n that contains more than a $1/q$ fraction of all vectors spans F^n. The sets $a\mathbf{v}+U$ in (a) are translates of U, which Section 4.6 studies as the elements of the quotient space F^n/U.

Solution to Problem 1.15

Uses. Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.6.2 (Bases are independent spanning lists), (1.6.1), the binomial theorem, continuity of polynomial functions
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For (a), expand $(1-t)^{n-k}$ by the binomial theorem.
Hint 2. For (b), if m is the smallest index with $c_m\neq 0$, divide the relation by t^m for $t\neq 0$, and let $t\to 0$.
Hint 3. For (c), write $t^j=t^j(t+(1-t))^{n-j}$ and expand.
Sketch. (a) $b_k(t)=\sum_i\binom{n-k}{i}(-1)^i t^{k+i}$. (b) If $\sum c_k b_k=0$ and m is the smallest index with $c_m\neq 0$, then $\sum_{k\ge m}c_k t^{k-m}(1-t)^{n-k}=0$ for $t\neq 0$; this is a polynomial function, continuous at 0, whose value at 0 is c_m. So $c_m=0$, a contradiction. (c) $t^j=\sum_i\binom{n-j}{i}b_{j+i}(t)$. (d) $1=\sum_k\binom{n}{k}b_k$ and $t=\sum_{k\ge 1}\binom{n-1}{k-1}b_k$.

(a) By the binomial theorem, for every real t,

$$b_k(t)=t^k\sum_{i=0}^{n-k}\binom{n-k}{i}(-t)^i=\sum_{i=0}^{n-k}\binom{n-k}{i}(-1)^i t^{k+i}.$$

So b_k is a linear combination of $t^k,\dots,t^n$, and $b_k\in V_n$.

(b) Suppose $c_0b_0+\cdots+c_nb_n=\mathbf{0}$, the zero function, and suppose for contradiction that some $c_k\neq 0$. Let m be the smallest index with $c_m\neq 0$. Then for every real t,

$$\sum_{k=m}^{n}c_k t^k(1-t)^{n-k}=0,$$

and dividing by t^m for $t\neq 0$ gives $g(t)=0$ for all $t\neq 0$, where

$$g(t)=\sum_{k=m}^{n}c_k t^{k-m}(1-t)^{n-k}=c_m(1-t)^{n-m}+\sum_{k=m+1}^{n}c_k t^{k-m}(1-t)^{n-k}.$$

The function g is a polynomial function, hence continuous, so $g(0)=\lim_{t\to 0}g(t)=0$. But every term of the second sum vanishes at $t=0$, since $k-m\ge 1$, so $g(0)=c_m$. Hence $c_m=0$, a contradiction. By Theorem 1.5.5, $b_0,\dots,b_n$ is independent.

(c) For $0\le j\le n$ and every real t, since $t+(1-t)=1$,

$$t^j=t^j(t+(1-t))^{n-j}=\sum_{i=0}^{n-j}\binom{n-j}{i}t^{j+i}(1-t)^{n-j-i}=\sum_{i=0}^{n-j}\binom{n-j}{i}b_{j+i}(t).$$

So each t^j lies in $\operatorname{span}(b_0,\dots,b_n)$, and $V_n\subseteq\operatorname{span}(b_0,\dots,b_n)$ by Proposition 1.4.14(b). With (a), the two spans are equal, so $b_0,\dots,b_n$ spans V_n; with (b), it is a basis of V_n (Proposition 1.6.2).

(d) Taking $j = 0$ and $j = 1$ in (c),

$$1 = \sum_{k=0}^{n} \binom{n}{k} b_k, \quad t = \sum_{k=1}^{n} \binom{n-1}{k-1} b_k.$$

By (1.6.1), the coordinates of 1 are $\left(\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}\right)$, and those of t are $\left(0, \binom{n-1}{0}, \binom{n-1}{1}, \ldots, \binom{n-1}{n-1}\right)$.

The proof of (b) is the extremal argument of Proposition 1.5.24 and Example 1.5.27, read at $t = 0$ instead of at $t = \infty$: near 0, the term with the smallest power of t dominates. These functions are the Bernstein basis of V_n.

Solution to Problem 1.16

Uses. Definition 1.1.28 (Characteristic), Exercise 1.1.48, Definition 1.1.32 (Subfield), Corollary 1.1.14 (Cancellation), Example 1.2.6 (Small fields), Proposition 1.2.19 (Restriction of scalars), Corollary 1.6.24 (Bases exist in finite dimensions), Proposition 1.6.16 (The coordinate map is a bijection), an injective map from a finite set to itself is surjective
Hint 1. For (a), the elements $m * 1$ for $m = 1, 2, 3, \ldots$ cannot all be different in a finite field.
Hint 2. For (b), every $m * 1$ with $m \geq 0$ equals $r * 1$ for the remainder r of m mod p. For inverses inside K, use the counting argument of Proposition 1.1.24: multiplication by $a \neq 0$ is injective on K.
Hint 3. For (c), regard F as a vector space over K. Why is it finite-dimensional, and what does a basis of length n say about the number of elements?
Sketch. (a) $m * 1 = m' * 1$ with $m < m'$ gives $(m' - m) * 1 = 0$, so $\operatorname{char} F > 0$, hence prime. (b) The p listed elements are distinct by minimality of p, and K is the set of all $m * 1$, which is closed under sums and products; $-(m * 1) = (p - m) * 1$; and $x \mapsto ax$ is an injection of the finite set K into itself, hence onto, so a has an inverse in K. (c) F is a K-vector space, spanned by the list of all its elements, so it has a basis of some length $n \geq 1$, and the coordinate map is a bijection $F \to K^n$.

(a) The elements $1 * 1, 2 * 1, 3 * 1, \ldots$ lie in the finite set F, so two of them coincide: $m * 1 = m' * 1$ with $1 \leq m < m'$. Subtracting $m * 1$ from both sides gives $(m' - m) * 1 = 0$ with $m' - m \geq 1$. So $\operatorname{char} F$ is positive (Definition 1.1.28), and by Exercise 1.1.48 it is a prime p.

(b) *K has p elements.* If $m * 1 = m' * 1$ with $0 \leq m < m' \leq p - 1$, then $(m' - m) * 1 = 0$ with $0 < m' - m < p$, contradicting the minimality of p. So the p listed elements are distinct.

K contains every $m * 1$. For an integer $m \geq 0$, write $m = sp + r$ with $0 \leq r \leq p - 1$. Grouping the m ones into s blocks of p and one block of r writes $m * 1$ as a sum of s copies of $p * 1$ plus $r * 1$. Since $p * 1 = 0$, this gives $m * 1 = r * 1 \in K$.

Closure. $0 = 0 * 1$ and $1 = 1 * 1$ lie in K. For $m, m' \geq 0$, $(m * 1) + (m' * 1) = (m + m') * 1 \in K$, and $(m * 1)(m' * 1) = (mm') * 1 \in K$, by distributivity (expand the product of the two sums of ones into mm' products $1 \cdot 1 = 1$). For $1 \leq m \leq p - 1$, $(m * 1) + ((p - m) * 1) = p * 1 = 0$, so $-(m * 1) = (p - m) * 1 \in K$, and $-0 = 0$. Finally, let $a \in K$ be nonzero. The map $K \to K$, $x \mapsto ax$, maps into K by closure, and it is injective by cancellation in F (Corollary 1.1.14). An injective map from a finite set to itself is surjective, so $ab = 1$ for some $b \in K$, and $a^{-1} = b \in K$. By Definition 1.1.32, K is a subfield of F.

(c) F is a vector space over itself (Example 1.2.6), hence over its subfield K (Proposition 1.2.19). Let $\mathcal{L}$ be a list whose entries are all the elements of F. Every $\alpha \in F$ equals $1 \cdot \alpha$, a linear combination of $\mathcal{L}$ over K, so $\mathcal{L}$ spans F over K, and F is finite-dimensional over K. By Corollary 1.6.24, F has a basis $\alpha_1, \ldots, \alpha_n$ over K, and $n \geq 1$ because $F \neq \{0\}$ (the empty list spans only $\{0\}$). By Proposition 1.6.16, the coordinate map is a bijection $F \to K^n$, so $|F| = |K^n| = p^n$.

Since 6 is not a power of a prime, no field has 6 elements. On the other hand, Exercise 1.1.56 constructs a field with $4 = 2^2$ elements; there, $K = \{0, 1\}$ and $1, \alpha$ is a basis over K.

Solution to Problem 1.17

Uses. Theorem 1.5.5 (Independence is decided at zero), Corollary 1.5.17 (Sublists), Proposition 1.5.15 (Short lists, zeros, and repeats), Definition 1.2.9 (Function spaces), Lemma 1.3.9 (Arithmetic of linear combinations)
Techniques. Induction on the length of a list (Technique 1.6.9), Uniqueness via two candidates (Technique 1.1.8)

Hint 1. The backward direction is immediate. For the forward direction, induct on n, applying the inductive hypothesis to $f_1, \dots, f_n$ inside $f_1, \dots, f_{n+1}$.
Hint 2. Given good points $t_1, \dots, t_n$ for $f_1, \dots, f_n$, look at the n equations in the $n+1$ unknowns $a_1, \dots, a_{n+1}$. Show that they have at most one solution with $a_{n+1} = 1$.
Hint 3. If every choice of t_{n+1} failed, each failure would produce a solution with $a_{n+1} = 1$, and by uniqueness they would all be the same solution $\mathbf{s}$. What relation among the f_i would $\mathbf{s}$ give?
Sketch. ($\Leftarrow$) A relation holds at every point, in particular at the t_j. ($\Rightarrow$) Induction on n. For $n = 1$, pick t_1 with $f_1(t_1) \neq 0$. For the step, take good points $t_1, \dots, t_n$ for $f_1, \dots, f_n$. A solution of the n equations in $n+1$ unknowns with $a_{n+1} = 0$ is 0, so two solutions with $a_{n+1} = 1$ agree (subtract them). If no t worked as t_{n+1}, each t would give a nontrivial solution, rescaled to $a_{n+1} = 1$, hence equal to one fixed $\mathbf{s}$; then $\sum s_i f_i(t) = 0$ for all t, a nontrivial relation.

($\Leftarrow$) Suppose such points exist, and let $a_1 f_1 + \cdots + a_n f_n = \mathbf{0}$. A relation among functions holds at every point (Definition 1.2.9), in particular at $t_1, \dots, t_n$, so all $a_i = 0$. The list is independent by Theorem 1.5.5.

($\Rightarrow$) We induct on n. For brevity, call points $t_1, \dots, t_n$ *good* for $f_1, \dots, f_n$ if the system above has only the zero solution.

Base case. If f_1 is independent, then $f_1 \neq \mathbf{0}$ (Proposition 1.5.15(b)), so $f_1(t_1) \neq 0$ for some $t_1 \in X$, and $a_1 f_1(t_1) = 0$ forces $a_1 = 0$.

Inductive step. Let $f_1, \dots, f_{n+1}$ be independent. Its sublist $f_1, \dots, f_n$ is independent (Corollary 1.5.17), so by the inductive hypothesis there are good points $t_1, \dots, t_n$ for it. Consider the system of n equations in $n+1$ unknowns

$$\sum_{i=1}^{n+1} a_i f_i(t_j) = 0 \quad (j = 1, \dots, n),$$

which we call (E). Two observations. First, a solution of (E) with $a_{n+1} = 0$ is zero, because then $a_1, \dots, a_n$ solves the system for $f_1, \dots, f_n$ at good points. Second, at most one solution has $a_{n+1} = 1$: if $\mathbf{a}$ and $\mathbf{a}'$ are two such solutions, their difference solves (E) (subtract the equations) and has last coordinate 0, so it is zero by the first observation.

Now suppose, for contradiction, that no point $t \in X$ is good as a last point, that is, for every $t \in X$ the points $t_1, \dots, t_n, t$ are not good for $f_1, \dots, f_{n+1}$. Then for every t there is a nonzero $\mathbf{a}(t) \in F^{n+1}$ solving (E) together with $\sum_i a_i f_i(t) = 0$. By the first observation, its last coordinate is nonzero, and dividing by it gives a solution with last coordinate 1 that still satisfies the equation at t. By the second observation, this rescaled solution is the same vector $\mathbf{s}$ for every t. (Here $X \neq \emptyset$, because $f_1 \neq \mathbf{0}$.) So $s_{n+1} = 1$ and

$$s_1 f_1(t) + \cdots + s_{n+1} f_{n+1}(t) = 0 \quad \text{for every } t \in X,$$

that is, $s_1 f_1 + \cdots + s_{n+1} f_{n+1} = \mathbf{0}$ is a nontrivial relation. This contradicts the independence of $f_1, \dots, f_{n+1}$. Hence some $t_{n+1} \in X$ makes $t_1, \dots, t_{n+1}$ good, completing the induction.

The uniqueness step is the subtraction pattern of Section 1.7.2: two normalized solutions differ by a solution of a problem that has only the zero solution.

Solution to Problem 1.18

Uses. Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations), Example 1.4.9 (Polynomials), Proposition 7.1.22 (Counting roots), Proposition 1.1.17 (No zero divisors), Definition 1.2.9 (Function spaces), Proposition 1.4.14 (The span is the smallest subspace containing a list), a surjection between finite sets of the same size is injective
Hint 1. The answer is: exactly the finite fields. For an infinite field, find a function that a polynomial cannot imitate, for instance one that vanishes at all but one point.
Hint 2. For a finite field, it suffices to produce, for each $c \in F$, the function δ_c that is 1 at c and 0 elsewhere. Try $t \mapsto \prod_{e \neq c}(t - e)$.
Hint 3. Products like this are polynomial functions: if g is the polynomial function of $\sum a_k x^k$, then $t \mapsto tg(t)$ is that of $\sum a_k x^{k+1}$. For uniqueness, count: how many polynomials of degree less than q, and how many functions?

Sketch. Infinite F: if δ_0 were the polynomial function of p, then p would have infinitely many roots, so $p = \mathbf{0}$, but $p(0) = 1$. Finite F with q elements: the functions of polynomials of degree at most d form a subspace P_d of F^F, and $g \mapsto (t \mapsto (t-e)g(t))$ maps P_d into P_{d+1}. So $h_c(t) = \prod_{e \neq c}(t-e)$ lies in P_{q-1}, $h_c(c) \neq 0$, and $\delta_c = h_c(c)^{-1}h_c$. Every $f = \sum_c f(c)\delta_c$. The map from $F[x]_{\leq q-1}$, with q^q elements, onto the q^q functions is then a bijection.

The answer: exactly the finite fields.

Infinite fields fail. Let F be infinite, and let $\delta_0 : F \to F$ be 1 at 0 and 0 elsewhere. Suppose δ_0 is the polynomial function of $p \in F[x]$. Then $p(c) = 0$ for every $c \neq 0$, so p has infinitely many roots. By Proposition 7.1.22, $p = \mathbf{0}$, and then $p(0) = 0 \neq 1 = \delta_0(0)$. So δ_0 is not a polynomial function.

Finite fields work. Let F have q elements, and for $d \geq 0$ let $P_d \subseteq F^F$ be the set of polynomial functions of polynomials in $F[x]_{\leq d}$.

Step 1: P_d is a subspace of F^F. It contains the zero function (of $\mathbf{0}$), and by Proposition 1.2.18(c), the sum of the functions of $p, p' \in F[x]_{\leq d}$ is the function of $p + p'$, and a times the function of p is the function of ap; both polynomials lie in $F[x]_{\leq d}$, which is a subspace (Example 1.4.9).

Step 2: multiplying by $t - d$. Let $g \in P_d$ be the function of $p = \sum_{k=0}^{d} a_k x^k$, and let $e \in F$. The polynomial $p^+ = \sum_{k=0}^{d} a_k x^{k+1}$ lies in $F[x]_{\leq d+1}$, and its value at t is $\sum_k a_k t^{k+1} = tg(t)$. So the function $t \mapsto (t-e)g(t) = tg(t) - eg(t)$ is the function of $p^+ - ep \in F[x]_{\leq d+1}$, by Proposition 1.2.18(c). That is, it lies in P_{d+1}.

Step 3: the functions δ_c. Fix $c \in F$, and let $h_c(t) = \prod_{e \neq c}(t-e)$, a product of $q-1$ factors. Starting from the constant function $1 \in P_0$ and applying Step 2 once for each factor, $h_c \in P_{q-1}$. At $t = c$ every factor is nonzero, so $h_c(c) \neq 0$ by Proposition 1.1.17; at $t \neq c$ the factor $t - t$ appears, so $h_c(t) = 0$. Hence $\delta_c = h_c(c)^{-1}h_c$, the function that is 1 at c and 0 elsewhere, lies in the subspace P_{q-1}.

Step 4: every function. For $f : F \to F$, the functions f and $\sum_{c \in F} f(c)\delta_c$ agree at every point t, since only the term with $c = t$ is nonzero there. So $f \in \operatorname{span}(\delta_c \mid c \in F) \subseteq P_{q-1}$ by Proposition 1.4.14(b). Every function is the polynomial function of some polynomial of degree at most $q-1$.

Step 5: uniqueness. The map $F[x]_{\leq q-1} \to F^F$ sending a polynomial to its function is surjective by Step 4. A polynomial of degree at most $q-1$ is determined by its q coefficients $a_0, \ldots, a_{q-1}$, so $F[x]_{\leq q-1}$ has q^q elements, and F^F has q^q elements as well. A surjection between finite sets of the same size is injective, so different polynomials of degree less than q have different functions.

For $F = \mathbb{F}_2$ this recovers Aside (Polynomials versus polynomial functions): $x^2 + x$ and $\mathbf{0}$ have the same function, and only one of them has degree less than 2.

Solution to Problem 1.19

Uses. Proposition 1.5.15 (Short lists, zeros, and repeats), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.30 (Additive arithmetic)
Hint 1. Translate the condition. Over $\mathbb{F}_2$ the only coefficients are 0 and 1, so a sublist $\mathbf{x}, \mathbf{y}, \mathbf{z}$ of distinct nonzero vectors is dependent exactly when $\mathbf{x} + \mathbf{y} + \mathbf{z} = \mathbf{0}$.
Hint 2. So the entries form a set S of nonzero vectors in which no sum of two different elements lies in S. Fix $\mathbf{s} \in S$ and compare S with $\mathbf{s} + S = \{\mathbf{s} + \mathbf{x} \mid \mathbf{x} \in S\}$.
Hint 3. For the construction, look at the vectors whose first coordinate is 1.
Sketch. The answer is 2^{n-1}. The condition says: the entries are nonzero, distinct, and $\mathbf{x} + \mathbf{y} \notin S$ for distinct $\mathbf{x}, \mathbf{y}$ in the set S of entries. For $\mathbf{s} \in S$, the sets S and $\mathbf{s} + S$ have k elements each and are disjoint, so $2k \leq 2^n$. The 2^{n-1} vectors with $x_1 = 1$ work, since a sum of two of them has $x_1 = 0$.

The answer is $k = 2^{n-1}$.

Translating the condition. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a list in $\mathbb{F}_2^n$. Its sublists of length 1 are independent exactly when every entry is nonzero (Proposition 1.5.15(b)). Given that, a sublist $\mathbf{v}_i, \mathbf{v}_j$ of length 2 is dependent exactly when $\mathbf{v}_i = \mathbf{v}_j$: over $\mathbb{F}_2$ the nontrivial relations are $\mathbf{v}_i = \mathbf{0}$, $\mathbf{v}_j = \mathbf{0}$, and $\mathbf{v}_i + \mathbf{v}_j = \mathbf{0}$, the last meaning $\mathbf{v}_i = \mathbf{v}_j$. Given that the entries are nonzero and distinct, a sublist $\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_l$ of length 3 is dependent exactly when $\mathbf{v}_i + \mathbf{v}_j + \mathbf{v}_l = \mathbf{0}$, since a nontrivial relation with one or two coefficients equal to 1 would make an entry

zero or two entries equal. So the condition holds exactly when the entries are nonzero and distinct, and the set S of entries, which has k elements, satisfies

$$\mathbf{x}+\mathbf{y}\notin S \quad \text{for all distinct } \mathbf{x},\mathbf{y}\in S.$$

(Indeed $\mathbf{x}+\mathbf{y}=\mathbf{z}$ is the same as $\mathbf{x}+\mathbf{y}+\mathbf{z}=\mathbf{0}$, and such a $\mathbf{z}\in S$ would differ from $\mathbf{x}$ and $\mathbf{y}$, as $\mathbf{x},\mathbf{y}\neq\mathbf{0}$.)

Upper bound. If $k=0$ there is nothing to prove. Otherwise fix $\mathbf{s}\in S$, and let $\mathbf{s}+S=\{\mathbf{s}+\mathbf{x}\mid\mathbf{x}\in S\}$. It has k elements, because $\mathbf{x}\mapsto\mathbf{s}+\mathbf{x}$ is injective (Proposition 1.2.30(b)). It is disjoint from S: suppose $\mathbf{s}+\mathbf{x}=\mathbf{y}$ with $\mathbf{x},\mathbf{y}\in S$. If $\mathbf{x}=\mathbf{s}$, then $\mathbf{y}=\mathbf{s}+\mathbf{s}=\mathbf{0}\notin S$. If $\mathbf{x}\neq\mathbf{s}$, then $\mathbf{s}+\mathbf{x}\notin S$ by the condition. Either way we have a contradiction. So S and $\mathbf{s}+S$ are disjoint subsets of $\mathbb{F}_2^n$ with k elements each, and $2k\le 2^n$, that is, $k\le 2^{n-1}$.

Construction. Let S be the set of 2^{n-1} vectors of $\mathbb{F}_2^n$ with first coordinate 1, listed in any order. They are nonzero and distinct. The sum of two of them has first coordinate $1+1=0$, so it is not in S. By the translation, every sublist of length at most 3 is independent.

So the largest k is 2^{n-1}. For $n=3$, the four vectors $(1,0,0),(1,1,0),(1,0,1),(1,1,1)$ have every three independent, while any five distinct nonzero vectors of $\mathbb{F}_2^3$ contain three with sum $\mathbf{0}$.

Section 2.1

Solution to Exercise 2.1.5

Uses. Theorem 2.1.3 (Checking linearity), Definition 1.2.1 (Vector space)
Hint 1. (a) is two of the vector space axioms of W.
Hint 2. (b) Every $x\in F$ is a scalar multiple of the single vector $1\in F$: $x=x\cdot 1$. Apply homogeneity.
Sketch. (a) $(x+y)\mathbf{w}=x\mathbf{w}+y\mathbf{w}$ and $(ax)\mathbf{w}=a(x\mathbf{w})$. (b) $T(x)=T(x\cdot 1)=xT(1)$.

(a) For $x,y,a\in F$, the vector space axioms of W give $(x+y)\mathbf{w}=x\mathbf{w}+y\mathbf{w}$ and $(ax)\mathbf{w}=a(x\mathbf{w})$. So the map is additive and homogeneous, hence linear by Theorem 2.1.3.

(b) Let $T\in\mathrm{L}(F,W)$ and $\mathbf{w}=T(1)$. For $x\in F$, the element $x=x\cdot 1$ is the scalar x times the vector $1\in F$, so homogeneity gives $T(x)=T(x\cdot 1)=xT(1)=x\mathbf{w}$. Taking $W=F$, every $T\in\mathrm{L}(F)$ is $x\mapsto ax$ with $a=T(1)$, and these maps are linear by (a).

Solution to Exercise 2.1.15

Uses. Example 2.1.7 (Maps $F^n\to F^m$ given by linear formulas), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Theorem 2.1.3 (Checking linearity)
Hint 1. Test $T(\mathbf{0})$ first, then $T(-\mathbf{v})$ or $T(2\mathbf{v})$. These catch most nonlinear formulas.
Hint 2. In (d) there are only two scalars, so homogeneity says very little. Test additivity on $\mathbf{e}_1$ and $\mathbf{e}_2$.
Sketch. (a) Linear: it is T_A with $A=\begin{pmatrix}0&1\\1&-3\end{pmatrix}$. (b) Not linear: $T(\mathbf{0})=(1,0)$. (c) Not linear: $T(-1)=(-1,1)\neq -T(1)$. (d) Homogeneous but not additive: $T(1,0)+T(0,1)=0\neq 1=T(1,1)$.

(a) Linear: $T=T_A$ for $A=\begin{pmatrix}0&1\\1&-3\end{pmatrix}$, so Example 2.1.7 applies.

(b) Not linear: $T(0,0)=(1,0)\neq\mathbf{0}$, contradicting Proposition 2.1.4.

(c) Not linear: $T(-1)=(-1,1)$, while $(-1)T(1)=(-1,-1)$, so T is not homogeneous. (Note that $T(0)=\mathbf{0}$, so the first test does not detect this one.)

(d) Not linear. Homogeneity does hold: the only scalars are 0 and 1, and $T(0\mathbf{v})=T(0,0)=0=0T(\mathbf{v})$ and $T(1\mathbf{v})=1T(\mathbf{v})$. But additivity fails: $T(1,0)+T(0,1)=0+0=0$, while $T(1,1)=1$.

Solution to Exercise 2.1.20

Uses. Proposition 2.1.17 (Compositions of linear maps are linear)
Hint. ST means "first T, then S", so it is a map $\mathbb{R}^2\to\mathbb{R}^2$. Substitute the coordinates of $T(u,v)$ into the formula for S.
Sketch. $ST(u,v)=(u+v,u+2v)$ and $TS(x,y,z)=(x-z,2y+2z,x+y)$.

$ST:\mathbb{R}^2\to\mathbb{R}^2$ is

$$(ST)(u,v)=S(u-v,2v,u)=((u-v)+2v,2v+u)=(u+v,u+2v),$$

and $TS : \mathbb{R}^3 \to \mathbb{R}^3$ is

$$(TS)(x, y, z) = T(x + y, y + z) = ((x + y) - (y + z), 2(y + z), x + y) = (x - z, 2y + 2z, x + y).$$

The two composites have different domains, so here the question whether $ST = TS$ does not even arise.

Solution to Exercise 2.1.25

Uses. Proposition 2.1.24 (Algebra of composition), Notation 2.1.18
Hint 1. Compute $T^2(x, y) = T(T(x, y))$ and add $T(x, y)$.
Hint 2. For the second part, expand $T(T + \mathrm{id})$ with Proposition 2.1.24.
Sketch. $T^2(x, y) = (x - y, 2y - x)$, and adding $(y, x - y)$ gives (x, y). Then $T(T + \mathrm{id}) = T^2 + T = \mathrm{id}$ and $(T + \mathrm{id})T = T^2 + T = \mathrm{id}$ by parts (b) and (c) of Proposition 2.1.24.

We have $T^2(x, y) = T(y, x - y) = (x - y, y - (x - y)) = (x - y, 2y - x)$, so

$$\big(T^2 + T\big)(x, y) = (x - y + y, 2y - x + x - y) = (x, y).$$

That is, $T^2 + T = \mathrm{id}$. By parts (b) and (c) of Proposition 2.1.24, $T(T + \mathrm{id}) = T^2 + T\,\mathrm{id} = T^2 + T = \mathrm{id}$ and $(T + \mathrm{id})T = T^2 + \mathrm{id}\,T = T^2 + T = \mathrm{id}$. So applying $T + \mathrm{id}$ undoes T, in either order; invertibility is the subject of Section 2.5.

Solution to Exercise 2.1.28

Uses. Definition 2.1.21 (Pointwise operations), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space), Theorem 2.1.3 (Checking linearity)
Hint. The vectors of $\mathrm{L}(V, W)$ are maps, and they are added and scaled pointwise. Evaluate $aS + T$ at $\mathbf{v}_0$.
Sketch. $(aS + T)(\mathbf{v}_0) = aS(\mathbf{v}_0) + T(\mathbf{v}_0)$ by the definition of the operations.

Write $E(T) = T(\mathbf{v}_0)$. For $S, T \in \mathrm{L}(V, W)$ and $a \in F$, the definition of the operations on $\mathrm{L}(V, W)$ gives

$$E(aS + T) = (aS + T)(\mathbf{v}_0) = aS(\mathbf{v}_0) + T(\mathbf{v}_0) = aE(S) + E(T).$$

So E is linear by Theorem 2.1.3. Note that the linearity of S and T played no role; the same computation shows that evaluation at a point is linear on W^X for any set X.

Solution to Exercise 2.1.29

Uses. Example 2.1.7 (Maps $F^n \to F^m$ given by linear formulas), Example 2.1.9 (Maps of polynomials), Non-example 2.1.14 (Complex conjugation)
Techniques. Assemble linear maps from linear pieces (Technique 2.1.26)
Hint 1. Entrywise operations make $F^{2\times 2}$ look like F^4. Which of the formulas are sums of constant multiples of the entries?
Hint 2. For (b), try two matrices with a single nonzero entry each, on the diagonal. For (e) over $\mathbb{C}$, compare $T(i \cdot 1)$ with $i \cdot T(1)$.
Sketch. (a), (c), and (d) are linear: (a) and (c) are linear formulas in the entries, and (d) is $\mathrm{id} + D$. (b) is not additive: the two diagonal matrices with a single 1 each map to 0, while their sum, the identity matrix, maps to 1. (e) is $\mathbb{R}$-linear, being $(a, b) \mapsto (b, a)$ in coordinates, but not $\mathbb{C}$-linear: $T(i) = 1$ while $iT(1) = -1$.

(a) Linear. The entries of $cA + B$ are $ca_{ij} + b_{ij}$, so the map sends $cA + B$ to $(ca_{11} + b_{11}) + (ca_{22} + b_{22}) = c(a_{11} + a_{22}) + (b_{11} + b_{22})$.

(b) Not linear, over any field. Let A have $a_{11} = 1$ and all other entries 0, and let B have $b_{22} = 1$ and all other entries 0. Both map to 0, but $A + B$ has diagonal entries 1, 1 and off-diagonal entries 0, so it maps to $1 \neq 0$. The map is not additive.

(c) Linear. Listing the entries as $(a_{11}, a_{12}, a_{21}, a_{22})$, the map is $(a, b, c, d) \mapsto (a, c, b, d)$, a linear formula as in Example 2.1.7 (each output coordinate is one input coordinate), and the operations of $F^{2\times 2}$ are those of F^4 in these coordinates.

(d) Linear: the map is $\mathrm{id}_{F[x]} + D$, a sum of linear maps (Technique 2.1.26).

(e) Write $T(z) = i\bar{z}$. For $z = a + bi$, $T(z) = i(a - bi) = b + ai$. Over $\mathbb{R}$, in the coordinates (a, b) this is $(a, b) \mapsto (b, a)$, which is linear (it is the reflection across the line $y = x$). Over $\mathbb{C}$ it is not homogeneous: $T(i \cdot 1) = T(i) = i \cdot (-i) = 1$, while $i \cdot T(1) = i \cdot i = -1$.

Solution to Exercise 2.1.30

Uses. Example 2.1.7 (Maps $F^n \to F^m$ given by linear formulas), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Hint 1. A function $\mathbb{F}_2^2 \to \mathbb{F}_2$ is a choice of one of two values at each of the four points of $\mathbb{F}_2^2$.
Hint 2. A linear T satisfies $T(x, y) = xT(\mathbf{e}_1) + yT(\mathbf{e}_2)$. How many choices are there for the pair $(T(\mathbf{e}_1), T(\mathbf{e}_2))$, and does each choice give a linear map?
Sketch. $2^4 = 16$ functions, of which $2^3 = 8$ send $\mathbf{0}$ to 0. A linear map is $(x, y) \mapsto \alpha x + \beta y$ with $(\alpha, \beta) = (T(\mathbf{e}_1), T(\mathbf{e}_2))$, and every such formula is linear, so there are 4: 0, x, y, and $x + y$.

$\mathbb{F}_2^2$ has 4 elements and $\mathbb{F}_2$ has 2, so there are $2^4 = 16$ functions $\mathbb{F}_2^2 \to \mathbb{F}_2$. Those with $T(\mathbf{0}) = 0$ are free only at the other three points, so there are $2^3 = 8$ of them.

If T is linear, then since $(x, y) = x\mathbf{e}_1 + y\mathbf{e}_2$,

$$T(x, y) = xT(\mathbf{e}_1) + yT(\mathbf{e}_2) = \alpha x + \beta y, \quad \text{where } \alpha = T(\mathbf{e}_1),\ \beta = T(\mathbf{e}_2).$$

So T is determined by $(\alpha, \beta) \in \mathbb{F}_2^2$, and there are at most 4 linear maps. Conversely, each formula $(x, y) \mapsto \alpha x + \beta y$ is linear by Example 2.1.7. So there are exactly 4 linear maps: $(x, y) \mapsto 0$, x, y, and $x + y$.

So half of the functions that fix $\mathbf{0}$ are not linear. They are $(x, y) \mapsto xy$, $x + xy$, $y + xy$, and $x + y + xy$.

Solution to Exercise 2.1.31

Uses. Proposition 2.1.16 (Linear maps preserve spans and linear relations)
Hint 1. Proposition 2.1.16 says that $T(\mathcal{L})$ spans $\operatorname{im}\left(T\restriction_{\operatorname{span}\mathcal{L}}\right)$. When is that all of W?
Hint 2. For counterexamples, try the zero map and a coordinate projection $\mathbb{R}^2 \to \mathbb{R}$.
Sketch. (a) False: the zero map $\mathbb{R} \to \mathbb{R}$ and $\mathcal{L} = 1$. (b) False: $(x, y) \mapsto x$ from $\mathbb{R}^2$ to $\mathbb{R}$ and $\mathcal{L} = \mathbf{e}_1$. (c) True: $W = \operatorname{span} T(\mathcal{L}) = \operatorname{im}\left(T\restriction_{\operatorname{span}\mathcal{L}}\right) \subseteq \operatorname{im} T \subseteq W$.

(a) False. Let $T : \mathbb{R} \to \mathbb{R}$ be the zero map and $\mathcal{L}$ the list 1, which spans $\mathbb{R}$. Then $T(\mathcal{L})$ is the list 0, which spans $\{0\} \neq \mathbb{R}$.

(b) False. Let $T : \mathbb{R}^2 \to \mathbb{R}$, $T(x, y) = x$, and $\mathcal{L} = \mathbf{e}_1$. Then $T(\mathcal{L})$ is the list 1, which spans $\mathbb{R}$, but $\mathcal{L}$ does not span $\mathbb{R}^2$: every element of $\operatorname{span}(\mathbf{e}_1)$ has second coordinate 0.

(c) True. By Proposition 2.1.16, $W = \operatorname{span} T(\mathcal{L}) = \operatorname{im}\left(T\restriction_{\operatorname{span}\mathcal{L}}\right) \subseteq \operatorname{im} T \subseteq W$, so $\operatorname{im} T = W$.

Solution to Exercise 2.1.32

Uses. Theorem 2.1.3 (Checking linearity), Definition 1.2.13 (Polynomials)
Hint 1. Write $cp + r$ in terms of the coefficients of p and r, using the same n for both.
Hint 2. Substituting q into $\sum_k (ca_k + b_k)x^k$ gives $\sum_k (ca_k + b_k)q^k$. Split it using the vector space axioms of $F[x]$.
Sketch. With $p = \sum a_k x^k$ and $r = \sum b_k x^k$, $cp + r = \sum (ca_k + b_k)x^k$, so $C_q(cp + r) = \sum (ca_k + b_k)q^k = c\sum a_k q^k + \sum b_k q^k$.

Let $p = \sum_{k=0}^n a_k x^k$ and $r = \sum_{k=0}^n b_k x^k$, padding with zero coefficients so that both use the same n. (Padding does not change $p(q)$, since the added terms are $0 \cdot q^k = 0$.) For $c \in F$, $cp + r = \sum_k (ca_k + b_k)x^k$, so

$$C_q(cp + r) = \sum_{k=0}^{n} (ca_k + b_k)q^k = c\sum_{k=0}^{n} a_k q^k + \sum_{k=0}^{n} b_k q^k = cC_q(p) + C_q(r),$$

by distributivity and the other vector space axioms of $F[x]$. By Theorem 2.1.3, C_q is linear.

Solution to Exercise 2.1.33

Uses. Theorem 2.1.3 (Checking linearity), Example 2.1.6 (Zero, identity, and scaling maps)
Hint 1. Apply L_q to the polynomial x, and to the polynomial $2x$. What does homogeneity say?
Hint 2. $L_q(x) = q$ and $L_q(2x) = q(2x)$. If $q = \sum b_k x^k$, then $q(2x) = \sum 2^k b_k x^k$. Compare coefficients with $2q$.
Sketch. L_q is linear exactly when $q = cx$ for some $c \in \mathbb{R}$. If L_q is linear, $q(2x) = L_q(2x) = 2L_q(x) = 2q$, so $\left(2^k - 2\right)b_k = 0$ for every k, and $b_k = 0$ unless $k = 1$. Conversely $L_{cx}(p) = cp$ is a scaling map.

The answer is $q = cx$ with $c \in \mathbb{R}$.

If $q = cx$, then $L_q(p) = cp$, which is linear (Example 2.1.6). Conversely, suppose L_q is linear, and write $q = b_0 + b_1x + \cdots + b_nx^n$. Then $L_q(x) = q$ and

$$L_q(2x) = q(2x) = \sum_{k=0}^{n} b_k(2x)^k = \sum_{k=0}^{n} 2^k b_k x^k.$$

By homogeneity, $L_q(2x) = 2L_q(x) = \sum_k 2b_kx^k$. Comparing coefficients, $\big(2^k - 2\big)b_k = 0$ for every k. For $k \neq 1$ we have $2^k - 2 \neq 0$ (it is -1 for $k = 0$ and at least 2 for $k \geq 2$), so $b_k = 0$. Hence $q = b_1x$.

The test used the scalar 2, which is 0 in $\mathbb{F}_2$, so over $\mathbb{F}_2$ the condition $L_q(2x) = 2L_q(x)$ says nothing. Indeed over $\mathbb{F}_2$ the answer is different: for $q = x^2$, the map $p \mapsto p^2$ turns out to be linear (try it).

Solution to Exercise 2.1.34

Uses. Example 2.1.9 (Maps of polynomials), Proposition 1.1.17 (No zero divisors), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
Hint 1. $D\big(\sum a_kx^k\big) = \sum ka_kx^{k-1}$, so $D(p) = 0$ if and only if $ka_k = 0$ for every $k \geq 1$.
Hint 2. In $\mathbb{F}_q$ for a prime q, $ka_k = (k \cdot 1)a_k$, and $k \cdot 1 = 0$ exactly when q divides k.
Sketch. $D(p) = 0$ if and only if $ka_k = 0$ for all $k \geq 1$. Over $\mathbb{F}_2$ this means $a_k = 0$ for odd k, so p is a polynomial in x^2. Over $\mathbb{F}_3$, p is a polynomial in x^3. Over $\mathbb{R}$, p is constant.

Let $p = a_0 + a_1x + \cdots + a_nx^n$. Then $D(p) = \sum_{k=1}^{n} ka_kx^{k-1}$, which is 0 if and only if $ka_k = 0$ for every $1 \leq k \leq n$. Here $ka_k = (k \cdot 1)a_k$, where $k \cdot 1$ is the sum of k ones in F. By Proposition 1.1.17, $ka_k = 0$ if and only if $k \cdot 1 = 0$ or $a_k = 0$.

In $\mathbb{F}_q$ for a prime q, the sum of k ones is $k \bmod q$, which is 0 exactly when q divides k. So over $\mathbb{F}_2$, $D(p) = 0$ if and only if $a_k = 0$ for every odd k, that is, $p = b_0 + b_1x^2 + b_2x^4 + \cdots$ for some $b_i \in \mathbb{F}_2$. For example $D\big(x^2 + 1\big) = 2x = 0$. Over $\mathbb{F}_3$, $D(p) = 0$ if and only if $a_k = 0$ whenever 3 does not divide k, that is, $p = b_0 + b_1x^3 + b_2x^6 + \cdots$. Over $\mathbb{R}$, $k \cdot 1 = k \neq 0$ for $k \geq 1$, so $D(p) = 0$ if and only if $p = a_0$ is constant.

Solution to Exercise 2.1.35

Uses. Theorem 2.1.3 (Checking linearity)
Hint. How many terms does the sum $x + x + \cdots + x$ have? For which c does that make sense?
Sketch. Writing cx as a sum of c copies of x only makes sense for positive integers c. The argument proves $T(nx) = nT(x)$ for $n = 1, 2, 3, \ldots$ and nothing about $c = 1/2$ or $c = \sqrt{2}$.

The first step writes cx as a sum of c copies of x, which only makes sense when c is a positive integer. For such $c = n$ the argument is correct: by induction on n, using additivity, $T(nx) = T((n-1)x + x) = T((n-1)x) + T(x) = nT(x)$. But it says nothing about $c = 1/2$ or $c = \sqrt{2}$, and homogeneity requires $T(cx) = cT(x)$ for every real c.

The gap is not a technicality. The conclusion $T(cx) = cT(x)$ can be rescued for rational c with a further idea (Exercise 2.1.37), but over $\mathbb{R}$ the claim is false: there are additive maps $\mathbb{R} \to \mathbb{R}$ that are not linear (Aside (Additive maps that are not linear)).

Solution to Exercise 2.1.36

Uses. Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Non-example 2.1.14 (Complex conjugation), Definition 1.1.19 (Integers modulo n)
Hint 1. In $\mathbb{F}_p$, every element is $1 + 1 + \cdots + 1$ for some number of terms.
Hint 2. If $c = 1 + \cdots + 1$ (n terms), then $c\mathbf{v} = \mathbf{v} + \cdots + \mathbf{v}$. Apply additivity $n - 1$ times. Do not forget $c = 0$.
Hint 3. For (b), look at the non-examples in this section.
Sketch. (a) True. Each $c \in \mathbb{F}_p$ is $n \cdot 1$ for some $0 \leq n < p$, so $c\mathbf{v}$ is a sum of n copies of $\mathbf{v}$ and additivity gives $T(c\mathbf{v}) = cT(\mathbf{v})$; the case $n = 0$ is $T(\mathbf{0}) = \mathbf{0}$, which holds for additive maps. (b) False: complex conjugation $\mathbb{C} \to \mathbb{C}$ is additive but not $\mathbb{C}$-linear.

(a) True. Let $c \in \mathbb{F}_p = \{0, 1, \ldots, p-1\}$, and let n be c regarded as an integer. In $\mathbb{F}_p$ the sum of n ones is $n \bmod p = c$. If $n = 0$, then $T(0\mathbf{v}) = T(\mathbf{0}) = \mathbf{0} = 0T(\mathbf{v})$, where $T(\mathbf{0}) = \mathbf{0}$ holds for additive maps by the second proof of Proposition 2.1.4. If $n \geq 1$, distributivity and $1\mathbf{v} = \mathbf{v}$ give $c\mathbf{v} = (1 + \cdots + 1)\mathbf{v} = \mathbf{v} + \cdots + \mathbf{v}$ (n terms), and additivity, applied $n - 1$ times, gives

$$T(c\mathbf{v}) = T(\mathbf{v} + \cdots + \mathbf{v}) = T(\mathbf{v}) + \cdots + T(\mathbf{v}) = (1 + \cdots + 1)T(\mathbf{v}) = cT(\mathbf{v}).$$

So T is homogeneous, hence linear.

(b) False. Complex conjugation $\mathbb{C} \to \mathbb{C}$ is additive but not $\mathbb{C}$-linear (Non-example 2.1.14).

Solution to Exercise 2.1.37

Uses. Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Exercise 2.1.35
Hint 1. First show $T(n\mathbf{v}) = nT(\mathbf{v})$ for every integer n, positive, zero, or negative.
Hint 2. For $c = m/n$ with $n \geq 1$, let $\mathbf{w} = c\mathbf{v}$. Then $n\mathbf{w} = m\mathbf{v}$. Apply T and use the integer case on both sides.
Sketch. For integers $n \geq 0$, $T(n\mathbf{v}) = nT(\mathbf{v})$ by induction, and $T(-\mathbf{u}) = -T(\mathbf{u})$ extends this to $n < 0$. If $c = m/n$, then $n(c\mathbf{v}) = m\mathbf{v}$, so $nT(c\mathbf{v}) = mT(\mathbf{v})$; divide by n.

Integers. We first show $T(n\mathbf{v}) = nT(\mathbf{v})$ for every $\mathbf{v} \in V$ and every integer n. For $n = 0$ this is $T(\mathbf{0}) = \mathbf{0}$, which holds for additive maps by the second proof of Proposition 2.1.4. If it holds for some $n \geq 0$, then $T((n+1)\mathbf{v}) = T(n\mathbf{v} + \mathbf{v}) = nT(\mathbf{v}) + T(\mathbf{v}) = (n+1)T(\mathbf{v})$. So it holds for all $n \geq 0$. For $n < 0$, the second proof of Proposition 2.1.4 also shows $T(-\mathbf{u}) = -T(\mathbf{u})$ for additive T, so $T(n\mathbf{v}) = T(-((-n)\mathbf{v})) = -((-n)T(\mathbf{v})) = nT(\mathbf{v})$.

Rationals. Let $c = m/n$ with $m, n \in \mathbb{Z}$ and $n \geq 1$, and put $\mathbf{w} = c\mathbf{v}$. Then $n\mathbf{w} = (nc)\mathbf{v} = m\mathbf{v}$, so by the integer case

$$nT(\mathbf{w}) = T(n\mathbf{w}) = T(m\mathbf{v}) = mT(\mathbf{v}).$$

Multiplying by $1/n \in \mathbb{Q}$ gives $T(c\mathbf{v}) = T(\mathbf{w}) = (m/n)T(\mathbf{v}) = cT(\mathbf{v})$. So T is homogeneous, hence linear.

Solution to Exercise 2.1.38

Uses. Proposition 1.1.17 (No zero divisors)
Hint 1. Homogeneity only relates $T(\mathbf{v})$ to the values of T on the line $\operatorname{span}(\mathbf{v})$. It says nothing about how different lines through $\mathbf{0}$ interact.
Hint 2. Let T be a linear function on the x-axis and 0 off the x-axis.
Sketch. $T(x, y) = x$ if $y = 0$, and $T(x, y) = 0$ if $y \neq 0$. Scaling by $c \neq 0$ keeps a point on or off the x-axis, which gives homogeneity. But $T(1,1) + T(0,-1) = 0 \neq 1 = T(1,0)$.

Define $T(x, y) = x$ if $y = 0$, and $T(x, y) = 0$ if $y \neq 0$.

Homogeneity. Let $c \in F$ and $(x, y) \in F^2$. If $c = 0$, then $T(c(x, y)) = T(0, 0) = 0 = cT(x, y)$. Let $c \neq 0$. If $y = 0$, then $T(cx, 0) = cx = cT(x, 0)$. If $y \neq 0$, then $cy \neq 0$ by Proposition 1.1.17, so $T(cx, cy) = 0 = c \cdot 0 = cT(x, y)$.

Additivity fails. $T(1,1) + T(0,-1) = 0 + 0 = 0$, since $1 \neq 0$ and $-1 \neq 0$, while $T((1,1) + (0,-1)) = T(1,0) = 1$.

Homogeneity is a condition on each line through $\mathbf{0}$ separately, and additivity is what ties different lines together.

Solution to Exercise 2.1.39

Uses. Theorem 2.1.3 (Checking linearity), Exercise 2.1.15
Hint 1. For homogeneity, use the pair of coefficients a and $1 - a$, with $\mathbf{0}$ as the second vector.
Hint 2. For additivity, write $\mathbf{u} + \mathbf{v} = (1+1)(h\mathbf{u} + h\mathbf{v})$ with $h = (1+1)^{-1}$, and note that $h + h = 1$.
Hint 3. For (b), what does the hypothesis say when the only scalars are 0 and 1?
Sketch. (a) $T(a\mathbf{u}) = T(a\mathbf{u} + (1-a)\mathbf{0}) = aT(\mathbf{u})$. With $h = (1+1)^{-1}$, homogeneity and the hypothesis with $a = b = h$ give $T(\mathbf{u} + \mathbf{v}) = (1+1)T(h\mathbf{u} + h\mathbf{v}) = (1+1)(hT(\mathbf{u}) + hT(\mathbf{v})) = T(\mathbf{u}) + T(\mathbf{v})$. (b) Over $\mathbb{F}_2$, $a + b = 1$ forces $\{a, b\} = \{0, 1\}$, and the hypothesis is empty. So $(x, y) \mapsto xy$ on $\mathbb{F}_2^2$ satisfies it but is not linear.

(a) *Homogeneity.* For $a \in F$ and $\mathbf{u} \in V$, the coefficients a and $1 - a$ add up to 1, so

$$T(a\mathbf{u}) = T(a\mathbf{u} + (1-a)\mathbf{0}) = aT(\mathbf{u}) + (1-a)T(\mathbf{0}) = aT(\mathbf{u}).$$

Additivity. Since $1 + 1 \neq 0$, it has an inverse $h = (1+1)^{-1}$, and $h + h = (1+1)h = 1$. For $\mathbf{u}, \mathbf{v} \in V$ we have $(1+1)(h\mathbf{u} + h\mathbf{v}) = \mathbf{u} + \mathbf{v}$, so by homogeneity and then the hypothesis with $a = b = h$,

$$T(\mathbf{u} + \mathbf{v}) = (1+1)T(h\mathbf{u} + h\mathbf{v}) = (1+1)(hT(\mathbf{u}) + hT(\mathbf{v})) = T(\mathbf{u}) + T(\mathbf{v}).$$

By Theorem 2.1.3, T is linear.

(b) In $\mathbb{F}_2$, $a+b=1$ holds only for $(a,b)=(1,0)$ and $(0,1)$, and then the hypothesis reads $T(\mathbf{u})=T(\mathbf{u})$ or $T(\mathbf{v})=T(\mathbf{v})$, which is always true. So every T with $T(\mathbf{0})=\mathbf{0}$ satisfies the hypotheses. For example $T:\mathbb{F}_2^2\to\mathbb{F}_2$, $T(x,y)=xy$, does, and it is not linear (Exercise 2.1.15).

Solution to Exercise 2.1.40

Uses. Example 2.1.7 (Maps $F^n\to F^m$ given by linear formulas)
Hint 1. $ST=0$ means that S sends every vector of the form $T(\mathbf{v})$ to $\mathbf{0}$. Choose T with all its values on a line, and S that kills that line.
Hint 2. Try $S(x,y)=(x,0)$, the projection onto the x-axis, and a T whose values lie on the y-axis.
Sketch. $S(x,y)=(x,0)$ and $T(x,y)=(0,x)$: then $ST(x,y)=S(0,x)=\mathbf{0}$, while $TS(x,y)=(0,x)$. The same T answers (b): $T^2(x,y)=T(0,x)=\mathbf{0}$.

Let $S(x,y)=(x,0)$ and $T(x,y)=(0,x)$; both are linear by Example 2.1.7.

(a) $(ST)(x,y)=S(0,x)=(0,0)$ for all (x,y), so $ST=0$. But $(TS)(x,y)=T(x,0)=(0,x)$, so $(TS)(\mathbf{e}_1)=\mathbf{e}_2\neq\mathbf{0}$ and $TS\neq 0$.

(b) The same T works: $T(\mathbf{e}_1)=\mathbf{e}_2\neq\mathbf{0}$, and $T^2(x,y)=T(0,x)=(0,0)$. Geometrically, T moves everything onto the y-axis and sends the y-axis to $\mathbf{0}$.

Solution to Exercise 2.1.41

Uses. Example 2.1.11 (Rotations, projections, and reflections), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Hint 1. Test the claim on a projection from Example 2.1.11.
Hint 2. Which step would be valid for numbers but not for linear maps? Can a linear map send a nonzero vector to $\mathbf{0}$ without being 0?
Sketch. The step "$P(\mathbf{w})=\mathbf{0}$ implies $P=0$ or $\mathbf{w}=\mathbf{0}$" is false. The projection $P(x,y)=(x,0)$ satisfies $P^2=P$ and is neither 0 nor id; there $P(P(\mathbf{v})-\mathbf{v})=P(0,-y)=\mathbf{0}$ with $(0,-y)\neq\mathbf{0}$ for $y\neq 0$.

The claim is false: $P(x,y)=(x,0)$ (the projection Pr_ℓ of Example 2.1.11 onto the x-axis) satisfies $P^2(x,y)=P(x,0)=(x,0)=P(x,y)$, and P is neither 0 nor $\mathrm{id}_{\mathbb{R}^2}$.

The error is the step "since P is linear, either $P=0$ or $P(\mathbf{v})-\mathbf{v}=\mathbf{0}$". It treats $P(\mathbf{w})=\mathbf{0}$ like a product of numbers equal to 0, but a linear map can send a nonzero vector to $\mathbf{0}$ without being the zero map. In the example, $P(\mathbf{v})-\mathbf{v}=(0,-y)$, which is nonzero when $y\neq 0$, and $P(0,-y)=\mathbf{0}$. The vectors that a linear map sends to $\mathbf{0}$ are studied in Section 2.6.

Solution to Exercise 2.1.42

Uses. Example 2.1.11 (Rotations, projections, and reflections), Example 2.1.27 (Maps assembled from linear pieces), Proposition 2.1.24 (Algebra of composition)
Techniques. Assemble linear maps from linear pieces (Technique 2.1.26)
Hint 1. Write $P(x,y)=s(\cos\theta,\sin\theta)$ with $s=x\cos\theta+y\sin\theta$, and apply P to the point $s(\cos\theta,\sin\theta)$.
Hint 2. Expand $(2P-\mathrm{id})(2P-\mathrm{id})$ as you would expand $(2p-1)^2$, but justify each step by a part of Proposition 2.1.24.
Sketch. (a) P applied to $s(\cos\theta,\sin\theta)$ gives $s(\cos^2\theta+\sin^2\theta)(\cos\theta,\sin\theta)=P(x,y)$. (b) $S^2=4P^2-2P-2P+\mathrm{id}=\mathrm{id}$ and $SP=2P^2-P=P=PS$. Reflecting twice returns every point; a point and its mirror image have the same foot on ℓ; points of ℓ are fixed by S.

(a) Let $\mathbf{u}=(\cos\theta,\sin\theta)$ and $s=x\cos\theta+y\sin\theta$, so $P(x,y)=s\mathbf{u}$. By linearity of P, $P(s\mathbf{u})=sP(\mathbf{u})$, and by the formula,

$$P(\mathbf{u})=(\cos\theta\cdot\cos\theta+\sin\theta\cdot\sin\theta)\mathbf{u}=\mathbf{u}.$$

So $P(P(x,y))=s\mathbf{u}=P(x,y)$, that is, $P^2=P$.

(b) Using parts (c), (d), and (b) of Proposition 2.1.24,

$$\begin{aligned}S^2=(2P-\mathrm{id})(2P-\mathrm{id})&=(2P)(2P)-(2P)\,\mathrm{id}-\mathrm{id}\,(2P)+\mathrm{id}\,\mathrm{id}\\&=4P^2-2P-2P+\mathrm{id}=4P-4P+\mathrm{id}=\mathrm{id},\end{aligned}$$

where we used $P^2=P$. Similarly,

$$SP = (2P - \mathrm{id})P = 2P^2 - P = P, \quad PS = P(2P - \mathrm{id}) = 2P^2 - P = P.$$

Geometrically, $S^2 = \mathrm{id}$ says that reflecting twice returns every point to where it started. $PS = P$ says that a point and its mirror image have the same foot of perpendicular on ℓ. $SP = P$ says that S fixes every point of ℓ, since the values of P are exactly the points of ℓ.

Solution to Exercise 2.1.43

Uses. Example 2.1.11 (Rotations, projections, and reflections), Proposition 2.1.17 (Compositions of linear maps are linear)
Hint 1. Substitute the formula for $R_\beta(x, y)$ into the formula for R_α, and collect the coefficients of x and y.
Hint 2. The coefficients come out as $\cos\alpha\cos\beta - \sin\alpha\sin\beta$ and similar expressions. Recognize them with the angle addition formulas. For the reflections, the angles are 2α and 2β.
Sketch. The first coordinate of $R_\alpha R_\beta(x, y)$ is $x(\cos\alpha\cos\beta - \sin\alpha\sin\beta) - y(\sin\alpha\cos\beta + \cos\alpha\sin\beta) = x\cos(\alpha+\beta) - y\sin(\alpha+\beta)$, and the second is similar. For reflections, the coefficients become $\cos 2\alpha\cos 2\beta + \sin 2\alpha\sin 2\beta = \cos(2\alpha - 2\beta)$ and $\sin 2\alpha\cos 2\beta - \cos 2\alpha\sin 2\beta = \sin(2\alpha - 2\beta)$.

Rotations. Write $(u, v) = R_\beta(x, y) = (x\cos\beta - y\sin\beta, x\sin\beta + y\cos\beta)$. Then $R_\alpha R_\beta(x, y) = (u\cos\alpha - v\sin\alpha, u\sin\alpha + v\cos\alpha)$. The first coordinate is

$$x(\cos\alpha\cos\beta - \sin\alpha\sin\beta) - y(\cos\alpha\sin\beta + \sin\alpha\cos\beta) = x\cos(\alpha+\beta) - y\sin(\alpha+\beta),$$

and the second is

$$x(\sin\alpha\cos\beta + \cos\alpha\sin\beta) + y(\cos\alpha\cos\beta - \sin\alpha\sin\beta) = x\sin(\alpha+\beta) + y\cos(\alpha+\beta).$$

So $R_\alpha R_\beta = R_{\alpha+\beta}$.

Reflections. Write $(u, v) = S_\beta(x, y) = (x\cos 2\beta + y\sin 2\beta, x\sin 2\beta - y\cos 2\beta)$. Then $S_\alpha S_\beta(x, y) = (u\cos 2\alpha + v\sin 2\alpha, u\sin 2\alpha - v\cos 2\alpha)$. The first coordinate is

$$x(\cos 2\alpha\cos 2\beta + \sin 2\alpha\sin 2\beta) + y(\cos 2\alpha\sin 2\beta - \sin 2\alpha\cos 2\beta) = x\cos(2\alpha - 2\beta) - y\sin(2\alpha - 2\beta),$$

and the second is

$$x(\sin 2\alpha\cos 2\beta - \cos 2\alpha\sin 2\beta) + y(\sin 2\alpha\sin 2\beta + \cos 2\alpha\cos 2\beta) = x\sin(2\alpha - 2\beta) + y\cos(2\alpha - 2\beta).$$

So $S_\alpha S_\beta = R_{2(\alpha-\beta)}$.

Geometrically: reflecting across one line and then across another is a rotation about their intersection point by twice the angle from the first line to the second. In particular, reflecting twice across the same line ($\alpha = \beta$) gives $R_0 = \mathrm{id}$. Also, two reflections never compose to a reflection: R_φ sends $\mathbf{e}_1, \mathbf{e}_2$ to $(\cos\varphi, \sin\varphi), (-\sin\varphi, \cos\varphi)$, and S_γ sends them to $(\cos 2\gamma, \sin 2\gamma), (\sin 2\gamma, -\cos 2\gamma)$, so $R_\varphi = S_\gamma$ would force $\cos\varphi = \cos 2\gamma = -\cos 2\gamma$ and $\sin\varphi = \sin 2\gamma = -\sin 2\gamma$, hence $\cos\varphi = \sin\varphi = 0$, which is impossible.

Solution to Exercise 2.1.44

Uses. Definition 1.4.1 (Subspace), Theorem 1.4.3 (Subspace criterion), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space)
Hint 1. A subspace contains the zero vector. What is the zero vector of $\mathrm{L}(V, W)$, and where does it send $\mathbf{v}_0$?
Hint 2. If $\mathbf{w}_0 = \mathbf{0}$, check closure directly, or note that the set is where the linear map $T \mapsto T(\mathbf{v}_0)$ of Exercise 2.1.28 takes the value $\mathbf{0}$.
Sketch. The zero map sends $\mathbf{v}_0$ to $\mathbf{0}$, so it lies in the set only if $\mathbf{w}_0 = \mathbf{0}$. If $\mathbf{w}_0 = \mathbf{0}$, then $(cS + T)(\mathbf{v}_0) = cS(\mathbf{v}_0) + T(\mathbf{v}_0) = \mathbf{0}$ for S, T in the set.

Call the set U. The zero vector of $\mathrm{L}(V, W)$ is the zero map (Proposition 2.1.23), which sends $\mathbf{v}_0$ to $\mathbf{0}$.

If $\mathbf{w}_0 \neq \mathbf{0}$, the zero map is not in U, so U is not a subspace.

If $\mathbf{w}_0 = \mathbf{0}$, the zero map is in U. For $S, T \in U$ and $c \in F$,

$$(cS + T)(\mathbf{v}_0) = cS(\mathbf{v}_0) + T(\mathbf{v}_0) = c\mathbf{0} + \mathbf{0} = \mathbf{0},$$

Taking $c = 1$ shows that U is closed under addition, and taking $T = 0$ shows that it is closed under scalar multiplication. So U is a subspace by Theorem 1.4.3.

Solution to Exercise 2.1.45

Uses. Proposition 2.1.24 (Algebra of composition), Theorem 1.4.3 (Subspace criterion), Notation 2.1.18
Hint 1. Every step is an application of Proposition 2.1.24. For (a), expand $A(cS+T)$ and $(cS+T)A$.
Hint 2. For (b), move A across ST one factor at a time, using associativity.
Hint 3. For (c), id_V and A commute with A. Use (b) and induction for the powers, then (a).
Sketch. (a) $A0 = 0 = 0A$, and $A(cS+T) = cAS + AT = cSA + TA = (cS+T)A$. (b) $AST = SAT = STA$. (c) $\mathrm{id}_V, A \in C(A)$, so every $A^k \in C(A)$ by (b), and linear combinations stay in $C(A)$ by (a).

(a) The zero map commutes with A: $(A0)(\mathbf{v}) = A(\mathbf{0}) = \mathbf{0} = (0A)(\mathbf{v})$ by Proposition 2.1.4. If $S, T \in C(A)$ and $c \in F$, then by parts (c) and (d) of Proposition 2.1.24,

$$A(cS+T) = c(AS) + AT = c(SA) + TA = (cS+T)A.$$

With $c = 1$ this shows that $C(A)$ is closed under addition, and with $T = 0$ that it is closed under scalar multiplication. So $C(A)$ is a subspace by Theorem 1.4.3.

(b) By associativity, $A(ST) = (AS)T = (SA)T = S(AT) = S(TA) = (ST)A$.

(c) $\mathrm{id}_V A = A = A\mathrm{id}_V$, and $AA = AA$, so $\mathrm{id}_V, A \in C(A)$. If $A^k \in C(A)$, then $A^{k+1} = A^k A \in C(A)$ by (b), so by induction every $A^k \in C(A)$. Since $C(A)$ is a subspace, it contains $a_0\mathrm{id}_V + a_1 A + \cdots + a_n A^n$.

Solution to Exercise 2.1.46

Uses. Example 2.1.9 (Maps of polynomials), Proposition 2.1.24 (Algebra of composition), Notation 2.1.18
Hint 1. (a) Apply both sides to $p = \sum_j a_j x^j$ and compare coefficients.
Hint 2. (b) Induct on k. First check, using only Proposition 2.1.24, that $DM^{k+1} - M^{k+1}D = \big(DM^k - M^kD\big)M + M^k(DM - MD)$.
Sketch. (a) $DM(p) = \sum (j+1)a_j x^j$ and $MD(p) = \sum j a_j x^j$, so the difference is p. (b) By the identity in the hint, the induction hypothesis, and (a), $DM^{k+1} - M^{k+1}D = kM^{k-1}M + M^k = (k+1)M^k$. Applied to p, (b) is the product rule $D\big(x^k p\big) = kx^{k-1}p + x^k D(p)$.

(a) Let $p = \sum_{j=0}^n a_j x^j$. Then $M(p) = \sum_{j=0}^n a_j x^{j+1}$, so $(DM)(p) = \sum_{j=0}^n (j+1)a_j x^j$, while $(MD)(p) = x\sum_{j=1}^n j a_j x^{j-1} = \sum_{j=0}^n j a_j x^j$ (the term $j=0$ is 0). Since $(j+1)a_j - j a_j = a_j$,

$$(DM - MD)(p) = \sum_{j=0}^n a_j x^j = p.$$

(b) We induct on k. For $k = 1$ this is (a), since $M^0 = \mathrm{id}$. Suppose $DM^k - M^kD = kM^{k-1}$ for some $k \geq 1$. By Proposition 2.1.24,

$$\big(DM^k - M^kD\big)M + M^k(DM - MD) = DM^{k+1} - M^kDM + M^kDM - M^{k+1}D = DM^{k+1} - M^{k+1}D.$$

By the induction hypothesis and (a), the left side is $kM^{k-1}M + M^k\,\mathrm{id} = kM^k + M^k = (k+1)M^k$, as desired.

Applied to p, (b) says $D\big(x^k p\big) - x^k D(p) = kx^{k-1}p$, that is, $D\big(x^k p\big) = kx^{k-1}p + x^k D(p)$: the product rule for the product of x^k and p. Exercise 2.1.49 proves the product rule in general.

Solution to Exercise 2.1.47

Uses. Non-example 2.1.14 (Complex conjugation), Exercise 2.1.5, Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space)
Techniques. Assemble linear maps from linear pieces (Technique 2.1.26)
Hint 1. An $\mathbb{R}$-linear map $T : \mathbb{C} \to \mathbb{C}$ satisfies $T(a+bi) = aT(1) + bT(i)$ for $a, b \in \mathbb{R}$. So it is determined by $T(1)$ and $T(i)$.
Hint 2. The map $z \mapsto \alpha z + \beta\overline{z}$ is $\mathbb{R}$-linear. Make its values at 1 and i equal $T(1)$ and $T(i)$: solve $\alpha + \beta = T(1)$ and $(\alpha - \beta)i = T(i)$.
Sketch. Both T and $z \mapsto \alpha z + \beta\overline{z}$ are $\mathbb{R}$-linear, so they agree everywhere as soon as they agree at 1 and i, which means $\alpha + \beta = T(1)$ and $\alpha - \beta = -iT(i)$. This system has the unique solution $\alpha = (T(1) - iT(i))/2$, $\beta = (T(1) + iT(i))/2$. Then $\beta = 0$ if and only if $T(i) = iT(1)$, which holds when T is $\mathbb{C}$-linear; and $\beta = 0$ makes $T(z) = \alpha z$ $\mathbb{C}$-linear.

(a) For $a, b \in \mathbb{R}$, $\mathbb{R}$-linearity gives $T(a+bi) = aT(1) + bT(i)$, so an $\mathbb{R}$-linear map $\mathbb{C} \to \mathbb{C}$ is determined by its values at 1 and i. For $\alpha, \beta \in \mathbb{C}$, the map $T_{\alpha,\beta}(z) = \alpha z + \beta\overline{z}$ is $\mathbb{R}$-linear. Indeed, multiplication by a complex number is $\mathbb{C}$-linear (Exercise 2.1.5), hence $\mathbb{R}$-linear, and conjugation is $\mathbb{R}$-linear (Non-example 2.1.14). So

$z \mapsto \beta\overline{z}$, the composite of conjugation and multiplication by β, is $\mathbb{R}$-linear by Proposition 2.1.17, and the sum of two $\mathbb{R}$-linear maps is $\mathbb{R}$-linear by Proposition 2.1.23. (Here β is not a scalar of the real vector space of $\mathbb{R}$-linear maps, so $z \mapsto \beta\overline{z}$ is a composite, not a scalar multiple.) Its values are $T_{\alpha,\beta}(1) = \alpha + \beta$ and $T_{\alpha,\beta}(i) = (\alpha - \beta)i$.

So $T = T_{\alpha,\beta}$ if and only if $\alpha + \beta = T(1)$ and $(\alpha - \beta)i = T(i)$, that is, $\alpha - \beta = -iT(i)$. Adding and subtracting these two equations shows that they hold if and only if

$$\alpha = \frac{T(1) - iT(i)}{2}, \quad \beta = \frac{T(1) + iT(i)}{2}.$$

So α and β exist and are unique.

(b) From the formula, $\beta = 0$ if and only if $T(1) + iT(i) = 0$; multiplying by $-i$, this says $T(i) = iT(1)$. If T is $\mathbb{C}$-linear, then $T(i) = T(i \cdot 1) = iT(1)$. Conversely, if $\beta = 0$, then $T(z) = \alpha z$, which is $\mathbb{C}$-linear by Exercise 2.1.5.

Solution to Exercise 2.1.48

Uses. Proposition 1.2.35 (Zero products), Theorem 1.5.5 (Independence is decided at zero), Definition 2.1.1 (Linear map)
Hint 1. For each nonzero $\mathbf{v}$ there is a scalar $\lambda_{\mathbf{v}}$ with $T(\mathbf{v}) = \lambda_{\mathbf{v}}\mathbf{v}$, and it is unique. The task is to show that $\lambda_{\mathbf{v}}$ does not depend on $\mathbf{v}$.
Hint 2. Take nonzero $\mathbf{u}, \mathbf{v}$. If the list $\mathbf{u}, \mathbf{v}$ is linearly independent, expand $T(\mathbf{u} + \mathbf{v})$ in two ways and compare coefficients.
Hint 3. If $\mathbf{u}, \mathbf{v}$ is dependent, then $\mathbf{v} = c\mathbf{u}$ for some c; compute $T(\mathbf{v})$ directly.
Sketch. Write $T(\mathbf{v}) = \lambda_{\mathbf{v}}\mathbf{v}$ for $\mathbf{v} \neq \mathbf{0}$. If $\mathbf{u}, \mathbf{v}$ is independent, then $\lambda_{\mathbf{u}+\mathbf{v}}(\mathbf{u} + \mathbf{v}) = \lambda_{\mathbf{u}}\mathbf{u} + \lambda_{\mathbf{v}}\mathbf{v}$, and comparing coefficients gives $\lambda_{\mathbf{u}} = \lambda_{\mathbf{u}+\mathbf{v}} = \lambda_{\mathbf{v}}$. If $\mathbf{v} = c\mathbf{u}$, then $T(\mathbf{v}) = c\lambda_{\mathbf{u}}\mathbf{u} = \lambda_{\mathbf{u}}\mathbf{v}$. So all $\lambda_{\mathbf{v}}$ are equal.

If $V = \{\mathbf{0}\}$, any λ works. Otherwise, for each $\mathbf{v} \neq \mathbf{0}$ choose $\lambda_{\mathbf{v}} \in F$ with $T(\mathbf{v}) = \lambda_{\mathbf{v}}\mathbf{v}$. It is unique: if $\lambda\mathbf{v} = \mu\mathbf{v}$, then $(\lambda - \mu)\mathbf{v} = \mathbf{0}$, so $\lambda = \mu$ by Proposition 1.2.35. We show $\lambda_{\mathbf{u}} = \lambda_{\mathbf{v}}$ for all nonzero $\mathbf{u}, \mathbf{v} \in V$.

Case 1: the list $\mathbf{u}, \mathbf{v}$ is linearly independent. Then $\mathbf{u} + \mathbf{v} \neq \mathbf{0}$ (otherwise $1\mathbf{u} + 1\mathbf{v} = \mathbf{0}$ would be a nontrivial relation), and

$$\lambda_{\mathbf{u}+\mathbf{v}}\mathbf{u} + \lambda_{\mathbf{u}+\mathbf{v}}\mathbf{v} = T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \lambda_{\mathbf{u}}\mathbf{u} + \lambda_{\mathbf{v}}\mathbf{v}.$$

So $(\lambda_{\mathbf{u}+\mathbf{v}} - \lambda_{\mathbf{u}})\mathbf{u} + (\lambda_{\mathbf{u}+\mathbf{v}} - \lambda_{\mathbf{v}})\mathbf{v} = \mathbf{0}$, and by Theorem 1.5.5 both coefficients are 0. Hence $\lambda_{\mathbf{u}} = \lambda_{\mathbf{u}+\mathbf{v}} = \lambda_{\mathbf{v}}$.

Case 2: the list $\mathbf{u}, \mathbf{v}$ is linearly dependent. By Theorem 1.5.5 there are $a, b \in F$, not both 0, with $a\mathbf{u} + b\mathbf{v} = \mathbf{0}$. If $b = 0$, then $a \neq 0$ and $a\mathbf{u} = \mathbf{0}$ would force $\mathbf{u} = \mathbf{0}$ by Proposition 1.2.35; so $b \neq 0$, and $\mathbf{v} = c\mathbf{u}$ with $c = -a/b$. Then

$$T(\mathbf{v}) = cT(\mathbf{u}) = c\lambda_{\mathbf{u}}\mathbf{u} = \lambda_{\mathbf{u}}(c\mathbf{u}) = \lambda_{\mathbf{u}}\mathbf{v},$$

so $\lambda_{\mathbf{v}} = \lambda_{\mathbf{u}}$ by uniqueness.

So all the $\lambda_{\mathbf{v}}$ equal a common value λ, and $T(\mathbf{v}) = \lambda\mathbf{v}$ for every $\mathbf{v} \neq \mathbf{0}$. Since also $T(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0}$, we get $T = \lambda \mathrm{id}_V$.

Solution to Exercise 2.1.49

Uses. Example 2.1.9 (Maps of polynomials), Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. For (a), fix q. Both sides of $D(pq) = D(p)q + pD(q)$ are linear in p, so it is enough to check them for $p = x^j$. Then fix $p = x^j$ and argue the same way in q.
Hint 2. For (b), first find $T(1)$ by applying the rule to $1 = 1 \cdot 1$. Then compute $T(x^n)$ by induction, writing $x^{n+1} = x^n \cdot x$.
Hint 3. You should get $T(x^n) = nx^{n-1}T(x)$. Two linear maps that agree on every x^n agree on every polynomial.
Sketch. (a) For a fixed q, $p \mapsto D(pq)$, $p \mapsto D(p)q$, and $p \mapsto pD(q)$ are linear (composites of D and multiplication maps), so it suffices to take $p = x^j$, and likewise $q = x^k$; there $(j + k)x^{j+k-1} = jx^{j-1}x^k + x^j kx^{k-1}$. (b) $T(1) = T(1) + T(1)$ gives $T(1) = 0$, and induction gives $T(x^n) = nx^{n-1}r$ with $r = T(x)$. So T and $p \mapsto rD(p)$ agree on monomials, hence everywhere. Conversely rD is a derivation by (a).

For $s \in F[x]$ let $M_s : F[x] \to F[x]$ be multiplication by s, $M_s(p) = sp$. It is linear, since $s(cp+q) = c(sp) + sq$ in $F[x]$. We use one observation twice: if two linear maps $F[x] \to F[x]$ agree on every monomial x^j, they agree on every $p = \sum_j a_j x^j$, since by linearity both send p to $\sum_j a_j$ times the common value at x^j. (This is the idea of Technique 2.2.7.)

(a) Fix q. The maps $p \mapsto D(pq) = \big(DM_q\big)(p)$, $p \mapsto D(p)q = \big(M_q D\big)(p)$, and $p \mapsto pD(q) = M_{D(q)}(p)$ are linear, so $p \mapsto D(p)q + pD(q)$ is linear too. Hence it suffices to prove $D\big(x^j q\big) = D\big(x^j\big)q + x^j D(q)$ for every $j \geq 0$. Now fix j; by the same argument with the roles of p and q exchanged, it suffices to take $q = x^k$. For $j, k \geq 0$,

$$D\big(x^j x^k\big) = (j+k)x^{j+k-1} = jx^{j-1}x^k + x^j \cdot kx^{k-1} = D\big(x^j\big)x^k + x^j D\big(x^k\big),$$

where a term with coefficient 0 is read as 0 (so that $jx^{j-1} = 0$ when $j = 0$, matching $D(1) = 0$).

(b) Let T be a derivation and $r = T(x)$. From $1 = 1 \cdot 1$, $T(1) = T(1) \cdot 1 + 1 \cdot T(1) = T(1) + T(1)$, so $T(1) = 0$. We claim $T(x^n) = nx^{n-1}r$ for all $n \geq 1$. For $n = 1$ this is the definition of r. If it holds for n, then

$$T\big(x^{n+1}\big) = T(x^n \cdot x) = T(x^n)x + x^n T(x) = nx^n r + x^n r = (n+1)x^n r.$$

So $T(x^n) = rD(x^n)$ for all $n \geq 0$ (for $n = 0$ both sides are 0). The maps T and $M_r D$ are linear and agree on every monomial, so $T(p) = rD(p)$ for all p.

Conversely, $p \mapsto rD(p)$ is linear, being $M_r D$, and by (a) and commutativity of multiplication in $F[x]$,

$$rD(pq) = r(D(p)q + pD(q)) = (rD(p))q + p(rD(q)).$$

So it is a derivation, with value $rD(x) = r$ at x.

Solution to Exercise 2.1.50

> *Uses.* Example 2.1.9 (Maps of polynomials), Proposition 2.1.24 (Algebra of composition), Exercise 2.1.45
> *Hint 1.* Work with $e_n = x^n/n!$ for $n \geq 0$. Then $D(e_n) = e_{n-1}$ for $n \geq 1$ and $D(e_0) = 0$, and every polynomial is a combination of the e_n in exactly one way.
> *Hint 2.* Write $T(e_n) = \sum_j b_{n,j} e_j$. Apply $DT = TD$ to e_0 to see that $T(e_0)$ is constant, and to e_n to get $b_{n,j+1} = b_{n-1,j}$.
> *Hint 3.* Let $c_k = b_{k,0}$, the constant term of $T(e_k)$. Show by induction on n that $T(e_n) = c_0 e_n + c_1 e_{n-1} + \cdots + c_n e_0$, and note that $e_{n-k} = D^k(e_n)$.
> *Sketch.* With $e_n = x^n/n!$, D shifts e_n to e_{n-1}. Writing $T(e_n) = \sum_j b_{n,j} e_j$, the relation $DT(e_n) = T(e_{n-1})$ gives $b_{n,j+1} = b_{n-1,j}$, and $DT(e_0) = 0$ gives $b_{0,j} = 0$ for $j \geq 1$. By induction, $b_{n,j} = c_{n-j}$ for $j \leq n$ and 0 for $j > n$, where $c_k = b_{k,0}$. So $T(e_n) = \sum_{k \leq n} c_k e_{n-k} = \sum_{k \leq n} c_k D^k(e_n)$, and linearity extends this to all p.

Setup. For $n \geq 0$ let $e_n = x^n/n! \in \mathbb{F}[x]$; dividing by $n!$ is allowed since $n! \neq 0$ in $\mathbb{R}$ and $\mathbb{C}$. Then $D(e_0) = 0$ and $D(e_n) = nx^{n-1}/n! = e_{n-1}$ for $n \geq 1$, so $D^k(e_n) = e_{n-k}$ for $k \leq n$ and $D^k(e_n) = 0$ for $k > n$. A polynomial $\sum_j a_j x^j$ equals $\sum_j \big(j! a_j\big) e_j$, and comparing coefficients of x^j shows that this is the only way to write it as a combination of the e_j.

For each n write $T(e_n) = \sum_{j \geq 0} b_{n,j} e_j$, with only finitely many $b_{n,j}$ nonzero, and let $c_k = b_{k,0}$. We claim that for every $n \geq 0$,

$$b_{n,j} = c_{n-j} \quad \text{for } 0 \leq j \leq n, \quad \text{and} \quad b_{n,j} = 0 \quad \text{for } j > n. \tag{C.1.2}$$

The case $n = 0$. Since $DT = TD$, $D(T(e_0)) = T(D(e_0)) = T(0) = 0$. But $D(T(e_0)) = \sum_{j \geq 1} b_{0,j} e_{j-1}$, and by uniqueness of coefficients $b_{0,j} = 0$ for all $j \geq 1$. Also $b_{0,0} = c_0$ by definition.

The induction step. Let $n \geq 1$ and suppose the claim holds for $n-1$. Then

$$\sum_{j \geq 0} b_{n,j+1} e_j = D(T(e_n)) = T(D(e_n)) = T(e_{n-1}) = \sum_{j \geq 0} b_{n-1,j} e_j,$$

so $b_{n,j+1} = b_{n-1,j}$ for all $j \geq 0$. For $0 \leq j \leq n-1$ this gives $b_{n,j+1} = c_{n-1-j} = c_{n-(j+1)}$, and for $j \geq n$ it gives $b_{n,j+1} = 0$. Together with $b_{n,0} = c_n$, this is (C.1.2) for n.

Conclusion. By (C.1.2),

$$T(e_m) = \sum_{j=0}^{m} c_{m-j} e_j = \sum_{k=0}^{m} c_k e_{m-k} = \sum_{k=0}^{n} c_k D^k(e_m) \quad \text{for every } n \geq m,$$

since $e_{m-k} = D^k(e_m)$ for $k \leq m$ and $D^k(e_m) = 0$ for $k > m$. Now let $\deg p \leq n$, and write $p = \sum_{m=0}^{n} a_m e_m$. By linearity of T and of each D^k,

$$T(p) = \sum_{m=0}^{n} a_m \sum_{k=0}^{n} c_k D^k(e_m) = \sum_{k=0}^{n} c_k D^k\left(\sum_{m=0}^{n} a_m e_m\right) = \sum_{k=0}^{n} c_k D^k(p).$$

Converse. Given scalars $c_0, c_1, \ldots$, define $T(p) = \sum_{k=0}^{n} c_k D^k(p)$ for any $n \geq \deg p$. For $p, q \in \mathbb{F}[x]$ and $a \in \mathbb{F}$, choose one n at least $\deg p$ and $\deg q$; then $T(ap+q) = \sum_{k \leq n} c_k D^k(ap+q) = aT(p) + T(q)$, because each D^k is linear. Since $\deg D(p) \leq \deg p$,

$$T(D(p)) = \sum_{k=0}^{n} c_k D^{k+1}(p) = D\left(\sum_{k=0}^{n} c_k D^k(p)\right) = D(T(p)).$$

So T is linear and commutes with D. (Compare Exercise 2.1.45: when $c_k = 0$ for large k, T is a polynomial in D.)

Solution to Exercise 2.1.51

> *Uses.* Exercise 2.1.37, every real number is a limit of a sequence of rational numbers
> *Hint 1.* Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. By Exercise 2.1.37, T is $\mathbb{Q}$-linear. What is $T(q)$ for rational q?
> *Hint 2.* Approximate a real x by rationals and use continuity.
> *Sketch.* T is $\mathbb{Q}$-linear, so $T(q) = T(q \cdot 1) = qT(1)$ for $q \in \mathbb{Q}$. If $q_n \to x$ with $q_n \in \mathbb{Q}$, continuity gives $T(x) = \lim T(q_n) = \lim q_n T(1) = xT(1)$.

Regard $\mathbb{R}$ as a vector space over $\mathbb{Q}$. Since T is additive, Exercise 2.1.37 shows that T is $\mathbb{Q}$-linear, so $T(q) = T(q \cdot 1) = qT(1)$ for every $q \in \mathbb{Q}$. Let $x \in \mathbb{R}$, and choose rationals q_n with $q_n \to x$. Since T is continuous,

$$T(x) = \lim_{n \to \infty} T(q_n) = \lim_{n \to \infty} q_n T(1) = xT(1).$$

So T is the map $x \mapsto ax$ with $a = T(1)$, which is $\mathbb{R}$-linear by Exercise 2.1.5.

Section 2.2

Solution to Exercise 2.2.5

> *Uses.* Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.1 (Basis)
> *Hint 1.* First check that $(1,1), (1,2)$ is a basis of $\mathbb{R}^2$ by solving $(x,y) = a(1,1) + b(1,2)$ for a and b.
> *Hint 2.* Then $T(x,y) = aT(1,1) + bT(1,2) = a$.
> *Sketch.* Solving gives $a = 2x - y$ and $b = y - x$, uniquely, so the list is a basis and $T(x,y) = a = 2x - y$. Then $T(3,5) = 1$.

The equations $a + b = x$ and $a + 2b = y$ give $b = y - x$ by subtraction, and then $a = x - b = 2x - y$. Conversely these values satisfy both equations, so every (x,y) is a combination of $(1,1), (1,2)$ in exactly one way, and the list is a basis of $\mathbb{R}^2$. By Theorem 2.2.1 there is exactly one such T, and by (2.2.1),

$$T(x,y) = (2x - y) \cdot 1 + (y - x) \cdot 0 = 2x - y.$$

Check: $T(1,1) = 1$ and $T(1,2) = 0$. Finally, $T(3,5) = 6 - 5 = 1$.

Solution to Exercise 2.2.9

> *Uses.* Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space), Proposition 2.1.17 (Compositions of linear maps are linear)
> *Techniques.* Check it on a basis (Technique 2.2.7)
> *Hint 1.* Both sides are linear maps, so it suffices to compare them at each $\mathbf{v}_i$.
> *Hint 2.* Compute $T^2(\mathbf{v}_i) = T(\mathbf{v}_i + \mathbf{v}_1) = T(\mathbf{v}_i) + T(\mathbf{v}_1)$.
> *Sketch.* $T^2(\mathbf{v}_i) = (\mathbf{v}_i + \mathbf{v}_1) + 2\mathbf{v}_1 = \mathbf{v}_i + 3\mathbf{v}_1$, and $(3T - 2\,\mathrm{id})(\mathbf{v}_i) = 3\mathbf{v}_i + 3\mathbf{v}_1 - 2\mathbf{v}_i$ is the same vector.

Both T^2 and $3T - 2\mathrm{id}_V$ are linear (Proposition 2.1.17, Proposition 2.1.23), so by Corollary 2.2.6 it suffices to show that they agree at each $\mathbf{v}_i$. For every i,

$$T^2(\mathbf{v}_i) = T(\mathbf{v}_i + \mathbf{v}_1) = T(\mathbf{v}_i) + T(\mathbf{v}_1) = (\mathbf{v}_i + \mathbf{v}_1) + 2\mathbf{v}_1 = \mathbf{v}_i + 3\mathbf{v}_1,$$

and

$$(3T - 2\mathrm{id}_V)(\mathbf{v}_i) = 3(\mathbf{v}_i + \mathbf{v}_1) - 2\mathbf{v}_i = \mathbf{v}_i + 3\mathbf{v}_1.$$

(For $i = 1$ both computations give $4\mathbf{v}_1$.) So $T^2 = 3T - 2\mathrm{id}_V$.

Solution to Exercise 2.2.21

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.1 (Basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations)
Hint 1. (a) Solve $(x, y) = a(1, 2) + b(2, 1)$ for a and b in $\mathbb{F}_5$. You will need 3^{-1} in $\mathbb{F}_5$.
Hint 2. (b) Compare $(2, 1)$ with $2 \cdot (1, 2)$ in $\mathbb{F}_3^2$.
Sketch. (a) Solving gives $b = 4x + 3y$ and $a = 3x + 4y$, uniquely, so the list is a basis and $T(x, y) = a(1, 0) + b(0, 1) = (3x + 4y, 4x + 3y)$. (b) In $\mathbb{F}_3^2$, $(2, 1) = 2(1, 2)$, so a linear T would have $T(2, 1) = (2, 0) \neq (0, 1)$.

(a) Consider the equations $a + 2b = x$ and $2a + b = y$ over $\mathbb{F}_5$. Subtracting twice the first from the second gives $-3b = y - 2x$, that is, $3b = 2x - y$. Since $3 \cdot 2 = 6 = 1$ in $\mathbb{F}_5$, $3^{-1} = 2$, so

$$b = 2(2x - y) = 4x - 2y = 4x + 3y, \quad a = x - 2b = x - 8x - 6y = 3x + 4y.$$

Conversely, these values satisfy both equations: $a + 2b = 11x + 10y = x$ and $2a + b = 10x + 11y = y$ in $\mathbb{F}_5$. So every vector has exactly one representation, and $(1, 2), (2, 1)$ is a basis of $\mathbb{F}_5^2$. By Theorem 2.2.1 there is exactly one such T, and by (2.2.1)

$$T(x, y) = a(1, 0) + b(0, 1) = (3x + 4y, 4x + 3y).$$

Check: $T(1, 2) = (11, 10) = (1, 0)$ and $T(2, 1) = (10, 11) = (0, 1)$ in $\mathbb{F}_5^2$.

(b) In $\mathbb{F}_3^2$, $2 \cdot (1, 2) = (2, 4) = (2, 1)$. A linear T would therefore satisfy $T(2, 1) = 2T(1, 2) = (2, 0)$, which is not $(0, 1)$. So no such T exists. Here $(1, 2), (2, 1)$ is linearly dependent, and the prescribed values do not satisfy the relation $(2, 1) - 2 \cdot (1, 2) = \mathbf{0}$.

Solution to Exercise 2.2.22

Uses. Algorithm 2.2.15 (Deciding whether a linear map with prescribed values exists), Proposition 2.2.14 (Prescribing values on any list), Proposition 2.1.16 (Linear maps preserve spans and linear relations)
Hint 1. Follow Algorithm 2.2.15: look for a linear relation among the given vectors, then ask whether they span the domain.
Hint 2. (a) $(2, 4) = 2(1, 2)$. (b) $(1, 1, 2)$ is the sum of the other two vectors; do these span $\mathbb{R}^3$? (c) $(3, 1) = 2(1, 1) + (1, -1)$. (d) Add the three vectors in $\mathbb{F}_2^3$.
Sketch. (a) No: $T(2, 4)$ would be $2T(1, 2) = (2, 0)$. (b) Exists, since the relation is respected, but is not unique, since the vectors span only a plane. (c) Exists and is unique: $T(x, y) = x + y$. (d) No: the three vectors add up to $\mathbf{0}$ in $\mathbb{F}_2^3$, but the values add up to 1.

(a) No. Since $(2, 4) = 2(1, 2)$, a linear T would have $T(2, 4) = 2T(1, 2) = (2, 0) \neq (2, 1)$.

(b) Exists, but is not unique. Let $\mathbf{u}_1 = (1, 0, 1)$ and $\mathbf{u}_2 = (0, 1, 1)$. They are independent, since $a\mathbf{u}_1 + b\mathbf{u}_2 = (a, b, a + b)$ is $\mathbf{0}$ only for $a = b = 0$. The third vector is $(1, 1, 2) = \mathbf{u}_1 + \mathbf{u}_2$, and the values satisfy the same relation: $(1, 1) + (2, 0) = (3, 1)$. The list does not span $\mathbb{R}^3$, since every vector $(a, b, a + b)$ of the span has third coordinate equal to the sum of the first two, and $\mathbf{e}_3$ does not. Adding $\mathbf{e}_3$ gives a basis: $(x, y, z) = a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{e}_3$ exactly when $a = x$, $b = y$, and $c = z - x - y$. By Proposition 2.2.14, a map exists and is not unique. Explicitly, for each $\mathbf{c} \in \mathbb{R}^2$,

$$T(x, y, z) = x(1, 1) + y(2, 0) + (z - x - y)\mathbf{c}$$

has the prescribed values and $T(\mathbf{e}_3) = \mathbf{c}$.

(c) Exists and is unique. The vectors $(1,1),(1,-1)$ form a basis of $\mathbb{R}^2$ (Example 2.2.3), and $(3,1) = 2(1,1) + (1,-1)$. The values satisfy the same relation, $2 \cdot 2 + 0 = 4$, so by Remark 2.2.12 a map exists, and it is unique because the list spans $\mathbb{R}^2$. Explicitly, $T(x,y) = x + y$, which gives $T(1,1) = 2$, $T(1,-1) = 0$, and $T(3,1) = 4$.

(d) No. In $\mathbb{F}_2^3$,

$$(1,1,0) + (0,1,1) + (1,0,1) = (2,2,2) = \mathbf{0},$$

so by Proposition 2.1.16 a linear T would satisfy $T(1,1,0) + T(0,1,1) + T(1,0,1) = 0$. But $1+1+1 = 1$ in $\mathbb{F}_2$.

Solution to Exercise 2.2.23

Uses. Proposition 2.2.20 (Linear maps out of $F[x]$), Example 2.1.10 (Integration), Example 2.1.9 (Maps of polynomials), Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 1.2.18 (Degree and evaluation versus the operations)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. In each part, show that the proposed map is linear and has the right values on monomials; then Proposition 2.2.20 says it is the map in question.
Hint 2. $\int_0^1 t^n \, dt = 1/(n+1)$, and $p \mapsto p'(1)$ is the composite of differentiation and evaluation at 1.
Sketch. (a) $p \mapsto \int_0^1 p$ is linear and sends x^n to $1/(n+1)$. (b) $p \mapsto p'(1)$ is linear, being evaluation at 1 after D, and sends x^n to $n \cdot 1^{n-1} = n$. By uniqueness in Proposition 2.2.20, these are T and U. $U\left(x^2 + x\right) = 2 + 1 = 3$, which is 3, 1, and 0 in $\mathbb{R}$, $\mathbb{F}_2$, and $\mathbb{F}_3$.

(a) Let $I(p) = \int_0^1 p(t) \, dt$. Evaluation respects the operations of $\mathbb{R}[x]$, $(cp+q)(t) = cp(t) + q(t)$ (Proposition 1.2.18), so by the linearity of the integral (Example 2.1.10), $I(cp+q) = cI(p) + I(q)$, and I is linear. Since $I(x^n) = \int_0^1 t^n \, dt = 1/(n+1)$, the maps I and T are linear and agree on every monomial, so $I = T$ by the uniqueness part of Proposition 2.2.20.

(b) Let $U'(p) = p'(1)$, that is, $U' = \mathrm{ev}_1 \circ D$ with ev_1 the evaluation at 1 and D the formal derivative. Both are linear (Example 2.1.9), so U' is linear (Proposition 2.1.17). We have $U'(1) = \mathrm{ev}_1(0) = 0$, and $U'(x^n) = \mathrm{ev}_1\left(nx^{n-1}\right) = n \cdot 1^{n-1} = n$ for $n \geq 1$. So U' and U agree on every monomial, and $U' = U$ by Proposition 2.2.20.

Finally, $U\left(x^2 + x\right) = 2 + 1 = 3 \cdot 1$. This is 3 over $\mathbb{R}$, 1 over $\mathbb{F}_2$, and 0 over $\mathbb{F}_3$. Over $\mathbb{F}_3$, indeed, $D\left(x^2 + x\right) = 2x + 1$ and $2 + 1 = 0$.

Solution to Exercise 2.2.24

Uses. Proposition 2.2.11 (Prescribing values on a spanning list), Non-example 2.2.10 (A spanning list that is not independent)
Hint 1. Does $T(\mathbf{v})$ depend on which scalars were chosen?
Hint 2. Try $V = W = \mathbb{R}$, the spanning list $1, 1$, and the values $0, 1$.
Sketch. $T(\mathbf{v})$ depends on the chosen representation, and the proof silently uses other representations: the trivial one for $\mathbf{v}_i$, and $\sum_i (ca_i + b_i)\mathbf{v}_i$ for $c\mathbf{u} + \mathbf{v}$. Counterexample: the list $1, 1$ in $\mathbb{R}$ with values $0, 1$. The claim becomes true under (2.2.2).

The definition of T involves a choice of scalars for each $\mathbf{v}$, and when $\mathbf{v}$ has several representations, different choices can give different values. The proof then treats $T(\mathbf{v})$ as if it could be computed from any representation. It computes $T(\mathbf{v}_i)$ from the representation with $a_i = 1$, which need not be the one chosen for $\mathbf{v}_i$, and it computes $T(c\mathbf{u} + \mathbf{v})$ from the representation $\sum_i (ca_i + b_i)\mathbf{v}_i$, which need not be the one chosen for $c\mathbf{u} + \mathbf{v}$. Both steps are valid only if the value does not depend on the representation, and that is exactly what can fail.

The claim is false. In $V = W = \mathbb{R}$, the list $1, 1$ spans $\mathbb{R}$. With $\mathbf{w}_1 = 0$ and $\mathbf{w}_2 = 1$, a linear T would need $T(1) = 0$ and $T(1) = 1$. The argument becomes correct under the extra hypothesis (2.2.2), which is exactly what makes the value independent of the representation; the first proof of Proposition 2.2.11 is the repaired version.

Solution to Exercise 2.2.25

Uses. Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Check it on a basis (Technique 2.2.7)

Hint 1. Technique 2.2.7 applies to equations between linear maps and to statements that reduce to them. Which of these statements are of that kind?
Hint 2. For (a) and (c), try $T(x,y) = (x, 2y)$ and $T(x,y) = x - y$ on $\mathbb{R}^2$ with the standard basis.
Sketch. (a) False: $T(x,y) = (x,2y)$ scales $\mathbf{e}_1$ and $\mathbf{e}_2$, but $T(1,1) = (1,2)$. (b) True: $T(V) = \operatorname{span} T(\mathcal{B}) \subseteq U$. (c) False: $(x,y) \mapsto x - y$ kills $(1,1)$. (d) True: $T\big(\sum_i a_i \mathbf{v}_i\big) = \big(\sum_i a_i\big) T(\mathbf{v}_1)$.

(a) False. On $\mathbb{R}^2$ with the standard basis, $T(x,y) = (x,2y)$ satisfies $T(\mathbf{e}_1) = \mathbf{e}_1$ and $T(\mathbf{e}_2) = 2\mathbf{e}_2$. But $T(1,1) = (1,2)$, which is not in $\operatorname{span}((1,1))$, since multiples of $(1,1)$ have equal coordinates. The condition "$T(\mathbf{v}) \in \operatorname{span}(\mathbf{v})$" is not an equation between linear maps, so it does not pass from a basis to all vectors.

(b) True. By Proposition 2.1.16, $T(V) = T(\operatorname{span} \mathcal{B}) = \operatorname{span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n))$. Since U is a subspace containing every $T(\mathbf{v}_i)$, it contains their span (Proposition 1.4.14).

(c) False. Let $T : \mathbb{R}^2 \to \mathbb{R}$, $T(x,y) = x - y$. Then $T(\mathbf{e}_1) = 1$ and $T(\mathbf{e}_2) = -1$ are nonzero, but $T(1,1) = 0$.

(d) True. Let $\mathbf{w}$ be the common value. For $\mathbf{v} = \sum_i a_i \mathbf{v}_i$, linearity gives $T(\mathbf{v}) = \sum_i a_i \mathbf{w} = (a_1 + \cdots + a_n)\mathbf{w} \in \operatorname{span}(\mathbf{w})$.

Solution to Exercise 2.2.26

Uses. Theorem 2.2.1 (Linear maps are determined by a basis)
Hint 1. The theorem gives three maps, one for each pair of vectors. Why should they be the same map?
Hint 2. Write $(1,1)$ as a combination of $(1,2)$ and $(2,1)$.
Sketch. The three maps given by the theorem need not coincide, and the last sentence does not follow. The linear map with $T(1,2) = T(2,1) = 1$ is unique, namely $T(x,y) = (x+y)/3$, so $T(1,1)$ is forced to be $2/3$. Hence $c = 2/3$ only, and the claim is false.

(a) The argument produces, for each of the three pairs, a linear map that takes the two prescribed values on that pair. These are three maps, and nothing in the argument shows that they are the same map, or that any one of them takes all three values. The first sentence is true, but the conclusion does not follow from it, and (b) shows that the claim is false.

(b) The vectors $(1,2), (2,1)$ form a basis of $\mathbb{R}^2$: the equations $a + 2b = x$ and $2a + b = y$ have the unique solution $a = (2y - x)/3$, $b = (2x - y)/3$. By Theorem 2.2.1, there is exactly one linear T_0 with $T_0(1,2) = T_0(2,1) = 1$, namely

$$T_0(x,y) = a + b = \frac{x+y}{3}.$$

Any T with $T(1,2) = T(2,1) = 1$ equals T_0, so $T(1,1) = T_0(1,1) = 2/3$. Hence $c = 2/3$ is the only possible value, and for it T_0 works. In particular no linear map has $T(1,1) = 1$. Equivalently, $(1,1) = \frac{1}{3}(1,2) + \frac{1}{3}(2,1)$, and the prescribed values violate this relation unless $c = 2/3$.

Solution to Exercise 2.2.27

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.1 (Basis), Proposition 1.2.19 (Restriction of scalars), Non-example 2.1.14 (Complex conjugation)
Hint 1. (a) Solve $(z_1, z_2) = a(1,i) + b(1,-i)$ for $a, b \in \mathbb{C}$.
Hint 2. (b) Over $\mathbb{C}$, how is $(i,-1)$ related to $(1,i)$?
Hint 3. (c) Over $\mathbb{R}$ the scalar i is not available. Extend $(1,i), (i,-1)$ to a basis of $\mathbb{C}^2$ over $\mathbb{R}$ by adding $(0,1)$ and $(0,i)$.
Sketch. (a) $a = (z_1 - iz_2)/2$ and $b = (z_1 + iz_2)/2$, uniquely, so $T(z_1,z_2) = ((z_1 - iz_2)/2, (z_1 + iz_2)/2)$. (b) $(i,-1) = i(1,i)$, so $T(i,-1)$ would be $(i,0)$. (c) $(1,i), (i,-1), (0,1), (0,i)$ is a basis over $\mathbb{R}$; apply Theorem 2.2.1 over $\mathbb{R}$.

(a) The equation $(z_1, z_2) = a(1,i) + b(1,-i)$ means $a + b = z_1$ and $i(a - b) = z_2$, that is, $a - b = -iz_2$. Adding and subtracting,

$$a = \frac{z_1 - iz_2}{2}, \quad b = \frac{z_1 + iz_2}{2},$$

and these values satisfy both equations. So $(1,i), (1,-i)$ is a basis of $\mathbb{C}^2$, and by Theorem 2.2.1 there is exactly one such T:

$$T(z_1, z_2) = a(1,0) + b(0,1) = \left(\frac{z_1 - iz_2}{2}, \frac{z_1 + iz_2}{2}\right).$$

Check: at $(1,i)$, $z_1 - iz_2 = 1 - i^2 = 2$ and $z_1 + iz_2 = 1 + i^2 = 0$, so $T(1,i) = (1,0)$; similarly $T(1,-i) = (0,1)$.

(b) Since $i \cdot i = -1$, we have $(i,-1) = i(1,i)$. A $\mathbb{C}$-linear T would satisfy $T(i,-1) = iT(1,i) = (i,0)$, which is not $(0,1)$.

(c) Over $\mathbb{R}$ we show that $(1,i),(i,-1),(0,1),(0,i)$ is a basis of $\mathbb{C}^2$. For $a,b,c,d \in \mathbb{R}$,

$$a(1,i) + b(i,-1) + c(0,1) + d(0,i) = (a + bi, (c-b) + (a+d)i).$$

This equals a given (z_1, z_2) exactly when $a + bi = z_1$ and $(c-b) + (a+d)i = z_2$. Comparing real and imaginary parts, the first equation determines a and b, and then the second determines c and d. So there is exactly one solution, and the list is a basis of $\mathbb{C}^2$ over $\mathbb{R}$. By Theorem 2.2.1 over the field $\mathbb{R}$, there is an $\mathbb{R}$-linear T with $T(1,i) = (1,0)$, $T(i,-1) = (0,1)$, and, say, $T(0,1) = T(0,i) = \mathbf{0}$. By (b), T is not $\mathbb{C}$-linear. As with complex conjugation (Non-example 2.1.14), whether a map is linear depends on the field of scalars.

Solution to Exercise 2.2.28

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
Hint 1. Find a basis of $\mathbb{R}^3$ consisting of two vectors of U and $(1,1,1)$. Try $(1,-1,0)$ and $(0,1,-1)$.
Hint 2. Express (x,y,z) in that basis. The coefficient of $(1,1,1)$ is $(x+y+z)/3$.
Hint 3. For $T^2 = T$, compare both sides on your basis.
Sketch. $\mathbf{u}_1 = (1,-1,0)$ and $\mathbf{u}_2 = (0,1,-1)$ span U, and with $\mathbf{u}_3 = (1,1,1)$ they form a basis of $\mathbb{R}^3$, in which (x,y,z) has $\mathbf{u}_3$-coefficient $s = (x+y+z)/3$. The conditions force $\mathbf{u}_1 \mapsto \mathbf{u}_1$, $\mathbf{u}_2 \mapsto \mathbf{u}_2$, $\mathbf{u}_3 \mapsto \mathbf{0}$, so $T(\mathbf{v}) = \mathbf{v} - s\mathbf{u}_3$. Both T^2 and T fix $\mathbf{u}_1, \mathbf{u}_2$ and kill $\mathbf{u}_3$.

Let $\mathbf{u}_1 = (1,-1,0)$, $\mathbf{u}_2 = (0,1,-1)$, and $\mathbf{u}_3 = (1,1,1)$. For $(x,y,z) \in \mathbb{R}^3$,

$$a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 = (a + c, -a + b + c, -b + c).$$

Setting this equal to (x,y,z) and adding the three coordinates gives $3c = x + y + z$. The first and third coordinates then give $a = x - c$ and $b = c - z$, and the second holds automatically, since $-a + b + c = 3c - x - z = y$. So the unique solution is

$$c = \frac{x+y+z}{3}, \quad a = x - c, \quad b = c - z,$$

and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is a basis of $\mathbb{R}^3$. Both $\mathbf{u}_1$ and $\mathbf{u}_2$ lie in U, and if $(x,y,z) \in U$ then $c = 0$, so $U = \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$.

Uniqueness. Any T with the required properties satisfies $T(\mathbf{u}_1) = \mathbf{u}_1$, $T(\mathbf{u}_2) = \mathbf{u}_2$, and $T(\mathbf{u}_3) = \mathbf{0}$, and by Theorem 2.2.1 only one linear map does.

Existence. Let T be that linear map. For $\mathbf{u} = a\mathbf{u}_1 + b\mathbf{u}_2 \in U$, $T(\mathbf{u}) = a\mathbf{u}_1 + b\mathbf{u}_2 = \mathbf{u}$, and $T(1,1,1) = \mathbf{0}$. For a general vector $\mathbf{v} = (x,y,z) = a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3$,

$$T(\mathbf{v}) = a\mathbf{u}_1 + b\mathbf{u}_2 = \mathbf{v} - c\mathbf{u}_3 = (x - s, y - s, z - s), \quad \text{where } s = \frac{x+y+z}{3}.$$

So T subtracts from each coordinate the average of the three coordinates.

$T^2 = T$. Both sides are linear. For $i = 1,2$, $T^2(\mathbf{u}_i) = T(\mathbf{u}_i) = \mathbf{u}_i$, and $T^2(\mathbf{u}_3) = T(\mathbf{0}) = \mathbf{0} = T(\mathbf{u}_3)$. By Corollary 2.2.6, $T^2 = T$.

Solution to Exercise 2.2.29

Uses. Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.24 (Algebra of composition), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. All three statements are identities between linear maps, or follow from one. Check them on $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
Hint 2. T moves each basis vector one step around the cycle $\mathbf{v}_1 \to \mathbf{v}_2 \to \cdots \to \mathbf{v}_n \to \mathbf{v}_1$. Where is $\mathbf{v}_i$ after k steps?

Hint 3. $N(\mathbf{v}_i)$ is the sum of the vectors that $\mathbf{v}_i$ visits in its first $n-1$ steps, together with $\mathbf{v}_i$ itself.
Sketch. With indices read mod n, $T^k(\mathbf{v}_i) = \mathbf{v}_{i+k}$. So $T^n(\mathbf{v}_i) = \mathbf{v}_i$, and $N(\mathbf{v}_i) = \mathbf{v}_i + \mathbf{v}_{i+1} + \cdots + \mathbf{v}_{i+n-1} = \mathbf{s}$, since these are all n basis vectors once each. Linearity gives (b). Since $T(\mathbf{s}) = \mathbf{s}$, TN and NT also send every $\mathbf{v}_i$ to $\mathbf{s}$.

For an integer j, write $\mathbf{v}_j$ for $\mathbf{v}_r$, where $r \in \{1, \ldots, n\}$ and $r \equiv j \pmod{n}$. With this convention $T(\mathbf{v}_j) = \mathbf{v}_{j+1}$ for every integer j, and by induction on k, $T^k(\mathbf{v}_i) = \mathbf{v}_{i+k}$ for all $k \geq 0$. All maps below are linear by Proposition 2.1.17 and Proposition 2.1.23.

(a) $T^n(\mathbf{v}_i) = \mathbf{v}_{i+n} = \mathbf{v}_i$ for each i. So T^n and id_V agree on a basis, and $T^n = \mathrm{id}_V$ by Corollary 2.2.6.

(b) For each i,

$$N(\mathbf{v}_i) = \sum_{k=0}^{n-1} T^k(\mathbf{v}_i) = \sum_{k=0}^{n-1} \mathbf{v}_{i+k} = \mathbf{s},$$

because the n consecutive integers $i, i+1, \ldots, i+n-1$ cover each residue mod n exactly once. By linearity, $N\left(\sum_i a_i \mathbf{v}_i\right) = \sum_i a_i \mathbf{s} = (a_1 + \cdots + a_n)\mathbf{s}$.

(c) $T(\mathbf{s}) = \sum_i T(\mathbf{v}_i) = \sum_i \mathbf{v}_{i+1} = \mathbf{s}$, the same n vectors in another order. So $TN(\mathbf{v}_i) = T(\mathbf{s}) = \mathbf{s} = N(\mathbf{v}_i)$, and $NT(\mathbf{v}_i) = N(\mathbf{v}_{i+1}) = \mathbf{s} = N(\mathbf{v}_i)$. By Corollary 2.2.6, $TN = NT = N$. (Alternatively, by Proposition 2.1.24 and (a), $TN = T + T^2 + \cdots + T^n = T + \cdots + T^{n-1} + \mathrm{id}_V = N$, and likewise $NT = T + T^2 + \cdots + T^n = N$.)

Solution to Exercise 2.2.30

Uses. Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.24 (Algebra of composition), Exercise 2.1.5
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. $\mathbf{e}_1$ and $R(\mathbf{e}_1) = \mathbf{e}_2$ form a basis of $\mathbb{R}^2$, so T is determined by $T(\mathbf{e}_1)$ and $T(R(\mathbf{e}_1))$.
Hint 2. If $TR = RT$, then $T(R(\mathbf{e}_1)) = R(T(\mathbf{e}_1))$. Write $T(\mathbf{e}_1) = (a, b)$ and compare T with $a\,\mathrm{id} + bR$ at $\mathbf{e}_1$ and $\mathbf{e}_2$.
Sketch. If $TR = RT$ and $T(\mathbf{e}_1) = (a, b)$, then $T(\mathbf{e}_2) = T(R\mathbf{e}_1) = R(a, b) = (-b, a)$. The map $a\,\mathrm{id} + bR$ also sends $\mathbf{e}_1 \mapsto (a, b)$ and $\mathbf{e}_2 \mapsto (-b, a)$, so it equals T. Conversely $a\,\mathrm{id} + bR$ commutes with R by Proposition 2.1.24. Under $(x, y) \leftrightarrow x + yi$, R is multiplication by i and $a\,\mathrm{id} + bR$ is multiplication by $a + bi$.

Suppose $TR = RT$, and let $T(\mathbf{e}_1) = (a, b)$. Since $R(\mathbf{e}_1) = \mathbf{e}_2$,

$$T(\mathbf{e}_2) = T(R(\mathbf{e}_1)) = R(T(\mathbf{e}_1)) = R(a, b) = (-b, a).$$

On the other hand, $(a\,\mathrm{id} + bR)(\mathbf{e}_1) = (a, 0) + b(0, 1) = (a, b)$ and $(a\,\mathrm{id} + bR)(\mathbf{e}_2) = (0, a) + b(-1, 0) = (-b, a)$. So T and $a\,\mathrm{id} + bR$ are linear maps that agree on the basis $\mathbf{e}_1, \mathbf{e}_2$, and $T = a\,\mathrm{id} + bR$ by Corollary 2.2.6.

Conversely, by Proposition 2.1.24, $(a\,\mathrm{id} + bR)R = aR + bR^2 = R(a\,\mathrm{id} + bR)$.

Identify $(x, y) \in \mathbb{R}^2$ with $x + yi \in \mathbb{C}$. Then $R(x, y) = (-y, x)$ corresponds to $-y + xi = i(x + yi)$, so R is multiplication by i, and $a\,\mathrm{id} + bR$ is multiplication by $a + bi$. So the $\mathbb{R}$-linear maps of the plane that commute with the quarter turn are exactly the multiplications by complex numbers, which are the $\mathbb{C}$-linear maps $\mathbb{C} \to \mathbb{C}$ (Exercise 2.1.5).

Solution to Exercise 2.2.31

Uses. Corollary 2.2.18 (Linear maps correspond to lists of vectors), Example 2.2.3 (Building maps from a convenient basis)
Hint 1. By Corollary 2.2.18, choosing T is the same as freely choosing the images of the vectors of a basis. Pick a basis that contains $(1, 1)$.
Hint 2. For (b), when is $(1, 1), (1, -1)$ a basis of $\mathbb{F}_p^2$? What happens when $p = 2$?
Sketch. (a) $(1, 1), (1, 0)$ is a basis; $T(1, 1)$ has p allowed values and $T(1, 0)$ has p^2, so there are p^3 maps. (b) For odd p, $(1, 1), (1, -1)$ is a basis, giving $p \cdot p = p^2$. For $p = 2$, $(1, -1) = (1, 1)$, the two conditions coincide, and the answer is $2^3 = 8$.

(a) The list $(1, 1), (1, 0)$ is a basis of $\mathbb{F}_p^2$, since $(x, y) = a(1, 1) + b(1, 0)$ exactly when $a = y$ and $b = x - y$. By Corollary 2.2.18, $T \mapsto (T(1, 1), T(1, 0))$ is a bijection from $\mathrm{L}\left(\mathbb{F}_p^2\right)$ onto the pairs of vectors in $\mathbb{F}_p^2$. The condition restricts $T(1, 1)$ to $\mathrm{span}((1, 1)) = \left\{(c, c) \;\middle|\; c \in \mathbb{F}_p\right\}$, which has p elements, and leaves $T(1, 0)$ free among p^2 vectors. So there are $p \cdot p^2 = p^3$ such maps.

(b) If p is odd, then $1+1 \neq 0$ in $\mathbb{F}_p$, so $(1,1),(1,-1)$ is a basis of $\mathbb{F}_p^2$ (Example 2.2.3). The conditions restrict $T(1,1)$ to the p vectors (c,c) and $T(1,-1)$ to the p vectors $(c,-c)$, which are distinct for distinct c. So there are p^2 such maps.

If $p=2$, then $-1=1$ in $\mathbb{F}_2$, so $(1,-1)=(1,1)$ and the second condition is the first one again. By (a) the answer is $2^3=8$.

Solution to Exercise 2.2.32

Uses. Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.35 (Zero products), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Hint 1. (a) Suppose $a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n=\mathbf{0}$. Choose the $\mathbf{w}_i$ so that applying T to this relation isolates a_j.
Hint 2. Take $\mathbf{w}_j$ to be a nonzero vector of W and all other $\mathbf{w}_i=\mathbf{0}$.
Hint 3. (b) $\mathcal{L}$ is a basis of span $\mathcal{L}$.
Sketch. (a) If $\sum_i a_i\mathbf{v}_i=\mathbf{0}$, pick $\mathbf{w}\neq\mathbf{0}$ and let T send $\mathbf{v}_j\mapsto\mathbf{w}$ and the other $\mathbf{v}_i\mapsto\mathbf{0}$. Then $\mathbf{0}=T\big(\sum_i a_i\mathbf{v}_i\big)=a_j\mathbf{w}$, so $a_j=0$. (b) $\mathcal{L}$ spans span $\mathcal{L}$ and is independent, so it is a basis; apply Theorem 2.2.1.

(a) Suppose $a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n=\mathbf{0}$, and fix j. Choose $\mathbf{w}\in W$ with $\mathbf{w}\neq\mathbf{0}$, and apply the hypothesis with $\mathbf{w}_j=\mathbf{w}$ and $\mathbf{w}_i=\mathbf{0}$ for $i\neq j$. The resulting linear T gives

$$\mathbf{0}=T(\mathbf{0})=T\left(\sum_{i=1}^{n}a_i\mathbf{v}_i\right)=\sum_{i=1}^{n}a_iT(\mathbf{v}_i)=a_j\mathbf{w}.$$

Since $\mathbf{w}\neq\mathbf{0}$, Proposition 1.2.35 gives $a_j=0$. As j was arbitrary, every $a_i=0$, and $\mathcal{L}$ is linearly independent by Theorem 1.5.5. (The hypothesis $W\neq\{\mathbf{0}\}$ is needed: every list has the zero map into $\{\mathbf{0}\}$.)

(b) The span of $\mathcal{L}$ is a subspace of V (Proposition 1.4.14), hence a vector space, and $\mathcal{L}$ spans it. If $\mathcal{L}$ is linearly independent, it is therefore a basis of span $\mathcal{L}$, and Theorem 2.2.1 gives the map. To get a map on all of V, one needs to extend from span $\mathcal{L}$ to V, which is possible under the hypothesis of Proposition 2.2.14.

Solution to Exercise 2.2.33

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.14 (Coordinates), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.23 (L(V,W) is a vector space)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. (a) is Theorem 2.2.1 with $W=F$. Use (2.2.1) to evaluate $\varphi_i(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n)$.
Hint 2. (c) Both sides are linear maps $V\to F$. Compare them on the basis.
Sketch. (a) $\varphi_i\big(\sum_j a_j\mathbf{v}_j\big)=\sum_j a_j\varphi_i(\mathbf{v}_j)=a_i$. (b) is the definition of coordinates, rewritten with (a). (c) Both sides send $\mathbf{v}_k$ to $\varphi(\mathbf{v}_k)$, since $\varphi_i(\mathbf{v}_k)$ is 1 for $i=k$ and 0 otherwise.

(a) By Theorem 2.2.1 with $W=F$, there is exactly one linear φ_i with the prescribed values on $\mathcal{B}$. By (2.2.1),

$$\varphi_i(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n)=\sum_{j=1}^{n}a_j\varphi_i(\mathbf{v}_j)=a_i,$$

so $\varphi_i(\mathbf{v})$ is the i-th coordinate of $\mathbf{v}$ (Definition 1.6.14).

(b) By definition of coordinates, $\mathbf{v}=c_1\mathbf{v}_1+\cdots+c_n\mathbf{v}_n$ with $(c_1,\ldots,c_n)=[\mathbf{v}]_{\mathcal{B}}$, and $c_i=\varphi_i(\mathbf{v})$ by (a).

(c) Let $\psi=\varphi(\mathbf{v}_1)\varphi_1+\cdots+\varphi(\mathbf{v}_n)\varphi_n$, which is linear by Proposition 2.1.23. For each k,

$$\psi(\mathbf{v}_k)=\sum_{i=1}^{n}\varphi(\mathbf{v}_i)\varphi_i(\mathbf{v}_k)=\varphi(\mathbf{v}_k),$$

since only the term $i=k$ survives. So ψ and φ agree on a basis, and $\psi=\varphi$ by Corollary 2.2.6.

Solution to Exercise 2.2.34

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Corollary 2.2.6 (Agreeing on a spanning list), Theorem 1.5.5 (Independence is decided at zero), Definition 2.1.21 (Pointwise operations)
Techniques. Check it on a basis (Technique 2.2.7)

Hint 1. Each E_{ij} exists and is unique by Theorem 2.2.1. For independence, apply a vanishing combination $\sum_{i,j} c_{ij}E_{ij}$ to a basis vector $\mathbf{v}_k$.
Hint 2. For spanning, given T, write $T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m$, and compare T with $\sum_{i,j} a_{ij}E_{ij}$ on the basis of V.
Sketch. $\left(\sum_{i,j} c_{ij}E_{ij}\right)(\mathbf{v}_k) = \sum_i c_{ik}\mathbf{w}_i$. If the combination is 0, independence of the $\mathbf{w}_i$ gives every $c_{ik} = 0$. Given T with $T(\mathbf{v}_j) = \sum_i a_{ij}\mathbf{w}_i$, the map $\sum_{i,j} a_{ij}E_{ij}$ agrees with T at every $\mathbf{v}_k$, so it equals T.

By Theorem 2.2.1 each E_{ij} exists and is unique. For scalars c_{ij} and each k, the operations on $\mathrm{L}(V, W)$ are pointwise (Definition 2.1.21), so

$$\left(\sum_{i,j} c_{ij}E_{ij}\right)(\mathbf{v}_k) = \sum_{i,j} c_{ij}E_{ij}(\mathbf{v}_k) = \sum_{i=1}^{m} c_{ik}\mathbf{w}_i, \tag{C.1.3}$$

because $E_{ij}(\mathbf{v}_k) = \mathbf{0}$ unless $j = k$.

Independence. Suppose $\sum_{i,j} c_{ij}E_{ij} = 0$. Evaluating at $\mathbf{v}_k$ and using (C.1.3) gives $\sum_i c_{ik}\mathbf{w}_i = \mathbf{0}$. Since $\mathbf{w}_1, \ldots, \mathbf{w}_m$ is linearly independent, $c_{ik} = 0$ for every i (Theorem 1.5.5). This holds for every k, so all $c_{ij} = 0$, and the list is linearly independent by Theorem 1.5.5.

Spanning. Let $T \in \mathrm{L}(V, W)$. Since the $\mathbf{w}_i$ span W, there are scalars a_{ij} with $T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m$ for each j. By (C.1.3), $\left(\sum_{i,j} a_{ij}E_{ij}\right)(\mathbf{v}_k) = \sum_i a_{ik}\mathbf{w}_i = T(\mathbf{v}_k)$ for every k. So $T = \sum_{i,j} a_{ij}E_{ij}$ by Corollary 2.2.6.

The coefficients a_{ij} in the last step are the entries of the matrix of T (Section 2.3).

Solution to Exercise 2.2.35

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Definition 1.6.14 (Coordinates), Corollary 1.5.17 (Sublists), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Decide what P and S should do to each basis vector, then apply Theorem 2.2.1.
Hint 2. Let P fix each $\mathbf{u}_i$ and kill each $\mathbf{x}_j$; let S kill each $\mathbf{u}_i$ and fix each $\mathbf{x}_j$.
Hint 3. For (b), compute $S\left(\sum_i a_i\mathbf{u}_i + \sum_j b_j\mathbf{x}_j\right)$, and use that $\mathbf{x}_1, \ldots, \mathbf{x}_r$ is linearly independent. For (c), use the basis $(1, 2, 3), \mathbf{e}_2, \mathbf{e}_3$.
Sketch. $P : \mathbf{u}_i \mapsto \mathbf{u}_i$, $\mathbf{x}_j \mapsto \mathbf{0}$ gives $P\left(\sum_i a_i\mathbf{u}_i + \sum_j b_j\mathbf{x}_j\right) = \sum_i a_i\mathbf{u}_i$. $S : \mathbf{u}_i \mapsto \mathbf{0}$, $\mathbf{x}_j \mapsto \mathbf{x}_j$ gives $S(\mathbf{v}) = \sum_j b_j\mathbf{x}_j$, which is $\mathbf{0}$ iff every $b_j = 0$ iff $\mathbf{v} \in U$. For $U = \mathrm{span}((1, 2, 3))$: $P(x, y, z) = (x, 2x, 3x)$ and $S(x, y, z) = (0, y - 2x, z - 3x)$.

Each $\mathbf{v} \in V$ has exactly one representation $\mathbf{v} = \sum_i a_i\mathbf{u}_i + \sum_j b_j\mathbf{x}_j$.

(a) Let P be the linear map with $P(\mathbf{u}_i) = \mathbf{u}_i$ and $P(\mathbf{x}_j) = \mathbf{0}$ (Theorem 2.2.1). Then $P(\mathbf{v}) = \sum_i a_i\mathbf{u}_i \in U$, so $P(V) \subseteq U$. If $\mathbf{u} \in U$, write $\mathbf{u} = \sum_i c_i\mathbf{u}_i$; then $P(\mathbf{u}) = \sum_i c_i\mathbf{u}_i = \mathbf{u}$. In particular every $\mathbf{u} \in U$ equals $P(\mathbf{u}) \in P(V)$, so $P(V) = U$.

(b) Let S be the linear map with $S(\mathbf{u}_i) = \mathbf{0}$ and $S(\mathbf{x}_j) = \mathbf{x}_j$. Then $S(\mathbf{v}) = \sum_j b_j\mathbf{x}_j$. If $S(\mathbf{v}) = \mathbf{0}$, then every $b_j = 0$, because $\mathbf{x}_1, \ldots, \mathbf{x}_r$ is linearly independent, being a sublist of a basis (Corollary 1.5.17, Theorem 1.5.5); so $\mathbf{v} = \sum_i a_i\mathbf{u}_i \in U$. Conversely, if $\mathbf{v} = \sum_i c_i\mathbf{u}_i \in U$, then $S(\mathbf{v}) = \sum_i c_i S(\mathbf{u}_i) = \mathbf{0}$. So the vectors sent to $\mathbf{0}$ are exactly those of U.

(c) The single vector $(1, 2, 3)$ is a basis of U: it spans U, and it is linearly independent because it is nonzero. Together with $\mathbf{e}_2$ and $\mathbf{e}_3$ it forms a basis of $\mathbb{R}^3$, since

$$(x, y, z) = a(1, 2, 3) + b\mathbf{e}_2 + c\mathbf{e}_3 \quad \Longleftrightarrow \quad a = x, \quad b = y - 2x, \quad c = z - 3x.$$

So the constructions give

$$P(x, y, z) = x(1, 2, 3) = (x, 2x, 3x), \quad S(x, y, z) = (y - 2x)\mathbf{e}_2 + (z - 3x)\mathbf{e}_3 = (0, y - 2x, z - 3x).$$

As a check, $S(x, y, z) = \mathbf{0}$ exactly when $y = 2x$ and $z = 3x$, that is, when $(x, y, z) = x(1, 2, 3)$.

Solution to Exercise 2.2.36

Uses. Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 1.4.14 (The span is the smallest subspace containing a list), Exercise 2.2.25
Techniques. Check it on a basis (Technique 2.2.7)

Hint 1. Let $U_k = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ for $0 \le k \le n$, with $U_0 = \{\mathbf{0}\}$. Show that $T(U_k) \subseteq U_{k-1}$.
Hint 2. For $T(U_k) \subseteq U_{k-1}$, check where T sends the spanning list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of U_k (part (b) of Exercise 2.2.25).
Hint 3. Then show by induction on k that T^k sends every vector of U_k to $\mathbf{0}$.
Sketch. With $U_k = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, each $T(\mathbf{v}_j)$ with $j \le k$ lies in $U_{j-1} \subseteq U_{k-1}$, so $T(U_k) \subseteq U_{k-1}$. Each application of T lowers the index by one, so $T^k(U_k) \subseteq U_0 = \{\mathbf{0}\}$, and $k = n$ gives $T^n = 0$.

For $0 \le k \le n$, let $U_k = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$, so that $U_0 = \{\mathbf{0}\}$, $U_0 \subseteq U_1 \subseteq \cdots \subseteq U_n = V$, and each U_k is a subspace.

Claim: $T(U_k) \subseteq U_{k-1}$ **for** $1 \le k \le n$**.** For $j \le k$, the hypothesis gives $T(\mathbf{v}_j) \in U_{j-1} \subseteq U_{k-1}$ (for $j = 1$, $T(\mathbf{v}_1) = \mathbf{0} \in U_0$). So U_{k-1} is a subspace containing $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$, and by Proposition 2.1.16 and Proposition 1.4.14,

$$T(U_k) = \operatorname{span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)) \subseteq U_{k-1}.$$

Claim: $T^k(\mathbf{u}) = \mathbf{0}$ **for every** $\mathbf{u} \in U_k$**, for** $0 \le k \le n$**.** We induct on k. For $k = 0$, $U_0 = \{\mathbf{0}\}$ and $T^0(\mathbf{0}) = \mathbf{0}$. Suppose the claim holds for some $k < n$, and let $\mathbf{u} \in U_{k+1}$. Then $T(\mathbf{u}) \in U_k$ by the first claim, so $T^{k+1}(\mathbf{u}) = T^k(T(\mathbf{u})) = \mathbf{0}$.

Taking $k = n$, $T^n(\mathbf{v}) = \mathbf{0}$ for every $\mathbf{v} \in U_n = V$, that is, $T^n = 0$. (In the language of Section 2.3, the matrix of T in this basis is strictly upper triangular.)

Solution to Exercise 2.2.37

Uses. Proposition 2.2.11 (Prescribing values on a spanning list), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Section A.1
Hint 1. For ($\Longrightarrow$), apply S to a relation of $\mathcal{L}$; for the reverse implication between relations, use that S is injective.
Hint 2. For ($\Longleftarrow$), Proposition 2.2.11 gives a linear $S : \operatorname{span}\mathcal{L} \to W$ with $S(\mathbf{v}_i) = \mathbf{w}_i$. What is its image? Why is it injective?
Hint 3. If $S(\sum_i a_i\mathbf{v}_i) = S(\sum_i b_i\mathbf{v}_i)$, then $\sum_i (a_i - b_i)\mathbf{w}_i = \mathbf{0}$.
Sketch. If S exists, relations of $\mathcal{L}$ map to relations of $\mathcal{M}$, and if $\sum_i a_i\mathbf{w}_i = \mathbf{0}$ then $S(\sum_i a_i\mathbf{v}_i) = \mathbf{0} = S(\mathbf{0})$, so injectivity gives $\sum_i a_i\mathbf{v}_i = \mathbf{0}$. Conversely, Proposition 2.2.11 gives S on $\operatorname{span}\mathcal{L}$; its image is $\operatorname{span}\mathcal{M}$, and if $S(\mathbf{u}) = S(\mathbf{u}')$, the difference of coefficient lists is a relation of $\mathcal{M}$, hence of $\mathcal{L}$, so $\mathbf{u} = \mathbf{u}'$. In the example both lists have exactly the relations $a = b = -c$, so S exists.

($\Longrightarrow$). Let S be such a map. If $\sum_i a_i\mathbf{v}_i = \mathbf{0}$, then $\sum_i a_i\mathbf{w}_i = \sum_i a_i S(\mathbf{v}_i) = S(\sum_i a_i\mathbf{v}_i) = S(\mathbf{0}) = \mathbf{0}$ (Proposition 2.1.4). If $\sum_i a_i\mathbf{w}_i = \mathbf{0}$, then $S(\sum_i a_i\mathbf{v}_i) = \mathbf{0} = S(\mathbf{0})$, and since S is injective, $\sum_i a_i\mathbf{v}_i = \mathbf{0}$.

($\Longleftarrow$). The subspace $\operatorname{span}\mathcal{L}$ is a vector space spanned by $\mathcal{L}$, and every relation of $\mathcal{L}$ is a relation of $\mathcal{M}$, which is (2.2.2). By Proposition 2.2.11, applied with $\operatorname{span}\mathcal{L}$ in place of V, there is a linear $S : \operatorname{span}\mathcal{L} \to W$ with $S(\mathbf{v}_i) = \mathbf{w}_i$. By Proposition 2.1.16, $S(\operatorname{span}\mathcal{L}) = \operatorname{span}(S(\mathbf{v}_1), \ldots, S(\mathbf{v}_k)) = \operatorname{span}\mathcal{M}$, so S is a surjective linear map $\operatorname{span}\mathcal{L} \to \operatorname{span}\mathcal{M}$. It is injective: if $\mathbf{u} = \sum_i a_i\mathbf{v}_i$ and $\mathbf{u}' = \sum_i b_i\mathbf{v}_i$ satisfy $S(\mathbf{u}) = S(\mathbf{u}')$, then

$$\sum_{i=1}^{k} (a_i - b_i)\mathbf{w}_i = S(\mathbf{u}) - S(\mathbf{u}') = \mathbf{0},$$

which is a relation of $\mathcal{M}$, hence of $\mathcal{L}$: $\sum_i (a_i - b_i)\mathbf{v}_i = \mathbf{0}$, that is, $\mathbf{u} = \mathbf{u}'$. So S is bijective (Section A.1).

The example. For $\mathcal{L}$: $a(1,0) + b(0,1) + c(1,1) = (a + c, b + c)$, which is $\mathbf{0}$ exactly when $a = b = -c$. For $\mathcal{M}$: $a(1,1) + b(1,-1) + c(2,0) = (a + b + 2c, a - b)$, which is $\mathbf{0}$ exactly when $a = b$ and $2a + 2c = 0$, that is, when $a = b = -c$. The two lists have the same relations, so such an S exists. Here $\operatorname{span}\mathcal{L} = \operatorname{span}\mathcal{M} = \mathbb{R}^2$, and S is the linear map with $S(\mathbf{e}_1) = (1,1)$ and $S(\mathbf{e}_2) = (1,-1)$; indeed $S(1,1) = (2,0)$.

Solution to Exercise 2.2.38

Uses. Proposition 2.2.11 (Prescribing values on a spanning list), Proposition 1.2.35 (Zero products), Definition 1.1.28 (Characteristic)
Hint 1. The list $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$ spans F^n, so Proposition 2.2.11 applies. Find all linear relations among $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$.
Hint 2. Every relation is a multiple of $\mathbf{s} - \mathbf{e}_1 - \cdots - \mathbf{e}_n = \mathbf{0}$. So T exists if and only if $\mathbf{v}_{\sigma(n+1)} = \mathbf{v}_{\sigma(1)} + \cdots + \mathbf{v}_{\sigma(n)}$.
Hint 3. If $\sigma(n+1) = k \le n$, then $\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(n)}$ are $\mathbf{s}$ and the $\mathbf{e}_j$ with $j \ne k$, in some order.
Sketch. $\sum_i a_i\mathbf{e}_i + b\mathbf{s} = \mathbf{0}$ iff $a_i = -b$ for all i, so all relations are multiples of one, and T exists iff $\mathbf{v}_{\sigma(n+1)} = \sum_{i \le n} \mathbf{v}_{\sigma(i)}$. If σ fixes $n+1$, the right side is $\mathbf{s}$: fine. If $\sigma(n+1) = k \le n$, the right side is $\mathbf{s} + \sum_{j \ne k} \mathbf{e}_j = \mathbf{e}_k + 2\sum_{j \ne k} \mathbf{e}_j$, which equals $\mathbf{e}_k$ iff $1 + 1 = 0$. So: all $(n+1)!$ permutations if $1 + 1 = 0$, otherwise exactly the $n!$ that fix $n+1$.

The relations. For scalars $a_1, \dots, a_n, b$,

$$a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n + b\mathbf{s} = (a_1 + b, \dots, a_n + b),$$

which is $\mathbf{0}$ exactly when $a_i = -b$ for every i. So the linear relations of the list are exactly the multiples, by $b \in F$, of the relation $\mathbf{s} - \mathbf{e}_1 - \cdots - \mathbf{e}_n = \mathbf{0}$.

The criterion. The list contains the standard basis, so it spans F^n. Given σ, put $\mathbf{w}_i = \mathbf{v}_{\sigma(i)}$. A list satisfies all multiples of a relation as soon as it satisfies the relation itself, so (2.2.2) holds exactly when $\mathbf{w}_{n+1} - \mathbf{w}_1 - \cdots - \mathbf{w}_n = \mathbf{0}$. By Proposition 2.2.11, T exists if and only if

$$\mathbf{v}_{\sigma(n+1)} = \mathbf{v}_{\sigma(1)} + \cdots + \mathbf{v}_{\sigma(n)}. \tag{C.1.4}$$

Case $\sigma(n+1) = n+1$. Then σ permutes $\{1, \dots, n\}$, and the right side of (C.1.4) is $\mathbf{e}_1 + \cdots + \mathbf{e}_n = \mathbf{s} = \mathbf{v}_{n+1}$. So T exists; it permutes the standard basis.

Case $\sigma(n+1) = k \le n$. Then $\{\sigma(1), \dots, \sigma(n)\} = \{1, \dots, n+1\} \setminus \{k\}$, so with $2 = 1 + 1$ the right side of (C.1.4) is

$$\mathbf{s} + \sum_{j \le n, j \ne k} \mathbf{e}_j = \mathbf{e}_k + 2 \sum_{j \le n, j \ne k} \mathbf{e}_j.$$

Since $n \ge 2$, there is some $j \ne k$, and the vector $\sum_{j \ne k} \mathbf{e}_j$ is nonzero. By Proposition 1.2.35, (C.1.4) holds if and only if $2 = 0$.

Answer. If $1 + 1 \ne 0$ in F, exactly the $n!$ permutations with $\sigma(n+1) = n+1$ are realized by linear maps. If $1 + 1 = 0$ (characteristic 2, Definition 1.1.28), all $(n+1)!$ permutations are. The reason is visible in the relation: when $1 + 1 = 0$ it reads $\mathbf{e}_1 + \cdots + \mathbf{e}_n + \mathbf{s} = \mathbf{0}$, which is symmetric in all $n + 1$ vectors.

Solution to Exercise 2.2.39

> *Uses.* Proposition 2.2.20 (Linear maps out of $F[x]$), Exercise 2.1.32, Proposition 1.1.17 (No zero divisors), Definition 1.2.16 (Degree)
> *Techniques.* Check it on a basis (Technique 2.2.7)
> *Hint 1.* Let $e = f(1)$. Apply the condition to $1 = 1 \cdot 1$. What can e be?
> *Hint 2.* If $e = 1$, put $q = f(x)$. Compute $f(x^n)$ and compare f with the substitution map $p \mapsto p(q)$ on monomials.
> *Hint 3.* To show that $p \mapsto p(q)$ is multiplicative, fix r and note that both sides of $(pr)(q) = p(q)r(q)$ are linear in p, so it suffices to take $p = x^j$. Then fix j and argue the same way in r.
> *Sketch.* $e = f(1)$ satisfies $e^2 = e$, and comparing degrees gives $e \in \{0, 1\}$. If $e = 0$, then $f(p) = f(p)e = 0$. If $e = 1$ and $q = f(x)$, then $f(x^n) = q^n$, so f is the substitution map $C_q : p \mapsto p(q)$ by Proposition 2.2.20. Conversely C_q is linear, and multiplicative because both sides of $C_q(pr) = C_q(p)C_q(r)$ are linear in each of p and r and agree on monomials. Answer: $f = 0$ or $f = C_q$ for some $q \in F[x]$.

For $q \in F[x]$ let $C_q : F[x] \to F[x]$, $p \mapsto p(q)$, be the substitution map, which is linear by Exercise 2.1.32 and satisfies $C_q(x^n) = q^n$. We show that the maps with the property are $f = 0$ and the maps C_q.

The value $e = f(1)$. We have $e = f(1 \cdot 1) = f(1)f(1) = e^2$. Suppose $e \ne 0$, with degree d and leading coefficient $c \ne 0$ (Definition 1.2.16). In e^2, the coefficient of x^{2d} is $c^2 \ne 0$ (Proposition 1.1.17) and all higher coefficients are 0, so $\deg e^2 = 2d$. From $e^2 = e$ we get $2d = d$, so $d = 0$, and $e = c$ is a nonzero constant with $c^2 = c$, that is, $c(c-1) = 0$. Hence $c = 1$. So $e = 0$ or $e = 1$.

If $e = 0$. For every p, $f(p) = f(p \cdot 1) = f(p)e = 0$, so $f = 0$, which is linear and multiplicative.

If $e = 1$. Let $q = f(x)$. Then $f(x^n) = q^n$ for every $n \ge 0$, by induction: $f(x^0) = f(1) = 1 = q^0$, and $f(x^{n+1}) = f(x^n x) = f(x^n)f(x) = q^n q = q^{n+1}$. So f and C_q are linear maps that agree on every monomial, and $f = C_q$ by Proposition 2.2.20.

Each C_q is multiplicative. Fix $r \in F[x]$. The maps $p \mapsto C_q(pr)$ and $p \mapsto C_q(p)C_q(r)$ are linear in p: the first is C_q composed with multiplication by r, the second is multiplication by $C_q(r)$ composed with C_q, and multiplication by a fixed polynomial is linear by distributivity. By Proposition 2.2.20 it suffices to compare

them at $p = x^j$. Now fix j. The maps $r \mapsto C_q(x^j r)$ and $r \mapsto q^j C_q(r)$ are linear in r for the same reason, so it suffices to compare them at $r = x^k$:

$$C_q\big(x^j x^k\big) = C_q\big(x^{j+k}\big) = q^{j+k} = q^j q^k = C_q\big(x^j\big) C_q\big(x^k\big).$$

So $C_q(pr) = C_q(p)C_q(r)$ for all p, r.

Answer. The linear maps f with $f(pq) = f(p)f(q)$ are the zero map and the substitution maps $p \mapsto p(q)$, one for each $q \in F[x]$. Different q give different maps, since $C_q(x) = q$.

Solution to Exercise 2.2.40

Uses. Theorem 1.5.5 (Independence is decided at zero), Definition 1.6.14 (Coordinates), Theorem 2.2.1 (Linear maps are determined by a basis), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Check it on a basis (Technique 2.2.7), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. (a) If $a_j = 0$, find a nontrivial linear relation among the n vectors other than $\mathbf{v}_j$.
Hint 2. (b) Write $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ and compute $T(\mathbf{v}_{n+1})$ in two ways. Compare coordinates in the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
Hint 3. (c) Write $\mathbf{w}_{n+1} = \sum_i b_i \mathbf{w}_i$ with all $b_i \neq 0$, and define T on the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ by $T(\mathbf{v}_i) = (b_i/a_i)\mathbf{w}_i$. For uniqueness, argue as in (b).
Sketch. (a) If $a_j = 0$, then $\mathbf{v}_{n+1} - \sum_{i \neq j} a_i \mathbf{v}_i = \mathbf{0}$ is a nontrivial relation among the n vectors other than $\mathbf{v}_j$, which form a basis. (b) $T(\mathbf{v}_{n+1}) = \sum_i a_i \lambda_i \mathbf{v}_i = \lambda_{n+1} \sum_i a_i \mathbf{v}_i$, so $a_i \lambda_i = a_i \lambda_{n+1}$, and $a_i \neq 0$ gives $\lambda_i = \lambda_{n+1}$. (c) The map $\mathbf{v}_i \mapsto (b_i/a_i)\mathbf{w}_i$ sends $\mathbf{v}_{n+1}$ to $\mathbf{w}_{n+1}$. If $T'(\mathbf{v}_i) = c_i' \mathbf{w}_i$, comparing coordinates of $T'(\mathbf{v}_{n+1})$ forces $c_i' = c_{n+1}' b_i/a_i$, so $T' = c_{n+1}' T$.

By hypothesis, deleting any one vector from $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$ leaves a basis of V; in particular $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis.

(a) Write $\mathbf{v}_{n+1} = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n$, with unique coefficients. Suppose $a_j = 0$ for some j. Then

$$1 \cdot \mathbf{v}_{n+1} - \sum_{i \leq n,\, i \neq j} a_i \mathbf{v}_i = \mathbf{0}$$

is a linear relation, with the nonzero coefficient 1, among the vectors of the list obtained by deleting $\mathbf{v}_j$. That list is a basis, hence linearly independent, which contradicts Theorem 1.5.5. So every $a_i \neq 0$.

(b) Write $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ for $i = 1, \ldots, n+1$. Then

$$\sum_{i=1}^{n} a_i \lambda_i \mathbf{v}_i = \sum_{i=1}^{n} a_i T(\mathbf{v}_i) = T(\mathbf{v}_{n+1}) = \lambda_{n+1} \mathbf{v}_{n+1} = \sum_{i=1}^{n} \lambda_{n+1} a_i \mathbf{v}_i.$$

By uniqueness of coordinates in the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (Definition 1.6.14), $a_i \lambda_i = a_i \lambda_{n+1}$, and since $a_i \neq 0$ by (a), $\lambda_i = \lambda_{n+1}$. With $\lambda = \lambda_{n+1}$, T and $\lambda \mathrm{id}_V$ agree on the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, so $T = \lambda \mathrm{id}_V$ by Corollary 2.2.6.

(c) By (a) applied to the $\mathbf{w}_i$, $\mathbf{w}_{n+1} = b_1 \mathbf{w}_1 + \cdots + b_n \mathbf{w}_n$ with every $b_i \neq 0$.

Existence. By Theorem 2.2.1, let $T : V \to W$ be the linear map with $T(\mathbf{v}_i) = (b_i/a_i)\mathbf{w}_i$ for $i = 1, \ldots, n$. Then

$$T(\mathbf{v}_{n+1}) = \sum_{i=1}^{n} a_i \left(\frac{b_i}{a_i}\right) \mathbf{w}_i = \sum_{i=1}^{n} b_i \mathbf{w}_i = \mathbf{w}_{n+1}.$$

So $T(\mathbf{v}_i) = c_i \mathbf{w}_i$ with $c_i = b_i/a_i \neq 0$ for $i \leq n$ and $c_{n+1} = 1$.

Uniqueness up to scalars. Let $T' : V \to W$ be linear with $T'(\mathbf{v}_i) = c_i' \mathbf{w}_i$, all $c_i' \neq 0$. Then

$$\sum_{i=1}^{n} a_i c_i' \mathbf{w}_i = T'(\mathbf{v}_{n+1}) = c_{n+1}' \mathbf{w}_{n+1} = \sum_{i=1}^{n} c_{n+1}' b_i \mathbf{w}_i,$$

and comparing coordinates in the basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$ gives $c_i' = c_{n+1}' b_i/a_i$ for $i \leq n$. So $T'(\mathbf{v}_i) = c_{n+1}' T(\mathbf{v}_i)$ for $i \leq n$, and $T' = c_{n+1}' T$ by Corollary 2.2.6. Every such map is therefore a nonzero multiple of T, and any two of them are multiples of each other.

In projective geometry, (c) says that any two frames of $n+1$ points in general position are related by a projective transformation, which is unique.

Section 2.3

Solution to Exercise 2.3.7

Uses. Definition 2.3.1 (Matrix of a linear map), Notation 2.3.2 (The standard basis $\mathcal{E}$), Example 2.1.9 (Maps of polynomials)
Hint 1. Each coordinate of $T(p)$ is an evaluation, so T is linear. The columns of $[T]_{\mathcal{E}\mathcal{B}}$ are $T(1)$ and $T(x)$.
Hint 2. For $\mathring{\mathcal{B}}$, compute $T(1-x)$ and $T(x)$. You should also check that $\mathring{\mathcal{B}}$ is a basis.
Sketch. $T(1)=(1,1)$ and $T(x)=(0,1)$, so $[T]_{\mathcal{E}\mathcal{B}}=\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$. $T(1-x)=(1,0)$ and $T(x)=(0,1)$, so $[T]_{\mathcal{E}\mathring{\mathcal{B}}}=I_2$.

The coordinates of $T(p)$ are $\mathrm{ev}_0(p)$ and $\mathrm{ev}_1(p)$, which are linear in p (Example 2.1.9), so $T(cp+q)=cT(p)+T(q)$ coordinate by coordinate, and T is linear. By Notation 2.3.2, the columns of $[T]_{\mathcal{E}\mathcal{B}}$ are $T(1)=(1,1)$ and $T(x)=(0,1)$:

$$[T]_{\mathcal{E}\mathcal{B}}=\begin{pmatrix}1&0\\1&1\end{pmatrix}.$$

The list $\mathring{\mathcal{B}}$ is a basis: $c(1-x)+dx=c+(d-c)x$ equals $a+bx$ exactly when $c=a$ and $d=a+b$. Since $T(1-x)=(1,0)$ and $T(x)=(0,1)$,

$$[T]_{\mathcal{E}\mathring{\mathcal{B}}}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=I_2.$$

The same map has two different matrices. In the second basis the coordinates of p are exactly its values at 0 and 1.

Solution to Exercise 2.3.14

Uses. Definition 2.3.8 (Matrix–vector product), Proposition 2.3.11 (Entries of $A\mathbf{x}$)
Hint 1. By columns: 2 times the first column, plus 1 times the second, plus 4 times the third. Reduce modulo 5 at the end.
Hint 2. By rows: each entry pairs one row with the vector. In (b), $i\cdot i=-1$.
Sketch. (a) Both ways give $(16,12)$, which is $(1,2)$ in $\mathbb{F}_5^2$. By rows, $2+2+12=16$ and $8+0+4=12$. (b) $i(1,-i)+(i,2)=(2i,3)$, and by rows $i+i=2i$ and $-i^2+2=3$.

(a) By columns, in $\mathbb{F}_5^2$,

$$2\begin{pmatrix}1\\4\end{pmatrix}+1\begin{pmatrix}2\\0\end{pmatrix}+4\begin{pmatrix}3\\1\end{pmatrix}=\begin{pmatrix}2+2+12\\8+0+4\end{pmatrix}=\begin{pmatrix}16\\12\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}.$$

By rows, the first entry is $1\cdot 2+2\cdot 1+3\cdot 4=16=1$ and the second is $4\cdot 2+0\cdot 1+1\cdot 4=12=2$.

(b) By columns,

$$i\begin{pmatrix}1\\-i\end{pmatrix}+1\begin{pmatrix}i\\2\end{pmatrix}=\begin{pmatrix}i\\1\end{pmatrix}+\begin{pmatrix}i\\2\end{pmatrix}=\begin{pmatrix}2i\\3\end{pmatrix},$$

using $i(-i)=-i^2=1$. By rows, the entries are $1\cdot i+i\cdot 1=2i$ and $(-i)\cdot i+2\cdot 1=1+2=3$.

Solution to Exercise 2.3.18

Uses. Proposition 2.3.17 (Every linear map $F^n\to F^m$ is some T_A), Definition 2.3.8 (Matrix–vector product)
Hint 1. By Proposition 2.3.17, the columns of A are $T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n)$. No formula for the entries is needed.
Hint 2. In (c) the target is $\mathbb{R}=\mathbb{R}^1$, so A has one row. In (d), $T(1,1)=A(1,1)$ is the sum of the two columns.
Sketch. (a) Columns $\mathbf{e}_2,\mathbf{e}_3,\mathbf{e}_1$. (b) Columns $\mathbf{e}_1,\mathbf{e}_2,\mathbf{0}$. (c) $\left(\begin{smallmatrix}1&1&1\end{smallmatrix}\right)$. (d) $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$, and $T(1,1)=(1,1)+(0,1)=(1,0)$.

By Proposition 2.3.17, A is the matrix with columns $T(\mathbf{e}_1),\ldots,T(\mathbf{e}_n)$.

(a) $T(\mathbf{e}_1)=(0,1,0)$, $T(\mathbf{e}_2)=(0,0,1)$, and $T(\mathbf{e}_3)=(1,0,0)$, so $A=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$.

(b) $T(\mathbf{e}_1)=\mathbf{e}_1$, $T(\mathbf{e}_2)=\mathbf{e}_2$, and $T(\mathbf{e}_3)=\mathbf{0}$, so $A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$.

(c) $T(\mathbf{e}_1)=T(\mathbf{e}_2)=T(\mathbf{e}_3)=1$, so A is the 1×3 matrix $\left(\begin{smallmatrix}1&1&1\end{smallmatrix}\right)$.

(d) $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. By Definition 2.3.8, $T(1,1) = A(1,1) = (1,1) + (0,1) = (1,2) = (1,0)$ in $\mathbb{F}_2^2$. So T swaps $(1,0)$ and $(1,1)$ and fixes $(0,1)$.

Solution to Exercise 2.3.21

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.20 (Linear maps correspond to matrices), Theorem 2.2.1 (Linear maps are determined by a basis)
Hint 1. The columns say what $T(1)$ and $T(x)$ are, in $\mathcal{C}$-coordinates. Turn them back into polynomials.
Hint 2. Then $T(a+bx) = aT(1) + bT(x)$. For $[T]_{\mathcal{BB}}$, write $T(1)$ and $T(x)$ in the basis $1, x$.
Sketch. $T(1) = 1\cdot 1 + 0\cdot(1+x) = 1$ and $T(x) = 2\cdot 1 + 1\cdot(1+x) = 3+x$. So $T(a+bx) = (a+3b) + bx$ and $[T]_{\mathcal{BB}} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$.

$\mathcal{C}$ is a basis, since $a + bx = (a-b)\cdot 1 + b(1+x)$, and these are the only possible coefficients. By Definition 2.3.1, the first column says $T(1) = 1\cdot 1 + 0\cdot(1+x) = 1$, and the second says $T(x) = 2\cdot 1 + 1\cdot(1+x) = 3 + x$. By linearity,

$$T(a+bx) = aT(1) + bT(x) = (a+3b) + bx.$$

(By Theorem 2.3.20, this is the only operator with the given matrix.) In the basis $\mathcal{B}$, $T(1) = 1$ has coordinates $(1,0)$ and $T(x) = 3 + x$ has coordinates $(3,1)$, so

$$[T]_{\mathcal{BB}} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

Solution to Exercise 2.3.26

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 2.3.5 (Identity matrix), Definition 1.6.14 (Coordinates)
Hint 1. The j-th column of $[\mathrm{id}_V]_{\mathcal{CB}}$ is $[\mathbf{v}_j]_{\mathcal{C}}$. When is it $\mathbf{e}_j$?
Hint 2. A vector with $\mathcal{C}$-coordinates $\mathbf{e}_j$ is $0\mathbf{w}_1 + \cdots + 1\mathbf{w}_j + \cdots + 0\mathbf{w}_n$.
Sketch. The j-th column is $[\mathbf{v}_j]_{\mathcal{C}}$, and $[\mathbf{v}_j]_{\mathcal{C}} = \mathbf{e}_j$ if and only if $\mathbf{v}_j = \mathbf{w}_j$. So all columns agree with those of I_n if and only if $\mathbf{v}_j = \mathbf{w}_j$ for all j.

By Definition 2.3.1, the j-th column of $[\mathrm{id}_V]_{\mathcal{CB}}$ is $[\mathbf{v}_j]_{\mathcal{C}}$, and the j-th column of I_n is $\mathbf{e}_j$. Since a vector is determined by its coordinates (Definition 1.6.14), $[\mathbf{v}_j]_{\mathcal{C}} = \mathbf{e}_j$ holds if and only if

$$\mathbf{v}_j = 0\mathbf{w}_1 + \cdots + 1\mathbf{w}_j + \cdots + 0\mathbf{w}_n = \mathbf{w}_j.$$

So $[\mathrm{id}_V]_{\mathcal{CB}} = I_n$ if and only if $\mathbf{v}_j = \mathbf{w}_j$ for every j.

Solution to Exercise 2.3.29

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.9 (The matrix of a map computes the map), Definition 2.3.8 (Matrix–vector product)
Hint 1. First find a formula for the $\mathcal{C}$-coordinates (a,b,c) of a vector (x,y,z) by solving for a, b, c.
Hint 2. The columns of $[T]_{\mathcal{CB}}$ are the $\mathcal{C}$-coordinates of $T(1,1)$ and $T(1,-1)$. For (b), $(3,1) = 2(1,1) + (1,-1)$.
Sketch. $[(x,y,z)]_{\mathcal{C}} = (x-y, y-z, z)$. $T(1,1) = (2,0,2)$ and $T(1,-1) = (0,2,2)$ give $[T]_{\mathcal{CB}} = \begin{pmatrix} 2 & -2 \\ -2 & 0 \\ 2 & 2 \end{pmatrix}$. For $\mathbf{v} = (3,1)$, $[\mathbf{v}]_{\mathcal{B}} = (2,1)$, and both sides equal $(2,-4,6)$.

Since

$$a(1,0,0) + b(1,1,0) + c(1,1,1) = (a+b+c, b+c, c),$$

the $\mathcal{C}$-coordinates of (x,y,z) are $c = z$, $b = y - z$, and $a = x - y$:

$$[(x,y,z)]_{\mathcal{C}} = (x-y, y-z, z).$$

(a) $T(1,1) = (2,0,2)$ has $\mathcal{C}$-coordinates $(2,-2,2)$, and $T(1,-1) = (0,2,2)$ has $\mathcal{C}$-coordinates $(-2,0,2)$. So

$$[T]_{\mathcal{CB}} = \begin{pmatrix} 2 & -2 \\ -2 & 0 \\ 2 & 2 \end{pmatrix}.$$

(b) Since $(3,1) = 2(1,1) + 1(1,-1)$, $[\mathbf{v}]_{\mathcal{B}} = (2,1)$, and

$$[T]_{\mathcal{CB}}[\mathbf{v}]_{\mathcal{B}} = 2\begin{pmatrix}2\\-2\\2\end{pmatrix} + 1\begin{pmatrix}-2\\0\\2\end{pmatrix} = \begin{pmatrix}2\\-4\\6\end{pmatrix}.$$

On the other side, $T(3,1) = (4,2,6)$, whose $\mathcal{C}$-coordinates are $(4-2, 2-6, 6) = (2,-4,6)$. The two sides agree.

Solution to Exercise 2.3.30

Uses. Theorem 2.3.9 (The matrix of a map computes the map), Example 2.3.24 (The identity in two bases)
Hint 1. Check the answer: is $4(1,1) + 2(1,-1)$ equal to $(3,1)$?
Hint 2. By Theorem 2.3.9, $[\mathrm{id}]_{\mathcal{EB}}$ multiplies $\mathcal{B}$-coordinates and produces $\mathcal{E}$-coordinates. Which matrix goes the other way?
Sketch. $[\mathrm{id}]_{\mathcal{EB}}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{E}}$, so this matrix turns $\mathcal{B}$-coordinates into standard coordinates, not the reverse. The correct matrix is $[\mathrm{id}]_{\mathcal{BE}} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}$, which gives $[\mathbf{v}]_{\mathcal{B}} = (2,1)$.

The answer is wrong: $4(1,1) + 2(1,-1) = (6,2) \neq (3,1)$. By Theorem 2.3.9, $[\mathrm{id}]_{\mathcal{EB}}[\mathbf{w}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{E}}$ for every $\mathbf{w}$, so $[\mathrm{id}]_{\mathcal{EB}}$ takes $\mathcal{B}$-coordinates as input and produces standard coordinates. The computation feeds it standard coordinates. What it actually produced is the vector whose $\mathcal{B}$-coordinates are $(3,1)$, namely $3(1,1) + (1,-1) = (4,2)$.

To go from standard coordinates to $\mathcal{B}$-coordinates, use $[\mathrm{id}]_{\mathcal{BE}} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}$ (Example 2.3.24):

$$[\mathbf{v}]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{BE}}[\mathbf{v}]_{\mathcal{E}} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}.$$

Check: $2(1,1) + (1,-1) = (3,1)$. The subscripts are a guide: in $[\mathrm{id}]_{\mathcal{CB}}[\mathbf{v}]_{\mathcal{B}}$ the inner two must match.

Solution to Exercise 2.3.31

Uses. Example 2.2.3 (Building maps from a convenient basis), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A)
Hint 1. When is $(1,1), (1,-1)$ a basis of F^2? Look at Example 2.2.3.
Hint 2. Over $\mathbb{F}_2$, $-1 = 1$. Try $A = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$.
Sketch. $(1,1), (1,-1)$ is a basis only when $1 + 1 \neq 0$. Over $\mathbb{F}_2$ the two vectors coincide, and $A = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ satisfies both hypotheses without being I_2. When $1 + 1 \neq 0$ the argument is correct.

The error is the claim that $(1,1), (1,-1)$ is a basis of F^2 for every field. It is a basis when $1 + 1 \neq 0$ (Example 2.2.3), but over $\mathbb{F}_2$ we have $-1 = 1$, so the two vectors are equal and the list is linearly dependent. Then the hypothesis says only that $A(1,1) = (1,1)$, and the claim fails: over $\mathbb{F}_2$,

$$\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix},$$

so $A = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ satisfies both hypotheses and $A \neq I_2$. In general $(1,-1) = (1,1)$ exactly when $1 + 1 = 0$. When $1 + 1 \neq 0$ the argument is correct as written, by Corollary 2.2.6 and Proposition 2.3.17.

Solution to Exercise 2.3.32

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.9 (The matrix of a map computes the map), Proposition 2.3.11 (Entries of $A\mathbf{x}$), Theorem 2.3.20 (Linear maps correspond to matrices)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Expand $(x+1)^k$ for $k = 0, 1, 2, 3$. Over $\mathbb{F}_2$, reduce the binomial coefficients modulo 2.
Hint 2. $p(x+1) = p(x)$ means $T(p) = p$, that is, $A[p]_{\mathcal{B}} = [p]_{\mathcal{B}}$ with $A = [T]_{\mathcal{BB}}$. Write out the four entries of $A(a,b,c,d)$ and compare.
Sketch. The columns are the coefficients of 1, $1 + x$, $1 + 2x + x^2$, $1 + 3x + 3x^2 + x^3$. Over $\mathbb{F}_2$ the condition $A(a,b,c,d) = (a,b,c,d)$ reads $b + c + d = 0$ and $d = 0$, so $p = a + b(x + x^2)$. Over $\mathbb{Q}$ it reads $b + c + d = 0$, $2c + 3d = 0$, $3d = 0$, so p is constant.

(a) $T(1) = 1$, $T(x) = 1 + x$, $T(x^2) = 1 + 2x + x^2$, and $T(x^3) = 1 + 3x + 3x^2 + x^3$. Their coordinate vectors are the columns, so over $\mathbb{Q}$ and over $\mathbb{F}_2$ (where $2 = 0$ and $3 = 1$) respectively

$$[T]_{\mathcal{BB}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad [T]_{\mathcal{BB}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The entries are the binomial coefficients of Pascal's triangle, reduced modulo 2 in the second case.

(b) Let A be the matrix over $\mathbb{F}_2$, and $[p]_{\mathcal{B}} = (a, b, c, d)$. By Theorem 2.3.9 and since coordinates determine vectors, $T(p) = p$ if and only if $A(a, b, c, d) = (a, b, c, d)$. By Proposition 2.3.11,

$$A \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a + b + c + d \\ b + d \\ c + d \\ d \end{pmatrix},$$

so the condition is $b + c + d = 0$ and $d = 0$, that is, $d = 0$ and $c = b$ (in $\mathbb{F}_2$, $b + c = 0$ means $c = b$). So the solutions are the four polynomials $p = a + b(x + x^2)$ with $a, b \in \mathbb{F}_2$. Check: $(x + 1) + (x + 1)^2 = x + 1 + x^2 + 1 = x + x^2$ over $\mathbb{F}_2$.

(c) Over $\mathbb{Q}$ the condition $A(a, b, c, d) = (a, b, c, d)$ reads

$$a + b + c + d = a, \quad b + 2c + 3d = b, \quad c + 3d = c, \quad d = d,$$

so $3d = 0$, then $2c = 0$, then $b = 0$: only the constant polynomials satisfy $p(x + 1) = p(x)$. The difference between the two fields is the entry 2 in row 2, column 3: over $\mathbb{Q}$ it forces $c = 0$, and over $\mathbb{F}_2$ it vanishes, which lets $x + x^2$ through.

Solution to Exercise 2.3.33

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.9 (The matrix of a map computes the map), Definition 1.6.1 (Basis), Definition 1.2.17 (Evaluation)
Hint 1. (a) The columns are $T(1)$, $T(x)$, $T(x^2)$. (b) The condition says $T(p_j) = \mathbf{e}_{j+1}$: each p_j is 1 at j and 0 at the other two points. Build it from the factors x, $x - 1$, $x - 2$.
Hint 2. To prove that p_0, p_1, p_2 is a basis, show that $q - (q(0)p_0 + q(1)p_1 + q(2)p_2)$ vanishes at $0, 1, 2$, and that the only $r = a + bx + cx^2$ vanishing at $0, 1, 2$ is $r = \mathbf{0}$.
Hint 3. (d) Over $\mathbb{F}_2$, $2 = 0$. Compare the first and third coordinates of $T(p)$.
Sketch. (a) $\left(\begin{smallmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{smallmatrix}\right)$. (b) $p_0 = (x - 1)(x - 2)/2$, $p_1 = -x(x - 2)$, $p_2 = x(x - 1)/2$. Every q equals $\sum_j q(j)p_j$, because the difference vanishes at $0, 1, 2$ and $r(0) = r(1) = r(2) = 0$ forces $r = \mathbf{0}$ when $2 \neq 0$; the coefficients are unique because evaluating $\sum c_j p_j$ at k gives c_k. (c) $[T(q)]_{\mathcal{E}} = I_3[q]_{\mathring{\mathcal{B}}}$. (d) Every value $T(p) = (p(0), p(1), p(0))$ has equal first and third coordinates, but $\mathbf{e}_1$ does not.

(a) $T(1) = (1, 1, 1)$, $T(x) = (0, 1, 2)$, and $T(x^2) = (0, 1, 4)$, so

$$[T]_{\mathcal{EB}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}.$$

(b) We need $T(p_j) = \mathbf{e}_{j+1}$, that is, $p_j(k) = 1$ if $k = j$ and 0 otherwise, for $j, k \in \{0, 1, 2\}$. Let

$$p_0 = \frac{1}{2}(x - 1)(x - 2) = \frac{1}{2}(x^2 - 3x + 2),$$
$$p_1 = -x(x - 2) = -x^2 + 2x,$$
$$p_2 = \frac{1}{2}x(x - 1) = \frac{1}{2}(x^2 - x).$$

Evaluating the expanded forms: p_0 takes the values $1, 0, 0$ at $0, 1, 2$ (at 2: $(4 - 6 + 2)/2 = 0$), p_1 takes $0, 1, 0$, and p_2 takes $0, 0, 1$.

Spanning. Let $q \in F[x]_{\leq 2}$ and $r = q - (q(0)p_0 + q(1)p_1 + q(2)p_2)$. Evaluation is linear, so $r(k) = q(k) - q(k) = 0$ for $k = 0, 1, 2$. Write $r = a + bx + cx^2$. Then $r(0) = a = 0$, $r(1) = b + c = 0$, and $r(2) = 2b + 4c = 0$. Substituting $b = -c$ into the last equation gives $2c = 0$, so $c = 0$ because $2 \neq 0$, and then $b = 0$. Hence $r = \mathbf{0}$ and $q = q(0)p_0 + q(1)p_1 + q(2)p_2$.

Uniqueness. If $q = c_0p_0 + c_1p_1 + c_2p_2$, evaluating at k gives $q(k) = c_k$. So the coefficients are determined by q, and $\mathring{\mathcal{B}}$ is a basis (Definition 1.6.1). By construction, $[T]_{\mathcal{E}\mathring{\mathcal{B}}} = I_3$.

(c) By Theorem 2.3.9 and Proposition 2.3.13(c),

$$(q(0), q(1), q(2)) = [T(q)]_{\mathcal{E}} = I_3[q]_{\mathring{\mathcal{B}}} = [q]_{\mathring{\mathcal{B}}}.$$

(This is also visible in the uniqueness argument.) In this basis, the coordinates of a polynomial are its values.

(d) Over $\mathbb{F}_2$, $2 = 0$, so $T(p) = (p(0), p(1), p(0))$ for every p: the first and third coordinates of every value of T agree. If $[T]_{\mathcal{E}\mathring{\mathcal{B}}} = I_3$, then $T(p_0) = \mathbf{e}_1 = (1, 0, 0)$, whose first and third coordinates differ. So no such basis exists. Correspondingly, the first and third rows of $[T]_{\mathcal{E}\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ are equal over $\mathbb{F}_2$.

Solution to Exercise 2.3.34

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.1 (Basis), Proposition 1.2.35 (Zero products)
Hint 1. Every value of T is a multiple of $(1, 2)$: $T(x, y) = (x + y)(1, 2)$. For (a) you need $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{0}$.
Hint 2. For (b), the matrix would force $T(\mathbf{v}_1) = \mathbf{v}_1$. Then $\mathbf{v}_1$ is a value of T. Compute T on a multiple of $(1, 2)$.
Sketch. (a) $\mathcal{B} = (1, 0), (1, -1)$ and $\mathcal{C} = (1, 2), (0, 1)$: $T(1, 0) = (1, 2)$ and $T(1, -1) = \mathbf{0}$. (b) $T(\mathbf{v}_1) = \mathbf{v}_1$ forces $\mathbf{v}_1 = c(1, 2)$, but $T(c(1, 2)) = 3c(1, 2)$, so $2\mathbf{v}_1 = \mathbf{0}$. (c) $\mathcal{B} = (1, 2), (1, -1)$ gives $\begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$.

Note that $T(x, y) = (x + y)(1, 2)$, so every value of T is a multiple of $(1, 2)$, and $T(1, 2) = 3(1, 2)$.

(a) We need $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{0}$. Take $\mathcal{B} = (1, 0), (1, -1)$, a basis since $(x, y) = (x + y)(1, 0) - y(1, -1)$ with no other coefficients possible. Then $T(1, 0) = (1, 2)$ and $T(1, -1) = (0, 0)$. Take $\mathcal{C} = (1, 2), (0, 1)$, a basis since $(x, y) = x(1, 2) + (y - 2x)(0, 1)$ with no other coefficients possible. Then $[T(1, 0)]_{\mathcal{C}} = (1, 0)$ and $[T(1, -1)]_{\mathcal{C}} = (0, 0)$, so $[T]_{\mathcal{C}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(b) Suppose $\mathcal{B} = \mathbf{v}_1, \mathbf{v}_2$ is a basis with $[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. The first column says $T(\mathbf{v}_1) = \mathbf{v}_1$. So $\mathbf{v}_1$ is a value of T, hence $\mathbf{v}_1 = c(1, 2)$ for some $c \in \mathbb{R}$, and then $T(\mathbf{v}_1) = cT(1, 2) = 3c(1, 2) = 3\mathbf{v}_1$. So $3\mathbf{v}_1 = \mathbf{v}_1$, that is, $2\mathbf{v}_1 = \mathbf{0}$, and $\mathbf{v}_1 = \mathbf{0}$ by Proposition 1.2.35. But a basis does not contain $\mathbf{0}$, a contradiction.

(c) Let $\mathcal{B} = (1, 2), (1, -1)$. It is a basis: $a(1, 2) + b(1, -1) = (x, y)$ exactly when $a = (x + y)/3$ and $b = (2x - y)/3$. Since $T(1, 2) = 3(1, 2)$ and $T(1, -1) = \mathbf{0}$,

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}.$$

With two bases the nonzero entry can be scaled to 1, as in (a); with one basis the factor 3 by which T stretches the line through $(1, 2)$ cannot be changed. Such factors are the eigenvalues of Section 7.2.

Solution to Exercise 2.3.35

Uses. Definition 2.3.1 (Matrix of a linear map), Exercise 2.1.5, Non-example 2.1.14 (Complex conjugation), Theorem 2.3.20 (Linear maps correspond to matrices)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. The columns are the $\mathcal{B}$-coordinates of $M_z(1) = z$ and $M_z(i) = zi$.
Hint 2. For (c), a $\mathbb{C}$-linear map $\mathbb{C} \to \mathbb{C}$ is multiplication by some complex number (Exercise 2.1.5), and it is in particular $\mathbb{R}$-linear, so it has a matrix with respect to $\mathcal{B}$.

Sketch. (a) $z = a + bi$ and $zi = -b + ai$ give $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. (b) $\overline{1} = 1$ and $\overline{i} = -i$ give $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. (c) A $\mathbb{C}$-linear T is M_z for $z = T(1)$, so its matrix has the form in (a); conversely $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = [M_{a+bi}]_{\mathcal{BB}}$.

(a) M_z is $\mathbb{R}$-linear (it is even $\mathbb{C}$-linear, Exercise 2.1.5), so it has a matrix with respect to $\mathcal{B}$. Since $M_z(1) = a + bi$ and $M_z(i) = (a + bi)i = -b + ai$,

$$[M_z]_{\mathcal{BB}} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

For $z = \cos\theta + i\sin\theta$ this is the rotation matrix of Example 2.3.3: multiplying by a complex number of modulus 1 rotates the plane.

(b) Conjugation is $\mathbb{R}$-linear (Non-example 2.1.14), with $\overline{1} = 1$ and $\overline{i} = -i$, so its matrix is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(c) Let $T : \mathbb{C} \to \mathbb{C}$ be $\mathbb{C}$-linear. By Exercise 2.1.5, $T = M_z$ with $z = T(1)$, and T is in particular $\mathbb{R}$-linear. Writing $z = a + bi$, part (a) gives $[T]_{\mathcal{BB}} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Conversely, $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = [M_{a+bi}]_{\mathcal{BB}}$ by (a), and M_{a+bi} is $\mathbb{C}$-linear. By Theorem 2.3.20, each such matrix is the matrix of exactly one $\mathbb{R}$-linear map, so the matrices of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ are exactly the matrices of the $\mathbb{C}$-linear maps. The matrix of conjugation in (b) is not of this form, consistent with Non-example 2.1.14.

Solution to Exercise 2.3.36

Uses. Theorem 2.3.20 (Linear maps correspond to matrices), Example 2.3.6 (The zero map and the identity map), Example 2.3.23 (Maps of the plane in adapted bases), Definition 2.3.1 (Matrix of a linear map)
Hint 1. Only (b) is true. For (a), (c), and (d), look at Example 2.3.23 and at maps that send one basis to another.
Hint 2. For (e), what would the two columns say about the vectors of $\mathcal{B}$?
Sketch. (a) False: $[S_\ell]_{\mathcal{UU}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is also the standard matrix of the reflection across the x-axis. (b) True: $[T]_{\mathcal{BB}} = [0]_{\mathcal{BB}}$, so $T = 0$ by Theorem 2.3.20. (c) False: $T = 2\,\mathrm{id}$ with $\mathcal{C} = 2\mathbf{v}_1, \ldots, 2\mathbf{v}_n$. (d) False: S_ℓ in $\mathcal{U}$ and in $\mathcal{E}$. (e) False: both columns equal means $\mathbf{v}_1 = \mathbf{v}_2$.

(a) False. Let ℓ be the line $y = 2x$ and $\mathcal{U}$ as in Example 2.3.23. Then $[S_\ell]_{\mathcal{UU}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which is also $[S]_{\mathcal{EE}}$ for the reflection $S(x, y) = (x, -y)$ across the x-axis. But $S_\ell \neq S$, since $S_\ell(1, 2) = (1, 2)$ and $S(1, 2) = (1, -2)$.

(b) True. The zero map also has the zero matrix (Example 2.3.6), so $[T]_{\mathcal{BB}} = [0]_{\mathcal{BB}}$, and $T = 0$ by Theorem 2.3.20(a).

(c) False. Take $V = \mathbb{R}$ with $\mathcal{B} = 1$ and $\mathcal{C} = 2$, and $T = 2\mathrm{id}_{\mathbb{R}}$. Then $T(1) = 2 = 1 \cdot 2$, so $[T]_{\mathcal{CB}} = (1) = I_1$, but $T \neq \mathrm{id}$. In general, $[T]_{\mathcal{CB}} = I_n$ says exactly that $T(\mathbf{v}_j) = \mathbf{w}_j$ for each j, which describes the map sending one basis to the other.

(d) False. With $\mathcal{U}$ as in (a), $[S_\ell]_{\mathcal{UU}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is diagonal, but $[S_\ell]_{\mathcal{EE}} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$ is not (Example 2.3.19).

(e) False. The columns of $[\mathrm{id}]_{\mathcal{CB}}$ are $[\mathbf{v}_1]_{\mathcal{C}}$ and $[\mathbf{v}_2]_{\mathcal{C}}$. If both are $(1, 1)$, then $\mathbf{v}_1 = \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{v}_2$, so $\mathcal{B}$ contains a vector twice and is linearly dependent, which is impossible for a basis.

Solution to Exercise 2.3.37

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.14 (Coordinates), Definition 1.6.1 (Basis)
Hint 1. Changing $\mathcal{B}$ changes which vectors T is applied to, so it acts on the columns. Changing $\mathcal{C}$ changes the coordinates of every value, so it acts on the rows.
Hint 2. For (d) and (e), take $T(\mathbf{v}_j) = \sum_l a_{lj}\mathbf{w}_l$ and rewrite the right side in the new basis. In (e), $\mathbf{w}_k = (\mathbf{w}_k + c\mathbf{w}_i) - c\mathbf{w}_i$.
Sketch. (a) Columns j and k swap. (b) Column j is multiplied by c. (c) Rows i and k swap. (d) Row i is multiplied by c^{-1}, since $a\mathbf{w}_i = (ac^{-1})(c\mathbf{w}_i)$. (e) Row i becomes row i minus c times row k, since $a\mathbf{w}_k = a(\mathbf{w}_k + c\mathbf{w}_i) - ac\mathbf{w}_i$.

The modified lists are bases. Reordering a list does not change which combinations represent a vector, only the order of the terms, so (a) and (c) give bases. For (b),

$$\sum_l a_l\mathbf{v}_l = \sum_{l \neq j} a_l\mathbf{v}_l + (a_j c^{-1})(c\mathbf{v}_j),$$

so a vector has a representation in the new list exactly when it has one in the old, and the coefficients correspond one to one; the same goes for (d). For (e), write $\mathbf{w}'_k = \mathbf{w}_k + c\mathbf{w}_i$ and $\mathbf{w}'_l = \mathbf{w}_l$ for $l \neq k$. Then

$$\sum_l b_l \mathbf{w}'_l = \sum_l b_l \mathbf{w}_l + c b_k \mathbf{w}_i,$$

so the combination with coefficients (b_l) in the new list equals the combination with coefficients (a_l) in the old list, where $a_l = b_l$ for $l \neq i$ and $a_i = b_i + cb_k$. This correspondence between coefficient lists is a bijection (its inverse is $b_l = a_l$ for $l \neq i$ and $b_i = a_i - ca_k$), so every vector has exactly one representation in the new list.

(a) The columns of the new matrix are $[T(\mathbf{v}_l)]_{\mathcal{C}}$ in the new order, so columns j and k of A are swapped.

(b) The new j-th column is $\big[T\big(c\mathbf{v}_j\big)\big]_{\mathcal{C}} = c\big[T\big(\mathbf{v}_j\big)\big]_{\mathcal{C}}$ by linearity of T and of coordinates, so column j is multiplied by c; the other columns are unchanged.

(c) The coordinates of a vector with respect to the reordered basis are its old coordinates reordered in the same way. So in every column, the entries in positions i and k are swapped: rows i and k of A are swapped.

(d) For each j,

$$T\big(\mathbf{v}_j\big) = \sum_{l \neq i} a_{lj}\mathbf{w}_l + \big(a_{ij}c^{-1}\big)(c\mathbf{w}_i),$$

so in the new basis the coordinates of $T\big(\mathbf{v}_j\big)$ are the old ones with the i-th entry multiplied by c^{-1}. Row i of A is multiplied by c^{-1}, not by c.

(e) For each j, substituting $\mathbf{w}_k = \mathbf{w}'_k - c\mathbf{w}_i$,

$$T\big(\mathbf{v}_j\big) = \sum_{l \neq k} a_{lj}\mathbf{w}_l + a_{kj}\big(\mathbf{w}'_k - c\mathbf{w}_i\big) = \sum_{l \neq i,k} a_{lj}\mathbf{w}'_l + \big(a_{ij} - ca_{kj}\big)\mathbf{w}'_i + a_{kj}\mathbf{w}'_k.$$

So row i of A is replaced by row i minus c times row k, and the other rows are unchanged.

In the language of Section 3.2: changing the basis of the target performs row operations on the matrix, and changing the basis of the domain performs column operations. Note that in (d) and (e) the row operation undoes the operation on the basis, with the roles of the indices exchanged in (e): adding c times $\mathbf{w}_i$ to $\mathbf{w}_k$ subtracts c times row k from row i.

Solution to Exercise 2.3.38

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.14 (Coordinates), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. Column k of $[T]_{\mathcal{BB}}$ is $[T(\mathbf{v}_k)]_{\mathcal{B}}$. Upper triangular means that its entries below position k vanish.
Hint 2. For (ii) $\Rightarrow$ (iii), $T(V_k)$ is spanned by $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$, and each of these lies in $V_j \subseteq V_k$ for some $j \leq k$.
Sketch. (i) $\Leftrightarrow$ (ii): $a_{ik} = 0$ for all $i > k$ says that $T(\mathbf{v}_k) = \sum_{i \leq k} a_{ik}\mathbf{v}_i \in V_k$, and conversely, by uniqueness of coordinates. (ii) $\Rightarrow$ (iii): $T(V_k) = \operatorname{span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)) \subseteq V_k$. (iii) $\Rightarrow$ (ii): $\mathbf{v}_k \in V_k$.

(i) $\Leftrightarrow$ (ii). Fix k. The k-th column of $[T]_{\mathcal{BB}}$ is $[T(\mathbf{v}_k)]_{\mathcal{B}} = (a_{1k}, \ldots, a_{nk})$, so $T(\mathbf{v}_k) = \sum_{i=1}^{n} a_{ik}\mathbf{v}_i$. If $a_{ik} = 0$ for all $i > k$, then $T(\mathbf{v}_k) = \sum_{i \leq k} a_{ik}\mathbf{v}_i \in V_k$. Conversely, if $T(\mathbf{v}_k) \in V_k$, then $T(\mathbf{v}_k) = \sum_{i \leq k} c_i\mathbf{v}_i$ for some scalars, which is a representation in terms of $\mathcal{B}$ with coefficients 0 for $i > k$. By uniqueness of coordinates (Definition 1.6.14), $a_{ik} = 0$ for $i > k$. Applying this for every k gives the equivalence.

(ii) $\Rightarrow$ (iii). For $j \leq k$, $T\big(\mathbf{v}_j\big) \in V_j \subseteq V_k$. By Proposition 2.1.16, $T(V_k) = \operatorname{span}(T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k))$, and this is contained in the subspace V_k by Proposition 1.4.14.

(iii) $\Rightarrow$ (ii). $\mathbf{v}_k \in V_k$, so $T(\mathbf{v}_k) \in T(V_k) \subseteq V_k$.

For example, $[D]_{\mathcal{BB}}$ in Example 2.3.4 is upper triangular, because D maps polynomials of degree at most k to polynomials of degree at most k. Upper-triangular matrices of operators return in Section 7.5.

Solution to Exercise 2.3.39

Uses. Definition 2.3.1 (Matrix of a linear map), Example 2.1.9 (Maps of polynomials), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
Hint 1. Read off what $[D]_{\mathcal{BB}} = N$ says about $D(\mathbf{v}_1), \ldots, D(\mathbf{v}_4)$. What is $D(D(D(\mathbf{v}_4)))$?
Hint 2. Over $\mathbb{F}_2$, compute $D\big(D\big(x^k\big)\big)$; over $\mathbb{F}_3$, compute $D\big(D\big(D\big(x^k\big)\big)\big)$. Products of two, or of three, consecutive integers are divisible by 2, or by 3.
Hint 3. Over $\mathbb{Q}$ and $\mathbb{F}_5$, try to scale the monomials so that each one differentiates to the previous one.
Sketch. $[D]_{\mathcal{BB}} = N$ means $D(\mathbf{v}_1) = \mathbf{0}$ and $D(\mathbf{v}_k) = \mathbf{v}_{k-1}$, so $D^3(\mathbf{v}_4) = \mathbf{v}_1 \neq \mathbf{0}$ and $D^2(\mathbf{v}_3) = \mathbf{v}_1 \neq \mathbf{0}$. Over $\mathbb{F}_2$, $D^2 = 0$ since $k(k-1)$ is even; over $\mathbb{F}_3$, $D^3 = 0$ since $k(k-1)(k-2)$ is divisible by 3. So no basis exists there. Over $\mathbb{Q}$ and $\mathbb{F}_5$, $6 \neq 0$ and $1, x, x^2/2, x^3/6$ works.

The k-th column of N is $\mathbf{0}$ for $k = 1$ and $\mathbf{e}_{k-1}$ for $k = 2, 3, 4$. So $[D]_{\mathcal{BB}} = N$ for $\mathcal{B} = \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ means exactly that

$$D(\mathbf{v}_1) = \mathbf{0}, \quad D(\mathbf{v}_2) = \mathbf{v}_1, \quad D(\mathbf{v}_3) = \mathbf{v}_2, \quad D(\mathbf{v}_4) = \mathbf{v}_3.$$

In that case $D(D(\mathbf{v}_3)) = \mathbf{v}_1$ and $D(D(D(\mathbf{v}_4))) = \mathbf{v}_1$, and $\mathbf{v}_1 \neq \mathbf{0}$ because it belongs to a basis.

$\mathbb{F}_2$: no. For $k \geq 2$, $D\big(D\big(x^k\big)\big) = k(k-1)x^{k-2}$, and $k(k-1)$ is even, so it is 0 in $\mathbb{F}_2$. Also $D(D(1)) = D(D(x)) = \mathbf{0}$. So the linear maps $D \circ D$ and 0 agree on the basis $1, x, x^2, x^3$, and $D \circ D = 0$ (Corollary 2.2.6). This contradicts $D(D(\mathbf{v}_3)) = \mathbf{v}_1 \neq \mathbf{0}$.

$\mathbb{F}_3$: no. For $k \geq 3$, $D\big(D\big(D\big(x^k\big)\big)\big) = k(k-1)(k-2)x^{k-3}$, and one of three consecutive integers is divisible by 3, so the coefficient is 0 in $\mathbb{F}_3$. For $k \leq 2$, $D\big(D\big(D\big(x^k\big)\big)\big) = \mathbf{0}$ as well. As before, $D \circ D \circ D = 0$, contradicting $D(D(D(\mathbf{v}_4))) = \mathbf{v}_1 \neq \mathbf{0}$.

$\mathbb{Q}$ and $\mathbb{F}_5$: yes. In both fields $2 \neq 0$ and $6 \neq 0$ (in $\mathbb{F}_5$, $6 = 1$). Let

$$\mathcal{B} = 1, x, \frac{1}{2}x^2, \frac{1}{6}x^3.$$

It is a basis of $F[x]_{\leq 3}$, since

$$a + bx + cx^2 + dx^3 = a \cdot 1 + bx + 2c \cdot \frac{1}{2}x^2 + 6d \cdot \frac{1}{6}x^3,$$

and these coefficients are the only possible ones. Also $D(1) = \mathbf{0}$, $D(x) = 1$, $D\big(\frac{1}{2}x^2\big) = x$, and $D\big(\frac{1}{6}x^3\big) = \frac{1}{2}x^2$, so $[D]_{\mathcal{BB}} = N$. Over $\mathbb{F}_5$, $\frac{1}{2} = 3$ and $\frac{1}{6} = 1$, so the basis is $1, x, 3x^2, x^3$; for instance $D\big(x^3\big) = 3x^2$.

So the answer depends on the field: the property $D \circ D \circ D \neq 0$, which every operator with matrix N has, holds over $\mathbb{Q}$ and $\mathbb{F}_5$ and fails over $\mathbb{F}_2$ and $\mathbb{F}_3$.

Solution to Exercise 2.3.40

Uses. Definition 2.3.1 (Matrix of a linear map), Example 2.3.6 (The zero map and the identity map), Theorem 2.3.20 (Linear maps correspond to matrices), Exercise 2.3.37
Hint 1. One direction is quick: what is $[\lambda \mathrm{id}_V]_{\mathcal{BB}}$? For the other, fix one basis $\mathcal{B}$ with matrix A, and compare A with the matrix in a cleverly modified basis.
Hint 2. For $i \neq j$, replace $\mathbf{v}_j$ by $\mathbf{v}_j + \mathbf{v}_i$. How do the coordinates of a vector change? (See Exercise 2.3.37.)
Hint 3. In the new basis, a vector's coordinates are the old ones except that the i-th becomes (old i-th) minus (old j-th). Compare the entries in row i of the two matrices, first in a column $k \neq j$, then in column j.
Sketch. $[\lambda\ \mathrm{id}]_{\mathcal{BB}} = \lambda I_n$ for every $\mathcal{B}$. Conversely, let $A = [T]_{\mathcal{BB}}$ and let $\mathcal{B}'$ be $\mathcal{B}$ with $\mathbf{v}_j$ replaced by $\mathbf{v}_j + \mathbf{v}_i$. The $\mathcal{B}'$-coordinates of a vector are its $\mathcal{B}$-coordinates with the i-th entry replaced by (entry i) $-$ (entry j). Entry (i,k) of $[T]_{\mathcal{B}'\mathcal{B}'}$ for $k \neq j$ is $a_{ik} - a_{jk}$, so $a_{jk} = 0$. Entry (i,j) is $a_{ij} + a_{ii} - a_{jj} - a_{ji}$, so $a_{ii} = a_{jj}$. Hence $A = \lambda I_n$ and $T = \lambda\ \mathrm{id}$.

If $T = \lambda \mathrm{id}_V$. Then $T\big(\mathbf{v}_j\big) = \lambda \mathbf{v}_j$, whose coordinate vector is $\lambda \mathbf{e}_j$, so $[T]_{\mathcal{BB}} = \lambda I_n$ for every basis $\mathcal{B}$ of length n.

Conversely, suppose all the matrices $[T]_{\mathcal{BB}}$ with $\mathcal{B}$ of length n are equal. Fix a basis $\mathcal{B} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ and let $A = [T]_{\mathcal{BB}}$. If $n = 1$, then $A = (\lambda)$ for some λ, so $T(\mathbf{v}_1) = \lambda \mathbf{v}_1$, and $T = \lambda \mathrm{id}_V$ because both are linear and agree on the basis.

Let $n \geq 2$, and fix $i \neq j$. Let $\mathcal{B}'$ be $\mathcal{B}$ with $\mathbf{v}_j$ replaced by $\mathbf{v}_j + \mathbf{v}_i$; it is a basis of length n by Exercise 2.3.37(e) (with $c = 1$). Since $\mathbf{v}_j = \big(\mathbf{v}_j + \mathbf{v}_i\big) - \mathbf{v}_i$, a vector $\mathbf{y} = \sum_l c_l \mathbf{v}_l$ equals

$$\mathbf{y} = \sum_{l \neq i,j} c_l \mathbf{v}_l + (c_i - c_j)\mathbf{v}_i + c_j(\mathbf{v}_j + \mathbf{v}_i).$$

So $[\mathbf{y}]_{\mathcal{B}'}$ is $[\mathbf{y}]_{\mathcal{B}}$ with the i-th entry replaced by $c_i - c_j$.

Column $k \neq j$. The k-th vector of $\mathcal{B}'$ is $\mathbf{v}_k$, and $T(\mathbf{v}_k) = \sum_l a_{lk}\mathbf{v}_l$. So the (i,k) entry of $[T]_{\mathcal{B}'\mathcal{B}'}$ is $a_{ik} - a_{jk}$. Since this matrix equals A, $a_{ik} - a_{jk} = a_{ik}$, that is, $a_{jk} = 0$. Since $n \geq 2$, for every j there is some $i \neq j$, so $a_{jk} = 0$ whenever $j \neq k$: A is diagonal.

Column j. The j-th vector of $\mathcal{B}'$ is $\mathbf{v}_j + \mathbf{v}_i$, and $T(\mathbf{v}_j + \mathbf{v}_i) = \sum_l (a_{lj} + a_{li})\mathbf{v}_l$. So the (i,j) entry of $[T]_{\mathcal{B}'\mathcal{B}'}$ is $(a_{ij} + a_{ii}) - (a_{jj} + a_{ji}) = a_{ii} - a_{jj}$, using that A is diagonal. It must equal $a_{ij} = 0$, so $a_{ii} = a_{jj}$.

So $A = \lambda I_n$ with $\lambda = a_{11}$. Since $[\lambda \mathrm{id}_V]_{\mathcal{B}\mathcal{B}} = \lambda I_n$ as well, $T = \lambda \mathrm{id}_V$ by Theorem 2.3.20.

Solution to Exercise 2.3.41

Uses. Definition 2.3.8 (Matrix–vector product), Proposition 2.3.11 (Entries of $A\mathbf{x}$)
Hint 1. Pressing button j adds the j-th column of A to $\mathbf{s}$. Over $\mathbb{F}_2$, adding a vector twice adds nothing.
Hint 2. Switching everything off means solving $A\mathbf{x} = \mathbf{s}$. Set $x_0 = x_{n+1} = 0$, so that every equation reads $x_{i-1} + x_i + x_{i+1} = s_i$. Once x_1 is chosen, the equations for $i = 1, \ldots, n-1$ determine $x_2, \ldots, x_n$ one at a time.
Hint 3. Compare the two sequences obtained from $x_1 = 0$ and $x_1 = 1$. Their sum satisfies $d_{i+1} = d_i + d_{i-1}$ with $d_0 = 0$, $d_1 = 1$, which is periodic: $0, 1, 1, 0, 1, 1, \ldots$.
Sketch. (a) Each press of button j adds column $\mathbf{a}_j$; addition in $\mathbb{F}_2^n$ is commutative and $\mathbf{a}_j + \mathbf{a}_j = \mathbf{0}$, so the total change is $\sum_j x_j \mathbf{a}_j = A\mathbf{x}$. (b) Solving $A\mathbf{x} = \mathbf{s}$ by choosing x_1 and propagating, the last equation becomes $x_{n+1} = 0$. The two choices of x_1 give values of x_{n+1} differing by d_{n+1}, where d is $0, 1, 1$ repeated. If $3 \nmid n+1$, exactly one choice works for each $\mathbf{s}$. If $3 \mid n+1$, either both or neither work, so each solvable $\mathbf{s}$ has two solutions $\mathbf{x}$, only 2^{n-1} of the 2^n configurations are solvable, and some configuration cannot be switched off.

(a) Pressing button j switches exactly the lamps i with $|i - j| \leq 1$, that is, it adds to $\mathbf{s}$ the vector $\mathbf{a}_j$ with 1 in these positions, which is the j-th column of A. A sequence of presses adds the corresponding columns, and since addition in $\mathbb{F}_2^n$ is commutative, the order does not matter. Since $\mathbf{a}_j + \mathbf{a}_j = \mathbf{0}$ in $\mathbb{F}_2^n$, pressing button j a total of m_j times adds $\mathbf{a}_j$ if m_j is odd and nothing if it is even. So the configuration becomes $\mathbf{s} + \sum_j x_j \mathbf{a}_j = \mathbf{s} + A\mathbf{x}$ (Definition 2.3.8).

(b) By (a), $\mathbf{s}$ can be switched off exactly when $\mathbf{s} + A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x} \in \mathbb{F}_2^n$, that is, when $A\mathbf{x} = \mathbf{s}$ (in $\mathbb{F}_2$, $-\mathbf{s} = \mathbf{s}$). By Proposition 2.3.11, with the convention $x_0 = x_{n+1} = 0$, this system reads

$$x_{i-1} + x_i + x_{i+1} = s_i \quad (i = 1, \ldots, n).$$

Propagation. Given $\mathbf{s}$ and a choice of $x_1 \in \mathbb{F}_2$, set $x_0 = 0$ and define $x_2, x_3, \ldots, x_{n+1}$ by $x_{i+1} = s_i + x_{i-1} + x_i$ for $i = 1, \ldots, n$. For $i \leq n-1$ this is forced by the i-th equation, and the last step defines a number x_{n+1}. The sequence $x_1, \ldots, x_n$ is a solution if and only if the n-th equation also holds with $x_{n+1} = 0$, that is, if and only if the propagated value x_{n+1} is 0. Every solution arises this way from its own x_1. So for each $\mathbf{s}$, the solutions correspond to those $x_1 \in \{0, 1\}$ whose propagated x_{n+1} is 0.

Comparing the two choices. Let x and x' be the sequences propagated from $x_1 = 0$ and $x'_1 = 1$, and let $d_i = x_i + x'_i$. Then $d_0 = 0$, $d_1 = 1$, and $d_{i+1} = d_{i-1} + d_i$, since the s_i cancel. So $d = 0, 1, 1, 0, 1, 1, 0, \ldots$, and $d_i = 0$ exactly when 3 divides i.

If $3 \nmid n+1$, then $x_{n+1} + x'_{n+1} = d_{n+1} = 1$, so exactly one of x_{n+1}, x'_{n+1} is 0, and $A\mathbf{x} = \mathbf{s}$ has exactly one solution. Every configuration can be switched off.

If $3 \mid n+1$, then $x_{n+1} = x'_{n+1}$, so for each $\mathbf{s}$ the system has either two solutions or none. Each of the 2^n vectors $\mathbf{x}$ solves the system for exactly one $\mathbf{s}$, namely $\mathbf{s} = A\mathbf{x}$. So the solvable configurations number $2^n/2 = 2^{n-1} < 2^n$, and some configuration cannot be switched off. Finally, 3 divides $n+1$ exactly when $n = 3k+2$ for some integer $k \geq 0$.

For example, for $n = 2$ the matrix is $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, both buttons switch both lamps, and the configuration $(1, 0)$ cannot be switched off.

Solution to Exercise 2.3.42

> *Uses.* Example 2.3.28 (Translating in both directions), Theorem 2.3.9 (The matrix of a map computes the map), Proposition 2.3.11 (Entries of $A\mathbf{x}$)
> *Techniques.* Translate between maps and matrices (Technique 2.3.27)
> *Hint 1.* $\Delta(x^3) = (x+1)^3 - x^3 = 1 + 3x + 3x^2$. The other columns are as in Example 2.3.28, with an extra zero row.
> *Hint 2.* Solve $[\Delta](a,b,c,d) = (0,0,1,0)$ from the bottom up.
> *Sketch.* The new column is $(1,3,3,0)$. The system $b+c+d=0$, $2c+3d=0$, $3d=1$ gives $d=1/3$, $c=-1/2$, $b=1/6$, a free, so $p = a + x(x-1)(2x-1)/6$. Telescoping gives $\sum_{k=0}^{n-1} k^2 = n(n-1)(2n-1)/6$.

With $\mathcal{B} = 1, x, x^2, x^3$, we have $\Delta(1) = \mathbf{0}$, $\Delta(x) = 1$, $\Delta(x^2) = 1 + 2x$, and $\Delta(x^3) = 1 + 3x + 3x^2$, so

$$[\Delta]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

For $p = a + bx + cx^2 + dx^3$, the equation $\Delta(p) = x^2$ becomes $[\Delta]_{\mathcal{B}\mathcal{B}}(a,b,c,d) = (0,0,1,0)$ (Theorem 2.3.9), that is,

$$b + c + d = 0, \quad 2c + 3d = 0, \quad 3d = 1.$$

So $d = 1/3$, $c = -1/2$, $b = 1/2 - 1/3 = 1/6$, and a is arbitrary:

$$p = a + \frac{x}{6} - \frac{x^2}{2} + \frac{x^3}{3} = a + \frac{x(x-1)(2x-1)}{6}.$$

(Indeed $x(x-1)(2x-1) = 2x^3 - 3x^2 + x$.)

For $q(t) = t(t-1)(2t-1)/6$ and every number t,

$$q(t+1) - q(t) = \frac{t((t+1)(2t+1) - (t-1)(2t-1))}{6} = \frac{t \cdot 6t}{6} = t^2.$$

Summing over $t = 0, 1, \ldots, n-1$, the right side telescopes:

$$\sum_{k=0}^{n-1} k^2 = q(n) - q(0) = \frac{n(n-1)(2n-1)}{6}.$$

For $n = 3$: $0 + 1 + 4 = 5 = (3 \cdot 2 \cdot 5)/6$.

Solution to Exercise 2.3.43

> *Uses.* Definition 2.3.8 (Matrix–vector product), Proposition 2.3.11 (Entries of $A\mathbf{x}$), Proposition 2.3.13 (Linearity of the product)
> *Hint 1.* The map s is linear. Apply it to $A\mathbf{x} = \sum_j x_j \mathbf{a}_j$.
> *Hint 2.* For the converse in (a), take $\mathbf{x} = \mathbf{e}_j$. For (b), use the row-times-column rule.
> *Sketch.* (a) $s(A\mathbf{x}) = \sum_j x_j s(\mathbf{a}_j)$, which is $s(\mathbf{x})$ when every $s(\mathbf{a}_j) = 1$; conversely $s(\mathbf{a}_j) = s(A\mathbf{e}_j) = s(\mathbf{e}_j) = 1$. (b) The i-th entry of $A(1,\ldots,1)$ is the i-th row sum. (c) The entries $\sum_j a_{ij} x_j$ are sums of products of nonnegative numbers, and they add up to 1 by (a).

The map $s : F^k \to F$ is linear, since

$$s(c\mathbf{y} + \mathbf{z}) = \sum_i (cy_i + z_i) = cs(\mathbf{y}) + s(\mathbf{z}).$$

(a) Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be the columns of A. By Definition 2.3.8 and linearity of s,

$$s(A\mathbf{x}) = s(x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n) = x_1 s(\mathbf{a}_1) + \cdots + x_n s(\mathbf{a}_n).$$

If $s(\mathbf{a}_j) = 1$ for every j, this is $x_1 + \cdots + x_n = s(\mathbf{x})$. Conversely, if $s(A\mathbf{x}) = s(\mathbf{x})$ for all $\mathbf{x}$, take $\mathbf{x} = \mathbf{e}_j$: by Proposition 2.3.13(c), $s(\mathbf{a}_j) = s(A\mathbf{e}_j) = s(\mathbf{e}_j) = 1$.

(b) By Proposition 2.3.11, the i-th entry of $A(1,\dots,1)$ is $a_{i1}+\cdots+a_{in}$, the sum of the entries of row i. So $A(1,\dots,1)=c(1,\dots,1)$ if and only if every row sum is c.

(c) By Proposition 2.3.11, the i-th entry of $A\mathbf{x}$ is $\sum_j a_{ij}x_j$, a sum of products of nonnegative numbers, hence nonnegative. By (a), the entries of $A\mathbf{x}$ add up to $s(\mathbf{x})=1$. Vectors with nonnegative entries adding up to 1 are probability distributions, and matrices as in (c) describe how they evolve in a Markov chain (Section 11.1).

Solution to Exercise 2.3.44

Uses. Proposition 2.2.20 (Linear maps out of $F[x]$), Definition 1.2.13 (Polynomials), Example 2.1.9 (Maps of polynomials)
Hint 1. Column j lists the coefficients of the polynomial $T(x^j)$. For the converse, use Proposition 2.2.20.
Hint 2. For (c), write $p=\sum_j c_j x^j$ and apply T term by term. For (d), try an evaluation, viewed as a map into constants.
Sketch. (a) $T(x^j)$ is a polynomial, so its coefficients are eventually 0. Given a column-finite array, $\mathbf{w}_j=\sum_i a_{ij}x^i$ is a polynomial, and Proposition 2.2.20 gives exactly one T with $T(x^j)=\mathbf{w}_j$. (b) D: $a_{j-1,j}=j$; M: $a_{j+1,j}=1$; all else 0. (c) Expand $T(p)=\sum_j c_j T(x^j)$ and collect the coefficient of each x^i. (d) $p\mapsto p(1)$, regarded as a constant polynomial, has row 0 equal to $1,1,1,\dots$.

(a) Column j consists of the coefficients of the polynomial $T(x^j)$, and only finitely many coefficients of a polynomial are nonzero (Definition 1.2.13). Conversely, let (a_{ij}) be an array whose columns have finitely many nonzero entries. Then $\mathbf{w}_j=\sum_i a_{ij}x^i$ is a polynomial for each j, and by Proposition 2.2.20 there is exactly one $T\in \mathrm{L}(F[x])$ with $T(x^j)=\mathbf{w}_j$ for every j. Its matrix is the given array, and any operator with this matrix sends each x^j to $\mathbf{w}_j$, so it is this T.

(b) $D(x^j)=jx^{j-1}$ for $j\geq 1$ and $D(1)=\mathbf{0}$, so the matrix of D has the entries $a_{j-1,j}=j$ for $j\geq 1$, just above the diagonal, and zeros elsewhere. Multiplication by x sends x^j to x^{j+1}, so its matrix has $a_{j+1,j}=1$ just below the diagonal, and zeros elsewhere.

(c) Let $p=\sum_{j=0}^{d} c_j x^j$, with $c_j=0$ for $j>d$. By linearity,

$$T(p)=\sum_{j=0}^{d} c_j T(x^j)=\sum_{j=0}^{d} c_j \sum_i a_{ij}x^i=\sum_i\left(\sum_{j=0}^{d} a_{ij}c_j\right)x^i.$$

So the i-th coefficient of $T(p)$ is $\sum_j a_{ij}c_j$, in which the terms with $j>d$ are 0. This is the row-times-column rule of Proposition 2.3.11 for infinite matrices whose columns are finite.

(d) Let $T(p)=p(1)$, regarded as a constant polynomial. It is linear (Example 2.1.9), and $T(x^j)=1$ for every j, so row 0 of its matrix is $1,1,1,\dots$. Rows can be infinite; only the columns must be finite.

Solution to Exercise 2.3.45

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.1 (Basis), Theorem 2.3.9 (The matrix of a map computes the map), Corollary 2.2.6 (Agreeing on a spanning list)
Hint 1. (a) The columns are $T(1,0)$ and $T(0,1)$.
Hint 2. (b) Compute $T(1,i)$ and $T(1,-i)$ and express them in terms of $(1,i)$ and $(1,-i)$.
Sketch. (a) $\begin{pmatrix}1 & i\\ i & -1\end{pmatrix}$. (b) $(z,w)=a(1,i)+b(1,-i)$ has the unique solution $a=(z-iw)/2$, $b=(z+iw)/2$. $T(1,i)=\mathbf{0}$ and $T(1,-i)=(2,2i)=2(1,i)$, so $[T]_{\mathcal{BB}}=\begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}$. So T sends $(1,-i)$ to a multiple of $(1,i)$ and $(1,i)$ to $\mathbf{0}$, hence $T\circ T=0$.

(a) $T(1,0)=(1,i)$ and $T(0,1)=(i,-1)$, so $[T]_{\mathcal{EE}}=\begin{pmatrix}1 & i\\ i & -1\end{pmatrix}$.

(b) The equation $(z,w)=a(1,i)+b(1,-i)$ means $a+b=z$ and $i(a-b)=w$, that is, $a-b=-iw$. Adding and subtracting, $a=(z-iw)/2$ and $b=(z+iw)/2$, and these values do satisfy both equations. So every vector has exactly one representation, and $\mathcal{B}$ is a basis. Now

$$T(1,i)=\left(1+i^2,i-i\right)=(0,0),\quad T(1,-i)=\left(1-i^2,i+i\right)=(2,2i)=2(1,i),$$

so

$$[T]_{\mathcal{BB}}=\begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}.$$

So T kills $(1,i)$ and sends $(1,-i)$ into $\operatorname{span}((1,i))$. Hence $T\big(T(\mathbf{v}_j)\big) = \mathbf{0}$ for both basis vectors, and $T \circ T = 0$ by Corollary 2.2.6. Directly,

$$T(T(z,w)) = (z + iw + i(iz - w), i(z + iw) - (iz - w)) = (0,0).$$

The standard matrix hides this; the basis $\mathcal{B}$ shows it.

Solution to Exercise 2.3.46

Uses. Example 2.1.11 (Rotations, projections, and reflections), Exercise 2.1.43, Definition 2.3.1 (Matrix of a linear map), Example 2.3.3 (Matrices with respect to standard bases)
Hint 1. (a) Compare $a\mathbf{u}_1 + b\mathbf{u}_2$ with the formula for $R_\varphi(a,b)$. For the basis property, R_φ is undone by $R_{-\varphi}$ (Exercise 2.1.43).
Hint 2. The formula for S_α shows that S_α sends the unit vector at angle β to the unit vector at angle $2\alpha - \beta$, and R_θ adds θ to the angle. Write down the angles of the images of $\mathbf{u}_1$ (angle φ) and $\mathbf{u}_2$ (angle $\varphi + \pi/2$), and use (a).
Sketch. (a) Expanding gives $a\mathbf{u}_1 + b\mathbf{u}_2 = R_\varphi(a,b)$, and since $R_{-\varphi}R_\varphi = \mathrm{id} = R_\varphi R_{-\varphi}$, $(a,b) = R_{-\varphi}(\mathbf{y})$ is the unique coefficient vector for $\mathbf{y}$; $R_{-\varphi}$ subtracts φ from the angle. (b) $R_\theta(\mathbf{u}_1)$ and $R_\theta(\mathbf{u}_2)$ have angles $\varphi + \theta$ and $\varphi + \pi/2 + \theta$, hence coordinates $(\cos\theta, \sin\theta)$ and $(-\sin\theta, \cos\theta)$. (c) $S_\alpha(\mathbf{u}_1)$ and $S_\alpha(\mathbf{u}_2)$ have angles $2\alpha - \varphi$ and $2\alpha - \varphi - \pi/2$, hence coordinates $(\cos 2\gamma, \sin 2\gamma)$ and $(\sin 2\gamma, -\cos 2\gamma)$ with $\gamma = \alpha - \varphi$.

(a) By the formula for R_φ in Example 2.1.11,

$$a\mathbf{u}_1 + b\mathbf{u}_2 = (a\cos\varphi - b\sin\varphi, a\sin\varphi + b\cos\varphi) = R_\varphi(a,b).$$

By Exercise 2.1.43, $R_\varphi R_{-\varphi} = R_{-\varphi}R_\varphi = R_0 = \mathrm{id}$. So for $\mathbf{y} \in \mathbb{R}^2$, the coefficients $(a,b) = R_{-\varphi}(\mathbf{y})$ work, since R_φ undoes $R_{-\varphi}$. They are the only ones: if $R_\varphi(a,b) = \mathbf{y}$, then applying $R_{-\varphi}$ gives $(a,b) = R_{-\varphi}(\mathbf{y})$. Hence $\mathcal{U}$ is a basis and $[\mathbf{y}]_\mathcal{U} = R_{-\varphi}(\mathbf{y})$. Rotation by $-\varphi$ subtracts φ from the polar angle (Example 2.1.11), so

$$[(r\cos\beta, r\sin\beta)]_\mathcal{U} = (r\cos(\beta - \varphi), r\sin(\beta - \varphi)).$$

(b) The vector $\mathbf{u}_1$ has angle φ and $\mathbf{u}_2 = (\cos(\varphi + \pi/2), \sin(\varphi + \pi/2))$ has angle $\varphi + \pi/2$, both of length 1. Rotation adds θ to the angle, so by (a),

$$[R_\theta(\mathbf{u}_1)]_\mathcal{U} = (\cos\theta, \sin\theta), \quad [R_\theta(\mathbf{u}_2)]_\mathcal{U} = (\cos(\theta + \pi/2), \sin(\theta + \pi/2)) = (-\sin\theta, \cos\theta).$$

These are the columns of $[R_\theta]_{\mathcal{E}\mathcal{E}}$ (Example 2.3.3).

(c) From the formula in Example 2.1.11 and the angle subtraction formulas,

$$S_\alpha(\cos\beta, \sin\beta) = (\cos\beta\cos 2\alpha + \sin\beta\sin 2\alpha, \cos\beta\sin 2\alpha - \sin\beta\cos 2\alpha) = (\cos(2\alpha - \beta), \sin(2\alpha - \beta)).$$

So $S_\alpha(\mathbf{u}_1)$ has angle $2\alpha - \varphi$ and $S_\alpha(\mathbf{u}_2)$ has angle $2\alpha - \varphi - \pi/2$. With $\gamma = \alpha - \varphi$, part (a) gives the coordinates

$$[S_\alpha(\mathbf{u}_1)]_\mathcal{U} = (\cos 2\gamma, \sin 2\gamma), \quad [S_\alpha(\mathbf{u}_2)]_\mathcal{U} = (\cos(2\gamma - \pi/2), \sin(2\gamma - \pi/2)) = (\sin 2\gamma, -\cos 2\gamma).$$

By the formula for S_γ, these are $S_\gamma(\mathbf{e}_1)$ and $S_\gamma(\mathbf{e}_2)$, the columns of $\big[S_\gamma\big]_{\mathcal{E}\mathcal{E}}$.

In words: in coordinates rotated by φ, a rotation looks the same, and the reflection across the line at angle α looks like the reflection across the line at angle $\alpha - \varphi$, the angle of the mirror measured from $\mathbf{u}_1$ instead of from $\mathbf{e}_1$. Taking $\varphi = \alpha$ puts $\mathbf{u}_1$ on the mirror, and then $[S_\alpha]_{\mathcal{U}\mathcal{U}} = [S_0]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, as in Example 2.3.23.

Solution to Exercise 2.3.47

Uses. Proposition 2.3.11 (Entries of $A\mathbf{x}$), the triangle inequality $|z + w| \le |z| + |w|$ in $\mathbb{C}$, Section B.1
Hint 1. Suppose $A\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \neq \mathbf{0}$. Look at the equation coming from a cleverly chosen row.
Hint 2. Choose i with $|x_i|$ as large as possible. The i-th entry of $A\mathbf{x}$ gives $a_{ii}x_i = -\sum_{j \neq i} a_{ij}x_j$. Estimate both sides.
Sketch. If $\mathbf{x} \neq \mathbf{0}$, pick i with $|x_i|$ maximal, so $|x_i| > 0$. Row i of $A\mathbf{x} = \mathbf{0}$ and the triangle inequality give $|a_{ii}||x_i| \le \sum_{j \neq i} |a_{ij}||x_j|$. Bounding each $|x_j|$ by $|x_i|$ makes the right side smaller than $|a_{ii}||x_i|$, a contradiction.

Suppose $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x} \neq \mathbf{0}$. Choose i with $|x_i| \ge |x_j|$ for all j; then $|x_i| > 0$. By Proposition 2.3.11, the i-th entry of $A\mathbf{x}$ is $\sum_j a_{ij}x_j = 0$, so $a_{ii}x_i = -\sum_{j \neq i} a_{ij}x_j$. Taking absolute values and using the triangle inequality,

$$|a_{ii}||x_i| = \left|\sum_{j\neq i} a_{ij}x_j\right| \leq \sum_{j\neq i}|a_{ij}||x_j| \leq \left(\sum_{j\neq i}|a_{ij}|\right)|x_i| < |a_{ii}||x_i|,$$

where the last step uses the hypothesis and $|x_i| > 0$. This is a contradiction, so $\mathbf{x} = \mathbf{0}$.

The proof chose an extremal coordinate, the largest one in absolute value, so that a single row of A could be played against all the others. In the language of Section 2.6 the conclusion says that T_A is injective, and since A is square it is then invertible (Corollary 3.5.25).

Solution to Exercise 2.3.48

Uses. Definition 2.3.1 (Matrix of a linear map), Definition 1.6.1 (Basis), Definition 1.2.16 (Degree), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Corollary 2.2.6 (Agreeing on a spanning list), binomial coefficients: $\binom{p}{k}$ is divisible by p for $0 < k < p$

Hint 1. (a) $\deg b_k = k$, which gives the basis property. For the matrix, prove the polynomial identity $b_k(x+1) = b_k + b_{k-1}$ by pulling out the common factor $x(x-1)\cdots(x-k+2)$.

Hint 2. (b), "if": for $k \leq n < p$, the integer $k!$ is not divisible by p, so it is invertible in $\mathbb{F}_p$ and (a) goes through. "Only if": what does $[T]_{\mathcal{B}\mathcal{B}} = J$ say about $T(\mathbf{v}_k)$? Compute $T^r(\mathbf{v}_s)$ by induction on r.

Hint 3. You should find $T^r(\mathbf{v}_s) = \sum_k \binom{r}{k}\mathbf{v}_{s-k}$. Take $r = s = p$, and compare with $T^p(f) = f(x+p)$.

Sketch. (a) The b_k have degrees $0, 1, \ldots, n$, so they form a basis. Pulling out the common factor $x(x-1)\cdots(x-k+2)$ shows $b_k(x+1) - b_k = b_{k-1}$, so $T(b_k) = b_k + b_{k-1}$ and $T(b_0) = b_0$: the matrix is J. (b) If $n < p$, the same construction works in $\mathbb{F}_p$. If $n \geq p$ and $[T]_{\mathcal{B}\mathcal{B}} = J$ with $\mathcal{B} = \mathbf{v}_0, \ldots, \mathbf{v}_n$, then $T(\mathbf{v}_k) = \mathbf{v}_k + \mathbf{v}_{k-1}$, and by induction $T^p(\mathbf{v}_p) = \sum_k \binom{p}{k}\mathbf{v}_{p-k} = \mathbf{v}_p + \mathbf{v}_0$ in characteristic p. But $T^p(f) = f(x+p) = f$, so $\mathbf{v}_0 = \mathbf{0}$, a contradiction.

Throughout, index the basis vectors from 0, so that column k of J (for $k = 0, \ldots, n$) is $\mathbf{e}_k$ for $k = 0$ and $\mathbf{e}_{k-1} + \mathbf{e}_k$ for $k \geq 1$, where $\mathbf{e}_0, \ldots, \mathbf{e}_n$ is the standard basis of F^{n+1}. So $[T]_{\mathcal{B}\mathcal{B}} = J$ for $\mathcal{B} = \mathbf{v}_0, \ldots, \mathbf{v}_n$ means

$$T(\mathbf{v}_0) = \mathbf{v}_0, \quad T(\mathbf{v}_k) = \mathbf{v}_k + \mathbf{v}_{k-1} \quad (1 \leq k \leq n). \tag{C.1.5}$$

Polynomials of degrees $0, 1, \ldots, n$. Let $q_0, \ldots, q_n \in F[x]_{\leq n}$ with $\deg q_k = k$. They span $F[x]_{\leq n}$: by induction on k, every x^k is in their span, since x^k minus a suitable multiple of q_k has degree less than k. They are linearly independent: in a relation $\sum_k c_k q_k = \mathbf{0}$ with some $c_k \neq 0$, let k be the largest index with $c_k \neq 0$; then the coefficient of x^k on the left is c_k times the leading coefficient of q_k, which is nonzero. So $q_0, \ldots, q_n$ is a basis (Definition 1.6.1).

(a) Over $\mathbb{Q}$, b_k is a product of k factors of degree 1 divided by $k!$, so $\deg b_k = k$, and $\mathcal{B}$ is a basis by the paragraph above. For $k \geq 1$, write $P = x(x-1)\cdots(x-k+2)$ (with $P = 1$ when $k = 1$), so that $b_k = P(x-k+1)/k!$ and $b_{k-1} = P/(k-1)!$. Substituting $x+1$ for x in b_k gives $(x+1)P/k!$. Hence

$$b_k(x+1) - b_k = \frac{P((x+1)-(x-k+1))}{k!} = \frac{kP}{k!} = \frac{P}{(k-1)!} = b_{k-1}.$$

So $T(b_k) = b_k + b_{k-1}$ for $k \geq 1$, and $T(b_0) = b_0$. This is (C.1.5), so $[T]_{\mathcal{B}\mathcal{B}} = J$.

(b) *If $n < p$.* For $k \leq n < p$, the integer $k!$ is a product of integers between 1 and $p-1$, none divisible by the prime p, so $k! \neq 0$ in $\mathbb{F}_p$ and it is invertible (Proposition 1.1.24). So $b_0, \ldots, b_n$ are defined in $\mathbb{F}_p[x]$, and the computation in (a) uses only the field operations and division by $k!$. It goes through unchanged, and gives a basis with $[T]_{\mathcal{B}\mathcal{B}} = J$.

If $n \geq p$. Suppose $\mathcal{B} = \mathbf{v}_0, \ldots, \mathbf{v}_n$ is a basis satisfying (C.1.5). Set $\mathbf{v}_k = \mathbf{0}$ for $k < 0$, so that $T(\mathbf{v}_k) = \mathbf{v}_k + \mathbf{v}_{k-1}$ for all $k \leq n$ (for $k < 0$ both sides are $\mathbf{0}$). We claim that for all $r \geq 0$ and $0 \leq s \leq n$,

$$T^r(\mathbf{v}_s) = \sum_{k=0}^{r}\binom{r}{k}\mathbf{v}_{s-k}.$$

For $r = 0$ both sides are $\mathbf{v}_s$. If it holds for r, then by linearity

$$T^{r+1}(\mathbf{v}_s) = \sum_{k=0}^{r} \binom{r}{k} (\mathbf{v}_{s-k} + \mathbf{v}_{s-k-1}) = \sum_{k=0}^{r+1} \left(\binom{r}{k} + \binom{r}{k-1} \right) \mathbf{v}_{s-k} = \sum_{k=0}^{r+1} \binom{r+1}{k} \mathbf{v}_{s-k},$$

with the conventions $\binom{r}{-1} = \binom{r}{r+1} = 0$ and Pascal's rule. Now take $r = s = p$, which is allowed since $p \leq n$. For $0 < k < p$, $\binom{p}{k}$ is divisible by p, so it is 0 in $\mathbb{F}_p$, and

$$T^p(\mathbf{v}_p) = \mathbf{v}_p + \mathbf{v}_0.$$

On the other hand, substituting $x + 1$ a total of p times gives $T^p(f) = f(x + p \cdot 1) = f(x)$, since $p \cdot 1 = 0$ in $\mathbb{F}_p$. So $T^p(\mathbf{v}_p) = \mathbf{v}_p$, and $\mathbf{v}_0 = \mathbf{0}$, which is impossible for a vector of a basis. Hence no such basis exists.

For example, over $\mathbb{F}_2$ with $n = 2$: $T^2 = \mathrm{id}$, while a basis as above would give $T^2(\mathbf{v}_2) = \mathbf{v}_2 + 2\mathbf{v}_1 + \mathbf{v}_0 = \mathbf{v}_2 + \mathbf{v}_0$.

Section 2.4

Solution to Exercise 2.4.6

Uses. Definition 2.4.1 (Matrix product), Definition 2.3.8 (Matrix–vector product)
Hint 1. Compute column by column: column j of AB is A times column j of B. Or use the entry formula: row i of A against column j of B.
Hint 2. Work with integers and reduce modulo 5 at the end. In (b), use $i^2 = -1$.
Sketch. (a) $AB = \begin{pmatrix} 2 & 7 \\ 10 & 10 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$ and $BA = \begin{pmatrix} 5 & 5 & 4 \\ 9 & 3 & 12 \\ 4 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 4 & 3 & 2 \\ 4 & 3 & 4 \end{pmatrix}$ in $\mathbb{F}_5$. A^2 needs $3 = 2$ and B^2 needs $2 = 3$, so neither is defined. (b) $\begin{pmatrix} 2i & 2 \\ 2 & -2i \end{pmatrix}$ and $\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$.

(a) By columns, $A(2, 0, 1) = 2(1, 3) + (0, 4) = (2, 10)$ and $A(1, 3, 1) = (1, 3) + 3(2, 1) + (0, 4) = (7, 10)$. Reducing modulo 5,

$$AB = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}.$$

The columns of BA are $B(1, 3) = (2 + 3, 0 + 9, 1 + 3)$, $B(2, 1) = (4 + 1, 0 + 3, 2 + 1)$, and $B(0, 4) = (0 + 4, 0 + 12, 0 + 4)$, so

$$BA = \begin{pmatrix} 5 & 5 & 4 \\ 9 & 3 & 12 \\ 4 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 4 & 3 & 2 \\ 4 & 3 & 4 \end{pmatrix} \quad \text{in } \mathbb{F}_5.$$

A product XY needs as many columns in X as rows in Y. A is 2×3, so AA is not defined, and B is 3×2, so BB is not defined either.

(b) By the entry formula, the first product has entries $1 \cdot i + i \cdot 1 = 2i$, $1 \cdot 1 + i(-i) = 2$, $0 \cdot i + 2 \cdot 1 = 2$, and $0 \cdot 1 + 2(-i) = -2i$, and the second has entries i, $i^2 + 2 = 1$, 1, and $i - 2i = -i$:

$$\begin{pmatrix} 1 & i \\ 0 & 2 \end{pmatrix} \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} = \begin{pmatrix} 2i & 2 \\ 2 & -2i \end{pmatrix},$$

$$\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} 1 & i \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}.$$

The two products differ.

Solution to Exercise 2.4.11

Uses. Proposition 2.4.8 (Four views of the product), Proposition 2.3.13 (Linearity of the product)
Hint 1. Use the column view for statements about columns of AB, and the row view for rows.
Hint 2. Column j of AB is $AB_{\cdot,j}$, and $\mathbf{x} \mapsto A\mathbf{x}$ is linear. For (c), you need a nonzero column $\mathbf{b}$ with $A\mathbf{b} = \mathbf{0}$.

Sketch. (a) $(AB)_{\cdot,j} = A\mathbf{0} = \mathbf{0}$, and $(AB)_{i,\cdot} = A_{i,\cdot}B$ is a combination of rows of B with all coefficients 0. (b) Apply the linear map $\mathbf{x} \mapsto A\mathbf{x}$ to the relation. (c) $A = \begin{pmatrix}1 & 1\end{pmatrix}$ and $B = \begin{pmatrix}1\\-1\end{pmatrix}$ give $AB = (0)$.

(a) By the column view (Proposition 2.4.8(b)), $(AB)_{\cdot,j} = AB_{\cdot,j} = A\mathbf{0} = \mathbf{0}$. By the row view (Proposition 2.4.8(c)), $(AB)_{i,\cdot} = a_{i1}B_{1,\cdot} + \cdots + a_{im}B_{m,\cdot}$, and every coefficient a_{ik} is 0, so this row is zero.

(b) By the column view and the linearity of $\mathbf{x} \mapsto A\mathbf{x}$ (Proposition 2.3.13),

$$\begin{aligned}(AB)_{\cdot,3} &= A\big(B_{\cdot,1} + 2B_{\cdot,2}\big) = AB_{\cdot,1} + 2AB_{\cdot,2}\\ &= (AB)_{\cdot,1} + 2(AB)_{\cdot,2}.\end{aligned}$$

In words: every linear relation among the columns of B also holds among the columns of AB.

(c) We need a nonzero column $\mathbf{b}$ with $A\mathbf{b} = \mathbf{0}$. Take $A = \begin{pmatrix}1 & 1\end{pmatrix} \in \mathbb{R}^{1\times 2}$ and $B = \begin{pmatrix}1\\-1\end{pmatrix} \in \mathbb{R}^{2\times 1}$. Then $AB = (1 - 1) = (0)$, a zero column, while B has none. A square example: $\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\-1 & -2\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}$.

Solution to Exercise 2.4.17

Uses. Proposition 2.4.14 (Algebra of matrices), Definition 1.2.7 (Matrices)
Hint 1. Expand $(A + B)(A + B)$ with both distributive laws. Do not reorder any product.
Hint 2. Subtract $A^2 + AB + B^2$ from both expressions and compare what is left.
Sketch. Distributivity gives $(A + B)^2 = A^2 + AB + BA + B^2$. This equals $A^2 + AB + AB + B^2$ exactly when $BA = AB$, by cancelling in the vector space $F^{n\times n}$.

By the two distributive laws of Proposition 2.4.14,

$$\begin{aligned}(A + B)(A + B) &= A(A + B) + B(A + B)\\ &= A^2 + AB + BA + B^2.\end{aligned}$$

So $(A + B)^2 = A^2 + 2AB + B^2$ holds if and only if $A^2 + AB + BA + B^2 = A^2 + AB + AB + B^2$. Adding $-\big(A^2 + AB + B^2\big)$ to both sides in the vector space $F^{n\times n}$, this holds if and only if $BA = AB$. The familiar identity from high school algebra secretly uses commutativity.

Solution to Exercise 2.4.20

Uses. Proposition 2.4.8 (Four views of the product), Definition 2.3.8 (Matrix–vector product)
Hint 1. By the column view, $AB = 0$ means $A\mathbf{b} = \mathbf{0}$ for each column $\mathbf{b}$ of B. Start from an A that kills some nonzero vector.
Hint 2. Try $A = \begin{pmatrix}1 & 2\\1 & 2\end{pmatrix}$. Which vectors does it kill? Build B from such columns, and then compute BA.
Sketch. $A = \begin{pmatrix}1 & 2\\1 & 2\end{pmatrix}$ kills $(2, -1)$, so $B = \begin{pmatrix}2 & 2\\-1 & -1\end{pmatrix}$ gives $AB = 0$, while $BA = \begin{pmatrix}4 & 8\\-2 & -4\end{pmatrix} \neq 0$.

By the column view (Proposition 2.4.8(b)), $AB = 0$ means that A sends each column of B to $\mathbf{0}$. Let $A = \begin{pmatrix}1 & 2\\1 & 2\end{pmatrix}$. By Definition 2.3.8, $A(x, y) = (x + 2y, x + 2y)$, so $A(2, -1) = \mathbf{0}$. Let $B = \begin{pmatrix}2 & 2\\-1 & -1\end{pmatrix}$, whose columns are both $(2, -1)$. Then $AB = 0$, and

$$\begin{aligned}BA &= \begin{pmatrix}2 + 2 & 4 + 4\\-1 - 1 & -2 - 2\end{pmatrix}\\ &= \begin{pmatrix}4 & 8\\-2 & -4\end{pmatrix} \neq 0.\end{aligned}$$

No entry of A or B is 0.

Solution to Exercise 2.4.25

Uses. Proposition 2.4.23 (Transpose of a product), Proposition 2.4.22 (Transpose is a linear involution)
Hint 1. Take the transpose of each expression with Proposition 2.4.23, and use $\big(A^{\mathsf{T}}\big)^{\mathsf{T}} = A$.
Hint 2. For symmetric A and B, $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}} = BA$.
Sketch. (a) $\big(A^{\mathsf{T}}A\big)^{\mathsf{T}} = A^{\mathsf{T}}\big(A^{\mathsf{T}}\big)^{\mathsf{T}} = A^{\mathsf{T}}A$, and likewise for AA^{T}. (b) $(AB)^{\mathsf{T}} = BA$, so transposing swaps AB and BA: $AB + BA$ is fixed, $AB - BA$ changes sign, and AB is fixed exactly when $AB = BA$.

(a) By Proposition 2.4.23 and Proposition 2.4.22, $\left(A^{\mathsf{T}}A\right)^{\mathsf{T}} = A^{\mathsf{T}}\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A^{\mathsf{T}}A$ and $\left(AA^{\mathsf{T}}\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^{\mathsf{T}}A^{\mathsf{T}} = AA^{\mathsf{T}}$.

(b) Since $A^{\mathsf{T}} = A$ and $B^{\mathsf{T}} = B$, Proposition 2.4.23 gives $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}} = BA$, and likewise $(BA)^{\mathsf{T}} = AB$. By linearity of transposition (Proposition 2.4.22),

$$(AB + BA)^{\mathsf{T}} = BA + AB,$$
$$(AB - BA)^{\mathsf{T}} = BA - AB = -(AB - BA).$$

Finally AB is symmetric if and only if $AB = (AB)^{\mathsf{T}} = BA$.

Solution to Exercise 2.4.32

Uses. Proposition 2.4.28 (Cyclicity of the trace), Lemma 2.4.24 (Matrix units)
Hint 1. Proposition 2.4.28 says $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, an equality of two numbers. Does it say $XY = YX$?
Hint 2. For a counterexample, try $A = \mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}$, $B = \mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}$, and $C = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ in $F^{2\times 2}$.
Sketch. The step $\operatorname{tr}((AB)C) = \operatorname{tr}((BA)C)$ replaces AB by BA inside a product, which needs $AB = BA$; cyclicity only gives $\operatorname{tr}(XY) = \operatorname{tr}(YX)$. With the matrix units above, $ABC = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ has trace 1 and $BAC = 0$.

The error is in the step $\operatorname{tr}((AB)C) = \operatorname{tr}((BA)C)$. Proposition 2.4.28 says that $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ for two matrices X and Y. With $X = AB$ and $Y = C$ it gives $\operatorname{tr}(ABC) = \operatorname{tr}(CAB)$, a rotation. It does not say that AB and BA are equal, or that they can be exchanged inside a longer product; the argument used $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ as if it were $AB = BA$.

The conclusion is false. In $F^{2\times 2}$, let $A = \mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}$, $B = \mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}$, and $C = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$. By Lemma 2.4.24(c), $AB = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ and $BA = \mathbf{e}_2\mathbf{e}_2^{\mathsf{T}}$, so

$$ABC = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}} = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}},$$
$$BAC = \mathbf{e}_2\mathbf{e}_2^{\mathsf{T}}\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}} = 0.$$

Hence $\operatorname{tr}(ABC) = 1 \neq 0 = \operatorname{tr}(BAC)$. In matrices: $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Solution to Exercise 2.4.37

Uses. Theorem 2.4.34 (Block multiplication), Proposition 2.4.14 (Algebra of matrices)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Multiply with Theorem 2.4.34, using sizes p, q for rows and columns of both factors.
Hint 2. Induct on k: $M^{k+1} = M^k M$.
Sketch. $\begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}\begin{pmatrix} I & Z \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & Z+Y \\ 0 & I \end{pmatrix}$ by block multiplication. Induction gives $M^k = \begin{pmatrix} I & kX \\ 0 & I \end{pmatrix}$, and $Y = X$, $Z = -X$ gives I.

Let $Y, Z \in F^{p\times q}$, and partition both factors below with sizes p, q. By Theorem 2.4.34 and Proposition 2.4.14(b), the blocks of the product are $I_pI_p + Y0 = I_p$, $I_pZ + YI_q = Z + Y$, $0I_p + I_q0 = 0$, and $0Z + I_qI_q = I_q$:

$$\begin{pmatrix} I_p & Y \\ 0 & I_q \end{pmatrix}\begin{pmatrix} I_p & Z \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} I_p & Z+Y \\ 0 & I_q \end{pmatrix}.$$

For $k = 1$ the claim is $M = M$. If $M^k = \begin{pmatrix} I_p & kX \\ 0 & I_q \end{pmatrix}$, then by the formula $M^{k+1} = M^kM = \begin{pmatrix} I_p & X+kX \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} I_p & (k+1)X \\ 0 & I_q \end{pmatrix}$. Taking $Y = X$ and $Z = -X$ gives $\begin{pmatrix} I_p & X \\ 0 & I_q \end{pmatrix}\begin{pmatrix} I_p & -X \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix} = I_{p+q}$. In the language of maps, M adds X times the last q coordinates to the first p coordinates, and the second factor subtracts it again.

Solution to Exercise 2.4.41

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.28 (Cyclicity of the trace), Example 2.1.9 (Maps of polynomials)
Hint 1. The columns of $[D]_{\mathcal{B}_2\mathcal{B}_3}$ are the coordinates of $D(1), D(x), D\left(x^2\right), D\left(x^3\right)$ in $\mathcal{B}_2$; the columns of $[M]_{\mathcal{B}_3\mathcal{B}_2}$ are those of x, x^2, x^3 in $\mathcal{B}_3$.

Hint 2. By Theorem 2.4.2, $[DM] = [D][M]$ and $[MD] = [M][D]$. Compute column by column. Then compute $\left(x \cdot x^k\right)'$ and $x\left(x^k\right)'$ directly.
Sketch. $[D] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ and $[M]$ has columns $\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$. Then $[DM] = \operatorname{diag}(1,2,3)$ and $[MD] = \operatorname{diag}(0,1,2,3)$, matching $\left(x^{k+1}\right)' = (k+1)x^k$ and $x\left(x^k\right)' = kx^k$. Both traces are 6, as Proposition 2.4.28 predicts for a 3×4 and a 4×3 matrix.

(a) $D(1) = \mathbf{0}$, $D(x) = 1$, $D\left(x^2\right) = 2x$, and $D\left(x^3\right) = 3x^2$, and $M(1) = x$, $M(x) = x^2$, $M\left(x^2\right) = x^3$, so

$$[D]_{\mathcal{B}_2\mathcal{B}_3} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},$$

$$[M]_{\mathcal{B}_3\mathcal{B}_2} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) By Theorem 2.4.2, $[DM]_{\mathcal{B}_2\mathcal{B}_2} = [D]_{\mathcal{B}_2\mathcal{B}_3}[M]_{\mathcal{B}_3\mathcal{B}_2}$. Its columns are $[D]$ times the columns $\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ of $[M]$, which are columns $2, 3, 4$ of $[D]$:

$$[DM]_{\mathcal{B}_2\mathcal{B}_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Similarly $[MD]_{\mathcal{B}_3\mathcal{B}_3} = [M]_{\mathcal{B}_3\mathcal{B}_2}[D]_{\mathcal{B}_2\mathcal{B}_3}$ has columns $[M]\mathbf{0} = \mathbf{0}$, $[M](1,0,0) = \mathbf{e}_2$, $[M](0,2,0) = 2\mathbf{e}_3$, and $[M](0,0,3) = 3\mathbf{e}_4$:

$$[MD]_{\mathcal{B}_3\mathcal{B}_3} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Directly, $DM\left(x^k\right) = \left(x^{k+1}\right)' = (k+1)x^k$ for $k = 0, 1, 2$, and $MD\left(x^k\right) = x \cdot kx^{k-1} = kx^k$ for $k = 0, 1, 2, 3$, which are the diagonal entries found above.

(c) Both traces are $1 + 2 + 3 = 6 = 0 + 1 + 2 + 3$, although the matrices have different sizes. This is Proposition 2.4.28 for the 3×4 matrix $[D]$ and the 4×3 matrix $[M]$. Note also that $[DM] - [MD]$ is not defined, since the sizes differ; on all of $F[x]$, $DM - MD = \mathrm{id}$ (Exercise 2.4.51).

Solution to Exercise 2.4.42

Uses. Example 2.4.10 (Diagonal matrices), Proposition 2.4.8 (Four views of the product), Proposition 1.1.17 (No zero divisors)
Hint 1. (a) is the row view and the column view of Example 2.4.10, read entry by entry.
Hint 2. $DX = XD$ means $(d_i - d_j)x_{ij} = 0$ for all i, j. When does this force $x_{ij} = 0$?
Sketch. (a) DX scales row i by d_i; XD scales column j by d_j. (b) $XD = DX$ iff $(d_i - d_j)x_{ij} = 0$ for all i, j; with distinct d's this forces $x_{ij} = 0$ for $i \neq j$. (c) The condition kills exactly the entries x_{ij} with $d_i \neq d_j$, so X is block diagonal with blocks of sizes $2, 1$.

(a) By Example 2.4.10, row i of DX is $d_i X_{i,\cdot}$ and column j of XD is $d_j X_{\cdot,j}$. Reading off the (i,j) entries gives $(DX)_{ij} = d_i x_{ij}$ and $(XD)_{ij} = x_{ij} d_j$.

(b) By (a), $DX = XD$ if and only if $d_i x_{ij} = x_{ij} d_j$, that is, $(d_i - d_j)x_{ij} = 0$, for all i, j. For $i = j$ this always holds. For $i \neq j$, $d_i - d_j \neq 0$, so it holds if and only if $x_{ij} = 0$ (Proposition 1.1.17). So X commutes with D exactly when all its off-diagonal entries vanish.

(c) Here $d_1 = d_2 = 1$ and $d_3 = 0$. The condition $(d_i - d_j)x_{ij} = 0$ is automatic when $d_i = d_j$, that is, for $i, j \in \{1, 2\}$ and for $i = j = 3$, and it forces $x_{ij} = 0$ when exactly one of i, j equals 3, since then $d_i - d_j = \pm 1 \neq 0$. So the matrices commuting with $\operatorname{diag}(1, 1, 0)$ are

$$X = \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{21} & x_{22} & 0 \\ 0 & 0 & x_{33} \end{pmatrix},$$

with the five remaining entries arbitrary. These are the block diagonal matrices with blocks of sizes 2 and 1, matching the blocks on which $\operatorname{diag}(1,1,0)$ is a scalar matrix.

Solution to Exercise 2.4.43

Uses. Lemma 2.4.24 (Matrix units), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Use only a few upper triangular matrices: the matrix units $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ with $i \le j$.
Hint 2. Commuting with $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ says $A_{\cdot,i}\mathbf{e}_i^{\mathsf{T}} = \mathbf{e}_i A_{i,\cdot}$. The left side is zero outside column i, the right side zero outside row i. What does that force?
Hint 3. Once A is diagonal, compare $A\big(\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}\big)$ with $\big(\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}\big)A$.
Sketch. Commuting with $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ gives $A_{\cdot,i}\mathbf{e}_i^{\mathsf{T}} = \mathbf{e}_i A_{i,\cdot}$; comparing entries outside position (i,i), column i and row i of A vanish off the diagonal, so A is diagonal. Then $A\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}} = a_{11}\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}$ and $\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}A = a_{jj}\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}$ give $a_{jj} = a_{11}$.

If. cI_n commutes with every matrix, since $(cI_n)X = cX = X(cI_n)$ by Proposition 2.4.14.

Only if. Suppose A commutes with every upper triangular matrix. Fix i. The matrix unit $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ is upper triangular, and by associativity and Lemma 2.4.24(b),

$$A\big(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\big) = A_{\cdot,i}\mathbf{e}_i^{\mathsf{T}}, \quad \big(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\big)A = \mathbf{e}_i A_{i,\cdot}.$$

The first matrix has (k,l) entry $a_{ki}\delta_{li}$: it is zero outside column i, where it equals column i of A. The second has (k,l) entry $\delta_{ki}a_{il}$: it is zero outside row i, where it equals row i of A. Comparing the entries in positions (k,i) with $k \ne i$ gives $a_{ki} = 0$, and comparing those in positions (i,l) with $l \ne i$ gives $a_{il} = 0$. As i was arbitrary, A is diagonal, so $A\mathbf{e}_j = a_{jj}\mathbf{e}_j$ and $\mathbf{e}_j^{\mathsf{T}}A = a_{jj}\mathbf{e}_j^{\mathsf{T}}$ for every j.

Now let $2 \le j \le n$. The matrix unit $\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}$ is upper triangular, and

$$A\big(\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}\big) = (A\mathbf{e}_1)\mathbf{e}_j^{\mathsf{T}} = a_{11}\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}},$$
$$\big(\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}\big)A = \mathbf{e}_1\big(\mathbf{e}_j^{\mathsf{T}}A\big) = a_{jj}\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}}.$$

These are equal, and $\mathbf{e}_1\mathbf{e}_j^{\mathsf{T}} \ne 0$, so $a_{jj} = a_{11}$. Hence $A = a_{11}I_n$.

Solution to Exercise 2.4.44

Uses. Proposition 2.4.8 (Four views of the product), Definition 2.4.21 (Transpose), Exercise 2.4.25, Proposition 2.4.36 (Products of upper triangular matrices), Example 2.4.18 ($AB \ne BA$), Proposition 1.1.17 (No zero divisors)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (a), (b), and (c) are false; look among 2×2 matrices with entries 0 and 1. For (a), think of a reflection.
Hint 2. (d) Compute the diagonal entries of $A^{\mathsf{T}}A$. (e) Use Proposition 2.4.36.
Sketch. (a) False: $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)^2 = I_2$. (b) False: $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. (c) False: with $A = \mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}$, $B = \mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}$, $(AB)^2 = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ but $A^2 = 0$. (d) True over $\mathbb{R}$: $\big(A^{\mathsf{T}}A\big)_{jj} = \sum_i a_{ij}^2$. False over $\mathbb{C}$ and $\mathbb{F}_2$: $A = \left(\begin{smallmatrix} 1 & 0 \\ i & 0 \end{smallmatrix}\right)$, resp. $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. (e) True: $\big(A^2\big)_{ii} = a_{ii}^2$.

We give the counterexamples for $n = 2$. For $n > 2$, those in (b), (c), and (d) can be placed in the top left corner of an $n \times n$ zero matrix (Theorem 2.4.34), and the one in (a) in the top left corner of I_n.

(a) False. $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ swaps $\mathbf{e}_1$ and $\mathbf{e}_2$, so $A^2 = I_2$, but $A \ne \pm I_2$ (in $\mathbb{F}_2$, where $-I_2 = I_2$, still $A \ne I_2$). Geometrically, A is the reflection across the line $y = x$. For $n > 2$, the matrix that swaps $\mathbf{e}_1$ and $\mathbf{e}_2$ and fixes the other $\mathbf{e}_j$ works.

(b) False. $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ are symmetric, but $AB = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ is not. By Exercise 2.4.25, AB is symmetric exactly when $AB = BA$.

(c) False. With $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ (Example 2.4.18), $AB = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, so $(AB)^2 = AB \ne 0$, while $A^2 = 0$ gives $A^2B^2 = 0$. In general $(AB)^2 = ABAB$, and this equals $AABB$ when $AB = BA$.

(d) True over $\mathbb{R}$. By Proposition 2.4.8(a) and Definition 2.4.21, the (j,j) entry of $A^{\mathsf{T}}A$ is $\sum_i \left(A^{\mathsf{T}}\right)_{ji} a_{ij} = \sum_i a_{ij}^2$. If $A^{\mathsf{T}}A = 0$, then for each j this sum of squares of real numbers is 0, so every $a_{ij} = 0$.

Over $\mathbb{C}$ it is false: for $A = \begin{pmatrix} 1 & 0 \\ i & 0 \end{pmatrix}$, $A^{\mathsf{T}}A = \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1+i^2 & 0 \\ 0 & 0 \end{pmatrix} = 0$. Over $\mathbb{F}_2$, $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ gives $A^{\mathsf{T}}A = \begin{pmatrix} 1+1 & 0 \\ 0 & 0 \end{pmatrix} = 0$. The real proof used that a sum of squares vanishes only if each term does, which fails in both fields. Over $\mathbb{C}$ the cure is to conjugate, using A^*A (Section 6.1).

(e) True. By Proposition 2.4.36, A^2 is upper triangular with diagonal entries a_{ii}^2. If $A^2 = 0$, then $a_{ii}^2 = 0$, so $a_{ii} = 0$ by Proposition 1.1.17.

Solution to Exercise 2.4.45

Uses. Proposition 2.4.8 (Four views of the product), Proposition 1.1.17 (No zero divisors)
Hint 1. Compute A^2. Its off-diagonal entries are $b(a+d)$ and $c(a+d)$.
Hint 2. If $a+d \neq 0$, then $b = c = 0$; what do the diagonal entries then force? For (b), count the pairs (b,c) with $bc = -a^2$, separately for $a = 0$ and $a \neq 0$.
Sketch. $A^2 = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix}$. If $a+d \neq 0$ then $b = c = 0$ and $a^2 = d^2 = 0$, so $a = d = 0$, a contradiction; so $d = -a$, and then all four conditions reduce to $a^2 + bc = 0$. (b) For $a = 0$: $bc = 0$ has $2p-1$ solutions. For each of the $p-1$ nonzero a: $bc = -a^2 \neq 0$ has $p-1$ solutions. Total $2p - 1 + (p-1)^2 = p^2$.

(a) By the entry formula,

$$A^2 = \begin{pmatrix} a^2+bc & b(a+d) \\ c(a+d) & d^2+bc \end{pmatrix}.$$

Suppose $A^2 = 0$. If $a+d \neq 0$, the off-diagonal entries give $b = c = 0$ (Proposition 1.1.17), and then the diagonal entries give $a^2 = d^2 = 0$, so $a = d = 0$, contradicting $a+d \neq 0$. Hence $a+d = 0$, and then $a^2 + bc = 0$ is the $(1,1)$ entry. Conversely, if $d = -a$ and $a^2 + bc = 0$, then the off-diagonal entries are 0, the $(1,1)$ entry is 0, and the $(2,2)$ entry is $d^2 + bc = a^2 + bc = 0$.

(b) By (a), such A correspond to triples $(a,b,c) \in \mathbb{F}_p^3$ with $bc = -a^2$, with $d = -a$. If $a = 0$, we need $bc = 0$, that is, $b = 0$ or $c = 0$: there are $p + p - 1 = 2p - 1$ such pairs. If $a \neq 0$, then $-a^2 \neq 0$, so $b \neq 0$ and $c = -a^2 b^{-1}$ is determined by b: there are $p-1$ pairs. With $p-1$ nonzero values of a, the total is

$$(2p-1) + (p-1)^2 = p^2.$$

For $p = 2$ the four matrices are 0, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Solution to Exercise 2.4.46

Uses. Notation 2.4.16 (Powers of a matrix), Proposition 2.4.8 (Four views of the product), Theorem 2.4.2 (The matrix of a composite), Example 2.4.5 (Composites through their matrices)
Hint 1. (a) Induct on k, computing $J^{k+1} = J^k J$ with the entry formula.
Hint 2. (b) Take $\lambda = 1$ and $k = p$. (c) Over $\mathbb{R}$, think of a rotation. Over $\mathbb{F}_2$, try $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.
Sketch. (a) $J^k J = \begin{pmatrix} \lambda^{k+1} & \lambda^k + k\lambda^k \\ 0 & \lambda^{k+1} \end{pmatrix}$. (b) $p * 1 = 0$ in $\mathbb{F}_p$. (c) Over $\mathbb{F}_2$, $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^3 = I_2$. Over $\mathbb{R}$, the rotation by $2\pi/3$, $\left[R_{2\pi/3}\right] = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$.

(a) For $k = 1$ the formula reads $J = J$, since $1 * 1 = 1$ and $\lambda^0 = 1$. If it holds for k, then

$$J^{k+1} = J^k J = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

By the entry formula, this product has diagonal entries λ^{k+1}, lower left entry 0, and upper right entry $\lambda^k \cdot 1 + k\lambda^{k-1} \cdot \lambda = (k+1)\lambda^k$, which is the formula for $k+1$.

(b) With $\lambda = 1$ and $k = p$, (a) gives $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^p = \begin{pmatrix} 1 & p*1 \\ 0 & 1 \end{pmatrix} = I_2$, since $p * 1 = 0$ in $\mathbb{F}_p$. The map $(x,y) \mapsto (x+y, y)$ of $\mathbb{F}_p^2$ adds y to the first coordinate, and doing so p times adds $p * y = 0$.

(c) Over $\mathbb{F}_2$, let $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Then, using $1 + 1 = 0$,

$$A^2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$
$$A^3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Over $\mathbb{R}$, let $A' = \left[R_{2\pi/3}\right]_{\mathcal{E}\mathcal{E}} = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$. By Theorem 2.4.2 and Example 2.4.5(b), $A'^3 = \left[R_{2\pi/3}^3\right] = [R_{2\pi}] = I_2$, and $A' \neq I_2$.

Solution to Exercise 2.4.47

Uses. Notation 2.4.16 (Powers of a matrix), Proposition 2.4.8 (Four views of the product), Proposition 2.4.14 (Algebra of matrices)
Hint 1. (a) Induct on k, with $Q^{k+1} = Q^k Q$.
Hint 2. (b) Compare the $(1,2)$ entries of Q^{m+n} and $Q^m Q^n$.
Sketch. (a) $Q^k Q = \begin{pmatrix} f_{k+1}+f_k & f_{k+1} \\ f_k+f_{k-1} & f_k \end{pmatrix}$. (b) $(Q^{m+n})_{12} = f_{m+n}$ and $(Q^m Q^n)_{12} = f_{m+1}f_n + f_m f_{n-1}$. Put $m = n$.

(a) For $k = 1$, $\begin{pmatrix} f_2 & f_1 \\ f_1 & f_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = Q$. If the formula holds for k, then

$$\begin{aligned} Q^{k+1} = Q^k Q &= \begin{pmatrix} f_{k+1} & f_k \\ f_k & f_{k-1} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} f_{k+1}+f_k & f_{k+1} \\ f_k + f_{k-1} & f_k \end{pmatrix} \\ &= \begin{pmatrix} f_{k+2} & f_{k+1} \\ f_{k+1} & f_k \end{pmatrix}. \end{aligned}$$

(b) By Notation 2.4.16, $Q^{m+n} = Q^m Q^n$. The $(1,2)$ entry of the left side is f_{m+n} by (a), and by the entry formula the $(1,2)$ entry of the right side is

$$\begin{aligned} &(Q^m)_{11}(Q^n)_{12} + (Q^m)_{12}(Q^n)_{22} \\ &\quad = f_{m+1}f_n + f_m f_{n-1}. \end{aligned}$$

With $m = n$ this is $f_{2n} = f_{n+1}f_n + f_n f_{n-1} = f_n(f_{n+1} + f_{n-1})$. For example, $f_{10} = f_5(f_6 + f_4) = 5(8+3) = 55$. Linear recurrences in general are treated with matrix powers in Section 15.2.

Solution to Exercise 2.4.48

Uses. Proposition 2.4.8 (Four views of the product), Proposition 2.4.23 (Transpose of a product)
Hint 1. The first three equalities are correct. In the entry formula $(XY)_{ij} = \sum_k x_{ik}y_{kj}$, which index of each factor is the summation index?
Hint 2. In $\left(A^{\mathsf{T}}\right)_{kj}\left(B^{\mathsf{T}}\right)_{ik}$, the summation index k is the *first* index of the A^{T} factor. Reorder the two factors.
Sketch. The last step misreads the index pattern. After reordering the factors, the sum is $\sum_k \left(B^{\mathsf{T}}\right)_{ik}\left(A^{\mathsf{T}}\right)_{kj}$, which is the (i,j) entry of $B^{\mathsf{T}}A^{\mathsf{T}}$, not of $A^{\mathsf{T}}B^{\mathsf{T}}$. The correct conclusion is $(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$.

Everything up to the last equality is correct. In a product $(XY)_{ij} = \sum_k x_{ik}y_{kj}$, the row index i belongs to the left factor, the column index j to the right factor, and the summation index is the column index of the left factor and the row index of the right one. In $\sum_k \left(A^{\mathsf{T}}\right)_{kj}\left(B^{\mathsf{T}}\right)_{ik}$, the factor carrying i is B^{T} and the factor carrying j is A^{T}. Since scalars commute,

$$\begin{aligned} \sum_k \left(A^{\mathsf{T}}\right)_{kj}\left(B^{\mathsf{T}}\right)_{ik} &= \sum_k \left(B^{\mathsf{T}}\right)_{ik}\left(A^{\mathsf{T}}\right)_{kj} \\ &= \left(B^{\mathsf{T}}A^{\mathsf{T}}\right)_{ij}. \end{aligned}$$

So the computation actually proves Proposition 2.4.23. The claimed identity is false: with $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ as in Example 2.4.19, we have $(AB)^{\mathsf{T}} = 0^{\mathsf{T}} = 0$, while $A^{\mathsf{T}}B^{\mathsf{T}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Solution to Exercise 2.4.49

Uses. Lemma 2.4.24 (Matrix units), Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.4.14 (Algebra of matrices)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. (a) Rotate inside the trace: $\operatorname{tr}\left(A\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right) = \operatorname{tr}\left(\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_j\right)$, a 1×1 matrix.
Hint 2. (c) Choose B so that $\operatorname{tr}(BX)$ and $\varphi(X)$ agree on the matrix units $X = \mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}$. Uniqueness is (b) applied to a difference.
Sketch. (a) By cyclicity, $\operatorname{tr}\left(A\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right) = \operatorname{tr}\left(\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_j\right) = a_{ij}$. (b) Take $X = \mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}$. (c) Let $b_{ij} = \varphi\left(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right)$; then $X \mapsto \operatorname{tr}(BX)$ is linear and agrees with φ on the matrix units, which span. If $\operatorname{tr}(BX) = \operatorname{tr}(B'X)$ for all X, then $\operatorname{tr}((B - B')X) = 0$ for all X, so $B = B'$ by (b).

(a) By Proposition 2.4.28 with the factors $A\mathbf{e}_j \in F^{m\times 1}$ and $\mathbf{e}_i^{\mathsf{T}} \in F^{1\times m}$, and by Lemma 2.4.24(b),

$$\operatorname{tr}\left((A\mathbf{e}_j)\mathbf{e}_i^{\mathsf{T}}\right) = \operatorname{tr}\left(\mathbf{e}_i^{\mathsf{T}}(A\mathbf{e}_j)\right) = \mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_j = a_{ij}.$$

(b) Taking $X = \mathbf{e}_j\mathbf{e}_i^{\mathsf{T}} \in F^{n\times m}$ in the hypothesis and using (a) gives $a_{ij} = 0$ for all i, j.

(c) Define $B \in F^{m\times n}$ by $b_{ij} = \varphi\left(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right)$. The map $\psi(X) = \operatorname{tr}(BX)$ is linear, by Proposition 2.4.14 and Proposition 2.4.27, and by (a) applied to B, $\psi\left(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right) = b_{ij} = \varphi\left(\mathbf{e}_j\mathbf{e}_i^{\mathsf{T}}\right)$. The matrix units span $F^{n\times m}$ (Lemma 2.4.24(a)), so $\psi = \varphi$ by Corollary 2.2.6. If also $\varphi(X) = \operatorname{tr}(B'X)$ for all X, then $\operatorname{tr}((B - B')X) = 0$ for all X, and $B = B'$ by (b). So the linear maps $F^{n\times m} \to F$ correspond exactly to the matrices of $F^{m\times n}$; this is a first example of duality (Section 13.1).

Solution to Exercise 2.4.50

Uses. Lemma 2.4.24 (Matrix units), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.4.27 (The trace is linear)
Techniques. Check it on a basis (Technique 2.2.7), Cyclicity of trace (Technique 2.4.29)
Hint 1. It is enough to compute φ on the matrix units $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$. Write each one as a product XY for which YX is simpler.
Hint 2. For $i \neq j$, $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} = \left(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right)\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right)$, and the product in the other order is 0. For the diagonal units, compare $\left(\mathbf{e}_i\mathbf{e}_1^{\mathsf{T}}\right)\left(\mathbf{e}_1\mathbf{e}_i^{\mathsf{T}}\right)$ with the product in the other order.
Sketch. For $i \neq j$, write $\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} = XY$ with $X = \mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ and $Y = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$; then $YX = 0$, so $\varphi\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right) = \varphi(YX) = 0$. Writing $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ and $\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ as the two products of $\mathbf{e}_i\mathbf{e}_1^{\mathsf{T}}$ and $\mathbf{e}_1\mathbf{e}_i^{\mathsf{T}}$ shows that φ takes the same value c on all diagonal units. So φ and $c\operatorname{tr}$ agree on the matrix units.

We compute φ on the matrix units, using Lemma 2.4.24(c) throughout. For $i \neq j$,

$$\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} = \left(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right)\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right),$$
$$\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right)\left(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right) = \delta_{ji}\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}} = 0,$$

so the hypothesis gives $\varphi\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right) = \varphi(0) = 0$. For the diagonal units,

$$\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}} = \left(\mathbf{e}_i\mathbf{e}_1^{\mathsf{T}}\right)\left(\mathbf{e}_1\mathbf{e}_i^{\mathsf{T}}\right),$$
$$\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}} = \left(\mathbf{e}_1\mathbf{e}_i^{\mathsf{T}}\right)\left(\mathbf{e}_i\mathbf{e}_1^{\mathsf{T}}\right),$$

so $\varphi\left(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right) = \varphi\left(\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}\right)$ for every i; call this value c. On the other hand $\operatorname{tr}\left(\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\right) = \delta_{ij}$. So the linear maps φ and $c\operatorname{tr}$ (Proposition 2.4.27) agree on every matrix unit. The matrix units span $F^{n\times n}$ (Lemma 2.4.24(a)), so $\varphi = c\operatorname{tr}$ by Corollary 2.2.6.

Solution to Exercise 2.4.51

Uses. Example 2.1.9 (Maps of polynomials), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space), Proposition 2.1.17 (Compositions of linear maps are linear), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.4.30 ($AB - BA \neq I$), Definition 1.2.16 (Degree)
Hint 1. (a) Both sides are linear, so check on the monomials x^k, which span $F[x]$.
Hint 2. (b) Proposition 2.4.30 is about $n \times n$ matrices. (c) If $U \neq \{\mathbf{0}\}$, take a nonzero $p \in U$ of the largest degree among a finite spanning list, and apply M.

Sketch. (a) $(DM - MD)\big(x^k\big) = (k+1)x^k - kx^k = x^k$. (b) $F[x]$ has no finite basis, so D and M are not given by square matrices; the trace argument needs a finite size. (c) Every element of U has degree at most the largest degree d in a finite spanning list; if $U \neq \{\mathbf{0}\}$ it contains an element of degree exactly d, and M of it has degree $d+1$.

(a) Both $DM - MD$ and id are linear (Proposition 2.1.23, Proposition 2.1.17). For $k \geq 0$,

$$\begin{aligned}(DM - MD)\big(x^k\big) &= D\big(x^{k+1}\big) - xD\big(x^k\big)\\ &= (k+1)x^k - kx^k = x^k,\end{aligned}$$

where for $k = 0$ the second term is $xD(1) = \mathbf{0}$. The monomials span $F[x]$, so $DM - MD = \mathrm{id}$ by Corollary 2.2.6. (This is the product rule $(xp)' = p + xp'$.)

(b) Proposition 2.4.30 is a statement about $n \times n$ matrices, that is, about operators on F^n, and its proof takes the trace, a finite sum of diagonal entries. The maps D and M act on $F[x]$, which is spanned by no finite list (every finite list of polynomials has a largest degree). So they have no $n \times n$ matrices, and there is no trace to take. Part (c) shows that we also cannot cut $F[x]$ down to a finite-dimensional piece on which M acts.

(c) Suppose $U = \operatorname{span}(p_1, \ldots, p_r)$ with $M(U) \subseteq U$ and $U \neq \{\mathbf{0}\}$. We may drop the p_i that are $\mathbf{0}$, so all are nonzero, and $r \geq 1$. Let d be the largest of $\deg p_1, \ldots, \deg p_r$. Every element of U is a combination of the p_i, so every nonzero element of U has degree at most d, and p_i has degree d for some i. But $M(p_i) = xp_i \in U$ has degree $d+1$, a contradiction. So $U = \{\mathbf{0}\}$. (The same argument shows that $F[x]$ itself is not spanned by finitely many polynomials.)

Solution to Exercise 2.4.52

Uses. Proposition 2.4.30 ($AB - BA \neq I$), Theorem 2.4.34 (Block multiplication), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.3.13 (Linearity of the product)
Techniques. Block decomposition (Technique 2.4.35), Cyclicity of trace (Technique 2.4.29)
Hint 1. (a) A and B are the matrices of differentiation and of multiplication by x on polynomials of degree less than p, in the basis $1, x, \ldots, x^{p-1}$, with x^p replaced by 0. Compute $(AB - BA)\mathbf{e}_k$, treating $k = p$ separately.
Hint 2. (b) One direction is Proposition 2.4.30: over $\mathbb{F}_p$, $n * 1 = 0$ exactly when $p \mid n$. For the other, put n/p copies of the matrices from (a) on the diagonal of a block diagonal matrix.
Sketch. (a) For $k < p$, $AB\mathbf{e}_k = A\mathbf{e}_{k+1} = k\mathbf{e}_k$ and $BA\mathbf{e}_k = (k-1)\mathbf{e}_k$, so the difference is $\mathbf{e}_k$. For $k = p$, $AB\mathbf{e}_p = \mathbf{0}$ and $BA\mathbf{e}_p = (p-1)\mathbf{e}_p$, so the difference is $(1-p)\mathbf{e}_p = \mathbf{e}_p$. (b) If $p \nmid n$ then $n * 1 \neq 0$ in $\mathbb{F}_p$. If $n = pq$, let $X = \operatorname{diag}(A, \ldots, A)$ and $Y = \operatorname{diag}(B, \ldots, B)$ (q blocks); block multiplication gives $XY - YX = \operatorname{diag}\big(I_p, \ldots, I_p\big) = I_n$.

(a) We compare $(AB - BA)\mathbf{e}_k$ with $\mathbf{e}_k$ for each k; by Proposition 2.3.13(c) these are the columns of the two matrices. Note that $BA\mathbf{e}_k = B((k-1)\mathbf{e}_{k-1}) = (k-1)\mathbf{e}_k$ for $2 \leq k \leq p$, and $BA\mathbf{e}_1 = B\mathbf{0} = \mathbf{0} = (1-1)\mathbf{e}_1$, so $BA\mathbf{e}_k = (k-1)\mathbf{e}_k$ for all k. For $1 \leq k \leq p-1$, $AB\mathbf{e}_k = A\mathbf{e}_{k+1} = k\mathbf{e}_k$, so

$$(AB - BA)\mathbf{e}_k = k\mathbf{e}_k - (k-1)\mathbf{e}_k = \mathbf{e}_k.$$

For $k = p$, $AB\mathbf{e}_p = A\mathbf{0} = \mathbf{0}$, so $(AB - BA)\mathbf{e}_p = -(p-1)\mathbf{e}_p$, which is $\mathbf{e}_p$ because $-(p-1) = 1 - p = 1$ in $\mathbb{F}_p$. So $AB - BA = I_p$. For $p = 2$ these are the matrices $A = \left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix}0 & 0\\ 1 & 0\end{smallmatrix}\right)$ of Remark 2.4.31.

(b) If $p \nmid n$, then $n * 1 = n \bmod p \neq 0$ in $\mathbb{F}_p$, and Proposition 2.4.30 says that no such X, Y exist. Conversely, suppose $n = pq$. Let X and Y be the block diagonal matrices with q diagonal blocks, all equal to A, respectively all equal to B, partitioned with sizes $p, \ldots, p$. By Theorem 2.4.34, XY and YX are block diagonal with diagonal blocks AB, respectively BA, so $XY - YX$ is block diagonal with diagonal blocks $AB - BA = I_p$. That is the matrix I_n.

Solution to Exercise 2.4.53

Uses. Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Proposition 2.4.14 (Algebra of matrices)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. It suffices to write C^k as a commutator $AX - XA$ for some X, since commutators have trace 0.
Hint 2. $C^k = C^{k-1}(AB - BA) = C^{k-1}AB - C^{k-1}BA$. Move A past C^{k-1} in the first term.

Sketch. Since C commutes with A, so does C^{k-1}. Expanding $C^k = C^{k-1}(AB - BA)$ and moving A to the front of the first term gives $C^k = AX - XA$ with $X = C^{k-1}B$, so $\operatorname{tr}\big(C^k\big) = 0$ by cyclicity.

Since $AC = CA$, induction gives $AC^{k-1} = C^{k-1}A$ for every $k \geq 1$ (for $k = 1$, $C^0 = I_n$). By distributivity and associativity (Proposition 2.4.14),

$$\begin{aligned} C^k &= C^{k-1}(AB - BA) \\ &= \big(C^{k-1}A\big)B - C^{k-1}BA \\ &= A\big(C^{k-1}B\big) - \big(C^{k-1}B\big)A. \end{aligned}$$

So with $X = C^{k-1}B$, Proposition 2.4.27 and Proposition 2.4.28 give $\operatorname{tr}\big(C^k\big) = \operatorname{tr}(AX) - \operatorname{tr}(XA) = 0$. Over $\mathbb{C}$, the vanishing of every $\operatorname{tr}\big(C^k\big)$ forces C to be nilpotent, a result known as Jacobson's lemma; that last step needs eigenvalues (Section 7.3).

Solution to Exercise 2.4.54

Uses. Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Proposition 2.4.23 (Transpose of a product), Proposition 2.4.8 (Four views of the product), Exercise 2.4.25
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. The equality case suggests looking at $C = AB - BA$. What is C^{T}?
Hint 2. Expand $\operatorname{tr}\big(C^2\big)$ and use cyclicity to bring each of the four terms to $\operatorname{tr}(ABAB)$ or $\operatorname{tr}\big(A^2B^2\big)$.
Hint 3. For a real skew-symmetric C, $\operatorname{tr}\big(C^2\big) = -\operatorname{tr}\big(C^{\mathsf{T}}C\big)$, and $\operatorname{tr}\big(C^{\mathsf{T}}C\big)$ is a sum of squares of entries.
Sketch. $C = AB - BA$ is skew-symmetric. Cyclicity gives $\operatorname{tr}\big(C^2\big) = 2\operatorname{tr}(ABAB) - 2\operatorname{tr}\big(A^2B^2\big)$. Also $\operatorname{tr}\big(C^2\big) = -\operatorname{tr}\big(C^{\mathsf{T}}C\big) = -\sum_{i,j} c_{ij}^2 \leq 0$, with equality iff $C = 0$.

Let $C = AB - BA$. By Exercise 2.4.25, C is skew-symmetric: $C^{\mathsf{T}} = -C$. Expanding,

$$C^2 = ABAB - ABBA - BAAB + BABA.$$

By Proposition 2.4.28, rotating the factors gives $\operatorname{tr}(ABBA) = \operatorname{tr}(AABB)$, $\operatorname{tr}(BAAB) = \operatorname{tr}(AABB)$, and $\operatorname{tr}(BABA) = \operatorname{tr}(ABAB)$. So by Proposition 2.4.27,

$$\operatorname{tr}\big(C^2\big) = 2\operatorname{tr}(ABAB) - 2\operatorname{tr}\big(A^2B^2\big).$$

On the other hand $C^2 = -C^{\mathsf{T}}C$, and by the entry formula the (j,j) entry of $C^{\mathsf{T}}C$ is $\sum_i c_{ij}^2$. Hence

$$\operatorname{tr}\big(C^2\big) = -\operatorname{tr}\big(C^{\mathsf{T}}C\big) = -\sum_{i,j} c_{ij}^2 \leq 0,$$

with equality if and only if every $c_{ij} = 0$, that is, $AB = BA$. Combining the two expressions for $\operatorname{tr}\big(C^2\big)$ gives $\operatorname{tr}(ABAB) \leq \operatorname{tr}\big(A^2B^2\big)$, with equality exactly when $AB = BA$.

Solution to Exercise 2.4.55

Uses. Theorem 2.4.34 (Block multiplication), Proposition 2.4.14 (Algebra of matrices), Notation 2.4.16 (Powers of a matrix)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. (a) Induct on k. Compute $N^{k+1} = N^kN$ by blocks; the top right block is $A^kX + X_kD$.
Hint 2. (b) $N^{2k} = N^kN^k$, and N^k has zero diagonal blocks.
Sketch. (a) By blocks, N^kN has diagonal blocks A^{k+1}, D^{k+1} and top right block $A^kX + X_kD$, and $X_kD = \sum_{i=0}^{k-1} A^iXD^{k-i}$, so the top right block is X_{k+1}. (b) $N^k = \left(\begin{smallmatrix} 0 & X_k \\ 0 & 0 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 0 & Y \\ 0 & 0 \end{smallmatrix}\right)^2 = 0$.

(a) For $k = 1$, $X_1 = A^0XD^0 = X$. Suppose the formula holds for k. By Theorem 2.4.34,

$$\begin{aligned} N^{k+1} = N^kN &= \begin{pmatrix} A^k & X_k \\ 0 & D^k \end{pmatrix}\begin{pmatrix} A & X \\ 0 & D \end{pmatrix} \\ &= \begin{pmatrix} A^{k+1} & A^kX + X_kD \\ 0 & D^{k+1} \end{pmatrix}, \end{aligned}$$

and by distributivity

$$A^kX + X_kD = A^kXD^0 + \sum_{i=0}^{k-1} A^iXD^{k-i}$$

$$= \sum_{i=0}^{k} A^iXD^{k-i} = X_{k+1}.$$

(b) If $A^k = 0$ and $D^k = 0$, then $N^k = \begin{pmatrix} 0 & X_k \\ 0 & 0 \end{pmatrix}$, and by block multiplication

$$N^{2k} = N^kN^k = \begin{pmatrix} 0 & X_k \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & X_k \\ 0 & 0 \end{pmatrix} = 0,$$

since each block of the product is a sum of terms with a zero factor. So a block upper triangular matrix whose diagonal blocks are nilpotent is nilpotent. The bound $2k$ cannot be improved to k in general: for $p = q = 1$, $A = D = 0$, and $X = 1$, $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has $A^1 = D^1 = 0$ but $N^1 \neq 0$.

Solution to Exercise 2.4.56

Uses. Definition 2.4.38 (Adjacency matrix), Proposition 2.4.39 (Powers count walks)
Hint 1. A has a 1 in positions $(1,2)$, $(2,3)$, $(3,1)$, and $(1,1)$.
Hint 2. By Proposition 2.4.39, the answers are $\left(A^3\right)_{11}$ and $\operatorname{tr}\left(A^3\right)$. Check them by listing walks.
Sketch. $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ and $A^3 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. The two walks of length 3 from 1 to 1 are $1,1,1,1$ and $1,2,3,1$, and $\operatorname{tr}\left(A^3\right) = 4$.

(a)

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

$$A^3 = A^2A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(b) By Proposition 2.4.39 there are $\left(A^3\right)_{11} = 2$ walks of length 3 from 1 to 1: they are $1,1,1,1$ (the loop three times) and $1,2,3,1$ (around the triangle). In all there are $\operatorname{tr}\left(A^3\right) = 2 + 1 + 1 = 4$ closed walks of length 3: the two above, and $2,3,1,2$ and $3,1,2,3$.

Solution to Exercise 2.4.57

Uses. Proposition 2.4.39 (Powers count walks), Proposition 2.4.8 (Four views of the product), Proposition 2.4.14 (Algebra of matrices), Definition 2.4.26 (Trace)
Hint 1. (a) Every entry of J^2 is a sum of n products $1 \cdot 1$.
Hint 2. (b) Induct on k. Expand $\left(a_kJ + (-1)^kI\right)(J - I)$ with (a), and check that the coefficient of J satisfies $a_{k+1} = (n-1)a_k + (-1)^k$.
Sketch. (a) Each entry of J^2 is a sum of n ones. (b) Multiplying $a_kJ + (-1)^kI$ by $J - I$ and using $J^2 = nJ$ gives coefficient $(n-1)a_k + (-1)^k$ on J and $(-1)^{k+1}$ on I; the formula for a_k satisfies this recursion. (c) Off the diagonal, $\left(A^k\right)_{ij} = a_k$; the trace is $na_k + n(-1)^k$, which simplifies to $(n-1)^k + (n-1)(-1)^k$.

(a) By the entry formula, $\left(J^2\right)_{ij} = \sum_{k=1}^{n} 1 \cdot 1 = n$, so $J^2 = nJ$.

(b) For $k = 1$, $a_1 = (n - 1 + 1)/n = 1$, and $A = J - I_n = a_1J + (-1)^1I_n$. If $A^k = a_kJ + (-1)^kI_n$, then by Proposition 2.4.14 and (a),

$$A^{k+1} = \left(a_kJ + (-1)^kI_n\right)(J - I_n)$$

$$= a_knJ - a_kJ + (-1)^kJ - (-1)^kI_n$$

$$= \Big((n-1)a_k + (-1)^k\Big)J + (-1)^{k+1}I_n.$$

It remains to show that the coefficient of J is a_{k+1}. Write $b_k = na_k = (n-1)^k - (-1)^k$. Then $n\Big((n-1)a_k + (-1)^k\Big) = (n-1)b_k + n(-1)^k$, and

$$\begin{aligned}&(n-1)b_k + n(-1)^k\\ &= (n-1)^{k+1} - (n-1)(-1)^k + n(-1)^k\\ &= (n-1)^{k+1} + (-1)^k = b_{k+1}.\end{aligned}$$

Dividing by n gives $(n-1)a_k + (-1)^k = a_{k+1}$.

(c) By Proposition 2.4.39 and (b), for $i \neq j$ there are $\big(A^k\big)_{ij} = a_k = \Big((n-1)^k - (-1)^k\Big)/n$ walks of length k from i to j, and the number of closed walks of length k is

$$\begin{aligned}\operatorname{tr}\big(A^k\big) &= na_k + n(-1)^k\\ &= (n-1)^k - (-1)^k + n(-1)^k\\ &= (n-1)^k + (n-1)(-1)^k.\end{aligned}$$

For instance, the triangle ($n = 3$) has $2^3 - 2 = 6$ closed walks of length 3, the two directions around it from each of the three vertices. As a check, a_k is an integer, since $n-1$ and -1 leave the same remainder modulo n, so n divides $(n-1)^k - (-1)^k$.

Section 2.5

Solution to Exercise 2.5.9

Uses. Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Proposition 2.4.8 (Four views of the product), Proposition 2.3.13 (Linearity of the product)
Hint 1. If column j is zero, then $A\mathbf{e}_j = \mathbf{0} = A\mathbf{0}$. What does that say about T_A?
Hint 2. If row i is zero, then row i of AC is zero for every C, by the row view. Compare with row i of I_m.
Sketch. (a) $A\mathbf{e}_j = \mathbf{0} = A\mathbf{0}$ with $\mathbf{e}_j \neq \mathbf{0}$, so T_A is not injective, and Proposition 2.5.4 rules out a left inverse. (b) Row i of AC is row i of A times C, which is zero, while row i of I_m is $\mathbf{e}_i^{\mathsf{T}} \neq 0$.

(a) Suppose column j of A is zero. By Proposition 2.3.13, $A\mathbf{e}_j$ is column j of A, so $A\mathbf{e}_j = \mathbf{0} = A\mathbf{0}$ although $\mathbf{e}_j \neq \mathbf{0}$. Hence T_A is not injective, and by Proposition 2.5.4, A has no left inverse. (Directly: for every $B \in F^{n\times m}$, $(BA)\mathbf{e}_j = B\big(A\mathbf{e}_j\big) = \mathbf{0}$, whereas $I_n\mathbf{e}_j = \mathbf{e}_j$.)

(b) Suppose row i of A is zero, and let $C \in F^{n\times m}$. By the row view (Proposition 2.4.8(c)), row i of AC is a combination of the rows of C with the entries of row i of A as coefficients, so it is zero. Row i of I_m has a 1 in position i, so $AC \neq I_m$. (In the language of Proposition 2.5.5: every vector $A\mathbf{x}$ has i-th entry 0, so $\mathbf{e}_i$ is not a value of T_A.)

Solution to Exercise 2.5.19

Uses. Definition 2.5.12 (Invertible), Proposition 2.5.17 (Algebra of inverses), Definition 2.4.1 (Matrix product)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Multiply out both products in (a).
Hint 2. By Proposition 2.5.17(b), $(AB)^{-1} = B^{-1}A^{-1}$, so it is enough to compute the two products of inverses and identify which one it is.
Sketch. (a) Both products are I_2 in each case. (b) $B^{-1}A^{-1} = \begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix} = (AB)^{-1}$, with $AB = \begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}$, while $A^{-1}B^{-1} = \begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix}$ is the inverse of $BA = \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$.

(a) $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & -1+1\\ 0 & 1\end{pmatrix} = I_2$ and $\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1-1\\ 0 & 1\end{pmatrix} = I_2$. Similarly $\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 1-1 & 1\end{pmatrix} = I_2$, and the product in the other order is I_2 as well.

(b) We have $AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and

$$\begin{aligned} B^{-1}A^{-1} &= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \\ A^{-1}B^{-1} &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

By Proposition 2.5.17(b), $(AB)^{-1} = B^{-1}A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$; as a check, $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2-1 & -2+2 \\ 1-1 & -1+2 \end{pmatrix} = I_2$. The product $A^{-1}B^{-1}$ is different; it is $(BA)^{-1}$, the inverse of $BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Solution to Exercise 2.5.22

Uses. Definition 2.3.1 (Matrix of a linear map), Proposition 2.5.20 (Inverses and matrices), Definition 2.4.1 (Matrix product)
Techniques. Translate between maps and matrices (Technique 2.3.27), Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. The columns of $[S]_{\mathcal{BB}}$ are the coordinates of $S(1) = 1$, $S(x) = x - 1$, and $S\left(x^2\right) = (x-1)^2 = x^2 - 2x + 1$.
Hint 2. Multiply $[T][S]$ and $[S][T]$. If both are I_3, then by Proposition 2.5.20(a), S is a two-sided inverse of T.
Sketch. $[S]_{\mathcal{BB}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$. Both products $[T][S]$ and $[S][T]$ are I_3, so by Proposition 2.5.20(a), S is a left and a right inverse of T.

From $S(1) = 1$, $S(x) = -1 + x$, and $S\left(x^2\right) = 1 - 2x + x^2$,

$$[S]_{\mathcal{BB}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

By the entry formula,

$$\begin{aligned} [T][S] &= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1+1 & 1-2+1 \\ 0 & 1 & -2+2 \\ 0 & 0 & 1 \end{pmatrix} = I_3, \end{aligned}$$

and

$$\begin{aligned} [S][T] &= \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1-1 & 1-2+1 \\ 0 & 1 & 2-2 \\ 0 & 0 & 1 \end{pmatrix} = I_3, \end{aligned}$$

where $[T]$ and $[S]$ abbreviate the matrices in the basis $\mathcal{B}$. So $[S]_{\mathcal{BB}}$ is a left and a right inverse of $[T]_{\mathcal{BB}}$. By Proposition 2.5.20(a), S is a left and a right inverse of T, so T is invertible with $T^{-1} = S$. Shifting by 1 and then by -1 changes nothing, as expected; the matrices let us confirm it without composing substitutions.

Solution to Exercise 2.5.30

Uses. Definition 2.5.1 (Left inverse), Definition 2.5.2 (Right inverse), Corollary 2.5.11 (Many one-sided inverses), Proposition 2.4.8 (Four views of the product)
Hint 1. A left inverse is a row $(b_1 \;\; b_2 \;\; b_3)$ with $b_1 + 2b_2 - b_3 = 1$. Solve for b_3.
Hint 2. For a right inverse $(c_1 \;\; c_2 \;\; c_3)$, the 3×3 product $\mathbf{a}(c_1 \;\; c_2 \;\; c_3)$ has (i,j) entry $a_i c_j$. Compare the entries in positions $(1,2)$ and $(2,2)$ with those of I_3.
Sketch. (a) The rows $(s \;\; t \;\; s+2t-1)$ for $s,t \in \mathbb{Q}$. (b) $\mathbf{a}(c_1 \;\; c_2 \;\; c_3) = I_3$ would need $c_2 = 0$ (position $(1,2)$) and $2c_2 = 1$ (position $(2,2)$). Alternatively, $(1 \;\; 0 \;\; 0)$ and $(0 \;\; 0 \;\; -1)$ are two different left inverses, so Corollary 2.5.11 applies.

(a) A left inverse is a matrix $B = (b_1 \;\; b_2 \;\; b_3) \in \mathbb{Q}^{1\times 3}$ with $B\mathbf{a} = I_1$, that is, $b_1 + 2b_2 - b_3 = 1$. For any $b_1 = s$ and $b_2 = t$ this determines $b_3 = s + 2t - 1$. So the left inverses of $\mathbf{a}$ are the rows

$$(s \;\; t \;\; s + 2t - 1) \quad (s, t \in \mathbb{Q}).$$

For instance $s = 1, t = 0$ gives $(1 \;\; 0 \;\; 0)$, $s = t = 0$ gives $(0 \;\; 0 \;\; -1)$, and $s = 1/6, t = 1/3$ gives $(1/6 \;\; 1/3 \;\; -1/6)$.

(b) A right inverse would be $C = (c_1 \;\; c_2 \;\; c_3) \in \mathbb{Q}^{1\times 3}$ with $\mathbf{a}C = I_3$. The product $\mathbf{a}C$ has (i,j) entry $a_i c_j$ (Proposition 2.4.8(a)). Position $(1,2)$ gives $1 \cdot c_2 = 0$, and position $(2,2)$ gives $2c_2 = 1$, which contradict each other. So there is no right inverse.

Alternatively, by (a), $(1 \;\; 0 \;\; 0)$ and $(0 \;\; 0 \;\; -1)$ are two different left inverses of $\mathbf{a}$, and by Corollary 2.5.11 a matrix with two different left inverses has no right inverse.

Solution to Exercise 2.5.31

Uses. Proposition 2.5.28 (Inverse of a 2×2 matrix)
Hint 1. Compute $\Delta = ad - bc$ in each case and apply Proposition 2.5.28.
Hint 2. In (b), $\Delta = 18 - 10 = 8$, which is 1 in $\mathbb{F}_7$. In (c), $\Delta = t^2 - 4$.
Sketch. (a) $\Delta = 2 - i^2 = 3$, so the inverse is $\frac{1}{3}\begin{pmatrix} 2 & -i \\ -i & 1 \end{pmatrix}$. (b) $\Delta = 1$, so the inverse is $\begin{pmatrix} 6 & -5 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 5 & 3 \end{pmatrix}$. (c) Over $\mathbb{R}$, $t \neq \pm 2$. Over $\mathbb{F}_5$, $t^2 \neq 4$, so $t \in \{0, 1, 4\}$.

(a) $\Delta = 1 \cdot 2 - i \cdot i = 2 + 1 = 3$, so by Proposition 2.5.28

$$\begin{pmatrix} 1 & i \\ i & 2 \end{pmatrix}^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -i \\ -i & 1 \end{pmatrix}.$$

Check: $\begin{pmatrix} 1 & i \\ i & 2 \end{pmatrix}\begin{pmatrix} 2 & -i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 2-i^2 & -i+i \\ 2i-2i & -i^2+2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$.

(b) $\Delta = 3 \cdot 6 - 5 \cdot 2 = 8 = 1$ in $\mathbb{F}_7$, so the inverse is $\begin{pmatrix} 6 & -5 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 5 & 3 \end{pmatrix}$. Check:

$$\begin{pmatrix} 3 & 5 \\ 2 & 6 \end{pmatrix}\begin{pmatrix} 6 & 2 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 18 + 25 & 6 + 15 \\ 12 + 30 & 4 + 18 \end{pmatrix} = \begin{pmatrix} 43 & 21 \\ 42 & 22 \end{pmatrix},$$

which is I_2 modulo 7 since $43 = 42 + 1$ and $22 = 21 + 1$.

(c) Here $\Delta = t^2 - 4$. By Proposition 2.5.28 the matrix is invertible exactly when $t^2 \neq 4$. Over $\mathbb{R}$ this means $t \neq \pm 2$, and then the inverse is $(t^2 - 4)^{-1}\begin{pmatrix} t & -1 \\ -4 & t \end{pmatrix}$. In $\mathbb{F}_5$ the squares of $0, 1, 2, 3, 4$ are $0, 1, 4, 4, 1$, so $t^2 = 4$ exactly for $t \in \{2, 3\}$, and the matrix is invertible for $t \in \{0, 1, 4\}$.

Solution to Exercise 2.5.32

Uses. Corollary 2.5.6 (One-sided cancellation), Proposition 2.5.17 (Algebra of inverses), Exercise 2.5.9, Notation 2.4.16 (Powers of a matrix)
Hint 1. (a) and (c) are cancellations: write the hypothesis as $AX = AY$.
Hint 2. (b) Powers of an invertible matrix are invertible. Is the zero matrix invertible?
Hint 3. (d) Compute AA^{k-1} and $A^{k-1}A$.
Sketch. (a) $AB = A0$, and A can be cancelled. (b) If A were invertible, so would be $A^k = 0$, but 0 has a zero column. (c) $AA = AI_n$; cancel. (d) $AA^{k-1} = A^{k-1}A = A^k = I_n$.

(a) $AB = 0 = A0$, and A is left-invertible, so $B = 0$ by Corollary 2.5.6.

(b) If A were invertible, then A^k would be invertible by Proposition 2.5.17(b) and induction. But $A^k = 0$, whose columns are all zero, and a matrix with a zero column has no left inverse (Exercise 2.5.9). So A is not invertible.

(c) $AA = A = AI_n$, so $A = I_n$ by Corollary 2.5.6.

(d) By Notation 2.4.16, $AA^{k-1} = A^k = I_n$ and $A^{k-1}A = A^k = I_n$ (for $k = 1$, $A^0 = I_n$ and $A = I_n$). So A^{k-1} is a two-sided inverse of A.

Solution to Exercise 2.5.33

Uses. Definition 2.5.12 (Invertible), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. (a) Fix $\mathbf{w} \in W$ and write it as $T(\mathbf{v})$. Compute $T(S(\mathbf{w}))$.
Hint 2. (b) Fix $\mathbf{v}$. Show that $T(S(T(\mathbf{v}))) = T(\mathbf{v})$, and use injectivity.
Sketch. (a) For $\mathbf{w} = T(\mathbf{v})$: $TS(\mathbf{w}) = T(ST(\mathbf{v})) = T(\mathbf{v}) = \mathbf{w}$. (b) $T(ST(\mathbf{v})) = (TS)(T(\mathbf{v})) = T(\mathbf{v})$, so $ST(\mathbf{v}) = \mathbf{v}$ by injectivity. (c) One direction is (a) or (b); the other is Proposition 2.5.5 or Proposition 2.5.4.

(a) Let $\mathbf{w} \in W$. Since T is surjective, $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V$. Then

$$\begin{aligned} T(S(\mathbf{w})) &= T(S(T(\mathbf{v}))) \\ &= T((ST)(\mathbf{v})) = T(\mathbf{v}) = \mathbf{w}. \end{aligned}$$

As $\mathbf{w}$ was arbitrary, $TS = \mathrm{id}_W$.

(b) Let $\mathbf{v} \in V$. Then $T(S(T(\mathbf{v}))) = (TS)(T(\mathbf{v})) = T(\mathbf{v})$. Since T is injective, $S(T(\mathbf{v})) = \mathbf{v}$. As $\mathbf{v}$ was arbitrary, $ST = \mathrm{id}_V$.

(c) Let S be a left inverse of T. If T is surjective, then by (a) S is also a right inverse, so T is invertible. Conversely, an invertible map has a right inverse and is therefore surjective (Proposition 2.5.5). The second statement follows in the same way from (b) and Proposition 2.5.4.

Solution to Exercise 2.5.34

Uses. Definition 2.5.12 (Invertible), Definition 2.5.2 (Right inverse), Theorem 2.5.10 (Left and right inverses agree), Example 2.5.27 (An invertible product of non-invertible factors)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. What does $AB = I_m$ say about B, according to the definitions? What does "$B = A^{-1}$" require?
Hint 2. Try $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
Hint 3. If A also has some left inverse, what does Theorem 2.5.10 say?
Sketch. $AB = I_m$ only says that B is a right inverse of A; the step "$B = A^{-1}$" assumes that A is invertible, which is what the claim was supposed to deliver. With $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $AB = I_1$ but $BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. If A also has a left inverse C, then $C = B$ by Theorem 2.5.10, so $BA = CA = I_n$.

The error is the step "$B = A^{-1}$". By Definition 2.5.2, $AB = I_m$ says only that B is a right inverse of A. The inverse A^{-1} is a matrix that is a left and a right inverse at the same time (Definition 2.5.12), and nothing in the hypothesis says that A has an inverse on the left. The argument assumes, in the notation A^{-1}, the very fact $BA = I_n$ that it set out to prove.

The conclusion is false: as in Example 2.5.27, $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ satisfy $AB = (1) = I_1$, but $BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq I_2$.

It becomes true if A is also assumed to be left-invertible. If $CA = I_n$, then $C = B$ by Theorem 2.5.10, so $BA = CA = I_n$. For square matrices no extra hypothesis is needed (Theorem 3.5.23), but that is a deeper fact, and the argument above does not prove it. For maps from a space to itself it fails without finite dimension: $DJ = \mathrm{id}$ but $JD \neq \mathrm{id}$ on $\mathbb{Q}[x]$ (Example 2.5.24).

Solution to Exercise 2.5.35

Uses. Definition 2.5.12 (Invertible), Theorem 2.5.10 (Left and right inverses agree), Proposition 2.5.17 (Algebra of inverses), Exercise 2.5.9
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. (a) Try $B = -A$. (b) From an inverse X of A^2, build a left and a right inverse of A separately.
Hint 2. (c) Use Proposition 2.5.17(d). (d) Write down a candidate left inverse that involves $\left(A^{\mathsf{T}}A\right)^{-1}$.
Hint 3. (e) The field matters. Try a column $A \in \mathbb{F}_2^{2\times 1}$ or $A \in \mathbb{C}^{2\times 1}$ with $A^{\mathsf{T}}A = 0$.
Sketch. (a) False: $I_n + (-I_n) = 0$. (b) True: if $XA^2 = A^2X = I_n$, then XA is a left inverse and AX a right inverse of A. (c) True: $\left(A^{-1}\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^{-1} = A^{-1}$. (d) True: $\left(A^{\mathsf{T}}A\right)^{-1}A^{\mathsf{T}}$ is a left inverse. (e) False: $A = \begin{pmatrix}1\\1\end{pmatrix}$ over $\mathbb{F}_2$, or $\begin{pmatrix}1\\i\end{pmatrix}$ over $\mathbb{C}$, is left-invertible with $A^{\mathsf{T}}A = 0$.

(a) False. I_n and $-I_n$ are invertible (each is its own inverse), but $I_n + (-I_n) = 0$ has a zero column and is not invertible (Exercise 2.5.9). This works over every field, including $\mathbb{F}_2$, where $-I_n = I_n$.

(b) True. Let $X = \left(A^2\right)^{-1}$, so $XA^2 = A^2X = I_n$. Then $(XA)A = XA^2 = I_n$, so XA is a left inverse of A, and $A(AX) = A^2X = I_n$, so AX is a right inverse of A. By Theorem 2.5.10, A is invertible (and $XA = AX = A^{-1}$).

(c) True. By Proposition 2.5.17(d) and $A^{\mathsf{T}} = A$, $\left(A^{-1}\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^{-1} = A^{-1}$.

(d) True. Let $X = \left(A^{\mathsf{T}}A\right)^{-1} \in F^{n\times n}$. Then $\left(XA^{\mathsf{T}}\right)A = X\left(A^{\mathsf{T}}A\right) = I_n$, so $XA^{\mathsf{T}} \in F^{n\times m}$ is a left inverse of A.

(e) False. Over $\mathbb{F}_2$, let $A = \begin{pmatrix}1\\1\end{pmatrix}$. It is left-invertible, with left inverse $\begin{pmatrix}1 & 0\end{pmatrix}$, but $A^{\mathsf{T}}A = 1 + 1 = 0$ is not invertible. Over $\mathbb{C}$, $A = \begin{pmatrix}1\\i\end{pmatrix}$ has the left inverse $\begin{pmatrix}1 & 0\end{pmatrix}$ and $A^{\mathsf{T}}A = 1 + i^2 = 0$. Over $\mathbb{R}$ the statement is true. By positivity of sums of squares, $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$ forces $A\mathbf{x} = \mathbf{0}$ (Section 6.5), hence $\mathbf{x} = \mathbf{0}$ when A is left-invertible. So the square matrix $A^{\mathsf{T}}A$ defines an injective map, and that makes it invertible (Section 3.5).

Solution to Exercise 2.5.36

Uses. Definition 2.5.1 (Left inverse), Proposition 2.4.8 (Four views of the product), Corollary 2.5.11 (Many one-sided inverses), Proposition 2.4.23 (Transpose of a product)
Hint 1. By the row view, $BA = I_2$ means: row 1 of B times A is $\begin{pmatrix}1 & 0\end{pmatrix}$ and row 2 of B times A is $\begin{pmatrix}0 & 1\end{pmatrix}$. Treat the two rows separately.
Hint 2. A row $\begin{pmatrix}r_1 & r_2 & r_3\end{pmatrix}$ times A is $\begin{pmatrix}r_1+r_3 & r_1+r_2\end{pmatrix}$.
Hint 3. For (c), transpose $BA = I_2$.
Sketch. $\begin{pmatrix}r_1 & r_2 & r_3\end{pmatrix}A = \begin{pmatrix}r_1+r_3 & r_1+r_2\end{pmatrix}$. Row 1 of B must be $\begin{pmatrix}t & -t & 1-t\end{pmatrix}$ and row 2 must be $\begin{pmatrix}s & 1-s & -s\end{pmatrix}$, with $s,t \in \mathbb{F}_p$ free: p^2 left inverses. Since $p^2 \geq 4 > 1$, there is no right inverse. C is a right inverse of A^{T} iff C^{T} is a left inverse of A, so A^{T} has p^2 right inverses.

(a) By the row view (Proposition 2.4.8(c)), $BA = I_2$ means that row 1 of B times A is $\begin{pmatrix}1 & 0\end{pmatrix}$ and row 2 of B times A is $\begin{pmatrix}0 & 1\end{pmatrix}$, and the two rows can be chosen independently. For a row $\mathbf{r} = \begin{pmatrix}r_1 & r_2 & r_3\end{pmatrix}$, the row view gives $\mathbf{r}A = r_1\begin{pmatrix}1 & 1\end{pmatrix} + r_2\begin{pmatrix}0 & 1\end{pmatrix} + r_3\begin{pmatrix}1 & 0\end{pmatrix} = \begin{pmatrix}r_1+r_3 & r_1+r_2\end{pmatrix}$. So $\mathbf{r}A = \begin{pmatrix}1 & 0\end{pmatrix}$ means $r_3 = 1 - r_1$ and $r_2 = -r_1$, giving the p rows $\begin{pmatrix}t & -t & 1-t\end{pmatrix}$ with $t \in \mathbb{F}_p$. And $\mathbf{r}A = \begin{pmatrix}0 & 1\end{pmatrix}$ means $r_3 = -r_1$ and $r_2 = 1 - r_1$, giving the p rows $\begin{pmatrix}s & 1-s & -s\end{pmatrix}$. Hence the left inverses are

$$B_{t,s} = \begin{pmatrix} t & -t & 1-t \\ s & 1-s & -s \end{pmatrix} \quad (s,t \in \mathbb{F}_p),$$

and there are p^2 of them.

(b) No. Since $p^2 \geq 4$, A has two different left inverses, so it has no right inverse by Corollary 2.5.11.

(c) By Proposition 2.4.23, $(BA)^{\mathsf{T}} = A^{\mathsf{T}}B^{\mathsf{T}}$, and $I_2^{\mathsf{T}} = I_2$. Since transposing is a bijection (Proposition 2.4.22), $BA = I_2$ if and only if $A^{\mathsf{T}}B^{\mathsf{T}} = I_2$. So C is a right inverse of A^{T} exactly when C^{T} is a left inverse of A, and A^{T} has p^2 right inverses.

Solution to Exercise 2.5.37

Uses. Theorem 2.5.10 (Left and right inverses agree), Definition 2.5.12 (Invertible)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. The equation $ST = \mathrm{id}_V$ also says something about S: on which side is T an inverse of S?

Hint 2. If $RS = \mathrm{id}_W$, then S has a left inverse R and a right inverse T. Apply Theorem 2.5.10 to S.
Sketch. $ST = \mathrm{id}_V$ says that T is a right inverse of S. If S also has a left inverse R, then $R = T$ by Theorem 2.5.10, so $TS = RS = \mathrm{id}_W$ and S is a two-sided inverse of T. Conversely, if T is invertible, then $S = T^{-1}$ and T is a left inverse of S.

Note first that $ST = \mathrm{id}_V$ says that T is a right inverse of S.

Suppose S has a left inverse $R \in \mathrm{L}(V, W)$, so $RS = \mathrm{id}_W$. Then S has the left inverse R and the right inverse T, so $R = T$ by Theorem 2.5.10, applied to S. Hence $TS = RS = \mathrm{id}_W$. Together with $ST = \mathrm{id}_V$, this shows that S is a two-sided inverse of T, so T is invertible and $T^{-1} = S$.

Conversely, suppose T is invertible. By Theorem 2.5.10, its only left inverse is T^{-1}, so $S = T^{-1}$. Then $TS = TT^{-1} = \mathrm{id}_W$, so T is a left inverse of S.

Solution to Exercise 2.5.38

Uses. Proposition 2.1.24 (Algebra of composition), Notation 2.1.18, Definition 2.5.12 (Invertible), Proposition 2.4.14 (Algebra of matrices)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Move $a_0\mathrm{id}_V$ to the other side and factor T out of what remains, on the left and on the right.
Hint 2. With $Q = a_kT^{k-1} + \cdots + a_1\mathrm{id}_V$, both TQ and QT equal $a_kT^k + \cdots + a_1T = -a_0\mathrm{id}_V$.
Hint 3. For A^3, multiply $A^2 - A + I_2 = 0$ by $A + I_2$.
Sketch. (a) With $Q = a_kT^{k-1} + \cdots + a_1\mathrm{id}_V$, distributivity and $TT^j = T^jT = T^{j+1}$ give $TQ = QT = -a_0\mathrm{id}_V$, so $-a_0^{-1}Q$ is a two-sided inverse. (b) $A^2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$, so $A^2 - A + I_2 = 0$, and (a) gives $A^{-1} = I_2 - A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. Then $0 = (A + I_2)\big(A^2 - A + I_2\big) = A^3 + I_2$.

(a) Let $Q = a_kT^{k-1} + \cdots + a_2T + a_1\mathrm{id}_V$. By the distributive laws in Proposition 2.1.24 and $TT^j = T^jT = T^{j+1}$ (Notation 2.1.18),

$$\begin{aligned} TQ = QT &= a_kT^k + \cdots + a_2T^2 + a_1T \\ &= -a_0\mathrm{id}_V, \end{aligned}$$

where the last step is the hypothesis. Multiplying by the scalar $-a_0^{-1}$ (Proposition 2.1.24(d)) gives

$$T\big(-a_0^{-1}Q\big) = \big(-a_0^{-1}Q\big)T = \mathrm{id}_V.$$

So $-a_0^{-1}Q$ is a two-sided inverse of T, and T is invertible with $T^{-1} = -a_0^{-1}Q$.

(b) By the entry formula, $A^2 = \left(\begin{smallmatrix} 1{-}1 & 1{+}0 \\ -1{+}0 & -1{+}0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$, so

$$\begin{aligned} A^2 - A + I_2 = &\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \\ &+ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 0. \end{aligned}$$

The relation has $a_2 = 1$, $a_1 = -1$, $a_0 = 1$, so by (a) (in its matrix form, with the same proof), $A^{-1} = -(A - I_2) = I_2 - A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. Check: $A(I_2 - A) = A - A^2 = \left(\begin{smallmatrix} 1{-}0 & 1{-}1 \\ -1{+}1 & 0{+}1 \end{smallmatrix}\right) = I_2$. Finally,

$$\begin{aligned} 0 &= (A + I_2)\big(A^2 - A + I_2\big) \\ &= A^3 - A^2 + A + A^2 - A + I_2 \\ &= A^3 + I_2, \end{aligned}$$

so $A^3 = -I_2$. In particular $A^6 = I_2$.

Solution to Exercise 2.5.39

Uses. Proposition 2.1.24 (Algebra of composition), Notation 2.1.18, Definition 2.5.12 (Invertible), Example 2.1.19 (Composites)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Multiply $(\mathrm{id} - N)\big(\mathrm{id} + N + \cdots + N^{k-1}\big)$ out. Most terms cancel in pairs.

Hint 2. For $\mathrm{id}+N$, apply the first part to $-N$. For (b), write the matrix as I_3+N with N strictly upper triangular, and compute N^2 and N^3.
Hint 3. For (c), $D^4=0$ on $F[x]_{\le 3}$.
Sketch. (a) The product telescopes to $\mathrm{id}-N^k=\mathrm{id}$, in either order. $\mathrm{id}+N=\mathrm{id}-(-N)$ and $(-N)^k=(-1)^kN^k=0$. (b) With $N=\begin{pmatrix}0&a&b\\0&0&c\\0&0&0\end{pmatrix}$, $N^2=\begin{pmatrix}0&0&ac\\0&0&0\\0&0&0\end{pmatrix}$ and $N^3=0$, so the inverse is $I_3-N+N^2=\begin{pmatrix}1&-a&ac-b\\0&1&-c\\0&0&1\end{pmatrix}$. (c) $(\mathrm{id}+D)^{-1}=\mathrm{id}-D+D^2-D^3$, which sends x^3 to x^3-3x^2+6x-6.

(a) Let $G=\mathrm{id}+N+\cdots+N^{k-1}$. By the distributive laws in Proposition 2.1.24,

$$\begin{aligned}(\mathrm{id}-N)G&=\left(\mathrm{id}+N+\cdots+N^{k-1}\right)\\&\quad-\left(N+N^2+\cdots+N^k\right)\\&=\mathrm{id}-N^k=\mathrm{id},\end{aligned}$$

and in the same way $G(\mathrm{id}-N)=\mathrm{id}-N^k=\mathrm{id}$. So $\mathrm{id}-N$ is invertible with inverse G. Since $(-N)^k=(-1)^kN^k=0$, the same applies to $-N$, so $\mathrm{id}+N=\mathrm{id}-(-N)$ is invertible, with inverse $\mathrm{id}-N+N^2-\cdots+(-1)^{k-1}N^{k-1}$.

(b) The matrix is I_3+N with $N=\begin{pmatrix}0&a&b\\0&0&c\\0&0&0\end{pmatrix}$. By the entry formula, $N^2=\begin{pmatrix}0&0&ac\\0&0&0\\0&0&0\end{pmatrix}$ and $N^3=N^2N=0$. By (a) applied to $-N$,

$$\begin{aligned}(I_3+N)^{-1}&=I_3-N+N^2\\&=\begin{pmatrix}1&-a&ac-b\\0&1&-c\\0&0&1\end{pmatrix}.\end{aligned}$$

Check, first row of the product of the matrix with this inverse: $(1,-a+a,(ac-b)-ac+b)=(1,0,0)$; the other rows are similar.

(c) Each application of D lowers the degree of a nonconstant polynomial and sends constants to 0 (Example 2.1.19), so $D^4=0$ on $F[x]_{\le 3}$. By (a), $\mathrm{id}+D$ is invertible with inverse $\mathrm{id}-D+D^2-D^3$. Applied to x^3:

$$(\mathrm{id}+D)^{-1}\left(x^3\right)=x^3-3x^2+6x-6.$$

Check: $(\mathrm{id}+D)\left(x^3-3x^2+6x-6\right)=\left(x^3-3x^2+6x-6\right)+\left(3x^2-6x+6\right)=x^3$.

Solution to Exercise 2.5.40

Uses. Theorem 2.4.34 (Block multiplication), Proposition 2.5.28 (Inverse of a 2×2 matrix), Definition 2.5.12 (Invertible), Proposition 2.4.14 (Algebra of matrices)
Techniques. Guess the inverse, then check both products (Technique 2.5.16), Block decomposition (Technique 2.4.35)
Hint 1. Multiply M by the candidate on both sides with Theorem 2.4.34, partitioning with sizes p,q.
Hint 2. To find the candidate yourself, solve $\begin{pmatrix}A&X\\0&D\end{pmatrix}\begin{pmatrix}P&Q\\0&S\end{pmatrix}=I$ block by block.
Hint 3. For (b), take $p=2$, $q=1$, $A=\begin{pmatrix}1&2\\1&3\end{pmatrix}$, $X=\begin{pmatrix}1\\0\end{pmatrix}$, $D=(2)$.
Sketch. By blocks, MN has diagonal blocks $AA^{-1}=I_p$ and $DD^{-1}=I_q$ and top right block $-AA^{-1}XD^{-1}+XD^{-1}=0$; NM has top right block $A^{-1}X-A^{-1}XD^{-1}D=0$. (b) $A^{-1}=\begin{pmatrix}3&-2\\-1&1\end{pmatrix}$, $D^{-1}=(1/2)$, and $-A^{-1}XD^{-1}=\begin{pmatrix}-3/2\\1/2\end{pmatrix}$.

(a) Let N be the proposed inverse, partitioned with sizes p,q like M. By Theorem 2.4.34, the four blocks of MN are

$$\begin{aligned}AA^{-1}+X\cdot 0&=I_p,\\A\left(-A^{-1}XD^{-1}\right)+XD^{-1}&=0,\\0\cdot A^{-1}+D\cdot 0&=0,\end{aligned}$$

$$0 \cdot \left(-A^{-1}XD^{-1}\right) + DD^{-1} = I_q,$$

so $MN = I_{p+q}$. In the same way, the blocks of NM are $A^{-1}A = I_p$, then $A^{-1}X + \left(-A^{-1}XD^{-1}\right)D = A^{-1}X - A^{-1}X = 0$, then 0, and $D^{-1}D = I_q$, so $NM = I_{p+q}$. Hence N is a two-sided inverse of M.

(b) Take $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, $X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and $D = (2)$. By Proposition 2.5.28, $\Delta = 3 - 2 = 1$ and $A^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$, and $D^{-1} = (1/2)$. Then

$$\begin{aligned} -A^{-1}XD^{-1} &= -\frac{1}{2}\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -3/2 \\ 1/2 \end{pmatrix}, \end{aligned}$$

so

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -2 & -3/2 \\ -1 & 1 & 1/2 \\ 0 & 0 & 1/2 \end{pmatrix}.$$

Check, row by row against the original: $(3 - 2, -2 + 2, -3/2 + 1 + 1/2) = (1, 0, 0)$ for the first row of the product, and similarly for the others.

Solution to Exercise 2.5.41

Uses. Example 2.1.9 (Maps of polynomials), Proposition 2.2.20 (Linear maps out of $F[x]$), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Corollary 2.2.6 (Agreeing on a spanning list)
Hint 1. M shifts coefficients up by one place, and L shifts them down, discarding a_0.
Hint 2. A left inverse S must satisfy $S\left(x^{k+1}\right) = S\left(M\left(x^k\right)\right) = x^k$ for all $k \geq 0$. What about $S(1)$?
Hint 3. A linear map out of $F[x]$ can take arbitrary values on the monomials (Proposition 2.2.20).
Sketch. (a) $L(M(p)) = p$ and $M(L(p)) = p - a_0$. M has no right inverse since 1 is not a value of M, and L has no left inverse since $L(1) = 0 = L(0)$. (b) A left inverse is determined by $S\left(x^{k+1}\right) = x^k$ and the free choice $S(1) = q \in F[x]$: the left inverses are $S_q(p) = L(p) + a_0 q$, one for each $q \in F[x]$.

(a) For $p = a_0 + a_1 x + \cdots + a_n x^n$, $M(p) = a_0 x + a_1 x^2 + \cdots + a_n x^{n+1}$, whose coefficients, shifted down by one place, are $a_0, \ldots, a_n$. So $L(M(p)) = p$, that is, $LM = \mathrm{id}$. In the other order, $M(L(p)) = a_1 x + \cdots + a_n x^n = p - a_0$, so $ML \neq \mathrm{id}$: it kills the constant term.

Every value of M has constant term 0, so 1 is not a value of M, and M is not surjective. By Proposition 2.5.5, M has no right inverse. Since $L(1) = 0 = L(0)$, L is not injective, so it has no left inverse by Proposition 2.5.4.

(b) Let $S \in \mathrm{L}(F[x])$ with $SM = \mathrm{id}$. Then $S\left(x^{k+1}\right) = S\left(M\left(x^k\right)\right) = x^k$ for every $k \geq 0$, so S agrees with L on $x, x^2, x^3, \ldots$, while $q = S(1)$ is unconstrained. Conversely, for any $q \in F[x]$, Proposition 2.2.20 gives a linear map S_q with $S_q(1) = q$ and $S_q\left(x^{k+1}\right) = x^k$, namely

$$S_q(p) = a_0 q + L(p),$$

where a_0 is the constant term of p. Then $S_q M$ and id agree on every monomial, since $S_q\left(M\left(x^k\right)\right) = S_q\left(x^{k+1}\right) = x^k$, so $S_q M = \mathrm{id}$ by Corollary 2.2.6. Conversely, a left inverse S agrees with S_q for $q = S(1)$ on every monomial, so $S = S_q$, again by Corollary 2.2.6. So the left inverses of M are exactly the maps S_q, one for each $q \in F[x]$, and $S_0 = L$. (Since there are many, Corollary 2.5.11 gives another proof that M has no right inverse.)

Solution to Exercise 2.5.42

Uses. Example 2.5.24 (Differentiation has a right inverse but no left inverse), Example 2.1.9 (Maps of polynomials), Proposition 2.5.5 (Right inverses give existence), Definition 1.2.17 (Evaluation), Proposition 2.1.17 (Compositions of linear maps are linear)

Hint 1. Start from J of Example 2.5.24. Adding a constant to $J(p)$ does not change its derivative. Which constant makes the value at 1 vanish?
Hint 2. For (b), the coefficient of x^{p-1} in $D\left(\sum a_k x^k\right)$ is pa_p, computed in $\mathbb{F}_p$.
Sketch. (a) $R(p) = J(p) - (J(p))(1)$, the antiderivative vanishing at 1; it is linear because evaluation at 1 is linear, and $DR = DJ = \mathrm{id}$. (b) The coefficient of x^{p-1} in $D\left(\sum a_k x^k\right) = \sum k a_k x^{k-1}$ is $p * a_p = 0$. So D is not surjective, and Proposition 2.5.5 rules out a right inverse.

(a) Let J be the antiderivative of Example 2.5.24 and define

$$R(p) = J(p) - (J(p))(1),$$

where the scalar $(J(p))(1)$ is subtracted as a constant polynomial. So $R(p)$ is the antiderivative of p that vanishes at 1. The map $p \mapsto (J(p))(1)$ is the composite of J with evaluation at 1, both linear (Example 2.1.9, Proposition 2.1.17), so R is linear. The subtracted term is a constant polynomial, so its derivative is 0, and $D(R(p)) = D(J(p)) = p$. So R is a right inverse of D, and $(R(p))(1) = (J(p))(1) - (J(p))(1) = 0$. For example, $R(x) = x^2/2 - 1/2$.

(b) By Example 2.1.9, $D\left(\sum_{k=0}^{n} a_k x^k\right) = \sum_{k=1}^{n} (k * a_k) x^{k-1}$, so the coefficient of x^{p-1} in $D(f)$ comes only from $k = p$ and equals $p * a_p = (p * 1)a_p = 0$ in $\mathbb{F}_p$ (it is 0 outright if $n < p$). So no polynomial has derivative x^{p-1}: D is not surjective, and by Proposition 2.5.5 it has no right inverse. Over $\mathbb{Q}$, $x^{p-1} = D(x^p/p)$, and the division by p is exactly what $\mathbb{F}_p$ cannot do.

Solution to Exercise 2.5.43

Uses. Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 1.1.17 (No zero divisors), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
Hint 1. Count the complement: quadruples (a, b, c, d) with $ad = bc$. Group them by the common value $t = ad = bc$.
Hint 2. The equation $ad = 0$ has $2p - 1$ solutions (a, d), and for $t \neq 0$ the equation $ad = t$ has $p - 1$ solutions.
Sketch. By Proposition 2.5.28, A is invertible iff $ad \neq bc$. The number of (a, b, c, d) with $ad = bc = t$ is $(2p-1)^2$ for $t = 0$ and $(p-1)^2$ for each of the $p - 1$ nonzero t. So $(2p-1)^2 + (p-1)^3$ matrices are not invertible, and p^4 minus this is $p^4 - p^3 - p^2 + p = (p^2 - 1)(p^2 - p)$.

By Proposition 2.5.28, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if and only if $ad \neq bc$. We count the matrices with $ad = bc$, grouping them by the common value t.

For $t = 0$: by Proposition 1.1.17, $ad = 0$ means $a = 0$ or $d = 0$, which holds for $p + p - 1 = 2p - 1$ pairs (a, d). Likewise there are $2p - 1$ pairs (b, c) with $bc = 0$, so $(2p-1)^2$ quadruples.

For $t \neq 0$: $ad = t$ forces $a \neq 0$, and then $d = ta^{-1}$ is determined, so there are $p - 1$ pairs (a, d), and likewise $p - 1$ pairs (b, c). That gives $(p-1)^2$ quadruples for each of the $p - 1$ nonzero values of t.

So $(2p-1)^2 + (p-1)^3$ matrices are not invertible. Expanding $(2p-1)^2 = 4p^2 - 4p + 1$ and $(p-1)^3 = p^3 - 3p^2 + 3p - 1$, the number of invertible ones is

$$\begin{aligned} & p^4 - (2p-1)^2 - (p-1)^3 \\ &= p^4 - p^3 - p^2 + p \\ &= (p^2 - 1)(p^2 - p). \end{aligned}$$

For $p = 2$ this is $3 \cdot 2 = 6$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

The factored form $(p^2 - 1)(p^2 - p)$ suggests another count: $p^2 - 1$ choices of a nonzero first column, then $p^2 - p$ choices of a second column that is not a multiple of the first. That count is justified by the

Fundamental Theorems (Theorem 2.7.8), which say that A is invertible exactly when its columns form a basis.

Solution to Exercise 2.5.44

Uses. Proposition 2.4.28 (Cyclicity of the trace), Definition 2.4.26 (Trace), Definition 1.1.28 (Characteristic), Definition 2.5.12 (Invertible)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Take the trace of $AA^{-1} = I_m$ and of $A^{-1}A = I_n$.
Hint 2. Proposition 2.4.28 applies to an $m \times n$ matrix times an $n \times m$ matrix. What is $\operatorname{tr} I_k$?
Sketch. (a) $m * 1 = \operatorname{tr} I_m = \operatorname{tr}(AA^{-1}) = \operatorname{tr}(A^{-1}A) = \operatorname{tr} I_n = n * 1$. In characteristic 0, $m * 1 = n * 1$ forces $m = n$. (b) If $BA = I_n$, then $n * 1 = \operatorname{tr}(BA) = \operatorname{tr}(AB) = m * 1$. (c) In $\mathbb{F}_2$, $m * 1 = n * 1$ whenever m and n have the same parity, so the trace cannot tell $m = 1$ from $n = 3$.

(a) The trace of I_k is $1 + \cdots + 1 = k * 1$. Since $A \in F^{m\times n}$ and $A^{-1} \in F^{n\times m}$, Proposition 2.4.28 gives

$$\begin{aligned} m * 1 = \operatorname{tr} I_m &= \operatorname{tr}\left(AA^{-1}\right) \\ &= \operatorname{tr}\left(A^{-1}A\right) = \operatorname{tr} I_n = n * 1. \end{aligned}$$

Suppose F has characteristic 0 and, say, $m > n$. Then $(m-n) * 1 = m * 1 - n * 1 = 0$ with $m - n \geq 1$, contradicting Definition 1.1.28. So $m = n$.

(b) If $BA = I_n$, then $n * 1 = \operatorname{tr}(BA) = \operatorname{tr}(AB) = \operatorname{tr} I_m = m * 1$ by Proposition 2.4.28, contradicting the hypothesis. So $BA \neq I_n$. For instance, over $\mathbb{Q}$, whenever $AB = I_m$ with A not square, BA is not the identity.

(c) In $\mathbb{F}_2$, $k * 1$ is 0 for even k and 1 for odd k, so the trace only sees the parity of the size. For $m = 1$ and $n = 3$ the argument gives $1 = 1$, no contradiction. Invertible matrices over $\mathbb{F}_2$, and over every field, are nevertheless square (Corollary 3.5.15), but the proof needs rank.

Solution to Exercise 2.5.45

Uses. Proposition 2.4.23 (Transpose of a product), Proposition 2.4.22 (Transpose is a linear involution), Proposition 2.5.17 (Algebra of inverses)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. Transpose the equation $BA = I_n$.
Hint 2. Transposing is a bijection, so $X = Y$ if and only if $X^{\mathsf{T}} = Y^{\mathsf{T}}$.
Sketch. $(BA)^{\mathsf{T}} = A^{\mathsf{T}}B^{\mathsf{T}}$ and $I_n^{\mathsf{T}} = I_n$, and transposing is injective, so $BA = I_n$ iff $A^{\mathsf{T}}B^{\mathsf{T}} = I_n$. The second claim follows since $B \mapsto B^{\mathsf{T}}$ is a bijection $F^{n\times m} \to F^{m\times n}$. Two-sided: apply both one-sided statements, and use $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$.

By Proposition 2.4.23 and Proposition 2.4.22, $(BA)^{\mathsf{T}} = A^{\mathsf{T}}B^{\mathsf{T}}$ and $I_n^{\mathsf{T}} = I_n$. Transposing is a bijection (Proposition 2.4.22), so two matrices are equal if and only if their transposes are. Hence

$$\begin{aligned} BA = I_n \quad &\Longleftrightarrow \quad (BA)^{\mathsf{T}} = I_n^{\mathsf{T}} \\ &\Longleftrightarrow \quad A^{\mathsf{T}}B^{\mathsf{T}} = I_n, \end{aligned}$$

that is, B is a left inverse of A if and only if B^{T} is a right inverse of A^{T}.

If A is left-invertible, with left inverse B, then B^{T} is a right inverse of A^{T}. Conversely, if C is a right inverse of A^{T}, then $B = C^{\mathsf{T}}$ satisfies $B^{\mathsf{T}} = C$, so B is a left inverse of A. The same argument with the roles of the two sides exchanged shows that A is right-invertible if and only if A^{T} is left-invertible. Combining the two, A is invertible if and only if A^{T} is; this is Proposition 2.5.17(d) together with its converse, which follows by applying it to A^{T}, since $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$.

Solution to Exercise 2.5.46

Uses. Proposition 2.4.8 (Four views of the product), Exercise 2.5.38, Proposition 1.1.17 (No zero divisors)
Hint 1. (a) is a direct computation. (b) The relation has constant term $ad - bc$.
Hint 2. (c) $A^{-1} = A$ means $A^2 = I_2$. Combine with (a): $(1 + \Delta)I_2 = (a + d)A$, where $\Delta = ad - bc$. What if $a + d \neq 0$?
Hint 3. If $a + d \neq 0$, then A is a scalar matrix. Otherwise $d = -a$ and $\Delta = -1$; count the triples (a, b, c) with $a^2 + bc = 1$, splitting according to whether $a^2 = 1$.

Sketch. (a) Compute A^2 and subtract. (b) With $\Delta \neq 0$, Exercise 2.5.38 gives $A^{-1} = -\Delta^{-1}(A-(a+d)I_2) = \Delta^{-1}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. (c) Besides $\pm I_2$, the solutions are exactly the A with $d=-a$ and $a^2+bc=1$. For $a=\pm 1$, $bc=0$ has $2p-1$ solutions; for the other $p-2$ values, $bc = 1-a^2 \neq 0$ has $p-1$. Total $2+2(2p-1)+(p-2)(p-1) = p^2+p+2$.

(a) By the entry formula,

$$A^2 = \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix},$$

$$(a+d)A = \begin{pmatrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{pmatrix},$$

so $A^2-(a+d)A = \begin{pmatrix} bc-ad & 0 \\ 0 & bc-ad \end{pmatrix} = -(ad-bc)I_2$, which is the claim.

(b) Let $\Delta = ad-bc \neq 0$ and $t = a+d$. The relation of (a) reads $A^2 - tA + \Delta I_2 = 0$, with $a_2 = 1$, $a_1 = -t$, and $a_0 = \Delta \neq 0$. By Exercise 2.5.38 (in matrix form),

$$\begin{aligned} A^{-1} &= -\Delta^{-1}(A - tI_2) = \Delta^{-1}(tI_2 - A) \\ &= \Delta^{-1}\begin{pmatrix} a+d-a & -b \\ -c & a+d-d \end{pmatrix} \\ &= \frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \end{aligned}$$

(c) $A^{-1} = A$ means $A^2 = I_2$. Let $t = a+d$ and $\Delta = ad-bc$. By (a), $A^2 = tA - \Delta I_2$, so $A^2 = I_2$ if and only if $tA = (1+\Delta)I_2$.

If $t \neq 0$, this says $A = t^{-1}(1+\Delta)I_2$ is a scalar matrix λI_2, and $A^2 = I_2$ means $\lambda^2 = 1$, so $\lambda = \pm 1$ (since $(\lambda-1)(\lambda+1)=0$, Proposition 1.1.17). Both I_2 and $-I_2$ satisfy $A^2 = I_2$, and they are different because p is odd.

If $t = 0$, the condition becomes $(1+\Delta)I_2 = 0$, that is, $\Delta = -1$. With $d=-a$, $\Delta = -a^2-bc$, so the condition is $a^2+bc=1$. Conversely every such A has $t=0$ and $\Delta=-1$, so $A^2 = tA - \Delta I_2 = I_2$. These matrices are not $\pm I_2$, whose traces are $\pm 2 \neq 0$.

It remains to count the triples $(a,b,c) \in \mathbb{F}_p^3$ with $bc = 1-a^2$. The equation $a^2=1$ has exactly the two solutions $a=\pm 1$; for each, $bc=0$ has $2p-1$ solutions. For each of the other $p-2$ values of a, $1-a^2 \neq 0$, so $b \neq 0$ and $c = (1-a^2)b^{-1}$: $p-1$ solutions. The total is

$$\begin{aligned} &2 + 2(2p-1) + (p-2)(p-1) \\ &= 2 + 4p - 2 + p^2 - 3p + 2 \\ &= p^2 + p + 2. \end{aligned}$$

For $p=3$ this gives 14 matrices. Over $\mathbb{F}_2$ the count is different, because $I_2 = -I_2$ there: of the six invertible matrices listed in Exercise 2.5.43, exactly four satisfy $A^2 = I_2$.

Solution to Exercise 2.5.47

Uses. Proposition 2.4.14 (Algebra of matrices), Definition 2.5.12 (Invertible)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Where does the formula come from? Pretend the matrices are numbers and expand $1/(1-ba)$ as a geometric series: $1+ba+baba+\cdots = 1+b(1+ab+abab+\cdots)a$. Then forget the heuristic and verify.
Hint 2. Write $C = (I_m - AB)^{-1}$. In $(I_n - BA)(I_n + BCA)$, collect the terms of the form $B(\cdots)A$.
Hint 3. The bracket you get is $C - I_m - ABC = (I_m - AB)C - I_m$.
Sketch. With $C = (I_m-AB)^{-1}$: $(I_n-BA)(I_n+BCA) = I_n + B[C - I_m - ABC]A = I_n + B[(I_m-AB)C - I_m]A = I_n$, and $(I_n+BCA)(I_n-BA) = I_n + B[C - I_m - CAB]A = I_n + B[C(I_m-AB) - I_m]A = I_n$. (b) is (a) with $m=1$, $B=\mathbf{u}$, $A=-\mathbf{v}^{\mathsf{T}}$.

(a) Let $C = (I_m - AB)^{-1}$, so $(I_m-AB)C = C(I_m-AB) = I_m$. By the distributive and associative laws (Proposition 2.4.14),

$$\begin{aligned}&(I_n - BA)(I_n + BCA)\\&= I_n + BCA - BA - BABCA\\&= I_n + B(C - I_m - ABC)A\\&= I_n + B((I_m - AB)C - I_m)A\\&= I_n + B(I_m - I_m)A = I_n,\end{aligned}$$

and similarly

$$\begin{aligned}&(I_n + BCA)(I_n - BA)\\&= I_n - BA + BCA - BCABA\\&= I_n + B(C - I_m - CAB)A\\&= I_n + B(C(I_m - AB) - I_m)A = I_n.\end{aligned}$$

So $I_n + BCA$ is a two-sided inverse of $I_n - BA$.

(b) Apply (a) with $m = 1$, $B = \mathbf{u} \in F^{n\times 1}$, and $A = -\mathbf{v}^{\mathsf{T}} \in F^{1\times n}$. Then $I_n - BA = I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$, and $I_1 - AB = 1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}$, a 1×1 matrix, which is invertible exactly when it is nonzero. By (a),

$$\begin{aligned}&\left(I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)^{-1}\\&= I_n + \mathbf{u}\left(1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}\right)^{-1}\left(-\mathbf{v}^{\mathsf{T}}\right)\\&= I_n - \left(1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}\right)^{-1}\mathbf{u}\mathbf{v}^{\mathsf{T}}.\end{aligned}$$

The point of (a) is that inverting the $n \times n$ matrix $I_n - BA$ reduces to inverting the $m \times m$ matrix $I_m - AB$, which in (b) is a single number.

Solution to Exercise 2.5.48

Uses. Proposition 2.1.24 (Algebra of composition), Theorem 2.5.10 (Left and right inverses agree), Example 2.5.24 (Differentiation has a right inverse but no left inverse)
Hint 1. Expand $(\mathrm{id} - S)(\mathrm{id} - T)$ and use the hypothesis.
Hint 2. By (a), $\mathrm{id} - T$ is a right inverse of $\mathrm{id} - S$. If $\mathrm{id} - S$ is invertible, then $\mathrm{id} - T$ is its inverse, so the product in the other order is also id. Expand that product.
Hint 3. For (c), you need $\mathrm{id} - S$ and $\mathrm{id} - T$ that are one-sided inverses of each other but not two-sided. Use D and J.
Sketch. (a) $(\mathrm{id} - S)(\mathrm{id} - T) = \mathrm{id} - S - T + ST = \mathrm{id}$. (b) Then $\mathrm{id} - T = (\mathrm{id} - S)^{-1}$ by Theorem 2.5.10, so $(\mathrm{id} - T)(\mathrm{id} - S) = \mathrm{id}$, which expands to $TS = S + T = ST$. (c) $S = \mathrm{id} - D$ and $T = \mathrm{id} - J$: $ST = 2\,\mathrm{id} - D - J + DJ = S + T$, while $TS - ST = JD - DJ = JD - \mathrm{id} \neq 0$.

(a) By Proposition 2.1.24 and the hypothesis,

$$\begin{aligned}(\mathrm{id} - S)(\mathrm{id} - T) &= \mathrm{id} - T - S + ST\\&= \mathrm{id} - T - S + S + T = \mathrm{id}.\end{aligned}$$

(b) By (a), $\mathrm{id} - T$ is a right inverse of $\mathrm{id} - S$. If $\mathrm{id} - S$ is invertible, then by Theorem 2.5.10 its right inverse $\mathrm{id} - T$ is its inverse, so also $(\mathrm{id} - T)(\mathrm{id} - S) = \mathrm{id}$. Expanding, $\mathrm{id} - S - T + TS = \mathrm{id}$, so $TS = S + T = ST$.

(c) Let $D, J \in \mathrm{L}(\mathbb{Q}[x])$ be as in Example 2.5.24, so $DJ = \mathrm{id}$ and $JD \neq \mathrm{id}$. Let $S = \mathrm{id} - D$ and $T = \mathrm{id} - J$. Then

$$\begin{aligned}ST &= (\mathrm{id} - D)(\mathrm{id} - J) = \mathrm{id} - J - D + DJ\\&= 2\,\mathrm{id} - D - J = S + T,\end{aligned}$$

while

$$\begin{aligned}TS &= (\mathrm{id} - J)(\mathrm{id} - D) = \mathrm{id} - D - J + JD\\&= ST + (JD - \mathrm{id}).\end{aligned}$$

Since $JD(1) = 0 \neq 1$, $JD - \mathrm{id} \neq 0$, so $TS \neq ST$. Here $\mathrm{id} - S = D$ has the right inverse $\mathrm{id} - T = J$ but is not invertible, which is exactly the hypothesis of (b) failing.

Solution to Exercise 2.5.49

Uses. Theorem 2.5.10 (Left and right inverses agree), Notation 2.4.16 (Powers of a matrix), Proposition 2.4.14 (Algebra of matrices), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Too many vectors force a dependence (Technique 4.1.30)
Hint 1. (a) First show $AB^k = B^{k-1}$ for $k \geq 1$, then induct on r.
Hint 2. (b) The matrices $I_n, B, \ldots, B^{n^2}$ are dependent. Take a nontrivial relation $\sum c_k B^k = 0$, let r be the smallest index with $c_r \neq 0$, and multiply the relation on the left by A^r.
Hint 3. After multiplying by A^r, the relation reads $c_r I_n + c_{r+1}B + c_{r+2}B^2 + \cdots = 0$. Solve for I_n and factor out B.
Hint 4. (c) B now has a left inverse and a right inverse.
Sketch. (a) $AB^k = (AB)B^{k-1} = B^{k-1}$, and induction on r. (b) Take a nontrivial relation $\sum_{k=r}^{N} c_k B^k = 0$ with $c_r \neq 0$ and multiply by A^r: $\sum_{k=r}^{N} c_k B^{k-r} = 0$, so $I_n = BQ$ with $Q = -c_r^{-1} \sum_{k=r+1}^{N} c_k B^{k-r-1}$. (c) A is a left inverse of B and Q a right inverse, so $A = Q$ by Theorem 2.5.10 and $BA = BQ = I_n$.

(a) For $k \geq 1$, $AB^k = (AB)B^{k-1} = B^{k-1}$ by associativity. We induct on r; for $r = 0$ the claim is $B^k = B^k$. If $A^r B^k = B^{k-r}$ for all $k \geq r$, and $k \geq r+1$, then

$$A^{r+1}B^k = A\left(A^r B^k\right) = AB^{k-r} = B^{k-r-1},$$

using $k - r \geq 1$.

(b) The $n^2 + 1$ matrices $I_n, B, B^2, \ldots, B^{n^2}$ are linearly dependent, so there are $c_0, \ldots, c_{n^2} \in F$, not all 0, with $\sum_{k=0}^{n^2} c_k B^k = 0$ (Theorem 1.5.5). Let r be the smallest index with $c_r \neq 0$, so the relation is $\sum_{k=r}^{n^2} c_k B^k = 0$. Multiplying on the left by A^r and using (a),

$$c_r I_n + c_{r+1}B + c_{r+2}B^2 + \cdots + c_{n^2}B^{n^2-r} = 0.$$

Let $Q = -c_r^{-1}\left(c_{r+1}I_n + c_{r+2}B + \cdots + c_{n^2}B^{n^2-r-1}\right)$, an empty sum being 0. Solving for I_n gives $I_n = BQ$. So Q is a right inverse of B, and it is a linear combination of powers of B.

(c) $AB = I_n$ says that A is a left inverse of B, and by (b), Q is a right inverse of B. By Theorem 2.5.10, $A = Q$, so $BA = BQ = I_n$.

This proves the square case of Theorem 3.5.23 using only that $F^{n\times n}$ is finite-dimensional, and finite dimension is essential. On $\mathbb{Q}[x]$, with $DJ = \mathrm{id}$, the argument breaks down at (b), because the powers of J are linearly independent.

Section 2.6

Solution to Exercise 2.6.10

Uses. Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Definition 1.3.10 (Span)
Hint 1. For the kernel, set the value equal to $\mathbf{0}$ and solve. For the image, write a general value as a linear combination of fixed vectors.
Hint 2. (a) Every value is a multiple of $(1, -2)$. (b) A value (a, b, c) of S satisfies $b = a + c$. (c) $\mathrm{ev}_1(a + bx) = a + b$.
Sketch. (a) $T(x, y) = (x - y)(1, -2)$: kernel $\mathrm{span}((1, 1))$, image $\mathrm{span}((1, -2))$, neither injective nor surjective. (b) $S(x, y) = \mathbf{0}$ forces $x = y = 0$; the image is $\{(a, b, c) \mid b = a + c\}$, which misses $(0, 1, 0)$; injective, not surjective. (c) Kernel $\mathrm{span}(1 - x)$, image F; surjective, not injective.

(a) $T(x, y) = (x - y, -2(x - y)) = (x - y)(1, -2)$. This is $\mathbf{0}$ if and only if $x = y$, so $\ker T = \mathrm{span}((1, 1))$, and $\mathrm{im}\, T = \{t(1, -2) \mid t \in \mathbb{R}\} = \mathrm{span}((1, -2))$, since $x - y$ takes every real value. As $(1, 1) \neq \mathbf{0}$ lies in the kernel, T is not injective, and as $(1, 0)$ is not a multiple of $(1, -2)$, T is not surjective.

(b) If $S(x, y) = \mathbf{0}$, then $x = 0$ and $y = 0$, so $\ker S = \{\mathbf{0}\}$ and S is injective (by definition: $S(x, y) = S(x', y')$ forces $x = x'$ and $y = y'$ from the first and last entries). Every value $(a, b, c) = (x, x + y, y)$ satisfies $b = a + c$, and conversely each $(a, a + c, c)$ equals $S(a, c)$. So $\mathrm{im}\, S = \left\{(a, b, c) \in F^3 \;\middle|\; b = a + c\right\} = \mathrm{span}((1, 1, 0), (0, 1, 1))$. It does not contain $(0, 1, 0)$, so S is not surjective.

(c) $\mathrm{ev}_1(a+bx) = a+b$. This is 0 if and only if $b = -a$, so $\ker \mathrm{ev}_1 = \{a - ax \mid a \in F\} = \mathrm{span}(1-x)$, and ev_1 is not injective. Every $c \in F$ equals $\mathrm{ev}_1(c)$, so $\operatorname{im} \mathrm{ev}_1 = F$ and ev_1 is surjective.

Solution to Exercise 2.6.16

Uses. Theorem 2.6.11 (Injectivity is decided at **0**), Definition 2.6.5 (Kernel)
Techniques. Test injectivity at **0** (Technique 2.6.12)
Hint 1. In each part, use the kernel: suppose the value is **0** and deduce that the input is **0**, or find a nonzero input with value **0**.
Hint 2. (b) If $p = a_0 + \cdots + a_d x^d$ with $a_d \neq 0$, the coefficient of x^{d+1} in $xp + p$ is a_d. (c) If $T \neq 0$, pick $\mathbf{v}$ with $T(\mathbf{v}) \neq \mathbf{0}$ and look at $T(\mathbf{v})$.
Sketch. (a) $T(x,y) = \mathbf{0}$ gives $x = 0$ and then $y = 0$. (b) For $p \neq 0$ of degree d, $T(p)$ has coefficient $a_d \neq 0$ at x^{d+1}. (c) If $T = 0$, any nonzero vector is in the kernel; otherwise $T(\mathbf{v}) \neq \mathbf{0}$ for some $\mathbf{v}$, and $T(T(\mathbf{v})) = \mathbf{0}$.

(a) If $T(x,y) = \mathbf{0}$, the first entry gives $x = 0$, and then the second gives $y = 0$. So $\ker T = \{\mathbf{0}\}$, and T is injective by Theorem 2.6.11. No division was needed, so this holds over every field.

(b) Let $p = a_0 + a_1 x + \cdots + a_d x^d$ with $a_d \neq 0$. Then $xp + p = a_0 + (a_0 + a_1)x + \cdots + (a_{d-1} + a_d)x^d + a_d x^{d+1}$, whose coefficient of x^{d+1} is $a_d \neq 0$. So $T(p) \neq \mathbf{0}$ for every $p \neq \mathbf{0}$, that is, $\ker T = \{\mathbf{0}\}$, and T is injective.

(c) Suppose $V \neq \{\mathbf{0}\}$. If $T = 0$, then every nonzero vector of V lies in $\ker T$. Otherwise there is $\mathbf{v}$ with $\mathbf{w} = T(\mathbf{v}) \neq \mathbf{0}$, and $T(\mathbf{w}) = T^2(\mathbf{v}) = \mathbf{0}$, so $\mathbf{w}$ is a nonzero vector of $\ker T$. In either case $\ker T \neq \{\mathbf{0}\}$, and T is not injective.

Solution to Exercise 2.6.26

Uses. Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Proposition 1.5.15 (Short lists, zeros, and repeats), Proposition 1.5.24 (Distinct degrees), Example 2.1.9 (Maps of polynomials)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. $x \cdot x^k = x^{k+1}$, whose derivative is $(k+1)x^k$.
Hint 2. So $T_c\left(x^k\right) = (k+1-c)x^k$. When is the list $(1-c)1, (2-c)x, (3-c)x^2$ independent, and when does it span?
Sketch. $T_c\left(x^k\right) = (k+1-c)x^k$. If $1-c$, $2-c$, $3-c$ are all nonzero, the values are nonzero multiples of $1, x, x^2$, so they form a basis and T_c is bijective. If one of them is 0, the list of values contains 0 (not injective) and its span misses the corresponding monomial (not surjective). Over $\mathbb{Q}$: bijective iff $c \notin \{1,2,3\}$, otherwise neither. Over $\mathbb{F}_2$ and $\mathbb{F}_3$ the set $\{1,2,3\}$ is all of F, so T_c is never injective or surjective; over $\mathbb{F}_5$ it is bijective iff $c \in \{0,4\}$.

(a) For $k = 0, 1, 2$, $D\left(x \cdot x^k\right) = D\left(x^{k+1}\right) = (k+1)x^k$, so $T_c\left(x^k\right) = (k+1-c)x^k$:

$$T_c(1) = (1-c)1, \quad T_c(x) = (2-c)x, \quad T_c\left(x^2\right) = (3-c)x^2.$$

(b) and **(c)**. Let $\mathcal{B} = 1, x, x^2$. Suppose first that $1-c$, $2-c$, and $3-c$ are all nonzero. Then $T_c(\mathcal{B})$ consists of nonzero multiples of 1, x, and x^2, which have distinct degrees, so it is linearly independent (Proposition 1.5.24). It also spans $F[x]_{\leq 2}$, since $x^k = (k+1-c)^{-1} T_c\left(x^k\right)$ for each k. By Proposition 2.6.22, T_c is injective and surjective.

Now suppose $k+1-c = 0$ for some $k \in \{0,1,2\}$. Then $T_c\left(x^k\right) = \mathbf{0}$, so $T_c(\mathcal{B})$ contains **0** and is linearly dependent (Proposition 1.5.15), and T_c is not injective. Every value of T_c is a combination of the other two monomials, so its coefficient of x^k is 0; hence $x^k \notin \operatorname{im} T_c$ and T_c is not surjective.

So T_c is bijective if $c \notin \{1,2,3\}$ and neither injective nor surjective if $c \in \{1,2,3\}$, where $1, 2, 3$ are read in F. Over $\mathbb{Q}$ these are three distinct values. Over $\mathbb{F}_2$, $\{1,2,3\} = \{1,0,1\} = \mathbb{F}_2$, and over $\mathbb{F}_3$, $\{1,2,3\} = \{1,2,0\} = \mathbb{F}_3$, so in both cases T_c is neither injective nor surjective for every c. Over $\mathbb{F}_5$, $\{1,2,3\}$ leaves $c \in \{0,4\}$, and these are the values for which T_c is bijective.

Solution to Exercise 2.6.33

Uses. Corollary 2.6.31 (Fibers are translates of the kernel), Example 2.6.19 (Computing a null space and a column space)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. The third column minus the first is $(2,5)$. Is there a simpler combination of columns that gives $(1,3)$?
Hint 2. Once you have one solution $\mathbf{x}_0$, every solution is $\mathbf{x}_0 + \mathbf{z}$ with $\mathbf{z} \in \ker T_A = \mathrm{span}((-2,1,0))$.
Sketch. $(1,3) = (3,7) - 2(1,2)$, so $\mathbf{x}_0 = (-2,0,1)$ is a solution. All solutions: $(-2-2t, t, 1)$ for $t \in F$.

The columns are $(1,2)$, $(2,4)$, and $(3,7)$, and $(3,7) - 2(1,2) = (1,3)$. So $\mathbf{x}_0 = (-2,0,1)$ satisfies $A\mathbf{x}_0 = (1,3)$. By Example 2.6.19(a), $\ker T_A = \operatorname{span}((-2,1,0))$, so by Corollary 2.6.31(b) the solutions are exactly

$$(-2,0,1) + t(-2,1,0) = (-2-2t, t, 1), \quad t \in F.$$

As a check, $(-2-2t) + 2t + 3 = 1$ and $2(-2-2t) + 4t + 7 = 3$.

Solution to Exercise 2.6.34

Uses. Definition 2.6.5 (Kernel), Proposition 2.6.18 (The column space is spanned by the columns), Corollary 1.3.20 (Operations that preserve the span), Corollary 2.6.31 (Fibers are translates of the kernel)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. $T(x,y,z) = x(1,0,1) + y(1,1,0) + z(0,1,-1)$. Is one of these three vectors a combination of the other two?
Hint 2. Every value (a,b,c) of T satisfies one linear equation. Find it by eliminating x, y, and z from $a = x + y$, $b = y + z$, $c = x - z$.
Sketch. The kernel is $\operatorname{span}((1,-1,1))$. Since $(0,1,-1) = (1,1,0) - (1,0,1)$, $\operatorname{im} T = \operatorname{span}((1,0,1),(1,1,0)) = \{(a,b,c) \mid a - b - c = 0\}$. So $(1,2,3) \notin \operatorname{im} T$ and $(2,1,1) = T(1,1,0) \in \operatorname{im} T$, with preimages $(1+t, 1-t, t)$.

(a) $T(x,y,z) = \mathbf{0}$ means $x + y = 0$, $y + z = 0$, and $x - z = 0$. The first two give $y = -x$ and $z = x$, and then the third holds. So $\ker T = \{(t,-t,t) \mid t \in \mathbb{R}\} = \operatorname{span}((1,-1,1))$.

T is T_A for the matrix with columns $\mathbf{a}_1 = (1,0,1)$, $\mathbf{a}_2 = (1,1,0)$, and $\mathbf{a}_3 = (0,1,-1) = \mathbf{a}_2 - \mathbf{a}_1$, so by Proposition 2.6.18 and Corollary 1.3.20(c), $\operatorname{im} T = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_2)$. Every value $(a,b,c) = (x+y, y+z, x-z)$ satisfies $a - b - c = 0$. Conversely, if $a - b - c = 0$, then

$$(a-b)\mathbf{a}_1 + b\mathbf{a}_2 = (a, b, a-b) = (a,b,c).$$

So $\operatorname{im} T = \left\{(a,b,c) \in \mathbb{R}^3 \;\middle|\; a - b - c = 0\right\}$, a plane.

(b) Since $1 - 2 - 3 = -4 \neq 0$, $(1,2,3) \notin \operatorname{im} T$. Since $2 - 1 - 1 = 0$, $(2,1,1) \in \operatorname{im} T$, and by the computation above $(2,1,1) = \mathbf{a}_1 + \mathbf{a}_2 = T(1,1,0)$. By Corollary 2.6.31, its preimages are

$$(1,1,0) + t(1,-1,1) = (1+t, 1-t, t), \quad t \in \mathbb{R}.$$

As a check, $T(1+t, 1-t, t) = (2,1,1)$.

Solution to Exercise 2.6.35

Uses. Proposition 2.6.18 (The column space is spanned by the columns), Definition 2.3.8 (Matrix–vector product), Proposition 1.2.35 (Zero products), Corollary 2.6.31 (Fibers are translates of the kernel)
Hint 1. $i(1,i) = (i,-1)$. What is $(1+i)(1,i)$?
Hint 2. Then $A\mathbf{x} = (x_1 + ix_2 + (1+i)x_3)(1,i)$, which is $\mathbf{0}$ exactly when the scalar factor is 0.
Sketch. The columns are $(1,i)$, $i(1,i)$, and $(1+i)(1,i)$, so $\operatorname{im} T_A = \operatorname{span}((1,i))$. $A\mathbf{x} = \mathbf{0}$ iff $x_1 = -ix_2 - (1+i)x_3$, so $\ker T_A = \operatorname{span}((-i,1,0),(-1-i,0,1))$. Neither injective nor surjective. $(1+i, i-1)$ is the third column, so the solutions are $\mathbf{e}_3 + \ker T_A$.

(a) The columns are $\mathbf{a}_1 = (1,i)$, $\mathbf{a}_2 = (i,-1) = i(1,i)$, and $\mathbf{a}_3 = (1+i, i-1) = (1+i)(1,i)$, since $(1+i)i = i + i^2 = i - 1$. By Proposition 2.6.18, $\operatorname{im} T_A = \operatorname{span}((1,i), i(1,i), (1+i)(1,i)) = \operatorname{span}((1,i))$.

(b) By Definition 2.3.8, $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = (x_1 + ix_2 + (1+i)x_3)(1,i)$. Since $(1,i) \neq \mathbf{0}$, this is $\mathbf{0}$ if and only if $x_1 = -ix_2 - (1+i)x_3$ (Proposition 1.2.35). So

$$\ker T_A = \{(-is - (1+i)t, s, t) \mid s,t \in \mathbb{C}\} = \operatorname{span}((-i,1,0),(-1-i,0,1)).$$

(c) T_A is not injective, since $(-i,1,0)$ is a nonzero vector of its kernel. It is not surjective: $(1,0) \notin \operatorname{span}((1,i))$, since $c(1,i) = (1,0)$ would need $c = 1$ and $ci = 0$. Finally, $(1+i, i-1) = \mathbf{a}_3 = A\mathbf{e}_3$, so by Corollary 2.6.31 the solutions of $A\mathbf{x} = (1+i, i-1)$ are

$$\mathbf{e}_3 + s(-i,1,0) + t(-1-i,0,1) = (-is - (1+i)t, s, 1+t), \quad s,t \in \mathbb{C}.$$

Solution to Exercise 2.6.36

Uses. Definition 2.1.21 (Pointwise operations), Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Proposition 1.4.24 (The sum of two subspaces)

Techniques. Fix an arbitrary element and chase it (Technique 1.3.13), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. $(S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$ by definition of the sum of maps.
Hint 2. For (c), try $T = -S$.
Sketch. (a) If $S(\mathbf{v}) = T(\mathbf{v}) = \mathbf{0}$, then $(S+T)(\mathbf{v}) = \mathbf{0}$. (b) $(S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}) \in \operatorname{im} S + \operatorname{im} T$. (c) With $V = W \neq \{\mathbf{0}\}$, $S = \mathrm{id}$, and $T = -\mathrm{id}$: $\ker S \cap \ker T = \{\mathbf{0}\}$ but $\ker(S+T) = V$, and $\operatorname{im}(S+T) = \{\mathbf{0}\}$ but $\operatorname{im} S + \operatorname{im} T = V$.

Recall that $(S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$ (Definition 2.1.21).

(a) If $\mathbf{v} \in \ker S \cap \ker T$, then $(S+T)(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.

(b) For every $\mathbf{v} \in V$, $(S+T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$ with $S(\mathbf{v}) \in \operatorname{im} S$ and $T(\mathbf{v}) \in \operatorname{im} T$, so it lies in $\operatorname{im} S + \operatorname{im} T$.

(c) Let $V = W$ be any nonzero space, $S = \mathrm{id}_V$, and $T = -\mathrm{id}_V$. Then $\ker S \cap \ker T = \{\mathbf{0}\}$, while $S + T = 0$ has kernel $V \neq \{\mathbf{0}\}$. And $\operatorname{im}(S+T) = \{\mathbf{0}\}$, while $\operatorname{im} S + \operatorname{im} T = V + V = V$.

Solution to Exercise 2.6.37

Uses. Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Definition 2.6.5 (Kernel), Definition 2.6.6 (Image)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
Hint 1. (a) Look for two different bijections of F^2. (b) Try $N(x,y) = (y,0)$.
Hint 2. (c) If $\mathbf{w} = P(\mathbf{v})$, then $P(\mathbf{w}) = P^2(\mathbf{v}) = \mathbf{w}$. (d) What is the kernel of the restriction? (e) Look at Example 2.6.4.
Sketch. (a) False: id and the swap $(x,y) \mapsto (y,x)$ on F^2. (b) False: for $N(x,y) = (y,0)$, $\ker N = \operatorname{im} N = \operatorname{span}(\mathbf{e}_1)$. (c) True: $\mathbf{w} \in \operatorname{im} P$ gives $P(\mathbf{w}) = \mathbf{w}$, and $\mathbf{w} \in \ker P$ gives $P(\mathbf{w}) = \mathbf{0}$. (d) True: $\ker(T|_U) = U \cap \ker T = \{\mathbf{0}\}$. (e) False: $P(x,y) = x$.

(a) False. On F^2, let $S = \mathrm{id}$ and $T(x,y) = (y,x)$. Both have kernel $\{\mathbf{0}\}$ and image F^2, since each is its own inverse ($T^2 = \mathrm{id}$), but $S(\mathbf{e}_1) = \mathbf{e}_1 \neq \mathbf{e}_2 = T(\mathbf{e}_1)$.

(b) False. Let $N : F^2 \to F^2$, $N(x,y) = (y,0)$. Then $N(x,y) = \mathbf{0}$ if and only if $y = 0$, so $\ker N = \operatorname{span}(\mathbf{e}_1)$, and the values of N are the vectors $(y,0)$, so $\operatorname{im} N = \operatorname{span}(\mathbf{e}_1)$ as well. Their intersection is $\operatorname{span}(\mathbf{e}_1) \neq \{\mathbf{0}\}$.

(c) True. Let $\mathbf{w} \in \ker P \cap \operatorname{im} P$, say $\mathbf{w} = P(\mathbf{v})$. Then $\mathbf{w} = P(\mathbf{v}) = P^2(\mathbf{v}) = P(\mathbf{w}) = \mathbf{0}$, the last step because $\mathbf{w} \in \ker P$.

(d) True. The restriction is linear, and its kernel is $\{\mathbf{u} \in U \mid T(\mathbf{u}) = \mathbf{0}\} = U \cap \ker T = U \cap \{\mathbf{0}\} = \{\mathbf{0}\}$, by Theorem 2.6.11 applied to T. Applying Theorem 2.6.11 to $T|_U$, it is injective. (Directly: if $\mathbf{u}, \mathbf{u}' \in U$ and $T(\mathbf{u}) = T(\mathbf{u}')$, then $\mathbf{u} = \mathbf{u}'$.)

(e) False. $P : F^2 \to F$, $P(x,y) = x$, is surjective but not injective (Example 2.6.4(b)).

Solution to Exercise 2.6.38

Uses. Definition 2.6.5 (Kernel), Definition 2.6.6 (Image)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Decide what T does to the standard basis. Which basis vectors should be sent to $\mathbf{0}$, and where should the others go so that their values fill up exactly the kernel?
Hint 2. In F^2: kill $\mathbf{e}_1$ and send $\mathbf{e}_2$ to $\mathbf{e}_1$. In F^4: kill $\mathbf{e}_1, \mathbf{e}_2$ and send $\mathbf{e}_3 \mapsto \mathbf{e}_1$, $\mathbf{e}_4 \mapsto \mathbf{e}_2$.
Hint 3. For (c), $T(\mathbf{v}) \in \operatorname{im} T$ for every $\mathbf{v}$. For the converse, try $T = 0$.
Sketch. (a) $T(x,y) = (y,0)$: kernel and image are $\operatorname{span}(\mathbf{e}_1)$. (b) $T(x_1,x_2,x_3,x_4) = (x_3,x_4,0,0)$: kernel and image are $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$. (c) $T(\mathbf{v}) \in \operatorname{im} T = \ker T$ gives $T^2(\mathbf{v}) = \mathbf{0}$; the zero map on F has $T^2 = 0$, kernel F, and image $\{0\}$.

(a) Let $T(x,y) = (y,0)$. As in Exercise 2.6.37(b), $\ker T = \operatorname{im} T = \operatorname{span}(\mathbf{e}_1)$.

(b) Let $T(x_1,x_2,x_3,x_4) = (x_3,x_4,0,0)$. Then $T(\mathbf{x}) = \mathbf{0}$ if and only if $x_3 = x_4 = 0$, so $\ker T = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$. The values of T are the vectors $(a,b,0,0)$ with $a, b \in F$, each of which occurs (take $x_3 = a$, $x_4 = b$), so $\operatorname{im} T = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ as well.

(c) Let $\ker T = \operatorname{im} T$ and $\mathbf{v} \in V$. Then $T(\mathbf{v}) \in \operatorname{im} T = \ker T$, so $T^2(\mathbf{v}) = T(T(\mathbf{v})) = \mathbf{0}$. Hence $T^2 = 0$. The converse fails: the zero map T on F satisfies $T^2 = 0$, but $\ker T = F$ and $\operatorname{im} T = \{0\}$.

No $T \in \mathrm{L}(F^3)$ has $\ker T = \operatorname{im} T$. Rank–nullity (Theorem 4.2.11) gives $\dim \ker T + \dim \operatorname{im} T = 3$, which is impossible if the two dimensions are equal.

Solution to Exercise 2.6.39

Uses. Proposition 2.6.15 (Kernels and images of composites), Proposition 1.4.24 (The sum of two subspaces), Proposition 2.6.8 (Kernels and images are subspaces)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. By Proposition 2.6.15, $\ker S \subseteq \ker(TS)$ and $\operatorname{im}(TS) \subseteq \operatorname{im} T$ always hold, so only one containment is at stake in each part.
Hint 2. (a) If $T(S(\mathbf{u})) = \mathbf{0}$, then $S(\mathbf{u})$ lies in $\ker T \cap \operatorname{im} S$. (b) If $T(\mathbf{v}) = T(S(\mathbf{u}))$, then $\mathbf{v} - S(\mathbf{u}) \in \ker T$.
Sketch. (a) "If": $\mathbf{u} \in \ker(TS)$ puts $S(\mathbf{u})$ in $\ker T \cap \operatorname{im} S = \{\mathbf{0}\}$. "Only if": for $\mathbf{v} = S(\mathbf{u}) \in \ker T$, $\mathbf{u} \in \ker(TS) = \ker S$, so $\mathbf{v} = \mathbf{0}$. (b) "If": for $\mathbf{v} = S(\mathbf{u}) + \mathbf{z}$ with $\mathbf{z} \in \ker T$, $T(\mathbf{v}) = TS(\mathbf{u})$. "Only if": $T(\mathbf{v}) = TS(\mathbf{u})$ for some $\mathbf{u}$, so $\mathbf{v} = S(\mathbf{u}) + (\mathbf{v} - S(\mathbf{u}))$. (c) Take $U = V$ and $S = T$.

(a) By Proposition 2.6.15(a), $\ker S \subseteq \ker(TS)$ always. Suppose $\ker T \cap \operatorname{im} S = \{\mathbf{0}\}$, and let $\mathbf{u} \in \ker(TS)$. Then $S(\mathbf{u}) \in \operatorname{im} S$ and $T(S(\mathbf{u})) = \mathbf{0}$, so $S(\mathbf{u}) \in \ker T \cap \operatorname{im} S = \{\mathbf{0}\}$, and $\mathbf{u} \in \ker S$. Conversely, suppose $\ker(TS) = \ker S$, and let $\mathbf{v} \in \ker T \cap \operatorname{im} S$, say $\mathbf{v} = S(\mathbf{u})$. Then $T(S(\mathbf{u})) = T(\mathbf{v}) = \mathbf{0}$, so $\mathbf{u} \in \ker(TS) = \ker S$, and $\mathbf{v} = S(\mathbf{u}) = \mathbf{0}$.

(b) By Proposition 2.6.15(b), $\operatorname{im}(TS) \subseteq \operatorname{im} T$ always. Suppose $\operatorname{im} S + \ker T = V$, and let $\mathbf{v} \in V$. Write $\mathbf{v} = S(\mathbf{u}) + \mathbf{z}$ with $\mathbf{u} \in U$ and $\mathbf{z} \in \ker T$. Then $T(\mathbf{v}) = T(S(\mathbf{u})) + T(\mathbf{z}) = T(S(\mathbf{u})) \in \operatorname{im}(TS)$. So $\operatorname{im} T \subseteq \operatorname{im}(TS)$. Conversely, suppose $\operatorname{im}(TS) = \operatorname{im} T$, and let $\mathbf{v} \in V$. Then $T(\mathbf{v}) \in \operatorname{im} T = \operatorname{im}(TS)$, so $T(\mathbf{v}) = T(S(\mathbf{u}))$ for some $\mathbf{u} \in U$. The vector $\mathbf{z} = \mathbf{v} - S(\mathbf{u})$ satisfies $T(\mathbf{z}) = \mathbf{0}$, and $\mathbf{v} = S(\mathbf{u}) + \mathbf{z} \in \operatorname{im} S + \ker T$. So $V \subseteq \operatorname{im} S + \ker T$, and the other containment holds because $\operatorname{im} S + \ker T$ is a subspace of V (Proposition 2.6.8, Proposition 1.4.24).

(c) Apply (a) and (b) with $U = V$ and $S = T$, so that $TS = T^2$.

Solution to Exercise 2.6.40

Uses. Proposition 2.6.22 (Testing injectivity and surjectivity on a basis), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Can a linear combination of nonzero vectors with nonzero coefficients be $\mathbf{0}$?
Hint 2. Try $T : F^2 \to F$, $T(x, y) = x + y$, with the standard basis.
Sketch. The step "none of the $T(\mathbf{v}_i)$ is $\mathbf{0}$, so every $a_i = 0$" confuses "nonzero entries" with "linearly independent". $T(x, y) = x + y$ has $T(\mathbf{e}_1) = T(\mathbf{e}_2) = 1 \neq 0$ but $T(1, -1) = 0$. Correct version: T is injective iff $T(\mathcal{B})$ is linearly independent (Proposition 2.6.22(a)).

The error is the step "none of the vectors $T(\mathbf{v}_i)$ is $\mathbf{0}$, so every coefficient a_i must be 0". A linear combination of nonzero vectors can vanish with nonzero coefficients; what the step needs is that $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ is linearly independent (Theorem 1.5.5), and nonzero entries do not guarantee that.

The claim is false. Let $T : F^2 \to F$, $T(x, y) = x + y$, with the standard basis. Then $T(\mathbf{e}_1) = T(\mathbf{e}_2) = 1 \neq 0$, but $T(1, -1) = 0$ with $(1, -1) \neq \mathbf{0}$, so T is not injective.

The correct statement is Proposition 2.6.22(a): T is injective if and only if $T(\mathcal{B})$ is linearly independent. The hypothesis of the claim only says that each one-entry list $T(\mathbf{v}_i)$ is linearly independent.

Solution to Exercise 2.6.41

Uses. Example 2.1.9 (Maps of polynomials), Definition 2.6.5 (Kernel), Definition 2.6.6 (Image)
Hint 1. Write $p = a + bx + cx^2 + dx^3$, so $D(p) = b + 2cx + 3dx^2$, and reduce the coefficients 2 and 3 in each field.
Hint 2. For (c), $D^2\left(x^k\right) = k(k-1)x^{k-2}$, and $k(k-1)$ is always even.
Sketch. (a) Over $\mathbb{F}_2$, $D(p) = b + dx^2$, so $\ker D = \operatorname{span}\left(1, x^2\right) = \operatorname{im} D$. (b) Over $\mathbb{F}_3$, $D(p) = b + 2cx$, so $\ker D = \operatorname{span}\left(1, x^3\right)$ and $\operatorname{im} D = \operatorname{span}(1, x)$; they are not equal. (c) $D^2\left(\sum a_k x^k\right) = \sum k(k-1)a_k x^{k-2}$, and $k(k-1) * 1 = 0$ in $\mathbb{F}_2$.

Write $p = a + bx + cx^2 + dx^3$, so that $D(p) = b + 2cx + 3dx^2$ (Example 2.1.9).

(a) In $\mathbb{F}_2$, $2 = 0$ and $3 = 1$, so $D(p) = b + dx^2$. This is $\mathbf{0}$ if and only if $b = d = 0$, so $\ker D = \left\{a + cx^2 \;\middle|\; a, c \in \mathbb{F}_2\right\} = \operatorname{span}\left(1, x^2\right)$. The values $b + dx^2$, with b, d arbitrary, form $\operatorname{span}\left(1, x^2\right)$ as well. So $\ker D = \operatorname{im} D$.

(b) In $\mathbb{F}_3$, $3 = 0$, so $D(p) = b + 2cx$. Since $2 \neq 0$ in $\mathbb{F}_3$, this is $\mathbf{0}$ if and only if $b = c = 0$, so $\ker D = \operatorname{span}\left(1, x^3\right)$. Every $e + fx$ is a value, namely $D\left(ex + 2fx^2\right)$, because $2 \cdot 2f = 4f = f$ in $\mathbb{F}_3$. So $\operatorname{im} D = \operatorname{span}(1, x)$. These are not equal, since $x^3 \in \ker D$ but $x^3 \notin \operatorname{im} D$. They share the constants.

(c) For $f = \sum_k a_k x^k \in \mathbb{F}_2[x]$, applying the formula for D twice gives $D^2(f) = \sum_{k\geq 2}(k(k-1) * 1)a_k x^{k-2}$. One of k and $k-1$ is even, so $k(k-1) * 1 = 0$ in $\mathbb{F}_2$, and $D^2(f) = 0$. In particular $\operatorname{im} D \subseteq \ker D$ on $\mathbb{F}_2[x]$, which is what made (a) possible.

Solution to Exercise 2.6.42

Uses. Definition 2.6.1 (Injective), Definition 2.6.6 (Image), Proposition 2.6.15 (Kernels and images of composites), Definition 2.1.1 (Linear map)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For each $\mathbf{u} \in U$, the vector $S(\mathbf{u})$ lies in $\operatorname{im} T$. How many $\mathbf{v} \in V$ satisfy $T(\mathbf{v}) = S(\mathbf{u})$?
Hint 2. Define $R(\mathbf{u})$ to be that vector. To prove that R is linear, apply T to $R(a\mathbf{u} + \mathbf{u}')$ and to $aR(\mathbf{u}) + R(\mathbf{u}')$, and use injectivity.
Hint 3. For (b), compute $\operatorname{im}(TR)$ with Proposition 2.6.15(b).
Sketch. Since T is injective and $S(\mathbf{u}) \in \operatorname{im} T$, there is exactly one $\mathbf{v}$ with $T(\mathbf{v}) = S(\mathbf{u})$; call it $R(\mathbf{u})$. Then $TR = S$, and $T(R(a\mathbf{u} + \mathbf{u}')) = S(a\mathbf{u} + \mathbf{u}') = T(aR(\mathbf{u}) + R(\mathbf{u}'))$, so R is linear by injectivity. If $TR' = S$, then $T(R'(\mathbf{u})) = S(\mathbf{u})$ forces $R'(\mathbf{u}) = R(\mathbf{u})$. (b) $TR = S$ gives $\operatorname{im} S = \operatorname{im}(TR) \subseteq \operatorname{im} T$.

(a) *Existence.* Let $\mathbf{u} \in U$. Since $S(\mathbf{u}) \in \operatorname{im} S \subseteq \operatorname{im} T$, there is some $\mathbf{v} \in V$ with $T(\mathbf{v}) = S(\mathbf{u})$, and since T is injective there is only one. Let $R(\mathbf{u})$ be this vector. This defines a function $R : U \to V$ with $T(R(\mathbf{u})) = S(\mathbf{u})$ for every $\mathbf{u}$, that is, $TR = S$. To see that R is linear, let $\mathbf{u}, \mathbf{u}' \in U$ and $a \in F$. Since T and S are linear,

$$T(aR(\mathbf{u}) + R(\mathbf{u}')) = aS(\mathbf{u}) + S(\mathbf{u}') = S(a\mathbf{u} + \mathbf{u}') = T(R(a\mathbf{u} + \mathbf{u}')).$$

As T is injective, $R(a\mathbf{u} + \mathbf{u}') = aR(\mathbf{u}) + R(\mathbf{u}')$.

Uniqueness. If $R' : U \to V$ also satisfies $TR' = S$, then $T(R'(\mathbf{u})) = S(\mathbf{u}) = T(R(\mathbf{u}))$ for every $\mathbf{u}$, so $R'(\mathbf{u}) = R(\mathbf{u})$ by injectivity. Hence $R' = R$.

(b) If $TR = S$ for some linear R, then $\operatorname{im} S = \operatorname{im}(TR) \subseteq \operatorname{im} T$ by Proposition 2.6.15(b). So without $\operatorname{im} S \subseteq \operatorname{im} T$ there is no such R at all.

Solution to Exercise 2.6.43

Uses. Definition 2.6.2 (Surjective), Definition 2.6.5 (Kernel), Proposition 2.6.15 (Kernels and images of composites), Definition 2.1.1 (Linear map)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. If $RT = S$, then $R(T(\mathbf{v})) = S(\mathbf{v})$, so there is only one candidate: for $\mathbf{w} \in W$, pick any $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{w}$ and set $R(\mathbf{w}) = S(\mathbf{v})$.
Hint 2. The danger is that $R(\mathbf{w})$ might depend on the choice of $\mathbf{v}$. If $T(\mathbf{v}) = T(\mathbf{v}') = \mathbf{w}$, where does $\mathbf{v} - \mathbf{v}'$ lie?
Hint 3. For linearity, if $T(\mathbf{v}) = \mathbf{w}$ and $T(\mathbf{v}') = \mathbf{w}'$, then $a\mathbf{v} + \mathbf{v}'$ is an allowed choice for $a\mathbf{w} + \mathbf{w}'$.
Sketch. For $\mathbf{w} \in W$ choose $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{w}$ and put $R(\mathbf{w}) = S(\mathbf{v})$. Another choice $\mathbf{v}'$ has $\mathbf{v} - \mathbf{v}' \in \ker T \subseteq \ker S$, so $S(\mathbf{v}) = S(\mathbf{v}')$: R is well defined. Choosing $a\mathbf{v} + \mathbf{v}'$ for $a\mathbf{w} + \mathbf{w}'$ shows that R is linear, and $R(T(\mathbf{v})) = S(\mathbf{v})$ by construction. Uniqueness: $R'T = RT$ and T surjective give $R' = R$. (b) If $RT = S$, then $\ker T \subseteq \ker(RT) = \ker S$.

(a) *Definition of R.* Let $\mathbf{w} \in W$. Since T is surjective, there is some $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$, and we would like to set $R(\mathbf{w}) = S(\mathbf{v})$. For this to make sense, the value must not depend on which preimage $\mathbf{v}$ we pick. If $T(\mathbf{v}) = T(\mathbf{v}') = \mathbf{w}$, then $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$, so $\mathbf{v} - \mathbf{v}' \in \ker T \subseteq \ker S$, and $S(\mathbf{v}) - S(\mathbf{v}') = S(\mathbf{v} - \mathbf{v}') = \mathbf{0}$. So $S(\mathbf{v})$ is the same for every preimage $\mathbf{v}$ of $\mathbf{w}$, and $R(\mathbf{w}) = S(\mathbf{v})$, for any such $\mathbf{v}$, defines a function $R : W \to U$.

R is linear. Let $\mathbf{w}, \mathbf{w}' \in W$ and $a \in F$, and choose $\mathbf{v}, \mathbf{v}'$ with $T(\mathbf{v}) = \mathbf{w}$ and $T(\mathbf{v}') = \mathbf{w}'$. Then $T(a\mathbf{v} + \mathbf{v}') = a\mathbf{w} + \mathbf{w}'$, so $a\mathbf{v} + \mathbf{v}'$ is a preimage of $a\mathbf{w} + \mathbf{w}'$, and

$$R(a\mathbf{w} + \mathbf{w}') = S(a\mathbf{v} + \mathbf{v}') = aS(\mathbf{v}) + S(\mathbf{v}') = aR(\mathbf{w}) + R(\mathbf{w}').$$

$RT = S$. For $\mathbf{v} \in V$, the vector $\mathbf{v}$ is a preimage of $T(\mathbf{v})$, so $R(T(\mathbf{v})) = S(\mathbf{v})$.

Uniqueness. If $R'T = S$ as well, then $R'(T(\mathbf{v})) = R(T(\mathbf{v}))$ for every $\mathbf{v} \in V$. Every $\mathbf{w} \in W$ is $T(\mathbf{v})$ for some $\mathbf{v}$, so $R' = R$.

(b) If $RT = S$ for some linear R, then $\ker T \subseteq \ker(RT) = \ker S$ by Proposition 2.6.15(a).

The construction is the general pattern for defining a map on W through a surjection T: define it on preimages, and check that the result does not depend on the choice. Since the preimages of $\mathbf{w}$ form the fiber $T^{-1}(\mathbf{w})$, a translate of $\ker T$ (Corollary 2.6.31), the check is exactly that S is constant on translates of $\ker T$, which is what $\ker T \subseteq \ker S$ says. The same pattern defines the maps out of quotient spaces in Section 4.6.

Solution to Exercise 2.6.44

Uses. Corollary 2.6.31 (Fibers are translates of the kernel)
Hint 1. By Corollary 2.6.31, V is the disjoint union of the fibers $T^{-1}(\mathbf{w})$ over $\mathbf{w} \in \operatorname{im} T$. How many elements does each fiber have?
Hint 2. For (b), $\left|\mathbb{F}_p^n\right| = p^n$. What are $|\ker T|$ and $|\operatorname{im} T|$ when T is injective? When it is surjective?
Sketch. (a) The fibers over the $|\operatorname{im} T|$ vectors of $\operatorname{im} T$ partition V, and each is in bijection with $\ker T$, so $|V| = |\operatorname{im} T| \cdot |\ker T|$. (b) Injective: $|\ker T| = 1$, so $p^n = |\operatorname{im} T| \leq p^m$. Surjective: $p^n = |\ker T| p^m \geq p^m$.

(a) By Corollary 2.6.31(c), V is the disjoint union of the fibers $T^{-1}(\mathbf{w})$ with $\mathbf{w} \in \operatorname{im} T$, and by part (b) of that corollary each of these fibers is in bijection with $\ker T$, so it has $|\ker T|$ elements. Since V is finite, so are $\ker T$ and $\operatorname{im} T$, and adding up the sizes of the fibers gives

$$|V| = \sum_{\mathbf{w} \in \operatorname{im} T} \left|T^{-1}(\mathbf{w})\right| = |\operatorname{im} T| \cdot |\ker T|.$$

(b) $\mathbb{F}_p^n$ has p^n elements. If T is injective, then $\ker T = \{\mathbf{0}\}$, so by (a) $|\operatorname{im} T| = p^n$. Since $\operatorname{im} T \subseteq \mathbb{F}_p^m$, $p^n \leq p^m$, and $n \leq m$. If T is surjective, then $\operatorname{im} T = \mathbb{F}_p^m$, so $p^n = |\ker T| \cdot p^m \geq p^m$, and $n \geq m$.

Part (b) is the "moreover" clause of the Fundamental Theorems over $\mathbb{F}_p$. Over other fields it needs a different argument (Section 3.5). Section 4.2 proves $\dim V = \dim \ker T + \dim \operatorname{im} T$ over every field; over $\mathbb{F}_p$ this is equivalent to (a), because a space of dimension d over $\mathbb{F}_p$ has p^d elements.

Solution to Exercise 2.6.45

Uses. Corollary 2.6.24 (The columns decide), Theorem 1.4.25 (The subspaces of F^2), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.3.16 (The matrix of T_A), Proposition 2.6.18 (The column space is spanned by the columns)
Hint 1. By Corollary 2.6.24, count the matrices $A \in \mathbb{F}_p^{2\times 3}$ whose three columns span $\mathbb{F}_p^2$. It is easier to count the complement.
Hint 2. By Theorem 1.4.25, if the columns do not span $\mathbb{F}_p^2$, their span is $\{\mathbf{0}\}$ or a line. How many lines through $\mathbf{0}$ does $\mathbb{F}_p^2$ have, and how many column triples lie in a given line?
Hint 3. Each line has $p-1$ nonzero vectors, distinct lines share only $\mathbf{0}$, and $\mathbb{F}_p^2$ has p^2-1 nonzero vectors.
Sketch. Count triples of columns that do not span: their span is $\{\mathbf{0}\}$ or one of the $p+1$ lines. All three columns lie in a given line in p^3 ways, and a triple other than the zero triple lies in at most one line. So $1 + (p+1)\left(p^3-1\right)$ triples fail to span, and $p^6 - 1 - (p+1)\left(p^3-1\right) = p^6 - p^4 - p^3 + p = \left(p^3-1\right)\left(p^3-p\right)$ succeed.

The maps T_A with $A \in \mathbb{F}_p^{2\times 3}$ are all the linear maps $\mathbb{F}_p^3 \to \mathbb{F}_p^2$, each exactly once. By Corollary 2.6.24(b), T_A is surjective if and only if the columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ span $\mathbb{F}_p^2$. There are p^6 matrices in all, so we count those whose columns do not span $\mathbb{F}_p^2$.

Lines. By Theorem 1.4.25, the span of the columns is $\{\mathbf{0}\}$, a line $\operatorname{span}(\mathbf{v})$ with $\mathbf{v} \neq \mathbf{0}$, or $\mathbb{F}_p^2$. A line $\operatorname{span}(\mathbf{v})$ consists of the p vectors $c\mathbf{v}$, $c \in \mathbb{F}_p$, which are distinct since $\mathbf{v} \neq \mathbf{0}$; so it has $p-1$ nonzero vectors. Two different lines share only $\mathbf{0}$: if $\mathbf{u} \neq \mathbf{0}$ lies on $\operatorname{span}(\mathbf{v})$, then $\mathbf{u} = c\mathbf{v}$ with $c \neq 0$, and $\operatorname{span}(\mathbf{u}) = \operatorname{span}(\mathbf{v})$ by Theorem 1.4.25. So the p^2-1 nonzero vectors of $\mathbb{F}_p^2$ are split among the lines, $p-1$ to a line, and there are $\left(p^2-1\right)/(p-1) = p+1$ lines.

Counting. The columns fail to span $\mathbb{F}_p^2$ exactly when all three lie in a common line (this includes the zero matrix, whose columns lie in every line). For a fixed line there are p^3 choices of three columns in it, one of which is the zero triple. A triple with a nonzero column $\mathbf{a}_j$ lies only in the line $\operatorname{span}(\mathbf{a}_j)$, so it is counted once. Hence the number of matrices whose columns do not span is

$$1 + (p+1)\left(p^3-1\right) = p^4 + p^3 - p,$$

and the number of surjective maps is

$$p^6 - p^4 - p^3 + p = \big(p^3 - 1\big)\big(p^3 - p\big).$$

For $p = 2$ this is $7 \cdot 6 = 42$ of the 64 linear maps $\mathbb{F}_2^3 \to \mathbb{F}_2^2$.

The factored form has its own meaning: $\big(p^3 - 1\big)\big(p^3 - p\big)$ is the number of ways to choose a nonzero first row and then a second row that is not a multiple of it, so it counts the 2×3 matrices with linearly independent rows. That the two counts agree is a first sign that row rank equals column rank (Section 4.3).

Solution to Exercise 2.6.46

Uses. Proposition 2.4.22 (Transpose is a linear involution), Definition 2.4.21 (Transpose), Definition 2.6.5 (Kernel), Definition 2.6.6 (Image)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Transposition is linear and $\big(A^{\mathsf{T}}\big)^{\mathsf{T}} = A$ (Proposition 2.4.22). For (b), compute $\Phi(A)^{\mathsf{T}}$, and for the converse try $A = 2^{-1}B$.
Hint 2. For (c), the (i, j) entry of $\Phi(A)$ over $\mathbb{F}_2$ is $a_{ij} + a_{ji}$. Given a symmetric B with zero diagonal, try A = the part of B strictly below the diagonal.
Sketch. (a) $\Phi(A) = 0$ iff $A^{\mathsf{T}} = A$. (b) $\Phi(A)^{\mathsf{T}} = -\Phi(A)$, and $\Phi(B/2) = B$ for skew-symmetric B. (c) Over $\mathbb{F}_2$, $\Phi(A) = A + A^{\mathsf{T}}$ is symmetric with diagonal entries $2a_{ii} = 0$; conversely $B = L + L^{\mathsf{T}} = \Phi(L)$ for the strictly lower part L of B. Over $\mathbb{F}_2$ skew-symmetric means symmetric, and I_n is symmetric but not in $\operatorname{im}\Phi$.

(a) By Proposition 2.4.22, $(cA + B)^{\mathsf{T}} = cA^{\mathsf{T}} + B^{\mathsf{T}}$, so

$$\Phi(cA + B) = cA + B - cA^{\mathsf{T}} - B^{\mathsf{T}} = c\Phi(A) + \Phi(B).$$

And $\Phi(A) = 0$ if and only if $A^{\mathsf{T}} = A$, that is, A is symmetric.

(b) For every A, $\Phi(A)^{\mathsf{T}} = A^{\mathsf{T}} - A = -\Phi(A)$, so $\operatorname{im}\Phi$ consists of skew-symmetric matrices. Conversely, let $B^{\mathsf{T}} = -B$. Since $2 = 1 + 1 \neq 0$,

$$\Phi\big(2^{-1}B\big) = 2^{-1}B - 2^{-1}B^{\mathsf{T}} = 2^{-1}B + 2^{-1}B = B.$$

So $\operatorname{im}\Phi$ is the set of skew-symmetric matrices.

(c) Over $\mathbb{F}_2$, $-1 = 1$, so $\Phi(A) = A + A^{\mathsf{T}}$, whose (i, j) entry is $a_{ij} + a_{ji}$. This is symmetric in i and j, and on the diagonal it is $a_{ii} + a_{ii} = 0$. Conversely, let B be symmetric with zero diagonal, and let L have entries $l_{ij} = b_{ij}$ for $i > j$ and $l_{ij} = 0$ for $i \leq j$. The (i, j) entry of $\Phi(L) = L + L^{\mathsf{T}}$ is $l_{ij} + l_{ji}$, which is b_{ij} if $i > j$, $b_{ji} = b_{ij}$ if $i < j$, and $0 = b_{ii}$ if $i = j$. So $\Phi(L) = B$, and $\operatorname{im}\Phi$ is exactly the set of symmetric matrices with zero diagonal.

Over $\mathbb{F}_2$, $B^{\mathsf{T}} = -B$ says the same thing as $B^{\mathsf{T}} = B$, so every symmetric matrix is skew-symmetric. In particular I_n is skew-symmetric, but its diagonal entries are 1, so $I_n \notin \operatorname{im}\Phi$. So the conclusion of (b) fails over $\mathbb{F}_2$, and the image is a proper subset of $\ker\Phi$.

Solution to Exercise 2.6.47

Uses. Proposition 2.4.14 (Algebra of matrices), Definition 2.6.5 (Kernel), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Hint 1. $\mathbf{x} \in \ker T_{I_n - BA}$ means $BA\mathbf{x} = \mathbf{x}$. Compute $AB(A\mathbf{x})$ by regrouping.
Hint 2. For (b), compose the two maps from (a) in both orders: $B(A\mathbf{x}) = BA\mathbf{x}$ and $A(B\mathbf{y}) = AB\mathbf{y}$.
Hint 3. For (c), a bijection between the two kernels matches $\{\mathbf{0}\}$ with $\{\mathbf{0}\}$.
Sketch. (a) If $BA\mathbf{x} = \mathbf{x}$, then $AB(A\mathbf{x}) = A(BA\mathbf{x}) = A\mathbf{x}$; symmetrically for $\mathbf{y}$. (b) On the kernels, $B(A\mathbf{x}) = BA\mathbf{x} = \mathbf{x}$ and $A(B\mathbf{y}) = AB\mathbf{y} = \mathbf{y}$, so the maps are mutually inverse. (c) The bijection sends $\mathbf{0}$ to $\mathbf{0}$, so one kernel is $\{\mathbf{0}\}$ iff the other is.

(a) Let $\mathbf{x} \in \ker T_{I_n - BA}$, that is, $\mathbf{x} - BA\mathbf{x} = \mathbf{0}$, or $BA\mathbf{x} = \mathbf{x}$. By associativity (Proposition 2.4.14),

$$AB(A\mathbf{x}) = A(BA\mathbf{x}) = A\mathbf{x},$$

so $(I_m - AB)(A\mathbf{x}) = \mathbf{0}$, and $A\mathbf{x} \in \ker T_{I_m - AB}$. In the same way, if $AB\mathbf{y} = \mathbf{y}$, then $BA(B\mathbf{y}) = B(AB\mathbf{y}) = B\mathbf{y}$, so $B\mathbf{y} \in \ker T_{I_n - BA}$.

(b) By (a), $f(\mathbf{x}) = A\mathbf{x}$ maps $K_1 = \ker T_{I_n - BA}$ into $K_2 = \ker T_{I_m - AB}$, and $g(\mathbf{y}) = B\mathbf{y}$ maps K_2 into K_1. For $\mathbf{x} \in K_1$, $g(f(\mathbf{x})) = BA\mathbf{x} = \mathbf{x}$, and for $\mathbf{y} \in K_2$, $f(g(\mathbf{y})) = AB\mathbf{y} = \mathbf{y}$. So f and g are inverse bijections between K_1 and K_2.

(c) The bijection f sends $\mathbf{0}$ to $\mathbf{0}$, so $K_1 = \{\mathbf{0}\}$ if and only if $K_2 = \{\mathbf{0}\}$. By Theorem 2.6.11, $T_{I_n - BA}$ is injective if and only if $T_{I_m - AB}$ is.

Compare Exercise 2.5.47, which produces an explicit inverse. In the language of Section 7.2, (b) says that AB and BA have "the same" eigenvectors for the eigenvalue 1; the same argument works for every nonzero eigenvalue.

Solution to Exercise 2.6.48

Uses. Example 2.6.7 (First kernels and images), Example 2.1.11 (Rotations, projections, and reflections), Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 2.5.20 (Inverses and matrices), Definition 2.5.12 (Invertible)
Hint 1. By Example 2.6.7(d), every $\mathbf{v} = (x, y)$ is $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{u}^\perp$ with $\alpha = x\cos\theta + y\sin\theta$, $\beta = y\cos\theta - x\sin\theta$, and $\mathbf{u}^\perp = (-\sin\theta, \cos\theta)$, and $\mathrm{Pr}_\ell(\mathbf{v}) = \alpha\mathbf{u}$.
Hint 2. So $(\mathrm{Pr}_\ell - \mathrm{id})(\mathbf{v}) = -\beta\mathbf{u}^\perp$. Express (b) and (c) through Pr_ℓ. For (d), use Proposition 2.5.28.
Sketch. With $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{u}^\perp$: $(\mathrm{Pr}_\ell - \mathrm{id})(\mathbf{v}) = -\beta\mathbf{u}^\perp$, so kernel ℓ and image $\ell^\perp$. (b) $S_\ell - \mathrm{id} = 2(\mathrm{Pr}_\ell - \mathrm{id})$: same. (c) $S_\ell + \mathrm{id} = 2\,\mathrm{Pr}_\ell$: kernel $\ell^\perp$, image ℓ. (d) $ad - bc = (\cos\varphi - 1)^2 + \sin^2\varphi = 2 - 2\cos\varphi \neq 0$, so $R_\varphi - \mathrm{id}$ is invertible: kernel $\{\mathbf{0}\}$, image $\mathbb{R}^2$.

Write $c = \cos\theta$, $s = \sin\theta$, and $\mathbf{u}^\perp = (-s, c)$. By Example 2.6.7(d), every $\mathbf{v} = (x, y)$ satisfies $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{u}^\perp$ with $\alpha = xc + ys$ and $\beta = yc - xs$, and $\mathrm{Pr}_\ell(\mathbf{v}) = \alpha\mathbf{u}$. Note that $\mathbf{v} = t\mathbf{u}$ gives $\beta = tsc - tcs = 0$, and $\mathbf{v} = t\mathbf{u}^\perp$ gives $\beta = tc^2 + ts^2 = t$.

(a) $(\mathrm{Pr}_\ell - \mathrm{id})(\mathbf{v}) = \alpha\mathbf{u} - \mathbf{v} = -\beta\mathbf{u}^\perp$. This is $\mathbf{0}$ if and only if $\beta = 0$, that is, $\mathbf{v} = \alpha\mathbf{u} \in \ell$; and every vector of ℓ has $\beta = 0$. So $\ker(\mathrm{Pr}_\ell - \mathrm{id}) = \ell$. Every value is a multiple of $\mathbf{u}^\perp$, and $(\mathrm{Pr}_\ell - \mathrm{id})\big(-t\mathbf{u}^\perp\big) = t\mathbf{u}^\perp$, so $\mathrm{im}(\mathrm{Pr}_\ell - \mathrm{id}) = \ell^\perp$.

(b) $S_\ell - \mathrm{id} = 2\,\mathrm{Pr}_\ell - 2\,\mathrm{id} = 2(\mathrm{Pr}_\ell - \mathrm{id})$. Multiplying a map by the nonzero scalar 2 changes neither its kernel nor its image, so the kernel is ℓ and the image is $\ell^\perp$: the reflection fixes ℓ and moves every other vector by a multiple of $\mathbf{u}^\perp$.

(c) $S_\ell + \mathrm{id} = 2\,\mathrm{Pr}_\ell$, so by Example 2.6.7(d) its kernel is $\ker \mathrm{Pr}_\ell = \ell^\perp$ and its image is $\mathrm{im}\,\mathrm{Pr}_\ell = \ell$.

(d) By Example 2.1.11, $R_\varphi - \mathrm{id}$ has matrix $\begin{pmatrix} \cos\varphi - 1 & -\sin\varphi \\ \sin\varphi & \cos\varphi - 1 \end{pmatrix}$, with $ad - bc = (\cos\varphi - 1)^2 + \sin^2\varphi = 2 - 2\cos\varphi$. Since φ is not a multiple of 2π, $\cos\varphi \neq 1$, so this is nonzero, and the matrix is invertible by Proposition 2.5.28, so $R_\varphi - \mathrm{id}$ is invertible (Proposition 2.5.20(c)). An invertible map is bijective, so $\ker\big(R_\varphi - \mathrm{id}\big) = \{\mathbf{0}\}$ and $\mathrm{im}\big(R_\varphi - \mathrm{id}\big) = \mathbb{R}^2$. Geometrically, a nontrivial rotation fixes only the origin.

Solution to Exercise 2.6.49

Uses. Definition 2.6.5 (Kernel), Proposition 2.1.24 (Algebra of composition), Proposition 1.2.35 (Zero products)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Look for the three pieces as combinations of $\mathbf{v}$, $T(\mathbf{v})$, and $T^2(\mathbf{v})$. The piece in $\ker T$ must be killed by T, and $T\big(\mathbf{v} - T^2(\mathbf{v})\big) = T(\mathbf{v}) - T^3(\mathbf{v}) = \mathbf{0}$.
Hint 2. Try $\mathbf{v}_\pm = 2^{-1}\big(T^2(\mathbf{v}) \pm T(\mathbf{v})\big)$. Where do these come from? When $2 \neq 0$, the identity $1 = \big(1 - x^2\big) + \big(x^2 + x\big)/2 + \big(x^2 - x\big)/2$ splits 1 along the roots $0, 1, -1$ of $x^3 - x$: each term is 1 at one root and 0 at the other two.
Hint 3. For (b), suppose $\mathbf{v}_0 + \mathbf{v}_+ + \mathbf{v}_- = \mathbf{0}$ and apply T, then T again.
Sketch. (a) $\mathbf{v}_0 = \mathbf{v} - T^2\mathbf{v}$, $\mathbf{v}_\pm = 2^{-1}\big(T^2\mathbf{v} \pm T\mathbf{v}\big)$; these add up to $\mathbf{v}$, and $T^3 = T$ shows $T\mathbf{v}_0 = \mathbf{0}$ and $T\mathbf{v}_\pm = \pm\mathbf{v}_\pm$. (b) If the three pieces add to $\mathbf{0}$, applying T gives $\mathbf{v}_+ - \mathbf{v}_- = \mathbf{0}$, and applying T twice gives $\mathbf{v}_+ + \mathbf{v}_- = \mathbf{0}$; as $2 \neq 0$, all pieces are $\mathbf{0}$. (c) $A^2 = I$ over $\mathbb{F}_2$, so $A^3 = A$, but $\ker T_A = \{\mathbf{0}\}$ and $\ker T_{A-I} = \ker T_{A+I} = \mathrm{span}(\mathbf{e}_1)$, which do not add up to $\mathbb{F}_2^2$.

(a) Write $T\mathbf{v}$ for $T(\mathbf{v})$, and let

$$\mathbf{v}_0 = \mathbf{v} - T^2\mathbf{v}, \quad \mathbf{v}_+ = 2^{-1}\big(T^2\mathbf{v} + T\mathbf{v}\big), \quad \mathbf{v}_- = 2^{-1}\big(T^2\mathbf{v} - T\mathbf{v}\big).$$

Then $\mathbf{v}_0 + \mathbf{v}_+ + \mathbf{v}_- = \mathbf{v} - T^2\mathbf{v} + T^2\mathbf{v} = \mathbf{v}$. Using $T^3 = T$,

$$T\mathbf{v}_0 = T\mathbf{v} - T^3\mathbf{v} = \mathbf{0},$$

$$T\mathbf{v}_+ = 2^{-1}\big(T^3\mathbf{v} + T^2\mathbf{v}\big) = 2^{-1}\big(T\mathbf{v} + T^2\mathbf{v}\big) = \mathbf{v}_+,$$
$$T\mathbf{v}_- = 2^{-1}\big(T^3\mathbf{v} - T^2\mathbf{v}\big) = 2^{-1}\big(T\mathbf{v} - T^2\mathbf{v}\big) = -\mathbf{v}_-.$$

So $\mathbf{v}_0 \in \ker T$, $\mathbf{v}_+ \in \ker(T - \mathrm{id})$, and $\mathbf{v}_- \in \ker(T + \mathrm{id})$.

(b) It suffices to show that if $\mathbf{z}_0 + \mathbf{z}_+ + \mathbf{z}_- = \mathbf{0}$ with $\mathbf{z}_0 \in \ker T$ and $T\mathbf{z}_\pm = \pm\mathbf{z}_\pm$, then all three are $\mathbf{0}$: the difference of two decompositions of $\mathbf{v}$ is such a decomposition of $\mathbf{0}$, since each of the three kernels is a subspace. Applying T gives $\mathbf{z}_+ - \mathbf{z}_- = \mathbf{0}$, and applying T once more gives $\mathbf{z}_+ + \mathbf{z}_- = \mathbf{0}$. Adding, $2\mathbf{z}_+ = \mathbf{0}$, so $\mathbf{z}_+ = \mathbf{0}$ because $2 \neq 0$ (Proposition 1.2.35). Then $\mathbf{z}_- = \mathbf{0}$, and $\mathbf{z}_0 = \mathbf{0}$.

(c) Over $\mathbb{F}_2$, $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$, so $A^3 = A$. Since A is invertible, $\ker T_A = \{\mathbf{0}\}$. Also $-1 = 1$, so $A - I_2 = A + I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, whose kernel is $\mathrm{span}(\mathbf{e}_1)$. Every sum $\mathbf{v}_0 + \mathbf{v}_+ + \mathbf{v}_-$ as in (a) therefore lies in $\mathrm{span}(\mathbf{e}_1)$, and $\mathbf{e}_2$ has no such decomposition.

The polynomial identity in the second hint is Lagrange interpolation (Proposition 3.4.27) at the points $0, 1, -1$. Decompositions of this kind, along the roots of a polynomial that T satisfies, are the subject of Section 7.4.

Solution to Exercise 2.6.50

> *Uses.* Proposition 1.4.20 (Unions of two subspaces), Proposition 2.6.8 (Kernels and images are subspaces), Proposition 1.2.35 (Zero products)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18), Induction on the length of a list (Technique 1.6.9)
> *Hint 1.* (a) asks for a vector outside $\ker S \cup \ker T$. When can a union of two subspaces be all of V?
> *Hint 2.* (b) $\mathbb{F}_2^2$ has three lines through $\mathbf{0}$.
> *Hint 3.* (c) Induct on k. Given $\mathbf{v}$ that works for $T_1, \ldots, T_{k-1}$ and $\mathbf{u}$ with $T_k(\mathbf{u}) \neq \mathbf{0}$, look at the line $\mathbf{v} + t\mathbf{u}$, $t \in F$. For how many t can $T_i(\mathbf{v} + t\mathbf{u}) = \mathbf{0}$?
> *Sketch.* (a) $\ker S$ and $\ker T$ are proper subspaces, and if their union were V it would be a subspace, so one contains the other (Proposition 1.4.20) and equals V. (b) x, y, $x + y$. (c) Induct: for each $i \leq k$, $T_i(\mathbf{v} + t\mathbf{u}) = T_i(\mathbf{v}) + tT_i(\mathbf{u})$ vanishes for at most one t, since two such t would give $T_i(\mathbf{u}) = \mathbf{0}$ and then $T_i(\mathbf{v}) = \mathbf{0}$. So at most k values of t are bad, and F is infinite.

(a) Since $S \neq 0$ and $T \neq 0$, the kernels $\ker S$ and $\ker T$ are subspaces of V (Proposition 2.6.8) different from V. Suppose $\ker S \cup \ker T = V$. Then the union is a subspace, so by Proposition 1.4.20 one kernel contains the other, and the union is the larger kernel. That kernel would be V, a contradiction. So some $\mathbf{v}$ lies in neither kernel.

(b) Let $T_1(x, y) = x$, $T_2(x, y) = y$, and $T_3(x, y) = x + y$. Then $(0, 0)$ lies in every kernel, $(1, 0) \in \ker T_2$, $(0, 1) \in \ker T_1$, and $(1, 1) \in \ker T_3$, since $1 + 1 = 0$. These are all the vectors of $\mathbb{F}_2^2$.

(c) We induct on k. For $k = 1$, $T_1 \neq 0$ means that $T_1(\mathbf{v}) \neq \mathbf{0}$ for some $\mathbf{v}$. Let $k \geq 2$, and suppose by induction that $\mathbf{v}$ satisfies $T_i(\mathbf{v}) \neq \mathbf{0}$ for $i = 1, \ldots, k-1$. Choose $\mathbf{u}$ with $T_k(\mathbf{u}) \neq \mathbf{0}$, and consider the vectors $\mathbf{v} + t\mathbf{u}$ for $t \in F$.

Claim: for each $i \in \{1, \ldots, k\}$, there is at most one $t \in F$ with $T_i(\mathbf{v} + t\mathbf{u}) = \mathbf{0}$. Suppose $t \neq t'$ both work. Then

$$\mathbf{0} = T_i(\mathbf{v} + t\mathbf{u}) - T_i(\mathbf{v} + t'\mathbf{u}) = (t - t')T_i(\mathbf{u}),$$

so $T_i(\mathbf{u}) = \mathbf{0}$ by Proposition 1.2.35, and then $T_i(\mathbf{v}) = T_i(\mathbf{v} + t\mathbf{u}) = \mathbf{0}$. For $i = k$ this contradicts $T_k(\mathbf{u}) \neq \mathbf{0}$, and for $i < k$ it contradicts $T_i(\mathbf{v}) \neq \mathbf{0}$.

So at most k values of t are excluded. Since F is infinite, some t is not excluded, and $\mathbf{v} + t\mathbf{u}$ is the required vector.

The proof of (c) only needs F to have more than k elements at the kth step, so the statement holds for k maps over any field with more than k elements; (b) shows that some such bound is needed. Every proper subspace U of a finite-dimensional space V is the kernel of a nonzero operator on V (Exercise 2.2.35).

Applied to such operators, (c) says that over an infinite field a finite-dimensional vector space is never the union of finitely many proper subspaces.

Solution to Exercise 2.6.51

Uses. Theorem 1.4.3 (Subspace criterion), Notation 2.6.30 (Preimages), Proposition 1.4.24 (The sum of two subspaces), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Techniques. Double containment (Technique 1.3.14), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. (a) is the proof of Proposition 2.6.8 with $\{\mathbf{0}\}$ replaced by U and V replaced by X.
Hint 2. (c) If $T(\mathbf{v}) = T(\mathbf{x})$ with $\mathbf{x} \in X$, then $\mathbf{v} - \mathbf{x} \in \ker T$.
Sketch. (a) $T(\mathbf{0}) = \mathbf{0}$, and $T(\mathbf{x} + \mathbf{x}') = T(\mathbf{x}) + T(\mathbf{x}')$, $T(a\mathbf{x}) = aT(\mathbf{x})$ give closure for both sets; $\ker T = T^{-1}(\mathbf{0}) \subseteq T^{-1}(U)$. (b) A value $T(\mathbf{v})$ with $T(\mathbf{v}) \in U$ lies in $U \cap \operatorname{im} T$, and $\mathbf{w} = T(\mathbf{v}) \in U$ comes from $\mathbf{v} \in T^{-1}(U)$. (c) $T(\mathbf{x} + \mathbf{z}) = T(\mathbf{x})$ for $\mathbf{z} \in \ker T$, and $T(\mathbf{v}) = T(\mathbf{x})$ gives $\mathbf{v} = \mathbf{x} + (\mathbf{v} - \mathbf{x})$.

(a) $\operatorname{im}(T\restriction_X)$ contains $T(\mathbf{0}) = \mathbf{0}$ (Proposition 2.1.4), and for $\mathbf{x}, \mathbf{x}' \in X$ and $a \in F$, $T(\mathbf{x}) + T(\mathbf{x}') = T(\mathbf{x} + \mathbf{x}')$ and $aT(\mathbf{x}) = T(a\mathbf{x})$ lie in $\operatorname{im}(T\restriction_X)$ because $\mathbf{x} + \mathbf{x}', a\mathbf{x} \in X$. So $\operatorname{im}(T\restriction_X)$ is a subspace by Theorem 1.4.3. Similarly, $T(\mathbf{0}) = \mathbf{0} \in U$, so $\mathbf{0} \in T^{-1}(U)$, and if $T(\mathbf{v}), T(\mathbf{v}') \in U$, then $T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') \in U$ and $T(a\mathbf{v}) = aT(\mathbf{v}) \in U$. So $T^{-1}(U)$ is a subspace. Since $\mathbf{0} \in U$, $\ker T = T^{-1}(\mathbf{0}) \subseteq T^{-1}(U)$.

(b) An element of $\operatorname{im}\left(T\restriction_{T^{-1}(U)}\right)$ has the form $T(\mathbf{v})$ with $T(\mathbf{v}) \in U$, so it lies in U and in $\operatorname{im} T$. Conversely, if $\mathbf{w} \in U \cap \operatorname{im} T$, then $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$, and $T(\mathbf{v}) = \mathbf{w} \in U$ means $\mathbf{v} \in T^{-1}(U)$, so $\mathbf{w} \in \operatorname{im}\left(T\restriction_{T^{-1}(U)}\right)$.

(c) If $\mathbf{x} \in X$ and $\mathbf{z} \in \ker T$, then $T(\mathbf{x} + \mathbf{z}) = T(\mathbf{x}) \in \operatorname{im}(T\restriction_X)$, so $\mathbf{x} + \mathbf{z} \in T^{-1}(\operatorname{im}(T\restriction_X))$. Conversely, if $T(\mathbf{v}) \in \operatorname{im}(T\restriction_X)$, then $T(\mathbf{v}) = T(\mathbf{x})$ for some $\mathbf{x} \in X$, so $\mathbf{z} = \mathbf{v} - \mathbf{x} \in \ker T$, and $\mathbf{v} = \mathbf{x} + \mathbf{z} \in X + \ker T$.

By (b), $\operatorname{im}\left(T\restriction_{T^{-1}(U)}\right) = U$ exactly when $U \subseteq \operatorname{im} T$, and by (c), $T^{-1}(\operatorname{im}(T\restriction_X)) = X$ exactly when $\ker T \subseteq X$. So taking images and preimages are inverse operations between the subspaces of V containing $\ker T$ and the subspaces of $\operatorname{im} T$.

Solution to Exercise 2.6.52

Uses. Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Proposition 2.1.24 (Algebra of composition)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For $\mathbf{v} \in \ker T$, compute $T(S(\mathbf{v}))$ using $TS = ST$. For $\mathbf{w} = T(\mathbf{v})$, rewrite $S(\mathbf{w})$ as a value of T.
Hint 2. For (b), T commutes with $T - \lambda \operatorname{id}$. For (c), use a swap and a projection on F^2.
Sketch. (a) $T(S(\mathbf{v})) = S(T(\mathbf{v})) = S(\mathbf{0}) = \mathbf{0}$, and $S(T(\mathbf{v})) = T(S(\mathbf{v})) \in \operatorname{im} T$. (b) $T(T - \lambda \operatorname{id}) = T^2 - \lambda T = (T - \lambda \operatorname{id})T$. (c) $S(x, y) = (y, x)$ and $T(x, y) = (x, 0)$: $\ker T = \operatorname{span}(\mathbf{e}_2)$, but $S(\mathbf{e}_2) = \mathbf{e}_1 \notin \ker T$.

(a) If $\mathbf{v} \in \ker T$, then $T(S(\mathbf{v})) = S(T(\mathbf{v})) = S(\mathbf{0}) = \mathbf{0}$, so $S(\mathbf{v}) \in \ker T$. If $\mathbf{w} = T(\mathbf{v}) \in \operatorname{im} T$, then $S(\mathbf{w}) = S(T(\mathbf{v})) = T(S(\mathbf{v})) \in \operatorname{im} T$.

(b) By Proposition 2.1.24, $T(T - \lambda \operatorname{id}) = T^2 - \lambda T = (T - \lambda \operatorname{id})T$. Apply (a) with $T - \lambda \operatorname{id}$ in place of T and T in place of S.

(c) On F^2, let $S(x, y) = (y, x)$ and $T(x, y) = (x, 0)$. Then $\ker T = \operatorname{span}(\mathbf{e}_2)$ and $\operatorname{im} T = \operatorname{span}(\mathbf{e}_1)$, but $S(\mathbf{e}_2) = \mathbf{e}_1 \notin \ker T$ and $S(\mathbf{e}_1) = \mathbf{e}_2 \notin \operatorname{im} T$. Indeed $ST(x, y) = (0, x)$ while $TS(x, y) = (y, 0)$.

In the language of Section 7.2, (b) says that eigenspaces of T are invariant under T, and (a) that they are invariant under everything that commutes with T.

Section 2.7

Solution to Exercise 2.7.6

Uses. Proposition 2.5.4 (Left inverses give uniqueness), Definition 2.6.1 (Injective), Definition 2.4.1 (Matrix product)
Techniques. Proving TFAE (Technique 2.7.1)
Hint 1. Prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i). The last step is Proposition 2.5.4.
Hint 2. For (ii) ⇒ (iii), some entry a_k of $\mathbf{a}$ is nonzero. Find a row $\mathbf{c}$ that reads off a_k and divides by it.
Sketch. (i) ⇒ (ii): $T_{\mathbf{a}}(1) = \mathbf{a}$ and $T_{\mathbf{a}}(0) = \mathbf{0}$, so $\mathbf{a} = \mathbf{0}$ would contradict injectivity. (ii) ⇒ (iii): if $a_k \neq 0$, the row $\mathbf{c} = a_k^{-1}\mathbf{e}_k^{\mathsf{T}}$ has $\mathbf{ca} = a_k^{-1}a_k = 1$. (iii) ⇒ (i): left-invertible maps are injective. The construction is (ii) ⇒ (iii).

We prove (i) ⇒ (ii) ⇒ (iii) ⇒ (i).

(i) ⇒ (ii). $T_{\mathbf{a}}(1) = \mathbf{a}$ and $T_{\mathbf{a}}(0) = \mathbf{0}$. If $\mathbf{a} = \mathbf{0}$, these are two different inputs with the same value, so $T_{\mathbf{a}}$ would not be injective (Definition 2.6.1).

(ii) ⇒ (iii). Since $\mathbf{a} \neq \mathbf{0}$, some entry a_k is nonzero. Let $\mathbf{c} \in F^{1\times m}$ be the row with entry a_k^{-1} in position k and 0 elsewhere. By Definition 2.4.1, $\mathbf{ca} = a_k^{-1}a_k = 1$.

(iii) ⇒ (i). If $\mathbf{ca} = 1$, then $T_{\mathbf{c}}$ is a left inverse of $T_{\mathbf{a}}$, since $\mathbf{c}(x\mathbf{a}) = x(\mathbf{ca}) = x$. So $T_{\mathbf{a}}$ is injective by Proposition 2.5.4.

The step (ii) ⇒ (iii) produces the left inverse, so it is the case $n = 1$ of the construction in part 3. Run on the single column $\mathbf{a}$, that construction appends every $\mathbf{e}_i$ except $\mathbf{e}_k$ for the *last* index k with $a_k \neq 0$, and it produces exactly the row $\mathbf{c} = a_k^{-1}\mathbf{e}_k^{\mathsf{T}}$.

Solution to Exercise 2.7.13

> *Uses.* Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 2.5.20 (Inverses and matrices), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.7.8 (Third Fundamental Theorem)
> *Hint 1.* Treat the two cases $ad - bc \neq 0$ and $ad - bc = 0$ separately. In each case, Proposition 2.5.28 tells you which inverses A has.
> *Hint 2.* If $ad - bc = 0$, A has no left inverse. Which condition of Theorem 2.7.2 fails, and what does that say about (I)?
> *Sketch.* If $\Delta = ad - bc \neq 0$, then A is invertible, so T_A is invertible and hence injective and surjective (Theorem 2.7.8). If $\Delta = 0$, then A has no left and no right inverse, so neither does T_A; (V) fails in Theorem 2.7.2 and Theorem 2.7.4, so (I) fails in both: T_A is neither injective nor surjective. In both cases the three statements have the same truth value.

Let $\Delta = ad - bc$. By Proposition 2.5.20(c), A has a left inverse, a right inverse, or an inverse exactly when T_A does.

If $\Delta \neq 0$, then A is invertible by Proposition 2.5.28(a), so T_A is invertible, and by Theorem 2.7.8 it is bijective. So all three statements hold.

If $\Delta = 0$, then by Proposition 2.5.28(b), A has neither a left nor a right inverse, so neither does T_A. Condition (V) of Theorem 2.7.2 fails, so (I) fails: T_A is not injective. Condition (V) of Theorem 2.7.4 fails, so T_A is not surjective. So all three statements fail.

The second case used (I) ⇒ (V) in both theorems, in contrapositive form. Without the constructions of this section, we would only know that A has no inverse on either side.

Solution to Exercise 2.7.15

> *Uses.* Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.4.2 (The matrix of a composite), Proposition 2.5.26 (One-sided inverses of composites)
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* Translate "the columns span" into a condition of Theorem 2.7.4 in which products are easy to handle.
> *Hint 2.* Use (I): $T_{AB} = T_A \circ T_B$. Or use (V): right inverses of a product.
> *Sketch.* By Theorem 2.7.4, T_A and T_B are surjective. $T_{AB} = T_A \circ T_B$ is then surjective: given $\mathbf{y}$, pick $\mathbf{z}$ with $A\mathbf{z} = \mathbf{y}$ and $\mathbf{x}$ with $B\mathbf{x} = \mathbf{z}$. So the columns of AB span F^m.

By Theorem 2.7.4, (III) ⇒ (I), $T_A : F^n \to F^m$ and $T_B : F^p \to F^n$ are surjective. Let $\mathbf{y} \in F^m$. There is $\mathbf{z} \in F^n$ with $A\mathbf{z} = \mathbf{y}$, and then $\mathbf{x} \in F^p$ with $B\mathbf{x} = \mathbf{z}$, so $(AB)\mathbf{x} = A(B\mathbf{x}) = \mathbf{y}$. Thus T_{AB} is surjective, and by (I) ⇒ (III) for AB, the columns of AB span F^m.

Alternatively, by (V), there are $C \in F^{n\times m}$ and $D \in F^{p\times n}$ with $AC = I_m$ and $BD = I_n$. Then $(AB)(DC) = A(BD)C = AC = I_m$ (Proposition 2.5.26(a)), so AB has a right inverse, which is (V) for AB.

Solution to Exercise 2.7.19

> *Uses.* Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.7.8 (Third Fundamental Theorem)
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* For injectivity, look for a nonzero vector in the kernel, or show that there is none. For surjectivity, try to solve $A\mathbf{x} = \mathbf{y}$ for a general $\mathbf{y}$, or find a condition that every value $A\mathbf{x}$ satisfies.
> *Hint 2.* (b) Compare the second entry of $A\mathbf{x}$ with i times the first. (d) Try $\mathbf{x} = (1, 1, 1)$.

Sketch. (a) Every column is a multiple of $(1,2)$: $(2,-1,0) \in \ker$ and $(1,0) \notin \operatorname{im}$, so no theorem applies. (b) $A(x,y) = (x+iy, ix-y, y)$ is $\mathbf{0}$ only for $x = y = 0$, and its second entry is always i times the first; only the first theorem applies. (c) $A\mathbf{x} = \mathbf{y}$ has the unique solution $x_1 = y_2$, $x_2 = y_1 + y_2$; all three apply. (d) $A\mathbf{x} = \mathbf{y}$ is solved by $(y_1 + y_2, y_2, 0)$, and $(1,1,1) \in \ker$; only the second applies.

(a) The columns are $(1,2)$, $2(1,2)$, and $3(1,2)$. So $A(2,-1,0) = 2(1,2) - 2(1,2) = \mathbf{0}$, and condition (IV) of Theorem 2.7.2 fails. Every value $A\mathbf{x}$ is a multiple of $(1,2)$, and $(1,0)$ is not ($t(1,2) = (1,0)$ forces $t = 1$ and $2t = 0$), so condition (I) of Theorem 2.7.4 fails. Since T_A is not injective, Theorem 2.7.8 fails as well.

(b) $A(x,y) = (x+iy, ix-y, y)$. If this is $\mathbf{0}$, the third entry gives $y = 0$ and then the first gives $x = 0$, so (IV) of Theorem 2.7.2 holds. The second entry is $ix - y = i(x+iy)$, i times the first, for every (x,y). The vector $(0,1,0)$ fails this, so it is not a value, and Theorem 2.7.4 and Theorem 2.7.8 fail.

(c) Over $\mathbb{F}_2$, $A\mathbf{x} = \mathbf{y}$ reads $x_1 + x_2 = y_1$, $x_1 = y_2$, whose only solution is $x_1 = y_2$, $x_2 = y_1 - y_2 = y_1 + y_2$. So (II) of Theorem 2.7.8 holds, and all three theorems apply. The inverse is $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$; check: $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = I_2$ in $\mathbb{F}_2$.

(d) Over $\mathbb{F}_3$, $A\mathbf{x} = \mathbf{y}$ reads $x_1 + 2x_2 = y_1$, $x_2 + 2x_3 = y_2$. Taking $x_3 = 0$ gives the solution $(y_1 - 2y_2, y_2, 0) = (y_1 + y_2, y_2, 0)$, so (II) of Theorem 2.7.4 holds. But $A(1,1,1) = (1+2, 1+2) = (0,0)$ in $\mathbb{F}_3$, so (IV) of Theorem 2.7.2 fails, and with it Theorem 2.7.8.

Solution to Exercise 2.7.20

Uses. Corollary 2.7.16 (The Fundamental Theorems for linear maps), Theorem 2.7.8 (Third Fundamental Theorem), Definition 2.3.1 (Matrix of a linear map)
Hint 1. Compute the matrix in the bases $1, x, x^2$ and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, or directly the values on $p = a + bx + cx^2$.
Hint 2. For (b), $(Dp)(1) = b + 2c$. What happens when $2 = 0$ in F?
Sketch. (a) $T(a + bx + cx^2) = (a, b, a+b+c)$, and $T(p) = (y_1, y_2, y_3)$ has the unique solution $a = y_1$, $b = y_2$, $c = y_3 - y_1 - y_2$; so $T^{-1}(\mathbf{y}) = y_1 + y_2 x + (y_3 - y_1 - y_2)x^2$. (b) $U(a + bx + cx^2) = (a, b, b + 2c)$. If $2 \neq 0$, solve uniquely with $c = 2^{-1}(y_3 - y_2)$; if $2 = 0$, then $U(x^2) = \mathbf{0}$, so U is not injective.

Write $p = a + bx + cx^2$, so that $Dp = b + 2cx$.

(a) $T(p) = (a, b, a+b+c)$, and the matrix of T in the bases $1, x, x^2$ and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ has columns $T(1) = (1,0,1)$, $T(x) = (0,1,1)$, $T(x^2) = (0,0,1)$. The equation $T(p) = (y_1, y_2, y_3)$, that is, $[T](a,b,c) = (y_1, y_2, y_3)$, has exactly one solution,

$$a = y_1, \quad b = y_2, \quad c = y_3 - y_1 - y_2.$$

No division occurred, so this holds over every field. By Theorem 2.7.8, (II) $\Rightarrow$ (III), the columns of $[T]$ form a basis of F^3, and by Corollary 2.7.16(c), T is invertible, with

$$T^{-1}(y_1, y_2, y_3) = y_1 + y_2 x + (y_3 - y_1 - y_2)x^2.$$

Check: this polynomial has value y_1 at 0, derivative y_2 at 0, and value $y_1 + y_2 + y_3 - y_1 - y_2 = y_3$ at 1.

(b) $U(p) = (a, b, b + 2c)$. If $2 \neq 0$ in F, then $U(p) = (y_1, y_2, y_3)$ has exactly one solution, $a = y_1$, $b = y_2$, $c = 2^{-1}(y_3 - y_2)$, and as in (a), U is invertible. If $2 = 0$, that is, if F has characteristic 2, then $U(x^2) = (0,0,0)$, so U is not injective (Theorem 2.6.11) and not invertible. So U is invertible exactly when $1 + 1 \neq 0$ in F.

Solution to Exercise 2.7.21

Uses. Corollary 2.6.24 (The columns decide), Example 2.7.5 (A tall matrix)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Proving TFAE (Technique 2.7.1)
Hint 1. Draw the four statements as points and the proved implications as arrows. Can you get from (i) to (iii)?
Hint 2. For a counterexample, look for a map that is injective but not surjective, such as the one in Example 2.7.5.
Sketch. The proof only links (i) with (ii) and (iii) with (iv); no implication connects the two pairs. The claim is false: $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ gives an injective $T_A(x) = (x, 0)$ that is not surjective.

The two equivalences are correct, but they form two separate groups: no proved implication leads from (i) or (ii) to (iii) or (iv), or back. So nothing has been shown about, say, (i) versus (iii).

The claim is in fact false. For $A = \begin{pmatrix}1\\0\end{pmatrix} \in F^{2\times 1}$, $T_A(x) = (x, 0)$ is injective, since $(x, 0) = \mathbf{0}$ forces $x = 0$, but not surjective, since $(0, 1)$ is not a value. Example 2.7.5 is another example. (For square matrices the claim is true, but its proof needs Section 3.5; see Corollary 3.5.25.)

Solution to Exercise 2.7.22

> *Uses.* Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Corollary 2.5.11 (Many one-sided inverses), Proposition 2.5.20 (Inverses and matrices), Proposition 2.6.15 (Kernels and images of composites), Theorem 2.4.2 (The matrix of a composite), Example 2.5.27 (An invertible product of non-invertible factors)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18)
> *Hint 1.* (a) and (c) are false; look for small non-square matrices.
> *Hint 2.* (b) Two different left inverses rule out a right inverse (Corollary 2.5.11). Which theorem turns "no right inverse" into a statement about surjectivity?
> *Hint 3.* (d) $T_{AB} = T_A \circ T_B$. What does Proposition 2.6.15(c) say?
> *Sketch.* (a) False: $A = \begin{pmatrix}1\\0\end{pmatrix}$. (b) True: no right inverse, so by Theorem 2.7.4, (I) $\Rightarrow$ (V), T_A is not surjective. (c) False: $A = \begin{pmatrix}1 & 0\end{pmatrix}$, $B = \begin{pmatrix}1\\0\end{pmatrix}$, $AB = (1)$. (d) True: $T_A \circ T_B$ bijective makes T_A surjective and T_B injective; translate with (I) $\Leftrightarrow$ (III).

(a) False. For $A = \begin{pmatrix}1\\0\end{pmatrix} \in F^{2\times 1}$, $Ax = (x, 0) = \mathbf{0}$ only for $x = 0$, but $Ax = (0, 1)$ has no solution.

(b) True. By Corollary 2.5.11, a matrix with two different left inverses has no right inverse, and by Proposition 2.5.20(c) neither does T_A. So condition (V) of Theorem 2.7.4 fails, and therefore so does (I): T_A is not surjective. This uses the implication (I) $\Rightarrow$ (V) of Theorem 2.7.4.

(c) False. With $A = \begin{pmatrix}1 & 0\end{pmatrix}$ and $B = \begin{pmatrix}1\\0\end{pmatrix}$, $AB = (1) = I_1$, so $T_{AB} = \mathrm{id}_F$ is bijective, but $T_A(x, y) = x$ is not injective and $T_B(x) = (x, 0)$ is not surjective (Example 2.5.27).

(d) True. $T_{AB} = T_A \circ T_B$ by Theorem 2.4.2. Since this composite is injective, T_B is injective, and since it is surjective, T_A is surjective (Proposition 2.6.15(c)). By Theorem 2.7.2 and Theorem 2.7.4, (I) $\Rightarrow$ (III), the columns of B are linearly independent and the columns of A span F^m.

Solution to Exercise 2.7.23

> *Uses.* Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Definition 2.5.12 (Invertible)
> *Techniques.* Proving TFAE (Technique 2.7.1)
> *Hint 1.* (iv) is the strongest statement, so make it the hub: prove that (i) and (ii) each imply (iv), and that (iv) implies the rest.
> *Hint 2.* For (i) $\Rightarrow$ (iv), compute $P(P\mathbf{v} - \mathbf{v})$. For (ii) $\Rightarrow$ (iv), write a vector as $P\mathbf{w}$ and apply P.
> *Sketch.* (iv) $\Rightarrow$ (iii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (ii) hold for every operator. (i) $\Rightarrow$ (iv): $P(P\mathbf{v} - \mathbf{v}) = P^2\mathbf{v} - P\mathbf{v} = \mathbf{0}$, so $P\mathbf{v} = \mathbf{v}$. (ii) $\Rightarrow$ (iv): if $\mathbf{v} = P\mathbf{w}$, then $P\mathbf{v} = P^2\mathbf{w} = P\mathbf{w} = \mathbf{v}$.

We prove (iv) $\Rightarrow$ (iii), (iii) $\Rightarrow$ (i), (iii) $\Rightarrow$ (ii), (i) $\Rightarrow$ (iv), and (ii) $\Rightarrow$ (iv). From each statement one can reach (iv) and from (iv) every statement, so this suffices.

(iv) $\Rightarrow$ (iii). id_V is invertible, with inverse id_V.

(iii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (ii). An invertible map has a left inverse and a right inverse (Definition 2.5.12), so it is injective (Proposition 2.5.4) and surjective (Proposition 2.5.5).

(i) $\Rightarrow$ (iv). Let $\mathbf{v} \in V$. Then $P(P\mathbf{v} - \mathbf{v}) = P^2\mathbf{v} - P\mathbf{v} = P\mathbf{v} - P\mathbf{v} = \mathbf{0}$, so $P\mathbf{v} - \mathbf{v} \in \ker P = \{\mathbf{0}\}$ (Theorem 2.6.11). Hence $P\mathbf{v} = \mathbf{v}$ for every $\mathbf{v}$, that is, $P = \mathrm{id}_V$.

(ii) $\Rightarrow$ (iv). Let $\mathbf{v} \in V$. Since P is surjective, $\mathbf{v} = P\mathbf{w}$ for some $\mathbf{w}$, and $P\mathbf{v} = P^2\mathbf{w} = P\mathbf{w} = \mathbf{v}$. So $P = \mathrm{id}_V$.

No basis was used, so the result holds in every vector space. Without the hypothesis $P^2 = P$, (i) and (ii) are not equivalent for operators on $F[x]$ (Example 2.6.4(e)).

Solution to Exercise 2.7.24

> *Uses.* Definition 2.5.1 (Left inverse), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.6.18 (The column space is spanned by the columns)

Hint 1. Compare the domain of R with the domain that Definition 2.5.1 requires of a left inverse of $T_A : F^n \to F^m$.
Hint 2. To repair the argument, R has to be extended from $\operatorname{im} T_A$ to a linear map on all of F^m. The columns of A form a basis of $\operatorname{im} T_A$. What does part 3 of the proof of Theorem 2.7.2 do with them?
Sketch. A left inverse of T_A must be a linear map $F^m \to F^n$, but R is defined only on the subspace $\operatorname{im} T_A$, which is all of F^m only when T_A is surjective. Repair: the columns form a basis of $\operatorname{im} T_A$; complete them to a basis of F^m and extend R by prescribing values on the added vectors, as in part 3 of the proof of Theorem 2.7.2.

Everything up to the last sentence is correct: T_A, regarded as a map onto its image, is a linear bijection, and its inverse $R : \operatorname{im} T_A \to F^n$ is linear. The error is the conclusion. By Definition 2.5.1, a left inverse of $T_A : F^n \to F^m$ is a linear map $F^m \to F^n$, defined on all of F^m, while R is defined only on $\operatorname{im} T_A$. These agree only when $\operatorname{im} T_A = F^m$, that is, when T_A is also surjective. For example, for $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, R is defined on the x-axis of F^2 and says nothing about $(0, 1)$.

Repair. The map R has to be extended linearly to F^m. Since T_A is injective, the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent (Theorem 2.7.2), and they span $\operatorname{im} T_A$ (Proposition 2.6.18), so they are a basis of $\operatorname{im} T_A$, and $R(\mathbf{a}_j) = \mathbf{e}_j$. Complete them to a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_t}$ of F^m as in part 3 of the proof of Theorem 2.7.2, and let $S : F^m \to F^n$ be the linear map with $S(\mathbf{a}_j) = \mathbf{e}_j$ and $S(\mathbf{e}_{i_s}) = \mathbf{0}$ (Theorem 2.2.1). Then S agrees with R on the basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of $\operatorname{im} T_A$, hence on all of $\operatorname{im} T_A$, and $S(T_A(\mathbf{x})) = R(T_A(\mathbf{x})) = \mathbf{x}$. So S is a left inverse.

Solution to Exercise 2.7.25

Uses. Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Exercise 2.5.45, Proposition 2.5.20 (Inverses and matrices), Proposition 2.4.22 (Transpose is a linear involution)
Hint 1. Which condition of the Fundamental Theorems behaves well under transposition?
Hint 2. Condition (V): $CA = I_n$ if and only if $A^{\mathsf{T}}C^{\mathsf{T}} = I_n$ (Exercise 2.5.45).
Sketch. (a) Columns of A independent $\Leftrightarrow$ A has a left inverse C (Theorem 2.7.2) $\Leftrightarrow$ A^{T} has the right inverse C^{T} $\Leftrightarrow$ the columns of A^{T} span F^n (Theorem 2.7.4). (b) Apply (a) to A^{T}, using $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$.

(a) By Theorem 2.7.2, (III) $\Leftrightarrow$ (V), the columns of A are linearly independent if and only if T_A has a left inverse, which by Proposition 2.5.20(c) means that $CA = I_n$ for some $C \in F^{n\times m}$. By Exercise 2.5.45, $CA = I_n$ if and only if $A^{\mathsf{T}}C^{\mathsf{T}} = I_n$. So A has a left inverse if and only if $A^{\mathsf{T}} \in F^{n\times m}$ has a right inverse, that is, if and only if $T_{A^{\mathsf{T}}}$ has a right inverse (Proposition 2.5.20(c)), which by Theorem 2.7.4, (V) $\Leftrightarrow$ (III), happens if and only if the columns of A^{T} span F^n. These columns are the rows of A.

(b) Apply (a) to A^{T}: the columns of A^{T} are linearly independent if and only if the columns of $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$ span F^m (Proposition 2.4.22). The columns of A^{T} are the rows of A.

Solution to Exercise 2.7.26

Uses. Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.3.16 (The matrix of T_A), Theorem 2.7.2 (First Fundamental Theorem), Proposition 1.5.14 (Appending a vector), Proposition 1.5.15 (Short lists, zeros, and repeats), Corollary 1.5.17 (Sublists), Proposition 1.2.35 (Zero products), Exercise 2.7.25
Hint 1. Every linear map $\mathbb{F}_p^2 \to \mathbb{F}_p^3$ is T_A for exactly one $A \in \mathbb{F}_p^{3\times 2}$. Count the matrices whose two columns are independent.
Hint 2. The first column must be nonzero, and the second must avoid the span of the first. How many vectors does $\operatorname{span}(\mathbf{a}_1)$ contain?
Hint 3. For (b), $A \mapsto A^{\mathsf{T}}$ is a bijection from $\mathbb{F}_p^{3\times 2}$ to $\mathbb{F}_p^{2\times 3}$.
Sketch. By Theorem 2.7.2, count pairs $\mathbf{a}_1, \mathbf{a}_2$ in $\mathbb{F}_p^3$ that are independent: $\mathbf{a}_1 \neq \mathbf{0}$ ($p^3 - 1$ choices) and $\mathbf{a}_2 \notin \operatorname{span}(\mathbf{a}_1)$, which has p elements ($p^3 - p$ choices). (b) By Exercise 2.7.25, T_A is injective iff $T_{A^{\mathsf{T}}}$ is surjective, and transposition is a bijection between the two sets of matrices.

(a) Every linear map $\mathbb{F}_p^2 \to \mathbb{F}_p^3$ is T_A for exactly one $A \in \mathbb{F}_p^{3\times 2}$ (Theorem 2.3.20, Proposition 2.3.16), and by Theorem 2.7.2, T_A is injective if and only if the columns $\mathbf{a}_1, \mathbf{a}_2$ are linearly independent. This holds if and only if $\mathbf{a}_1 \neq \mathbf{0}$ and $\mathbf{a}_2 \notin \operatorname{span}(\mathbf{a}_1)$: if the pair is independent, so is its sublist $\mathbf{a}_1$ (Corollary 1.5.17), so $\mathbf{a}_1 \neq \mathbf{0}$ (Proposition 1.5.15(b)), and then $\mathbf{a}_2 \notin \operatorname{span}(\mathbf{a}_1)$ by Proposition 1.5.14; the converse is Proposition 1.5.15(b) and Proposition 1.5.14 again. There are $p^3 - 1$ choices of $\mathbf{a}_1 \neq \mathbf{0}$. For each, $\operatorname{span}(\mathbf{a}_1) = \left\{ t\mathbf{a}_1 \;\middle|\; t \in \mathbb{F}_p \right\}$ has exactly p elements, since $t\mathbf{a}_1 = t'\mathbf{a}_1$ gives $(t - t')\mathbf{a}_1 = \mathbf{0}$ and hence $t = t'$ (Proposition 1.2.35). So there are $p^3 - p$ choices of $\mathbf{a}_2$, and $\left(p^3 - 1\right)\left(p^3 - p\right)$ injective maps in all.

(b) By Exercise 2.7.25(a), the columns of $A \in \mathbb{F}_p^{3\times 2}$ are independent if and only if the columns of $A^{\mathsf{T}} \in \mathbb{F}_p^{2\times 3}$ span $\mathbb{F}_p^2$, that is, T_A is injective if and only if $T_{A^{\mathsf{T}}}$ is surjective. Since $A \mapsto A^{\mathsf{T}}$ is a bijection $\mathbb{F}_p^{3\times 2} \to \mathbb{F}_p^{2\times 3}$ (it is its own inverse), it matches the injective maps $\mathbb{F}_p^2 \to \mathbb{F}_p^3$ with the surjective maps $\mathbb{F}_p^3 \to \mathbb{F}_p^2$ one to one. So the two counts agree.

Solution to Exercise 2.7.27

Uses. Theorem 2.7.2 (First Fundamental Theorem), Proposition 2.4.23 (Transpose of a product), Definition 2.4.1 (Matrix product), Exercise 2.4.44
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. By Theorem 2.7.2, the claim is that $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = \mathbf{0}$. Multiply $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$ on the left by $\mathbf{x}^{\mathsf{T}}$.
Hint 2. $\mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}A\mathbf{x} = (A\mathbf{x})^{\mathsf{T}}(A\mathbf{x})$ is the sum of the squares of the entries of $A\mathbf{x}$. When does such a sum vanish?
Sketch. (a) True: if $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$, then $0 = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}A\mathbf{x} = \sum_i (A\mathbf{x})_i^2$, so $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x} = \mathbf{0}$. (b) False: $A = \begin{pmatrix}1\\ i\end{pmatrix}$ gives $A^{\mathsf{T}}A = \begin{pmatrix}1 + i^2\end{pmatrix} = (0)$. (c) False: $A = \begin{pmatrix}1\\ 1\end{pmatrix}$ gives $A^{\mathsf{T}}A = (1 + 1) = (0)$.

By Theorem 2.7.2, (III) $\Leftrightarrow$ (IV), the columns of $A^{\mathsf{T}}A$ are independent if and only if $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$, and likewise for A.

(a) True. Let $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$, and put $\mathbf{z} = A\mathbf{x} \in \mathbb{R}^m$. By Proposition 2.4.23, $\mathbf{x}^{\mathsf{T}}A^{\mathsf{T}} = (A\mathbf{x})^{\mathsf{T}} = \mathbf{z}^{\mathsf{T}}$, so

$$0 = \mathbf{x}^{\mathsf{T}}\left(A^{\mathsf{T}}A\mathbf{x}\right) = \mathbf{z}^{\mathsf{T}}\mathbf{z} = z_1^2 + \cdots + z_m^2,$$

where the left side is the 1×1 matrix $\mathbf{x}^{\mathsf{T}}\mathbf{0} = (0)$, read as a number. A sum of squares of real numbers is 0 only if every term is, so $\mathbf{z} = A\mathbf{x} = \mathbf{0}$, and $\mathbf{x} = \mathbf{0}$ because the columns of A are independent.

(b) False. Let $A = \begin{pmatrix}1\\ i\end{pmatrix} \in \mathbb{C}^{2\times 1}$. Its single column is nonzero, hence independent, but $A^{\mathsf{T}}A = (1 \cdot 1 + i \cdot i) = (0)$, whose column is $\mathbf{0}$.

(c) False. Let $A = \begin{pmatrix}1\\ 1\end{pmatrix} \in \mathbb{F}_2^{2\times 1}$. Then $A^{\mathsf{T}}A = (1 + 1) = (0)$.

The proof of (a) breaks down exactly where sums of squares can vanish without each term vanishing, as in Exercise 2.4.44(d). Over $\mathbb{C}$, replacing A^{T} by the conjugate transpose repairs it (Section 6.5).

Solution to Exercise 2.7.28

Uses. Theorem 2.7.8 (Third Fundamental Theorem), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem)
Hint 1. Solve $A_c\mathbf{x} = (a, b, d)$ by hand: express x_2 through x_3 using the second equation, then x_1 through x_3 using the first, and substitute into the third.
Hint 2. You should reach $\left(1 + c^3\right)x_3 = d - ca + c^2b$. If $1 + c^3 = 0$, try $x_3 = 1$ for a vector in the kernel.
Sketch. The system is equivalent to $x_2 = b - cx_3$, $x_1 = a - cb + c^2x_3$, $\left(1 + c^3\right)x_3 = d - ca + c^2b$. If $1 + c^3 \neq 0$ there is exactly one solution, so A_c is invertible. If $1 + c^3 = 0$, then $\left(c^2, -c, 1\right)$ is a nonzero kernel vector and $(0, 0, 1)$ is not a value. Not invertible exactly for: $c = -1$ over $\mathbb{R}$; $c = -1, \left(1 \pm i\sqrt{3}\right)/2$ over $\mathbb{C}$; $c = 3, 5, 6$ over $\mathbb{F}_7$.

(a) The system $A_c\mathbf{x} = (a, b, d)$ is

$$x_1 + cx_2 = a, \quad x_2 + cx_3 = b, \quad cx_1 + x_3 = d.$$

The second equation says $x_2 = b - cx_3$; given this, the first says $x_1 = a - cx_2 = a - cb + c^2x_3$; given both, the third says $c\left(a - cb + c^2x_3\right) + x_3 = d$, that is,

$$\left(1 + c^3\right)x_3 = d - ca + c^2b.$$

So the system is equivalent to these three equations for x_2, x_1, and x_3.

If $1 + c^3 \neq 0$, the last equation determines x_3, and then x_2 and x_1 are determined. So the system has exactly one solution for every right-hand side, which is (II) of Theorem 2.7.8, and A_c is invertible.

If $1 + c^3 = 0$, take $(a, b, d) = (0, 0, 1)$: the last equation reads $0 = 1$, so $(0, 0, 1)$ is not a value, and T_{A_c} is not surjective. Take $(a, b, d) = \mathbf{0}$ and $x_3 = 1$: then $\mathbf{x} = \left(c^2, -c, 1\right)$ is a nonzero solution of $A_c\mathbf{x} = \mathbf{0}$ (check: $c^2 - c^2 = 0$, $-c + c = 0$, $c^3 + 1 = 0$), so T_{A_c} is not injective.

(b) Over $\mathbb{R}$, $c^3 = -1$ only for $c = -1$. Over $\mathbb{C}$, $c^3 + 1 = (c+1)\left(c^2 - c + 1\right)$, so $c = -1$ or $c = \left(1 \pm i\sqrt{3}\right)/2$. Over $\mathbb{F}_7$, the cubes of $0, 1, \ldots, 6$ are $0, 1, 1, 6, 1, 6, 6$ (for instance $3^3 = 27 = 3 \cdot 7 + 6$), so $c^3 = -1 = 6$ exactly for $c = 3, 5, 6$.

Solution to Exercise 2.7.29

Uses. Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.7.8 (Third Fundamental Theorem), Theorem 2.5.10 (Left and right inverses agree), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.5.20 (Inverses and matrices), Definition 2.4.1 (Matrix product), Exercise 2.5.45
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. If A is invertible, every one-sided inverse equals A^{-1} (Theorem 2.5.10). For the converse in (a), show that a left-invertible A with T_A not surjective has at least two left inverses.
Hint 2. Use the construction in part 3 of the proof of Theorem 2.7.2. If the columns do not span F^m, at least one standard basis vector $\mathbf{e}_{i_1}$ is appended, and the value of the left inverse there can be chosen in two ways.
Hint 3. For (b), if $AB = I_m$ and $A\mathbf{z} = \mathbf{0}$ with $\mathbf{z} \neq \mathbf{0}$, change one column of B by $\mathbf{z}$.
Sketch. If A is invertible, every one-sided inverse equals A^{-1}. (a) If A has exactly one left inverse, then T_A is injective. If it were not surjective, completing the columns to a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_t}$ would append $t \geq 1$ vectors, and $S\left(\mathbf{e}_{i_1}\right) = \mathbf{0}$ or $\mathbf{e}_1$ would give two left inverses. So T_A is bijective. (b) If $AB = I_m$ and $\mathbf{z} \neq \mathbf{0}$ lies in $\ker T_A$, then adding $\mathbf{z}$ to the first column of B gives a second right inverse. So a unique right inverse forces T_A to be injective as well as surjective.

By Proposition 2.5.20(c), $C \mapsto T_C$ matches the left (right) inverses of A with the left (right) inverses of T_A, so we may count either.

If A is invertible, then every left inverse C equals A^{-1} by Theorem 2.5.10, since A^{-1} is a right inverse; likewise every right inverse equals A^{-1}. This proves "if" in (a) and (b).

(a) Suppose A has exactly one left inverse. Then T_A has a left inverse, so by Theorem 2.7.2 the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent. Suppose, for contradiction, that T_A is not surjective. Complete the columns to a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_t}$ of F^m as in part 3 of the proof of Theorem 2.7.2. Since the columns do not span F^m (Theorem 2.7.4), $t \geq 1$. By Theorem 2.2.1 there are linear maps $S, S' : F^m \to F^n$ with

$$S\left(\mathbf{a}_j\right) = S'\left(\mathbf{a}_j\right) = \mathbf{e}_j \quad (j = 1, \ldots, n),$$
$$S\left(\mathbf{e}_{i_s}\right) = S'\left(\mathbf{e}_{i_s}\right) = \mathbf{0} \quad (s \geq 2),$$

and $S\left(\mathbf{e}_{i_1}\right) = \mathbf{0}$, $S'\left(\mathbf{e}_{i_1}\right) = \mathbf{e}_1$. Both are left inverses of T_A, by the computation in part 3 of the proof of Theorem 2.7.2, and $S \neq S'$ since $\mathbf{e}_1 \neq \mathbf{0}$. This contradicts uniqueness. So T_A is injective and surjective, and by Theorem 2.7.8, T_A and hence A are invertible.

(b) Suppose A has exactly one right inverse B. Then T_A is surjective (Theorem 2.7.4). Suppose, for contradiction, that T_A is not injective, and let $\mathbf{z} \neq \mathbf{0}$ with $A\mathbf{z} = \mathbf{0}$. Let B' be B with $\mathbf{z}$ added to its first column. By Definition 2.4.1, the columns of AB' are those of $AB = I_m$, except the first, which is $A(\mathbf{b}_1 + \mathbf{z}) = A\mathbf{b}_1 + \mathbf{0} = \mathbf{e}_1$. So $AB' = I_m$ and $B' \neq B$, a contradiction. So T_A is bijective, and A is invertible by Theorem 2.7.8.

Transposing gives a second route to (a): $C \mapsto C^{\mathsf{T}}$ matches the left inverses of A with the right inverses of A^{T} (Exercise 2.5.45), and A is invertible if and only if A^{T} is, so (a) for A is (b) for A^{T}. Done directly, the two parts are not symmetric in difficulty: a second right inverse comes from a single vector of the kernel, while a second left inverse needs the columns to be completed to a basis.

Solution to Exercise 2.7.30

Uses. Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Theorem 2.7.8 (Third Fundamental Theorem), Theorem 2.5.10 (Left and right inverses agree), Proposition 2.5.20 (Inverses and matrices)
Hint 1. $AB = I_m$ says that A is a left inverse of B and B is a right inverse of A. What does each of these say about injectivity or surjectivity?
Hint 2. In (a), T_B is injective and surjective. Apply Theorem 2.7.8, and then Theorem 2.5.10 to compare A with B^{-1}.
Sketch. (a) A is a left inverse of B, so T_B is injective; with surjectivity, B is invertible by Theorem 2.7.8, and the left inverse A equals B^{-1}, so $BA = I_n$. (b) B is a right inverse of A, so T_A is surjective; with injectivity, A is invertible and $B = A^{-1}$.

(a) Since $AB = I_m$, T_A is a left inverse of T_B, so T_B is injective (Proposition 2.5.4). It is also surjective, so by Theorem 2.7.8 it is invertible, and by Proposition 2.5.20(c) so is B. Now A is a left inverse and B^{-1} a right inverse of B, so $A = B^{-1}$ by Theorem 2.5.10, and $BA = BB^{-1} = I_n$.

(b) Since $AB = I_m$, T_B is a right inverse of T_A, so T_A is surjective (Proposition 2.5.5). It is also injective, so A is invertible as in (a). Then B is a right inverse and A^{-1} a left inverse of A, so $B = A^{-1}$, and $BA = A^{-1}A = I_n$.

Solution to Exercise 2.7.31

Uses. Definition 2.4.1 (Matrix product), Proposition 2.4.14 (Algebra of matrices), Definition 2.5.12 (Invertible), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. (a) Every entry of J^2 is a sum of n products $1 \cdot 1$.
Hint 2. (b) Multiply $(J - I_n)\big((n-1)^{-1}J - I_n\big)$ out with (a). Both factors are combinations of J and I_n, so they commute.
Hint 3. (c) Apply A to the vector $\mathbf{1}$ with all entries 1. For surjectivity, add up the entries of $A\mathbf{x}$.
Sketch. (a) Each entry of J^2 is $\sum_k 1 \cdot 1 = n$. (b) $(J - I)\big((n-1)^{-1}J - I\big) = (n-1)^{-1}(n-1)J - J + I = I$, and the factors commute. (c) $A\mathbf{1} = (n-1)\mathbf{1} = \mathbf{0}$, and the entries of $A\mathbf{x}$ add up to $(n-1)(x_1 + \cdots + x_n) = 0$, so $\mathbf{e}_1$ is not a value. (d) $n - 1 = 4$, so A is invertible over $\mathbb{F}_p$ exactly for $p \neq 2$.

(a) By Definition 2.4.1, the (i, j) entry of J^2 is $\sum_{k=1}^{n} 1 \cdot 1 = n$, which is the (i, j) entry of nJ.

(b) Let $B = (n-1)^{-1}J - I_n$. By Proposition 2.4.14 and (a),

$$\begin{aligned} AB &= (n-1)^{-1}J^2 - J - (n-1)^{-1}J + I_n \\ &= (n-1)^{-1}(n-1)J - J + I_n = I_n. \end{aligned}$$

The same expansion gives $BA = I_n$, since J commutes with itself and with I_n. So A is invertible with $A^{-1} = B$ (Definition 2.5.12).

(c) Let $\mathbf{1} \in F^n$ have every entry 1. Each entry of $A\mathbf{1}$ is the sum of the $n - 1$ off-diagonal entries of a row, which is $n - 1 = 0$. So $\mathbf{1}$ is a nonzero vector of $\ker T_A$, and T_A is not injective (Theorem 2.7.2). For $\mathbf{x} \in F^n$ with $s = x_1 + \cdots + x_n$, the ith entry of $A\mathbf{x}$ is $s - x_i$, so the entries of $A\mathbf{x}$ add up to $ns - s = (n-1)s = 0$. The vector $\mathbf{e}_1$, whose entries add up to 1, is therefore not a value, and T_A is not surjective (Theorem 2.7.4).

(d) Here $n - 1 = 4$, which is 0 in $\mathbb{F}_p$ exactly when $p = 2$. So A is invertible over $\mathbb{F}_p$ for every odd prime p, by (b), and not over $\mathbb{F}_2$, by (c).

Solution to Exercise 2.7.32

Uses. Theorem 1.5.5 (Independence is decided at zero), Corollary 1.5.17 (Sublists), Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Proposition 2.5.4 (Left inverses give uniqueness), induction on m
Hint 1. For the inductive step, look at the last entries $c_1, \ldots, c_{m+1}$ of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{m+1} \in F^m$. If they are all 0, the vectors live in a copy of F^{m-1}.
Hint 2. If some last entry is nonzero, say $c_{m+1} \neq 0$, subtract multiples of $\mathbf{v}_{m+1}$ from the others to make their last entries 0. Apply the inductive hypothesis to m of the resulting vectors, viewed in F^{m-1}.
Hint 3. For (b), a surjective T_A has a right inverse B, and T_B is injective.
Sketch. (a) $m = 1$: two scalars a, b satisfy a nontrivial relation. Step: if all last entries vanish, drop them and apply the hypothesis to $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Otherwise, with $c_{m+1} \neq 0$, the vectors $\mathbf{u}_i = \mathbf{v}_i - (c_i/c_{m+1})\mathbf{v}_{m+1}$, $i \leq m$, have last entry 0, so they satisfy a nontrivial relation, which expands to one among the $\mathbf{v}_i$ with some coefficient of $\mathbf{v}_i$, $i \leq m$, nonzero. (b) Independent columns in F^m number at most m. A right inverse $B \in F^{n \times m}$ of A has A as a left inverse, so $T_B : F^m \to F^n$ is injective and $m \leq n$.

(a) *Base case* $m = 1$. Two vectors of $F^1 = F$ are scalars a, b. If $a = 0$, then $1 \cdot a + 0 \cdot b = 0$ is a nontrivial relation; otherwise $b \cdot a + (-a) \cdot b = 0$ is one, since $-a \neq 0$.

Inductive step. Let $m \geq 2$, assume the claim for $m - 1$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_{m+1} \in F^m$. Let c_i be the last entry of $\mathbf{v}_i$. For a vector of F^m with last entry 0, write $\overline{\mathbf{u}} \in F^{m-1}$ for the vector of its first $m - 1$ entries. If $\mathbf{u}_1, \ldots, \mathbf{u}_k \in F^m$ all have last entry 0, then a relation $\sum_i a_i \overline{\mathbf{u}}_i = \mathbf{0}$ in F^{m-1} implies $\sum_i a_i \mathbf{u}_i = \mathbf{0}$ in F^m, since the last entries vanish anyway.

Case 1: all $c_i = 0$. The m vectors $\overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_m \in F^{m-1}$ satisfy a nontrivial relation by the inductive hypothesis, hence so do $\mathbf{v}_1, \ldots, \mathbf{v}_m$, and adding $0 \cdot \mathbf{v}_{m+1}$ gives a nontrivial relation among all $m+1$ vectors.

Case 2: some $c_k \neq 0$. Reordering the list does not affect dependence, so we may assume $c_{m+1} \neq 0$. For $i = 1, \ldots, m$ let $\mathbf{u}_i = \mathbf{v}_i - (c_i/c_{m+1})\mathbf{v}_{m+1}$, whose last entry is $c_i - c_i = 0$. By the inductive hypothesis applied to $\overline{\mathbf{u}}_1, \ldots, \overline{\mathbf{u}}_m \in F^{m-1}$, there are $a_1, \ldots, a_m$, not all 0, with $\sum_{i=1}^{m} a_i \mathbf{u}_i = \mathbf{0}$. Expanding,

$$a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m - \left(\sum_{i=1}^{m} a_i c_i/c_{m+1}\right)\mathbf{v}_{m+1} = \mathbf{0},$$

a relation among $\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}$ in which some a_i with $i \leq m$ is nonzero. So the list is dependent (Theorem 1.5.5).

(b) If T_A is injective, its n columns are linearly independent in F^m (Theorem 2.7.2). If $n > m$, the first $m+1$ columns would be an independent sublist (Corollary 1.5.17) of $m+1$ vectors in F^m, contradicting (a). So $n \leq m$.

If T_A is surjective, it has a right inverse by Theorem 2.7.4, (I) $\Rightarrow$ (V), that is, $AB = I_m$ for some $B \in F^{n\times m}$. Then A is a left inverse of B, so $T_B : F^m \to F^n$ is injective (Proposition 2.5.4), and by the first part, $m \leq n$.

If A is invertible, T_A is both, so $n \leq m$ and $n \geq m$.

Solution to Exercise 2.7.33

Uses. Theorem 2.7.2 (First Fundamental Theorem), Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.7.8 (Third Fundamental Theorem), Proposition 2.5.20 (Inverses and matrices), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Use condition (V). If $CA = I_n$, find a left inverse of GAH built from C, G^{-1}, and H^{-1}.
Hint 2. For the converse, write $A = G^{-1}(GAH)H^{-1}$ and apply the first direction with G^{-1} and H^{-1}.
Sketch. If $CA = I_n$, then $\big(H^{-1}CG^{-1}\big)(GAH) = H^{-1}(CA)H = I_n$; if $AC = I_m$, then $(GAH)\big(H^{-1}CG^{-1}\big) = I_m$. By (I) $\Leftrightarrow$ (V) of the first two theorems, injectivity and surjectivity pass from A to GAH, and back since $A = G^{-1}(GAH)H^{-1}$. Then apply Theorem 2.7.8.

By Proposition 2.5.20(c) and Theorem 2.7.2, (I) $\Leftrightarrow$ (V), T_A is injective if and only if $CA = I_n$ for some $C \in F^{n\times m}$. If so, then by Proposition 2.4.14,

$$\begin{aligned}\big(H^{-1}CG^{-1}\big)(GAH) &= H^{-1}C\big(G^{-1}G\big)AH \\ &= H^{-1}(CA)H = H^{-1}H = I_n,\end{aligned}$$

so GAH has a left inverse and T_{GAH} is injective. Likewise, if $AC = I_m$, then $(GAH)\big(H^{-1}CG^{-1}\big) = G(AC)G^{-1} = I_m$, so by Theorem 2.7.4, (V) $\Rightarrow$ (I), T_{GAH} is surjective whenever T_A is.

For the converses, apply what we just proved to the matrix GAH and the invertible matrices G^{-1} and H^{-1}: since $A = G^{-1}(GAH)H^{-1}$, injectivity and surjectivity pass from GAH back to A. Finally, by Theorem 2.7.8 and Proposition 2.5.20(c), a matrix is invertible if and only if its map is injective and surjective, so GAH is invertible if and only if A is.

Solution to Exercise 2.7.34

Uses. Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.8 (Four views of the product), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Theorem 2.7.8 (Third Fundamental Theorem), Proposition 2.7.7 (Bijective linear maps are invertible), Proposition 2.5.4 (Left inverses give uniqueness), Proposition 2.5.5 (Right inverses give existence), Proposition 2.5.20 (Inverses and matrices)
Hint 1. By the column view of the product, column k of AX is $A\mathbf{x}_k$, where $\mathbf{x}_k$ is column k of X. So $AX = Y$ is p separate systems $A\mathbf{x}_k = \mathbf{y}_k$.
Hint 2. For "only if" in (b), turn a nonzero $\mathbf{v} \in \ker T_A$ into a nonzero matrix X with $AX = 0$. For "only if" in (c), solve $AX = Y$ for a Y whose first column is a given $\mathbf{y}$.
Sketch. (a) $A(cX + Y) = cAX + AY$. (b) If T_A is injective and $AX = 0$, every column of X is in $\ker T_A = \{\mathbf{0}\}$. If $A\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$, the matrix with first column $\mathbf{v}$ and other columns $\mathbf{0}$ is a nonzero element of $\ker L_A$. (c) Solve column by column;

conversely, the first column of a solution of $AX = (\mathbf{y}, \mathbf{0}, \ldots, \mathbf{0})$ solves $A\mathbf{x} = \mathbf{y}$. (d) Combine (b), (c), Theorem 2.7.8, and Proposition 2.7.7.

Write $\mathbf{x}_1, \ldots, \mathbf{x}_p$ for the columns of $X \in F^{n\times p}$. By Proposition 2.4.8, column k of AX is $A\mathbf{x}_k$.

(a) By Proposition 2.4.14, $A(cX + Y) = cAX + AY$ for $X, Y \in F^{n\times p}$ and $c \in F$, so L_A is linear.

(b) Suppose T_A is injective, and $AX = 0$. Then $A\mathbf{x}_k = \mathbf{0}$ for every k, so every $\mathbf{x}_k = \mathbf{0}$, and $X = 0$. So $\ker L_A = \{0\}$, and L_A is injective (Theorem 2.6.11). Conversely, suppose T_A is not injective, and let $\mathbf{v} \neq \mathbf{0}$ with $A\mathbf{v} = \mathbf{0}$. Let X have first column $\mathbf{v}$ and all other columns $\mathbf{0}$ (this uses $p \geq 1$). Then $X \neq 0$ and $AX = 0$, so L_A is not injective.

(c) Suppose T_A is surjective, and let $Y \in F^{m\times p}$ have columns $\mathbf{y}_1, \ldots, \mathbf{y}_p$. Choose $\mathbf{x}_k$ with $A\mathbf{x}_k = \mathbf{y}_k$, and let X have columns $\mathbf{x}_1, \ldots, \mathbf{x}_p$; then $AX = Y$. Conversely, suppose L_A is surjective, and let $\mathbf{y} \in F^m$. Some X satisfies $AX = Y$, where Y has first column $\mathbf{y}$ and all other columns $\mathbf{0}$. Comparing first columns, $A\mathbf{x}_1 = \mathbf{y}$. So T_A is surjective.

(d) A linear map is invertible if and only if it is bijective (Proposition 2.7.7, Proposition 2.5.4, Proposition 2.5.5). By (b) and (c), L_A is bijective if and only if T_A is, which by Theorem 2.7.8 and Proposition 2.5.20(c) holds if and only if A is invertible.

Solution to Exercise 2.7.35

Uses. Theorem 2.7.2 (First Fundamental Theorem), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
Hint 1. By Theorem 2.7.2, it suffices to show that $A\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \in \mathbb{Q}^n$ forces $\mathbf{x} = \mathbf{0}$. Suppose not. Can you make $\mathbf{x}$ an integer vector?
Hint 2. Clear denominators, then divide by the largest power of p that divides every entry. The result is an integer solution with some entry not divisible by p. Now reduce modulo p.
Hint 3. (b) Take $p = 2$. (c) A 1×1 example suffices.
Sketch. If $A\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \in \mathbb{Q}^n$ nonzero, scale $\mathbf{x}$ to an integer vector, not all of whose entries are divisible by p. Reducing each equation $\sum_j a_{ij}x_j = 0$ modulo p (Lemma 1.1.21) gives $\overline{A}\overline{\mathbf{x}} = \mathbf{0}$ with $\overline{\mathbf{x}} \neq \mathbf{0}$, contradicting Theorem 2.7.2 over $\mathbb{F}_p$. (b) With $p = 2$, $\overline{A} = I_n$. (c) $A = (p)$.

(a) By Theorem 2.7.2, (III) $\Leftrightarrow$ (IV), over $\mathbb{Q}$ and over $\mathbb{F}_p$, it suffices to show: if $\overline{A}\mathbf{z} = \mathbf{0}$ has only the solution $\mathbf{z} = \mathbf{0}$ in $\mathbb{F}_p^n$, then $A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$ in $\mathbb{Q}^n$.

Suppose, for contradiction, that $A\mathbf{x} = \mathbf{0}$ for some nonzero $\mathbf{x} \in \mathbb{Q}^n$. Multiplying $\mathbf{x}$ by the product of the denominators of its entries, we may assume that $\mathbf{x} \in \mathbb{Z}^n$, still nonzero. Some entry x_k is a nonzero integer, and $p^e \mid x_k$ only for $p^e \leq |x_k|$, so there is a largest $e \geq 0$ such that p^e divides every entry of $\mathbf{x}$. Dividing by p^e, we may assume that $\mathbf{x} \in \mathbb{Z}^n$, $A\mathbf{x} = \mathbf{0}$, and some entry of $\mathbf{x}$ is not divisible by p.

For each i, $\sum_{j=1}^n a_{ij}x_j = 0$ in $\mathbb{Z}$. Reduction modulo p respects sums and products (Lemma 1.1.21(a)), so $\sum_j \overline{a}_{ij}\overline{x}_j = \overline{0} = 0$ in $\mathbb{F}_p$, where $\overline{a}$ is the remainder of a modulo p. That is, $\overline{A}\overline{\mathbf{x}} = \mathbf{0}$, where $\overline{\mathbf{x}} \in \mathbb{F}_p^n$ has entries $\overline{x}_j$. Since some x_j is not divisible by p, $\overline{x}_j \neq 0$, so $\overline{\mathbf{x}} \neq \mathbf{0}$. This contradicts the hypothesis on $\overline{A}$.

(b) With $p = 2$, an odd integer has remainder 1 and an even one remainder 0, so $\overline{A} = I_n$, whose columns $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are independent over $\mathbb{F}_2$. By (a), the columns of A are independent over $\mathbb{Q}$.

(c) $A = (p) \in \mathbb{Z}^{1\times 1}$ has a nonzero column, which is independent over $\mathbb{Q}$, but $\overline{A} = (0)$.

For (b), Section 3.5 shows that a square matrix with independent columns is invertible (Corollary 3.5.25), so these matrices are invertible over $\mathbb{Q}$.

Section 2.8

Solution to Exercise 2.8.5

Uses. Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.22 (Transpose is a linear involution), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space), Proposition 2.1.17 (Compositions of linear maps are linear), Exercise 2.1.32, Definition 2.4.26 (Trace), Lemma 2.4.24 (Matrix units)

Techniques. Assemble linear maps from linear pieces (Technique 2.1.26), Try small cases and small fields (Technique 1.2.38)
Hint 1. (a) and (b) are built from maps you already know to be linear. For (c), the formula is a product of entries, which suggests testing homogeneity with a scalar c. Which scalars satisfy $c^2 = c$?
Hint 2. For (c) over $\mathbb{F}_2$, expand $\operatorname{tr}(X^2) = \sum_i \sum_k x_{ik}x_{ki}$ and pair the term for (i,k) with the term for (k,i).
Sketch. (a) A difference of the composite $X \mapsto AX \mapsto AXB$ and the transpose, all linear. (b) $C_{x+1} - 2\,\mathrm{id} + C_{x-1}$, with the substitution maps of Exercise 2.1.32. (c) Over $\mathbb{F}_2$, the off-diagonal terms of $\operatorname{tr}(X^2)$ cancel in pairs and $x_{ii}^2 = x_{ii}$, so $\operatorname{tr}(X^2) = \operatorname{tr} X$, which is linear. Over any other field pick $c \notin \{0,1\}$; then $\operatorname{tr}\left(\left(c\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}\right)^2\right) = c^2 \neq c$.

(a) Linear. The maps $X \mapsto AX$ and $X \mapsto XB$ are linear by Proposition 2.4.14(c) and (d), so their composite $X \mapsto AXB$ is linear (Proposition 2.1.17). The transpose is linear by Proposition 2.4.22, and a difference of linear maps is linear (Proposition 2.1.23).

(b) Linear. For $q \in F[x]$, the substitution $C_q : p \mapsto p(q)$ is linear (Exercise 2.1.32), so the map is $C_{x+1} - 2\mathrm{id}_{F[x]} + C_{x-1}$, a combination of linear maps.

(c) Write $\varphi(X) = \operatorname{tr}(X^2)$. By the entry formula,

$$\varphi(X) = \sum_{i=1}^{n} \left(X^2\right)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} x_{ik}x_{ki} = \sum_{i=1}^{n} x_{ii}^2 + \sum_{i<k}(x_{ik}x_{ki} + x_{ki}x_{ik}).$$

Over $\mathbb{F}_2$. Each pair in the last sum is $(1+1)x_{ik}x_{ki} = 0$, and $c^2 = c$ for both elements c of $\mathbb{F}_2$. So $\varphi(X) = \sum_i x_{ii} = \operatorname{tr} X$, and φ is linear by Proposition 2.4.27.

Over any other field. F has an element c with $c \neq 0$ and $c \neq 1$. Let $X = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$, the matrix unit with a single 1 in position $(1,1)$, so $X^2 = X$ by Lemma 2.4.24(c). Then $\varphi(cX) = \operatorname{tr}(c^2X^2) = c^2$, while $c\varphi(X) = c$. Since $c^2 - c = c(c-1) \neq 0$, φ is not homogeneous, so it is not linear.

Over $\mathbb{F}_2$ the squares collapse, since $c^2 = c$ for both scalars, and the cross terms cancel in pairs; a product of two different inputs, such as $(x,y) \mapsto xy$, is still not linear there (Exercise 2.1.15(d)). So the rule of thumb in Technique 2.1.26, that a product of inputs is a candidate for a counterexample, has exceptions over small fields, which is a reason to test them separately (Technique 1.2.38).

Solution to Exercise 2.8.6

Uses. Lemma 2.4.24 (Matrix units), Proposition 2.4.27 (The trace is linear), Proposition 2.4.14 (Algebra of matrices), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Check it on a basis (Technique 2.2.7), Compute entry by entry (Technique 2.4.15)
Hint 1. Both sides of (a) are linear functions of A. By Lemma 2.4.24(c), $U_{ij}U_{kl} = \delta_{jk}U_{il}$.
Hint 2. Entry by entry: $U_{ij}AU_{ji} = \mathbf{e}_i\left(\mathbf{e}_j^{\mathsf{T}}A\mathbf{e}_j\right)\mathbf{e}_i^{\mathsf{T}}$, and the middle factor is a scalar. For (b), move A to the right in each term.
Sketch. (a) On units: $U_{ij}U_{kl}U_{ji} = \delta_{jk}\delta_{lj}U_{ii}$, which sums to $\delta_{kl}I_n = (\operatorname{tr} U_{kl})I_n$; both sides are linear in A and the units form a basis. Entrywise: $U_{ij}AU_{ji} = a_{jj}U_{ii}$, which sums to $\left(\sum_j a_{jj}\right)\left(\sum_i U_{ii}\right) = (\operatorname{tr} A)I_n$. (b) If A commutes with U_{ij}, the term is $U_{ij}U_{ji}A = U_{ii}A$, and the sum is n copies of I_nA.

(a) *On the matrix units.* The map $A \mapsto \sum_{i,j} U_{ij}AU_{ji}$ is a sum of maps $A \mapsto U_{ij}AU_{ji}$, each linear by Proposition 2.4.14(c) and (d), and $A \mapsto (\operatorname{tr} A)I_n$ is linear by Proposition 2.4.27. The matrix units form a basis of $F^{n\times n}$ (Lemma 2.4.24(a)), so by Corollary 2.2.6 it suffices to take $A = U_{kl}$. By Lemma 2.4.24(c), applied twice,

$$U_{ij}U_{kl}U_{ji} = \delta_{jk}U_{il}U_{ji} = \delta_{jk}\delta_{lj}U_{ii}.$$

Summing over j leaves only $j = k$, with the factor δ_{lk}, and summing over i then gives $\delta_{kl}\sum_i U_{ii} = \delta_{kl}I_n$. On the other side, $\operatorname{tr} U_{kl} = \delta_{kl}$. The two sides agree.

Entry by entry. By associativity and Lemma 2.4.24(b),

$$U_{ij}AU_{ji} = \mathbf{e}_i\left(\mathbf{e}_j^{\mathsf{T}}A\mathbf{e}_j\right)\mathbf{e}_i^{\mathsf{T}} = a_{jj}U_{ii}.$$

So the sum is $\sum_i \sum_j a_{jj}U_{ii} = \left(\sum_j a_{jj}\right)\left(\sum_i U_{ii}\right) = (\operatorname{tr} A)I_n$.

The second proof is shorter here, because each term collapses to a single entry of A. The first uses nothing but the multiplication table of the matrix units, and it is the version to use when the terms do not collapse so visibly.

(b) If A commutes with every matrix, then $U_{ij}AU_{ji} = U_{ij}U_{ji}A = U_{ii}A$, since $U_{ij}U_{ji} = \delta_{jj}U_{ii} = U_{ii}$. By (a),

$$(\operatorname{tr} A)I_n = \sum_{j=1}^{n}\sum_{i=1}^{n} U_{ii}A = \sum_{j=1}^{n} I_nA = (n * 1)A.$$

If $n * 1 \neq 0$, then $A = (n * 1)^{-1}(\operatorname{tr} A)I_n$ is a scalar multiple of I_n. Conversely cI_n commutes with every matrix, since $(cI_n)X = cX = X(cI_n)$ by Proposition 2.4.14(b) and (d).

Solution to Exercise 2.8.7

Uses. Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Definition 2.6.5 (Kernel), Definition 2.6.6 (Image), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A)
Techniques. Construct a map by choosing it on a basis (Technique 2.2.4), Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
Hint 1. Complete $(1,1,0,0), (0,0,1,1)$ to a basis of $\mathbb{Q}^4$ with two standard basis vectors. Send the first two vectors to $\mathbf{0}$ and the other two to the spanning vectors of L.
Hint 2. For (b), look for a nonzero vector in $K \cap L$. What does $T^2 = T$ do to a vector that is both a value of T and in $\ker T$?
Sketch. (a) $\mathbf{k}_1 = (1,1,0,0)$, $\mathbf{k}_2 = (0,0,1,1)$, $\mathbf{e}_1, \mathbf{e}_3$ is a basis; let T kill $\mathbf{k}_1, \mathbf{k}_2$ and send $\mathbf{e}_1, \mathbf{e}_3$ to $\mathbf{l}_1 = (1,0,1,0)$, $\mathbf{l}_2 = (0,1,0,1)$. Then $T(x) = (x_1 - x_2)\mathbf{l}_1 + (x_3 - x_4)\mathbf{l}_2$, a vector is killed exactly when its coefficients on $\mathbf{e}_1$ and $\mathbf{e}_3$ vanish, so $\ker T = K$, and $\operatorname{im} T = \operatorname{span}(\mathbf{l}_1, \mathbf{l}_2) = L$. (b) $\mathbf{u} = (1,1,1,1) = \mathbf{k}_1 + \mathbf{k}_2 = \mathbf{l}_1 + \mathbf{l}_2$. If $\mathbf{u} = T(\mathbf{w})$ and $T^2 = T$, then $\mathbf{u} = T(T(\mathbf{w})) = T(\mathbf{u}) = \mathbf{0}$.

(a) Let $\mathbf{k}_1 = (1,1,0,0)$, $\mathbf{k}_2 = (0,0,1,1)$, $\mathbf{l}_1 = (1,0,1,0)$, and $\mathbf{l}_2 = (0,1,0,1)$. The list $\mathbf{k}_1, \mathbf{k}_2, \mathbf{e}_1, \mathbf{e}_3$ is a basis of $\mathbb{Q}^4$: the equation

$$(x_1, x_2, x_3, x_4) = a\mathbf{k}_1 + b\mathbf{k}_2 + c\mathbf{e}_1 + d\mathbf{e}_3 = (a + c, a, b + d, b)$$

has exactly one solution, $a = x_2$, $b = x_4$, $c = x_1 - x_2$, $d = x_3 - x_4$. By Theorem 2.2.1 there is a linear map T with

$$T(\mathbf{k}_1) = T(\mathbf{k}_2) = \mathbf{0}, \quad T(\mathbf{e}_1) = \mathbf{l}_1, \quad T(\mathbf{e}_3) = \mathbf{l}_2,$$

and (2.2.1) gives

$$T(x_1, x_2, x_3, x_4) = (x_1 - x_2)\mathbf{l}_1 + (x_3 - x_4)\mathbf{l}_2 = (x_1 - x_2, x_3 - x_4, x_1 - x_2, x_3 - x_4).$$

Kernel. T kills $\mathbf{k}_1$ and $\mathbf{k}_2$, so it kills K. Conversely, if $T(\mathbf{v}) = \mathbf{0}$ for $\mathbf{v} = a\mathbf{k}_1 + b\mathbf{k}_2 + c\mathbf{e}_1 + d\mathbf{e}_3$, then $c\mathbf{l}_1 + d\mathbf{l}_2 = (c, d, c, d) = \mathbf{0}$, so $c = d = 0$ and $\mathbf{v} \in K$.

Image. By Proposition 2.1.16, $\operatorname{im} T$ is spanned by the values on the basis, $\mathbf{0}, \mathbf{0}, \mathbf{l}_1, \mathbf{l}_2$, so $\operatorname{im} T = L$.

The standard matrix has the columns $T(\mathbf{e}_j)$:

$$[T]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

(b) The vector $\mathbf{u} = (1,1,1,1)$ is $\mathbf{k}_1 + \mathbf{k}_2 \in K$ and $\mathbf{l}_1 + \mathbf{l}_2 \in L$. Let T have kernel K and image L, and suppose $T^2 = T$. Since $\mathbf{u} \in \operatorname{im} T$, $\mathbf{u} = T(\mathbf{w})$ for some $\mathbf{w}$, and then

$$\mathbf{u} = T(\mathbf{w}) = T(T(\mathbf{w})) = T(\mathbf{u}) = \mathbf{0},$$

because $\mathbf{u} \in \ker T$. This is a contradiction. The construction in (a) needed nothing about how K and L sit relative to each other, but an operator with $T^2 = T$ fixes its image pointwise, so its kernel and image can share only $\mathbf{0}$.

Solution to Exercise 2.8.8

Uses. Example 2.1.9 (Maps of polynomials), Proposition 2.2.20 (Linear maps out of $F[x]$), Notation 2.4.16 (Powers of a matrix), Theorem 2.3.20 (Linear maps correspond to matrices), Proposition 2.1.24 (Algebra of composition)
Techniques. Check it on a basis (Technique 2.2.7), Translate between maps and matrices (Technique 2.3.27)
Hint 1. D^p is linear, so it is enough to compute $D^p\left(x^k\right)$. Show by induction on j that $D^j\left(x^k\right) = k(k-1)\cdots(k-j+1)x^{k-j}$ for $j \le k$, and $D^j\left(x^k\right) = \mathbf{0}$ for $j > k$.
Hint 2. Among p consecutive integers, one is divisible by p. For (b), translate $D^p = 0$ into a statement about matrices with Notation 2.4.16. For (c), expand $(\mathrm{id} + D)^p$.
Sketch. (a) For $k < p$, $D^p\left(x^k\right) = \mathbf{0}$ because D kills constants; for $k \ge p$, $D^p\left(x^k\right) = k(k-1)\cdots(k-p+1)x^{k-p}$, and the integer coefficient is a product of p consecutive integers, hence divisible by p. So D^p vanishes on the monomials. (b) D maps $\mathbb{F}_p[x]_{\le n}$ into itself, and N^p is the matrix of the restriction of D^p, which is 0. Over $\mathbb{Q}$, $D^n(x^n) = n! \neq 0$. (c) $(\mathrm{id} + D)^p = \mathrm{id} + D^p = \mathrm{id}$, so $(\mathrm{id} + D)^{p-1}$ is a two-sided inverse.

(a) By Example 2.1.9, $D\left(x^k\right) = (k * 1)x^{k-1}$ for $k \ge 1$ and $D(1) = \mathbf{0}$. By induction on j,

$$D^j\left(x^k\right) = (k(k-1)\cdots(k-j+1) * 1)x^{k-j} \quad (j \le k), \quad D^j\left(x^k\right) = \mathbf{0} \quad (j > k):$$

one more application of D multiplies by the current exponent $k-j$ and lowers it by one, and once the constant $D^k\left(x^k\right) = k! * 1$ is reached, the next application gives $\mathbf{0}$. Now take $j = p$. If $k < p$, then $D^p\left(x^k\right) = \mathbf{0}$. If $k \ge p$, the integer $k(k-1)\cdots(k-p+1)$ is a product of p consecutive integers, one of which is divisible by p, so its image in $\mathbb{F}_p$ is 0 and $D^p\left(x^k\right) = \mathbf{0}$. So the linear map D^p agrees with 0 on every monomial, and $D^p = 0$ by the uniqueness in Proposition 2.2.20.

(b) D lowers degrees, so it maps $V = \mathbb{F}_p[x]_{\le n}$ into itself; call the restriction D_V. Then D_V^p is the restriction of $D^p = 0$, so $D_V^p = 0$. By Notation 2.4.16, $N^p = \left[D_V^p\right]_{\mathcal{BB}} = [0]_{\mathcal{BB}} = 0$, where $\mathcal{B} = 1, x, \ldots, x^n$.

Over $\mathbb{Q}$, the computation in (a) gives $D^n(x^n) = n! \neq 0$, so $D_V^n \neq 0$, and its matrix N^n is not 0 by Theorem 2.3.20. (In both cases $N^{n+1} = 0$, since D_V^{n+1} kills every polynomial of degree at most n.) So over $\mathbb{F}_p$ with $p \le n$, the matrix N dies earlier than the size forces it to.

(c) The operators id and D commute, so the binomial theorem gives

$$(\mathrm{id} + D)^p = \sum_{k=0}^{p} \binom{p}{k} D^k = \mathrm{id} + D^p = \mathrm{id},$$

because the coefficients $\binom{p}{k}$ with $0 < k < p$ are 0 in $\mathbb{F}_p$, and $D^p = 0$ by (a). Hence

$$(\mathrm{id} + D)(\mathrm{id} + D)^{p-1} = (\mathrm{id} + D)^{p-1}(\mathrm{id} + D) = (\mathrm{id} + D)^p = \mathrm{id},$$

and by Technique 2.5.16, $\mathrm{id} + D$ is invertible with inverse $(\mathrm{id} + D)^{p-1}$. This agrees with the inverse $\mathrm{id} - D + D^2 - \cdots$ found in Example 2.8.4, as it must, since an inverse is unique: over $\mathbb{F}_p$ the series stops at D^{p-1} for every polynomial.

Solution to Exercise 2.8.9

Uses. Theorem 2.4.34 (Block multiplication), Exercise 2.5.39, Proposition 2.4.14 (Algebra of matrices)
Techniques. Block decomposition (Technique 2.4.35), Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. $N = \begin{pmatrix} 0 & A & B \\ 0 & 0 & C \\ 0 & 0 & 0 \end{pmatrix}$. Only one block of N^2 is nonzero.
Hint 2. Since $N^3 = 0$, the geometric series $I - N + N^2$ is a candidate for M^{-1}. Check the sizes of AC and CA.
Sketch. N^2 has the single nonzero block AC in position $(1,3)$, and $N^3 = 0$. So $(I+N)\left(I - N + N^2\right) = I + N^3 = I$ in both orders, and $M^{-1} = \begin{pmatrix} I_p & -A & AC-B \\ 0 & I_q & -C \\ 0 & 0 & I_r \end{pmatrix}$. The product CA is defined only when $r = p$, and the corner of M^{-1} is $p \times r$, which AC is.

(a) $M = I + N$ with

$$N = \begin{pmatrix} 0 & A & B \\ 0 & 0 & C \\ 0 & 0 & 0 \end{pmatrix}.$$

By Theorem 2.4.34, the (I, J) block of N^2 is $\sum_K N_{IK} N_{KJ}$, and a term is nonzero only if $I < K < J$. So the only nonzero block of N^2 is in position $(1, 3)$, and it equals $N_{12} N_{23} = AC$:

$$N^2 = \begin{pmatrix} 0 & 0 & AC \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The only nonzero block row of N^2 is the first, and its only nonzero block sits in block column 3, which multiplies the third block row of N, which is zero. So $N^3 = N^2 N = 0$. Since I commutes with N,

$$(I + N)\big(I - N + N^2\big) = I - N + N^2 + N - N^2 + N^3 = I,$$

and the product in the other order is the same expansion. By Technique 2.5.16 (this is Exercise 2.5.39 with $k = 3$),

$$M^{-1} = I - N + N^2 = \begin{pmatrix} I_p & -A & AC - B \\ 0 & I_q & -C \\ 0 & 0 & I_r \end{pmatrix}.$$

(b) $AC - B$ is correct. The product CA is defined only when $r = p$, since C has r columns and A has p rows, and even then it is not the right block. As a check, the $(1, 3)$ block of MM^{-1} is

$$I_p(AC - B) + A(-C) + BI_r = AC - B - AC + B = 0,$$

which works because AC appears twice in the same order. In the scalar formula the order did not matter, and that is exactly the information the scalar formula loses.

Solution to Exercise 2.8.10

> *Uses.* Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 1.2.18 (Degree and evaluation versus the operations), Example 2.6.29 (Evaluation at 0, and multiplication by x), Example 2.1.19 (Composites), Proposition 2.1.24 (Algebra of composition)
> *Techniques.* Test injectivity at $\mathbf{0}$ (Technique 2.6.12), Guess the inverse, then check both products (Technique 2.5.16)
> *Hint 1.* (a) Compare degrees for injectivity. For surjectivity, evaluate a value $(M - c\,\mathrm{id})(q)$ at c.
> *Hint 2.* (b) $D - c\,\mathrm{id} = -c\big(\mathrm{id} - c^{-1}D\big)$. Expand $1/(1 - t)$ as a geometric series, and use that D^{d+1} kills polynomials of degree at most d.
> *Sketch.* (a) For $q \neq \mathbf{0}$ of degree d, xq has degree $d + 1$ and cq degree at most d, so $(M - c\,\mathrm{id})(q) \neq \mathbf{0}$. Every value vanishes at c, and 1 does not. (b) $S = -\sum_{k\geq 0} c^{-k-1} D^k$, a finite sum on each polynomial; $(D - c\,\mathrm{id})S$ and $S(D - c\,\mathrm{id})$ telescope to id. (c) Both maps are injective by the kernel test, but only one is surjective.

(a) *Injective.* Let $q \neq \mathbf{0}$ have degree d. Then xq has degree $d + 1$ and cq has degree at most d, so their degrees differ and $(M - c\,\mathrm{id})(q) = xq - cq$ has degree $d + 1$ (Proposition 1.2.18(a)). In particular it is not $\mathbf{0}$, and $M - c\,\mathrm{id}$ is injective by Theorem 2.6.11.

Not surjective. Every value $(M - c\,\mathrm{id})(q) = xq - cq$ vanishes at c (Example 2.6.29). The constant polynomial 1 does not, so it is not a value.

(b) Since $D - c\,\mathrm{id} = -c\big(\mathrm{id} - c^{-1}D\big)$ and $1/(1 - t) = 1 + t + t^2 + \cdots$, the candidate is

$$S = -c^{-1} \sum_{k\geq 0} c^{-k} D^k = -\sum_{k\geq 0} c^{-k-1} D^k.$$

On a polynomial of degree at most d only the terms with $k \leq d$ can be nonzero, since D^{d+1} kills it (Example 2.1.19(a)). So define $S(f) = -\sum_{k=0}^{d} c^{-k-1} D^k(f)$ for any $d \geq \deg f$; the choice of d does not matter, and on $F[x]_{\leq d}$, S agrees with a linear combination of powers of D, so S is linear. For $\deg f \leq d$,

$$(D - c\,\mathrm{id})(S(f)) = -\sum_{k=0}^{d} c^{-k-1} D^{k+1}(f) + \sum_{k=0}^{d} c^{-k} D^{k}(f)$$
$$= f - c^{-d-1} D^{d+1}(f) = f,$$

since the two sums cancel except for the term $k = 0$ of the second and the term $k = d$ of the first. The same computation gives $S((D - c\,\mathrm{id})(f)) = f$, because $(D - c\,\mathrm{id})(f)$ also has degree at most d. By Technique 2.5.16, $D - c\,\mathrm{id}$ is invertible and $(D - c\,\mathrm{id})^{-1} = -\sum_{k\geq 0} c^{-k-1} D^{k}$.

(c) The kernel test proves that $D - c\,\mathrm{id}$ is injective for $c \neq 0$: if $D(f) = cf$ with $f \neq \mathbf{0}$, the left side has smaller degree than the right. But it proves the same about $M - c\,\mathrm{id}$, which is not surjective. On $F[x]$ injectivity does not imply surjectivity, so a second argument was needed, and the guessed inverse supplied both halves at once.

Solution to Exercise 2.8.11

Uses. Proposition 2.4.14 (Algebra of matrices), Lemma 2.4.24 (Matrix units), Example 2.3.28 (Translating in both directions), Exercise 2.3.40, Proposition 2.3.16 (The matrix of T_A), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Example 2.3.6 (The zero map and the identity map), Proposition 2.4.12 (Matrices of sums and scalar multiples)
Techniques. Proving TFAE (Technique 2.7.1), Compute entry by entry (Technique 2.4.15)
Hint 1. Use (i) as a hub. Which spokes are already in the chapter? Look at Example 2.3.28(b) and Exercise 2.3.40.
Hint 2. For (iii) ⇒ (i), compute both sides of $AU_{ij} = U_{ij}A$: one has only column j nonzero, the other only row i.
Sketch. Hub at (i). The cycle (i) ⇒ (ii) ⇒ (iii) ⇒ (i): scalars commute with everything, (iii) is a special case of (ii), and AU_{ij} has column j equal to column i of A while $U_{ij}A$ has row i equal to row j of A, which forces $a_{ki} = 0$ for $k \neq i$ and $a_{ii} = a_{jj}$. (i) ⇔ (iv) is Example 2.3.28(b) plus a one-line converse. (i) ⇔ (v): since $[T_A]_{\mathcal{E}\mathcal{E}} = A$, (v) says that the matrix of T_A is the same in every basis, which is Exercise 2.3.40.

We use (i) as a hub. Each spoke is either a citation or a short computation.

(i) ⇒ (ii) ⇒ (iii) ⇒ (i). If $A = \lambda I_n$, then $AX = \lambda X = XA$ by Proposition 2.4.14(b) and (d). (iii) is the special case $X = U_{ij}$ of (ii). For (iii) ⇒ (i), fix i and j. By associativity and Lemma 2.4.24(b),

$$AU_{ij} = (A\mathbf{e}_i)\mathbf{e}_j^{\mathsf{T}}, \quad U_{ij}A = \mathbf{e}_i\big(\mathbf{e}_j^{\mathsf{T}}A\big).$$

The first matrix has column j equal to column i of A and every other column zero. The second has row i equal to row j of A and every other row zero. If they are equal, then column j of the first, $A_{\cdot,i}$, is zero outside row i, so $a_{ki} = 0$ for $k \neq i$; and comparing the entries in position (i, j) gives $a_{ii} = a_{jj}$. As i and j range over all indices, A is diagonal with all diagonal entries equal to a_{11}, so $A = a_{11}I_n$.

(i) ⇔ (iv). If $A = \lambda I_n$, then $A\mathbf{x} = \lambda\mathbf{x} \in \mathrm{span}(\mathbf{x})$. The converse is Example 2.3.28(b).

(i) ⇔ (v). By Proposition 2.3.16, $[T_A]_{\mathcal{E}\mathcal{E}} = A$, so (v) says that $[T_A]_{\mathcal{B}\mathcal{B}}$ is the same matrix for every basis $\mathcal{B}$ of length n. By Exercise 2.3.40, this holds if and only if $T_A = \lambda\,\mathrm{id}$ for some λ, that is, $T_A = T_{\lambda I_n}$, which by Proposition 2.3.17 means $A = \lambda I_n$. (For the direction (i) ⇒ (v) one can also compute directly: $[\lambda\,\mathrm{id}]_{\mathcal{B}\mathcal{B}} = \lambda I_n$ by Example 2.3.6 and Proposition 2.4.12.)

Every condition is connected to the hub in both directions, so all five are equivalent. Only one arrow, (iii) ⇒ (i), needed new work; the others were earlier results read as statements about A.

Solution to Exercise 2.8.12

Uses. Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Definition 2.4.26 (Trace), Definition 2.4.21 (Transpose)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Take the trace of both sides. The left side has trace 0.
Hint 2. $\mathrm{tr}\big(AA^{\mathsf{T}}\big) = \sum_{j,k} a_{jk}^2$. Over $\mathbb{R}$ this is a sum of squares. Over $\mathbb{C}$ and $\mathbb{F}_2$, find A with $AA^{\mathsf{T}} = -I_n$.
Sketch. (a) $\mathrm{tr}(AB - BA) = 0$ by cyclicity, while $\mathrm{tr}\big(AA^{\mathsf{T}} + I_n\big) = \sum_{j,k} a_{jk}^2 + n \geq n > 0$. (b) $A = iI_n$ over $\mathbb{C}$ and $A = I_n$ over $\mathbb{F}_2$ give $AA^{\mathsf{T}} = -I_n$, so with $B = 0$ both sides are 0. (c) Over $\mathbb{R}$ a sum of squares is 0 only if every term is. Over $\mathbb{C}$, $A = \left(\begin{smallmatrix} 1 & i \\ 0 & 0 \end{smallmatrix}\right)$; over $\mathbb{F}_2$, $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

The (j,j) entry of AA^{T} is $\sum_k a_{jk}\left(A^{\mathsf{T}}\right)_{kj} = \sum_k a_{jk}^2$, so

$$\operatorname{tr}\left(AA^{\mathsf{T}}\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}^2.$$

(a) By Proposition 2.4.27 and Proposition 2.4.28, $\operatorname{tr}(AB-BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. Over $\mathbb{R}$, every $a_{jk}^2 \geq 0$, so $\operatorname{tr}\left(AA^{\mathsf{T}} + I_n\right) = \operatorname{tr}\left(AA^{\mathsf{T}}\right) + n \geq n > 0$. The two sides have different traces, so they are not equal.

(b) Over $\mathbb{C}$, let $A = iI_n$ and $B = 0$. Then $AB - BA = 0$ and $AA^{\mathsf{T}} + I_n = i^2 I_n + I_n = 0$. Over $\mathbb{F}_2$, let $A = I_n$ and $B = 0$; then $AA^{\mathsf{T}} + I_n = (1+1)I_n = 0 = AB - BA$.

(c) Over $\mathbb{R}$, a sum of squares of real numbers is 0 only if every square is 0, so $\operatorname{tr}\left(AA^{\mathsf{T}}\right) = 0$ gives every $a_{jk} = 0$. Over $\mathbb{C}$, $A = \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}$ has $\operatorname{tr}\left(AA^{\mathsf{T}}\right) = 1^2 + i^2 = 0$. Over $\mathbb{F}_2$, $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ has $\operatorname{tr}\left(AA^{\mathsf{T}}\right) = 1 + 1 = 0$.

The trace argument used one fact beyond cyclicity, the positivity of sums of real squares, and (c) shows exactly where that fact lives. The same positivity drives the inner products of Section 6.1, where $\operatorname{tr}\left(A^{\mathsf{T}}A\right)$ is the squared length of A.

Solution to Exercise 2.8.13

Uses. Theorem 2.4.2 (The matrix of a composite), Definition 2.3.1 (Matrix of a linear map), Example 2.5.21 (Change of coordinates is invertible)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Write $T = \mathrm{id} \circ T \circ \mathrm{id}$ and apply Theorem 2.4.2 with the bases chosen so that the inner subscripts match.
Hint 2. Which way does $P = [\mathrm{id}]_{\mathcal{EB}}$ convert coordinates? Check the answer by computing $T(\mathbf{b}_1)$ and $T(\mathbf{b}_2)$ in $\mathcal{B}$-coordinates.
Sketch. $[T]_{\mathcal{BB}} = [\mathrm{id}]_{\mathcal{BE}}[T]_{\mathcal{EE}}[\mathrm{id}]_{\mathcal{EB}} = P^{-1}[T]_{\mathcal{EE}}P = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}$. The computation multiplied in the wrong order: P converts $\mathcal{B}$-coordinates to standard ones, so it must act first. Directly, $T(\mathbf{b}_1) = (0,1) = -\mathbf{b}_1 + \mathbf{b}_2$ and $T(\mathbf{b}_2) = \mathbf{b}_2$.

The error. By Theorem 2.4.2, applied to $T = \mathrm{id} \circ T \circ \mathrm{id}$ with the bases arranged so that the inner subscripts match,

$$[T]_{\mathcal{BB}} = [\mathrm{id}]_{\mathcal{BE}}[T]_{\mathcal{EE}}[\mathrm{id}]_{\mathcal{EB}} = P^{-1}[T]_{\mathcal{EE}}P,$$

where $[\mathrm{id}]_{\mathcal{BE}} = P^{-1}$ by Example 2.5.21. The matrix $P = [\mathrm{id}]_{\mathcal{EB}}$ converts $\mathcal{B}$-coordinates into standard coordinates, so it is applied first, on the right; the computation put it on the left. In $P[T]_{\mathcal{EE}}P^{-1}$ the subscripts do not match: the factor $P = [\mathrm{id}]_{\mathcal{EB}}$ expects $\mathcal{B}$-coordinates, but it receives the standard coordinates produced by $[T]_{\mathcal{EE}}$.

The correct matrix.

$$P^{-1}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}P = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}.$$

As a check, $T(\mathbf{b}_1) = (0,1) = -\mathbf{b}_1 + \mathbf{b}_2$ and $T(\mathbf{b}_2) = (1,1) = \mathbf{b}_2$, whose $\mathcal{B}$-coordinates $(-1,1)$ and $(0,1)$ are the columns found.

The wrong answer is not meaningless: $P[T]_{\mathcal{EE}}P^{-1}$ is the matrix of T in the basis $(1,0),(-1,1)$ formed by the columns of P^{-1}. That is one way to see that the error produces a plausible-looking matrix, which is why checking one column directly is worth the effort.

Solution to Exercise 2.8.14

Uses. Proposition 2.5.17 (Algebra of inverses), Proposition 2.4.14 (Algebra of matrices), Exercise 2.5.19
Techniques. Guess the inverse, then check both products (Technique 2.5.16), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Associativity moves parentheses. Does it move factors past each other?
Hint 2. $(AB)^{-1} = B^{-1}A^{-1}$ always. When is $B^{-1}A^{-1} = A^{-1}B^{-1}$? Invert both sides.

Sketch. Associativity gives $(AB)\left(A^{-1}B^{-1}\right) = ABA^{-1}B^{-1}$; replacing BA^{-1} by $A^{-1}B$ needs commutativity. With $A = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right)$, $A^{-1}B^{-1} = \left(\begin{smallmatrix}2 & -1\\ -1 & 1\end{smallmatrix}\right) \neq \left(\begin{smallmatrix}1 & -1\\ -1 & 2\end{smallmatrix}\right) = (AB)^{-1}$. The conclusion holds exactly when $AB = BA$: $(AB)^{-1} = B^{-1}A^{-1}$, and $B^{-1}A^{-1} = A^{-1}B^{-1}$ iff, inverting both sides, $AB = BA$.

The error. Associativity regroups a product but keeps the order of its factors: $(AB)\left(A^{-1}B^{-1}\right) = ABA^{-1}B^{-1}$. The step to $AA^{-1}BB^{-1}$ moves A^{-1} past B, which requires $BA^{-1} = A^{-1}B$.

The claim is false. Take $A = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right)$, as in Exercise 2.5.19. There, $(AB)^{-1} = B^{-1}A^{-1} = \left(\begin{smallmatrix}1 & -1\\ -1 & 2\end{smallmatrix}\right)$, while $A^{-1}B^{-1} = \left(\begin{smallmatrix}2 & -1\\ -1 & 1\end{smallmatrix}\right)$.

When it holds. By Proposition 2.5.17(d), the matrix version of (b), $(AB)^{-1} = B^{-1}A^{-1}$ always. So the conclusion holds if and only if $B^{-1}A^{-1} = A^{-1}B^{-1}$. By the same result together with Proposition 2.5.17(a), both sides are invertible, with inverses AB and BA, and two invertible matrices are equal exactly when their inverses are. So the conclusion holds if and only if $AB = BA$. In particular it holds for every pair when $n = 1$, and fails for some pair whenever $n \geq 2$ (put the example above in the top left corner).

Solution to Exercise 2.8.15

Uses. Corollary 2.5.6 (One-sided cancellation), Exercise 2.5.41, Exercise 2.5.33, Theorem 3.5.23 (One-sided inverses of square matrices)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Which property of T allows cancelling it on the right (Corollary 2.5.6)? Does T have it?
Hint 2. Try $V = F[x]$ and $T = M$, multiplication by x, with the left inverse L of Exercise 2.5.41.
Sketch. Cancelling T on the right requires T to be right-invertible, which is what the argument is trying to prove. On $F[x]$, $LM = \mathrm{id}$ but $ML(1) = \mathbf{0}$, and M is not surjective, so it has no right inverse. If T is surjective, then $(TS - \mathrm{id})T = 0$ says that $TS - \mathrm{id}$ vanishes on $\operatorname{im} T = V$, and the argument is correct (Exercise 2.5.33(a)).

The error. By Corollary 2.5.6(b), T can be cancelled on the right when T is right-invertible. That is exactly what the argument is trying to prove, so the cancellation assumes the conclusion. The first two steps are fine; only the cancellation is not.

The claim is false. Let $V = F[x]$, let M be multiplication by x, and let $L(a_0 + a_1x + \cdots + a_nx^n) = a_1 + a_2x + \cdots + a_nx^{n-1}$. Then $LM = \mathrm{id}_{F[x]}$ (Exercise 2.5.41), so M has a left inverse. But M is not surjective, since every value xq has constant term 0, so M has no right inverse (Proposition 2.5.5) and is not invertible. Indeed $(ML - \mathrm{id})(1) = M(\mathbf{0}) - 1 = -1 \neq \mathbf{0}$, although $(ML - \mathrm{id})M = M(LM) - M = 0$.

A repair. Suppose in addition that T is surjective. Then $(TS - \mathrm{id}_V)T = 0$ says that $TS - \mathrm{id}_V$ vanishes at every $T(\mathbf{v})$, that is, on $\operatorname{im} T = V$, so $TS = \mathrm{id}_V$. This is Exercise 2.5.33(a). For $V = F^n$ the claim itself is true (Theorem 3.5.23), but the argument above does not prove it, because it never uses that V is finite-dimensional.

Problems for Chapter 2

Solution to Problem 2.1

Uses. Definition 2.6.1 (Injective), Definition 2.6.5 (Kernel), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.5.15 (Short lists, zeros, and repeats), Definition 2.1.1 (Linear map)
Techniques. Test injectivity at $\mathbf{0}$ (Technique 2.6.12), Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For the theorem, $T(\mathbf{v}) = T(\mathbf{v}')$ should become a statement about $\mathbf{v} - \mathbf{v}'$. For the proposition, pull a relation among the $T(\mathbf{v}_i)$ back through T.
Hint 2. For (c), list the facts about T that the proof used: $T(\mathbf{0}) = \mathbf{0}$ and $T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}')$. Over $\mathbb{F}_2$, a function with $T(\mathbf{0}) = 0$ is automatically homogeneous.
Sketch. (b) If T is injective and $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$, then $\mathbf{v} = \mathbf{0}$; if $\ker T = \{\mathbf{0}\}$ and $T(\mathbf{v}) = T(\mathbf{v}')$, then $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}$. For independence: $\sum a_iT(\mathbf{v}_i) = \mathbf{0}$ gives $\sum a_i\mathbf{v}_i \in \ker T$; conversely a nonzero $\mathbf{v} \in \ker T$ is an independent list with dependent image. (c) Additivity alone gives $T(\mathbf{0}) = \mathbf{0}$ and $T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}')$. Over $\mathbb{F}_2$, $T(\mathbf{0}) = 0$ and $T(\mathbf{v}) = 1$ otherwise is homogeneous, has only $\mathbf{0}$ in its kernel, and has $T(\mathbf{e}_1) = T(\mathbf{e}_2)$.

(a) Theorem 2.6.11: a linear map $T \in \mathrm{L}(V,W)$ is injective if and only if $\ker T = \{\mathbf{0}\}$. Proposition 2.6.14: $T \in \mathrm{L}(V,W)$ is injective if and only if T sends every linearly independent list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V to a linearly independent list $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)$.

(b) *The theorem.* Suppose T is injective, and let $\mathbf{v} \in \ker T$. Then $T(\mathbf{v}) = \mathbf{0} = T(\mathbf{0})$ by Proposition 2.1.4, so $\mathbf{v} = \mathbf{0}$. Conversely, suppose $\ker T = \{\mathbf{0}\}$, and let $T(\mathbf{v}) = T(\mathbf{v}')$. By Proposition 2.1.4, $T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}') = \mathbf{0}$, so $\mathbf{v} - \mathbf{v}' \in \ker T$, that is, $\mathbf{v} = \mathbf{v}'$.

The proposition. Suppose T is injective, let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be independent, and let $\sum_i a_i T(\mathbf{v}_i) = \mathbf{0}$. By linearity, $T\left(\sum_i a_i \mathbf{v}_i\right) = \mathbf{0}$, so $\sum_i a_i \mathbf{v}_i \in \ker T = \{\mathbf{0}\}$, and independence gives every $a_i = 0$. So the image list is independent (Theorem 1.5.5). Conversely, suppose T is not injective. By the theorem there is $\mathbf{v} \neq \mathbf{0}$ in $\ker T$. The list $\mathbf{v}$ is independent (Proposition 1.5.15(b)), but its image is the list $\mathbf{0}$, which is dependent (Proposition 1.5.15(c)).

(c) *Additive maps.* The proof of the theorem used two facts: $T(\mathbf{0}) = \mathbf{0}$, and $T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}')$. Both follow from additivity alone. First, $T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, and adding $-T(\mathbf{0})$ gives $T(\mathbf{0}) = \mathbf{0}$. Next, $T(\mathbf{v}') + T(-\mathbf{v}') = T(\mathbf{0}) = \mathbf{0}$, so $T(-\mathbf{v}') = -T(\mathbf{v}')$, and then $T(\mathbf{v} - \mathbf{v}') = T(\mathbf{v}) + T(-\mathbf{v}') = T(\mathbf{v}) - T(\mathbf{v}')$. With these two facts the proof runs word for word, so an additive map is injective if and only if it sends only $\mathbf{0}$ to $\mathbf{0}$.

Homogeneous maps. Define $T : \mathbb{F}_2^2 \to \mathbb{F}_2$ by $T(\mathbf{0}) = 0$ and $T(\mathbf{v}) = 1$ for $\mathbf{v} \neq \mathbf{0}$. The only scalars are 0 and 1, and $T(0\mathbf{v}) = T(\mathbf{0}) = 0 = 0 \cdot T(\mathbf{v})$ and $T(1\mathbf{v}) = T(\mathbf{v}) = 1 \cdot T(\mathbf{v})$, so T is homogeneous. By construction $T(\mathbf{v}) = 0$ only for $\mathbf{v} = \mathbf{0}$. But $T(\mathbf{e}_1) = T(\mathbf{e}_2) = 1$ with $\mathbf{e}_1 \neq \mathbf{e}_2$, so T is not injective. It is not additive either, since $T(\mathbf{e}_1) + T(\mathbf{e}_2) = 0 \neq 1 = T(\mathbf{e}_1 + \mathbf{e}_2)$, and additivity is the half of linearity that the kernel test needs.

Solution to Problem 2.2

Uses. Definition 2.3.1 (Matrix of a linear map), Theorem 2.3.9 (The matrix of a map computes the map), Definition 2.4.1 (Matrix product), Proposition 2.3.16 (The matrix of T_A), Proposition 2.1.24 (Algebra of composition), Example 2.5.21 (Change of coordinates is invertible), Exercise 2.5.32, Exercise 2.3.26
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. For (b), compare the j-th columns. For associativity, read each matrix as the standard matrix of T_A, T_B, T_C.
Hint 2. For (c), try $T = \mathrm{id}_V$. Then $P = [\mathrm{id}_V]_{\mathcal{C}\mathcal{B}}$ satisfies $P^2 = P$, and P is invertible.
Sketch. (b) Column j of $[TS]_{\mathcal{C}\mathcal{A}}$ is $\left[T\left(S(\mathbf{u}_j)\right)\right]_{\mathcal{C}} = [T]_{\mathcal{C}\mathcal{B}}\left[S(\mathbf{u}_j)\right]_{\mathcal{B}}$, which is $[T]_{\mathcal{C}\mathcal{B}}$ times column j of $[S]_{\mathcal{B}\mathcal{A}}$. With standard bases, $(AB)C$ and $A(BC)$ are the matrices of $(T_A T_B)T_C$ and $T_A(T_B T_C)$, which are the same function. (c) If $\mathcal{B} = \mathcal{C}$ this is (a). Conversely, $T = \mathrm{id}$ gives $P^2 = P$ for the invertible $P = [\mathrm{id}]_{\mathcal{C}\mathcal{B}}$, so $P = I_n$, so $\mathcal{B} = \mathcal{C}$.

(a) Let U, V, W have bases $\mathcal{A} = \mathbf{u}_1, \ldots, \mathbf{u}_n$, $\mathcal{B}$, and $\mathcal{C}$. For $S \in \mathrm{L}(U,V)$ and $T \in \mathrm{L}(V,W)$,

$$[TS]_{\mathcal{C}\mathcal{A}} = [T]_{\mathcal{C}\mathcal{B}}[S]_{\mathcal{B}\mathcal{A}}.$$

(b) By Definition 2.3.1, column j of $[TS]_{\mathcal{C}\mathcal{A}}$ is $\left[T\left(S(\mathbf{u}_j)\right)\right]_{\mathcal{C}}$. By Theorem 2.3.9, applied to T at the vector $S(\mathbf{u}_j)$, this is $[T]_{\mathcal{C}\mathcal{B}}\left[S(\mathbf{u}_j)\right]_{\mathcal{B}}$, and $\left[S(\mathbf{u}_j)\right]_{\mathcal{B}}$ is column j of $[S]_{\mathcal{B}\mathcal{A}}$. By Definition 2.4.1, $[T]_{\mathcal{C}\mathcal{B}}$ times column j of $[S]_{\mathcal{B}\mathcal{A}}$ is column j of $[T]_{\mathcal{C}\mathcal{B}}[S]_{\mathcal{B}\mathcal{A}}$. The two matrices have the same columns.

Associativity. Let $A \in F^{\ell \times m}$, $B \in F^{m \times n}$, $C \in F^{n \times p}$, and use standard bases throughout, so that $[T_X]_{\mathcal{E}\mathcal{E}} = X$ for every matrix X (Proposition 2.3.16). Applying (a) twice,

$$[(T_A T_B)T_C]_{\mathcal{E}\mathcal{E}} = [T_A T_B]_{\mathcal{E}\mathcal{E}} C = (AB)C, \quad [T_A(T_B T_C)]_{\mathcal{E}\mathcal{E}} = A[T_B T_C]_{\mathcal{E}\mathcal{E}} = A(BC).$$

Composition of functions is associative (Proposition 2.1.24(a)), so $(T_A T_B)T_C = T_A(T_B T_C)$, and the two matrices are equal.

(c) If $\mathcal{B} = \mathcal{C}$, then (a) with $S = T$ and all three bases equal to $\mathcal{B}$ gives $[T]_{\mathcal{B}\mathcal{B}}[T]_{\mathcal{B}\mathcal{B}} = \left[T^2\right]_{\mathcal{B}\mathcal{B}}$.

Conversely, suppose the identity holds for every T, and take $T = \mathrm{id}_V$. With $P = [\mathrm{id}_V]_{\mathcal{C}\mathcal{B}}$ it reads $P^2 = [\mathrm{id}_V]_{\mathcal{C}\mathcal{B}} = P$. The matrix P is invertible, with inverse $[\mathrm{id}_V]_{\mathcal{B}\mathcal{C}}$ (Example 2.5.21), so $P^2 = P$ forces $P = I_n$ (Exercise 2.5.32(c)). By Exercise 2.3.26, $\mathcal{B} = \mathcal{C}$.

In the product $[T]_{\mathcal{C}\mathcal{B}}[T]_{\mathcal{C}\mathcal{B}}$, the right factor produces $\mathcal{C}$-coordinates and the left factor expects $\mathcal{B}$-coordinates. Part (c) says that this mismatch is harmless only when the two bases coincide.

Solution to Problem 2.3

> *Uses.* Theorem 2.7.4 (Second Fundamental Theorem), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.4.8 (Four views of the product), Theorem 2.4.2 (The matrix of a composite), Non-example 2.7.10 (Choices that are not linear), Corollary 2.6.31 (Fibers are translates of the kernel), Exercise 2.6.44
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4), Check it on a basis (Technique 2.2.7)
> *Hint 1.* For (b), pick $\mathbf{x}_i$ with $A\mathbf{x}_i = \mathbf{e}_i$, and let B have columns $\mathbf{x}_1, \dots, \mathbf{x}_m$.
> *Hint 2.* For (c), a right inverse is a choice of one solution of $A\mathbf{x} = \mathbf{e}_i$ for each i. How many solutions does each of these equations have? Use Corollary 2.6.31 and Exercise 2.6.44.
> *Sketch.* (b) With $A\mathbf{x}_i = \mathbf{e}_i$, the matrix B with columns $\mathbf{x}_i$ has $AB = I_m$ by the column view, so T_B is a right inverse. Choices made one $\mathbf{y}$ at a time need not fit together linearly (Non-example 2.7.10(b)); choices on a basis always extend. (c) $AC = I_m$ means $A\mathbf{c}_i = \mathbf{e}_i$ for each column. Each such equation has $|\ker T_A| = p^n/p^m$ solutions, and the m columns are chosen independently: $(p^{n-m})^m$.

(a) Theorem 2.7.4: for $A \in F^{m\times n}$ the following are equivalent. (I) T_A is surjective. (II) For every $\mathbf{y} \in F^m$, $A\mathbf{x} = \mathbf{y}$ has at least one solution. (III) The columns of A span F^m. (IV) $\operatorname{im} T_A = F^m$. (V) T_A has a right inverse, a linear map $S : F^m \to F^n$ with $T_A \circ S = \mathrm{id}_{F^m}$.

(b) Assume (II). For $i = 1, \dots, m$ choose $\mathbf{x}_i \in F^n$ with $A\mathbf{x}_i = \mathbf{e}_i$, and let $B \in F^{n\times m}$ have columns $\mathbf{x}_1, \dots, \mathbf{x}_m$. By the column view (Proposition 2.4.8(b)), column i of AB is $A\mathbf{x}_i = \mathbf{e}_i$, so $AB = I_m$. By Theorem 2.4.2(b), $T_A \circ T_B = T_{AB} = \mathrm{id}_{F^m}$, so $S = T_B$ is a right inverse. In the language of Theorem 2.2.1, S is the linear map with $S(\mathbf{e}_i) = \mathbf{x}_i$, and $T_A \circ S$ agrees with id on the standard basis.

The choices are made on a basis because only there do they automatically fit together. A choice of one solution for each $\mathbf{y}$ separately is a function $F^m \to F^n$, but usually not a linear one (Non-example 2.7.10(b)). Once $\mathbf{x}_1, \dots, \mathbf{x}_m$ are chosen, the extension sends $\mathbf{y} = \sum_i y_i \mathbf{e}_i$ to $\sum_i y_i \mathbf{x}_i$, which is a solution of $A\mathbf{x} = \mathbf{y}$ by linearity: no further choice is made.

(c) By the column view, $C \in \mathbb{F}_p^{n\times m}$ satisfies $AC = I_m$ if and only if its columns satisfy $A\mathbf{c}_i = \mathbf{e}_i$ for $i = 1, \dots, m$. The solutions of $A\mathbf{x} = \mathbf{e}_i$ form the fiber $T_A^{-1}(\mathbf{e}_i)$, which is nonempty since T_A is surjective, and is a translate of $\ker T_A$ (Corollary 2.6.31(b)); so it has $|\ker T_A|$ elements. By Exercise 2.6.44(a),

$$p^n = \left|\mathbb{F}_p^n\right| = |\ker T_A| \cdot |\operatorname{im} T_A| = |\ker T_A| \cdot p^m,$$

so $|\ker T_A| = p^{n-m}$. The m columns are chosen independently, so there are $(p^{n-m})^m = p^{m(n-m)}$ right inverses.

In Exercise 2.5.36(c), T_{A^T} is surjective for the matrix $A^\mathsf{T} \in \mathbb{F}_p^{2\times 3}$ there, and it has p^2 right inverses; the formula gives $p^{2(3-2)} = p^2$. A square surjective A has exactly one right inverse, as it must (Theorem 3.5.23).

Solution to Problem 2.4

> *Uses.* Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.23 (Transpose of a product), Proposition 2.4.22 (Transpose is a linear involution), Lemma 2.4.24 (Matrix units), Definition 2.3.1 (Matrix of a linear map), Proposition 2.4.12 (Matrices of sums and scalar multiples), Proposition 2.4.27 (The trace is linear), Exercise 2.2.34
> *Techniques.* Compute entry by entry (Technique 2.4.15), Check it on a basis (Technique 2.2.7)
> *Hint 1.* The diagonal entry of $[\Phi]_{ll}$ belonging to U_{ij} is the U_{ij}-coordinate of $\Phi(U_{ij})$, which is the (i,j) entry of $\Phi(U_{ij})$.
> *Hint 2.* For (c), compute $L_{U_{ij},U_{kl}}(X) = U_{ij}XU_{kl}$. Which matrix unit does it send where?
> *Sketch.* (a) Distributivity, associativity, and $\left(AX^\mathsf{T}B\right)^\mathsf{T} = B^\mathsf{T}XA^\mathsf{T}$. (b) $\left(AU_{ij}B\right)_{ij} = \left(\mathbf{e}_i^\mathsf{T}A\mathbf{e}_i\right)\left(\mathbf{e}_j^\mathsf{T}B\mathbf{e}_j\right) = a_{ii}b_{jj}$, which sums to $\operatorname{tr} A \operatorname{tr} B$; $\tau(U_{ij}) = U_{ji}$ contributes δ_{ij}, so $\operatorname{tr}\tau = n * 1$; $C_A = L_{A,I_n} - L_{I_n,A}$ has trace 0. (c) $U_{ij}XU_{kl} = x_{jk}U_{il}$, so $L_{U_{ij},U_{kl}}$ sends U_{jk} to U_{il} and the other units to 0; these n^4 maps form the basis of Exercise 2.2.34. In particular $X^\mathsf{T} = \sum_{i,j} U_{ij}XU_{ij}$.

(a) $L_{A,B}$ is linear by Proposition 2.4.14(c) and (d). For every X, associativity gives

$$L_{A,B}\big(L_{C,D}(X)\big) = A(CXD)B = (AC)X(DB) = L_{AC,DB}(X),$$

and by Proposition 2.4.23 and Proposition 2.4.22,

$$\tau\big(L_{A,B}(\tau(X))\big) = \big(AX^{\mathsf{T}}B\big)^{\mathsf{T}} = B^{\mathsf{T}}\big(X^{\mathsf{T}}\big)^{\mathsf{T}}A^{\mathsf{T}} = B^{\mathsf{T}}XA^{\mathsf{T}}.$$

(b) By Lemma 2.4.24(a), the coordinates of a matrix with respect to $\mathcal{U}$ are its entries. So the diagonal entry of $[\Phi]_{\mathcal{U}\mathcal{U}}$ in the position of U_{ij} is the (i,j) entry of $\Phi\big(U_{ij}\big)$, whatever order $\mathcal{U}$ is listed in, and $\operatorname{tr}\Phi$ is the sum of these entries over all (i,j). For $\Phi = L_{A,B}$, by Lemma 2.4.24(b),

$$\big(AU_{ij}B\big)_{ij} = \mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}B\mathbf{e}_j = a_{ii}b_{jj},$$

and $\sum_{i,j} a_{ii}b_{jj} = (\operatorname{tr}A)(\operatorname{tr}B)$.

For τ, $\tau\big(U_{ij}\big) = U_{ji}$, whose (i,j) entry is δ_{ij}. So $\operatorname{tr}\tau = \sum_{i,j}\delta_{ij} = n * 1$.

For $C_A = L_{A,I_n} - L_{I_n,A}$, Proposition 2.4.12 and Proposition 2.4.27 give $\operatorname{tr}C_A = (\operatorname{tr}A)(n*1) - (n*1)(\operatorname{tr}A) = 0$.

(c) For all indices, associativity and Lemma 2.4.24(b) give

$$L_{U_{ij},U_{kl}}(X) = \mathbf{e}_i\big(\mathbf{e}_j^{\mathsf{T}}X\mathbf{e}_k\big)\mathbf{e}_l^{\mathsf{T}} = x_{jk}U_{il}.$$

So $L_{U_{ij},U_{kl}}$ sends U_{jk} to U_{il} and every other matrix unit to 0. These are the maps of Exercise 2.2.34 for $V = W = F^{n\times n}$ with the basis $\mathcal{U}$ on both sides: as (i,j,k,l) runs over all quadruples, the pair $\big(U_{jk},U_{il}\big)$ runs over all pairs of matrix units exactly once. By that exercise they form a basis of $\mathrm{L}(F^{n\times n})$. Every Φ is therefore a combination $\sum c_{ijkl}L_{U_{ij},U_{kl}} = \sum L_{c_{ijkl}U_{ij},U_{kl}}$, a sum of maps of the form $L_{A,B}$.

The transpose sends U_{jk} to U_{kj}, so it is the sum of the maps with $i = k$ and $l = j$:

$$X^{\mathsf{T}} = \sum_{j=1}^{n}\sum_{k=1}^{n} U_{kj}XU_{kj}.$$

Directly, $U_{kj}XU_{kj} = x_{jk}U_{kj}$, and summing gives the matrix with (k,j) entry x_{jk}. As a check against (b), $\operatorname{tr}\tau = \sum_{j,k}\big(\operatorname{tr}U_{kj}\big)^2 = \sum_k 1 = n*1$. So the transpose is a sum of n^2 maps $X\mapsto AXB$, although for $n\geq 2$ it is not one of them: if $X^{\mathsf{T}} = AXB$ for all X, then $U_{11} = (A\mathbf{e}_1)\big(\mathbf{e}_1^{\mathsf{T}}B\big)$ and $U_{21} = (A\mathbf{e}_1)\big(\mathbf{e}_2^{\mathsf{T}}B\big)$, by taking $X = U_{11}$ and $X = U_{12}$. Comparing first columns, $A\mathbf{e}_1 \neq \mathbf{0}$ would be a multiple of both $\mathbf{e}_1$ and $\mathbf{e}_2$, which is impossible since they are independent.

Solution to Problem 2.5

> *Uses.* Example 2.1.9 (Maps of polynomials), Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space), Proposition 2.1.17 (Compositions of linear maps are linear), Exercise 2.1.32
> *Techniques.* Check it on a basis (Technique 2.2.7), Guess the inverse, then check both products (Technique 2.5.16)
> *Hint 1.* Both sides of (a) are linear on $\mathbb{Q}[x]_{\leq n}$, so compare them on x^m. Compute $D^k(x^m)$ first.
> *Hint 2.* $D^k(x^m)/k! = \binom{m}{k}x^{m-k}$ for $k\leq m$. For (c) with $n\geq p$, apply both sides to x^p.
> *Sketch.* (a) $D^k(x^m) = m(m-1)\cdots(m-k+1)x^{m-k}$ for $k\leq m$ and $\mathbf{0}$ for $k>m$, so $\sum_k c^kD^k(x^m)/k! = \sum_{k\leq m}\binom{m}{k}c^kx^{m-k} = (x+c)^m$. (b) $S_1S_{-1} = S_{-1}S_1 = \mathrm{id}$, so $S_1^{-1} = S_{-1} = \sum(-1)^kD^k/k!$. (c) For $k\leq n<p$, $k!$ is invertible in $\mathbb{F}_p$ and (a) goes through. For $n\geq p$, $D(x^p) = \mathbf{0}$, so the right side sends x^p to a_0x^p, while $S_1(x^p) = (x+1)^p = x^p+1$.

Throughout, by induction on k, $D^k(x^m) = (m(m-1)\cdots(m-k+1)*1)x^{m-k}$ for $k\leq m$, and $D^k(x^m) = \mathbf{0}$ for $k>m$ (Example 2.1.9).

(a) S_c maps $\mathbb{Q}[x]_{\leq n}$ into itself, since $(x+c)^m$ has degree m, and so does every D^k. Both sides of the claim are linear (Proposition 2.1.23, Proposition 2.1.17), so by Corollary 2.2.6 it suffices to compare them on x^m for $0\leq m\leq n$. For $k\leq m$, $m(m-1)\cdots(m-k+1)/k! = \binom{m}{k}$, and the terms with $k>m$ vanish, so

$$\sum_{k=0}^{n}\frac{c^k}{k!}D^k(x^m) = \sum_{k=0}^{m}\binom{m}{k}c^kx^{m-k} = (x+c)^m = S_c(x^m).$$

(b) $S_1S_{-1} = S_{-1}S_1 = S_0 = \mathrm{id}$, and both maps preserve $\mathbb{Q}[x]_{\leq n}$. By Technique 2.5.16, S_1 is invertible on $\mathbb{Q}[x]_{\leq n}$ with inverse S_{-1}, and by (a) with $c=-1$,

$$S_1^{-1} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} D^k.$$

(c) *The case $n < p$.* For $k \le n < p$, the scalar $k! * 1$ is a product of the nonzero elements $1, 2, \ldots, k$ of $\mathbb{F}_p$, so it is invertible. The integer identity $m(m-1)\cdots(m-k+1) = k!\binom{m}{k}$ holds in $\mathbb{F}_p$ after reduction, so $(k!)^{-1} D^k(x^m) = \binom{m}{k} x^{m-k}$, and the computation of (a) with $c = 1$ goes through unchanged.

The case $n \ge p$. Suppose $S_1 = \sum_{k=0}^{r} a_k D^k$ on $\mathbb{F}_p[x]_{\le n}$. The polynomial x^p lies in this space. Since $D(x^p) = (p * 1)x^{p-1} = \mathbf{0}$, every $D^k(x^p)$ with $k \ge 1$ is $\mathbf{0}$, and the right side sends x^p to $a_0 x^p$. On the other hand, the coefficients $\binom{p}{k}$ with $0 < k < p$ are 0 in $\mathbb{F}_p$, so

$$S_1(x^p) = (x+1)^p = x^p + 1.$$

So $x^p + 1 = a_0 x^p$, and the constant terms give $1 = 0$, a contradiction.

Over $\mathbb{F}_p$ the derivative cannot see x^p, but the shift moves it. So Taylor's formula is a statement about fields in which the factorials can be inverted, and it fails exactly when the space contains x^p.

Solution to Problem 2.6

> *Uses.* Proposition 2.2.20 (Linear maps out of $F[x]$), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Example 2.6.27 (Differentiation on $F[x]$), Exercise 2.1.49
> *Techniques.* Check it on a basis (Technique 2.2.7)
> *Hint 1.* If $TM = MT$, then $TM^k = M^k T$. Apply this to the polynomial 1.
> *Hint 2.* For (c), write T as multiplication by q and expand $D(qf)$ with the product rule. What does $TD = DT$ say about $D(q)$?
> *Sketch.* (a) $T\big(x^k\big) = T\big(M^k(1)\big) = M^k(T(1)) = x^k q$, and two linear maps that agree on the monomials are equal; conversely $q(xf) = x(qf)$. (b) $qf = \mathbf{0}$ with $q, f \ne \mathbf{0}$ is impossible by degrees; if T is onto, $qg = 1$ forces $\deg q = 0$. (c) $D(qf) = D(q)f + qD(f)$, so T commutes with D iff $D(q)f = \mathbf{0}$ for all f, iff $D(q) = \mathbf{0}$. So $T = c$ id in characteristic 0, and over $\mathbb{F}_p$, T is multiplication by a polynomial in x^p.

(a) If $T(f) = qf$ for all f, then $T(1) = q$ and $T(Mf) = qxf = xqf = M(T(f))$ by commutativity. Conversely, suppose $TM = MT$. By induction, $TM^k = M^k T$ for every $k \ge 0$: if $TM^k = M^k T$, then $TM^{k+1} = \big(TM^k\big)M = M^k TM = M^k MT$. With $q = T(1)$,

$$T\big(x^k\big) = T\big(M^k(1)\big) = M^k(T(1)) = x^k q.$$

Multiplication by q is linear, by distributivity and $c(qf) = q(cf)$, and it agrees with T on every monomial. By the uniqueness in Proposition 2.2.20, $T(f) = qf$ for all f.

(b) If $q = \mathbf{0}$, then $T = 0$, which is not injective. If $q \ne \mathbf{0}$ and $f \ne \mathbf{0}$, then $\deg(qf) = \deg q + \deg f \ge 0$, so $qf \ne \mathbf{0}$; by Theorem 2.6.11, T is injective.

If $q = c$ is a nonzero constant, then $T = c$ id is surjective. Conversely, if T is surjective, then $1 = qg$ for some g. Both q and g are nonzero, so $\deg q + \deg g = \deg 1 = 0$, which forces $\deg q = 0$: q is a nonzero constant. For instance, multiplication by x is injective but not surjective, as in Example 2.6.4(e).

(c) By (a), T commutes with M exactly when it is multiplication by some q. For such T, the product rule gives $D(T(f)) = D(qf) = D(q)f + qD(f) = D(q)f + T(D(f))$. So $DT = TD$ if and only if $D(q)f = \mathbf{0}$ for every f. Taking $f = 1$ shows that this forces $D(q) = \mathbf{0}$, and conversely $D(q) = \mathbf{0}$ makes it hold. So the operators commuting with M and D are the multiplications by the polynomials in $\ker D$, which Example 2.6.27 computed.

- If F has characteristic 0, $\ker D$ consists of the constants, so $T = c$ id for some $c \in F$.
- If $F = \mathbb{F}_p$, $\ker D$ consists of the polynomials $b_0 + b_1 x^p + b_2 x^{2p} + \cdots$, so T is multiplication by a polynomial in x^p. For instance $f \mapsto x^p f$ commutes with D, because $D(x^p) = (p * 1)x^{p-1} = \mathbf{0}$.

Solution to Problem 2.7

Uses. Proposition 3.4.27 (Lagrange's formula), Definition 3.4.26 (Lagrange polynomials), Proposition 1.2.18 (Degree and evaluation versus the operations), Example 2.1.9 (Maps of polynomials), Proposition 2.1.23 ($\mathrm{L}(V,W)$ is a vector space), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. Lagrange's formula (Proposition 3.4.27) writes every $f \in V$ as $\sum f(c_i)L_i$. For the second claim, compare both sides on the basis $L_0, \dots, L_n$.
Hint 2. In (b), the scalars are $w_i = \varphi(L_i)$ for $\varphi(f) = f'(0)$, and $f'(0)$ is the coefficient of x in f. Find L_0, L_1, L_2 by checking values. For (c), find a cubic that vanishes at $0, 1, 2$.
Sketch. (a) $f = \sum f(c_i)L_i$ by Proposition 3.4.27, and evaluating a relation $\sum a_i L_i = \mathbf{0}$ at c_k gives $a_k = 0$. Both φ and $\sum \varphi(L_i)\,\mathrm{ev}_{c_i}$ send L_k to $\varphi(L_k)$, so they agree; a relation $\sum a_i\,\mathrm{ev}_{c_i} = 0$ applied to L_k gives $a_k = 0$. (b) $L_0 = 1 - \frac{3}{2}x + \frac{1}{2}x^2$, $L_1 = 2x - x^2$, $L_2 = -\frac{1}{2}x + \frac{1}{2}x^2$, so $f'(0) = -\frac{3}{2}f(0) + 2f(1) - \frac{1}{2}f(2)$; over $\mathbb{F}_3$, $f'(0) = 2f(1) + f(2)$. (c) $f = x^3 - 3x^2 + 2x$ vanishes at $0, 1, 2$ and has $f'(0) = 2$.

(a) *Basis.* Let $f \in V$. By Proposition 3.4.27 with $d_i = f(c_i)$, the polynomial $\sum_i f(c_i)L_i$ is the unique element of V with the values $f(c_0), \dots, f(c_n)$, and f is such an element, so $f = \sum_i f(c_i)L_i$. Hence the L_i span V. If $\sum_i a_i L_i = \mathbf{0}$, then evaluating at c_k, which is linear (Proposition 1.2.18(c)), gives $a_k = 0$. So $L_0, \dots, L_n$ is a basis of V, and the coordinates of f in it are $(f(c_0), \dots, f(c_n))$.

The formula. Each ev_{c_i} is linear (Example 2.1.9), so $\psi = \sum_i \varphi(L_i)\,\mathrm{ev}_{c_i}$ is linear (Proposition 2.1.23). On a basis vector,

$$\psi(L_k) = \sum_{i=0}^{n} \varphi(L_i)L_k(c_i) = \varphi(L_k),$$

since only the term $i = k$ survives. By Corollary 2.2.6, $\psi = \varphi$. So the evaluations span $\mathrm{L}(V,F)$. If $\sum_i a_i\,\mathrm{ev}_{c_i}$ is the zero map, applying it to L_k gives $a_k = 0$. So $\mathrm{ev}_{c_0}, \dots, \mathrm{ev}_{c_n}$ is a basis of $\mathrm{L}(V,F)$, and the coefficients of φ in it are the values $\varphi(L_i)$.

(b) Since $2 \neq 0$, the points $0, 1, 2$ are distinct. The polynomials

$$L_0 = 1 - \frac{3}{2}x + \frac{1}{2}x^2, \quad L_1 = 2x - x^2, \quad L_2 = -\frac{1}{2}x + \frac{1}{2}x^2$$

have the required values: L_0 takes the values $1, 1 - \frac{3}{2} + \frac{1}{2} = 0, 1 - 3 + 2 = 0$ at $0, 1, 2$; L_1 takes $0, 1, 4 - 4 = 0$; and L_2 takes $0, 0, -1 + 2 = 1$. By the uniqueness in Definition 3.4.26, these are the Lagrange polynomials. The map $\varphi(f) = f'(0) = \mathrm{ev}_0(D(f))$ is linear, and for $f = a + bx + cx^2$, $D(f) = b + 2cx$, so $f'(0) = b$ is the coefficient of x. By (a), the formula holds with $w_i = \varphi(L_i)$, and with no other scalars, since the evaluations are independent:

$$f'(0) = -\frac{3}{2}f(0) + 2f(1) - \frac{1}{2}f(2).$$

Over $\mathbb{F}_3$, $2^{-1} = 2$, so $-\frac{3}{2} = 0$ and $-\frac{1}{2} = -2 = 1$, and the formula reads $f'(0) = 2f(1) + f(2)$. As a check over $\mathbb{F}_3$: for $f = 1, x, x^2$ the right side is $2 + 1 = 0$, $2 + 2 = 1$, $2 + 4 = 0$, which are the coefficients of x.

(c) Let $f = x^3 - 3x^2 + 2x$. Then $f(0) = 0$, $f(1) = 1 - 3 + 2 = 0$, $f(2) = 8 - 12 + 4 = 0$, and $f'(0) = 2$, the coefficient of x. Any formula $f'(0) = w_0 f(0) + w_1 f(1) + w_2 f(2)$ would give $2 = 0$. The three values cannot see a cubic that vanishes at all three points; on $\mathbb{Q}[x]_{\leq 3}$ the evaluations at three points span only a part of $\mathrm{L}\big(\mathbb{Q}[x]_{\leq 3}, \mathbb{Q}\big)$.

Solution to Problem 2.8

Uses. Exercise 2.5.46, Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 2.4.14 (Algebra of matrices)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. Write $t = a + d$ and $\Delta = ad - bc$, so $A^2 = tA - \Delta I_2$. If $A^2 = A$, then $(t-1)A = \Delta I_2$. Split into $t \neq 1$ and $t = 1$.
Hint 2. For (b), count the triples (a,b,c) with $bc = a(1-a)$, separating $a \in \{0,1\}$ from the other $p-2$ values. For (c), the inverse map is $S \mapsto (S + I_n)/2$.
Sketch. (a) If $t \neq 1$, A is a scalar λI_2 with $\lambda^2 = \lambda$; if $t = 1$, $\Delta I_2 = 0$. (b) $bc = a(1-a)$ has $2p - 1$ solutions for $a \in \{0,1\}$ and $p - 1$ for each other a: $2(2p-1) + (p-2)(p-1) + 2 = p^2 + p + 2$. (c) $(2P - I_n)^2 = 4\big(P^2 - P\big) + I_n$ and $((S + I_n)/2)^2 =$

> $(S+I_n)/2$ when $S^2=I_n$. (d) Over $\mathbb{F}_2$, $A^2=tA+\Delta I_2$ with $t,\Delta\in\mathbb{F}_2$; four involutions; $\Delta=1$ invertible, $(t,\Delta)=(0,0)$ square-zero, $(1,0)$ idempotent. Over $\mathbb{F}_3$, $\left(\begin{smallmatrix}2&0\\0&0\end{smallmatrix}\right)$ is none of these.

Write $t=a+d$ and $\Delta=ad-bc$, so that $A^2=tA-\Delta I_2$.

(a) Suppose $A^2=A$. Then $A=tA-\Delta I_2$, that is, $(t-1)A=\Delta I_2$. If $t\neq 1$, then $A=(t-1)^{-1}\Delta I_2=\lambda I_2$ for a scalar λ, and $A^2=A$ gives $\lambda^2=\lambda$, so $\lambda\in\{0,1\}$ and $A=0$ or $A=I_2$. If $t=1$, then $\Delta I_2=0$, so $\Delta=0$. Conversely, 0 and I_2 are idempotent, and if $t=1$ and $\Delta=0$, then $A^2=A-0=A$. (The three cases do not overlap: $\operatorname{tr}0=0$ and $\operatorname{tr}I_2=1+1$, and neither is 1.)

(b) By (a), we count the A with $d=1-a$ and $bc=a(1-a)$, and add 2. Fix $a\in\mathbb{F}_p$.

- If $a\in\{0,1\}$, the condition is $bc=0$, which has $2p-1$ solutions ($b=0$ with c arbitrary, or $c=0$ with $b\neq 0$).
- Otherwise $k=a(1-a)\neq 0$, and $bc=k$ has $p-1$ solutions: b is any nonzero element and $c=kb^{-1}$.

The total is $2(2p-1)+(p-2)(p-1)+2=4p-2+p^2-3p+2+2=p^2+p+2$. For $p=2$ there are no values of the second kind, and the count is 8.

(c) Write $2=1+1$. If $P^2=P$, then by Proposition 2.4.14

$$(2P-I_n)^2=4P^2-4P+I_n=I_n.$$

If $S^2=I_n$, then $P=2^{-1}(S+I_n)$ satisfies $P^2=4^{-1}\left(S^2+2S+I_n\right)=4^{-1}(2S+2I_n)=P$. The maps $P\mapsto 2P-I_n$ and $S\mapsto 2^{-1}(S+I_n)$ are inverse to each other, so the first is a bijection. For an odd prime p, (b) therefore counts the $S\in\mathbb{F}_p^{2\times 2}$ with $S^2=I_2$, which are the matrices with $S^{-1}=S$, and Exercise 2.5.46(c) found p^2+p+2 of them as well.

(d) Over $\mathbb{F}_2$, $-1=1$, so $A^2=tA+\Delta I_2$. If $S^2=I_2$, then $tS=(1+\Delta)I_2$. If $t=1$, then $S=(1+\Delta)I_2$ is 0 or I_2, and both have trace 0, a contradiction. So $t=0$, and then $(1+\Delta)I_2=0$ gives $\Delta=1$. Conversely, $t=0$ and $\Delta=1$ give $S^2=I_2$. These are the matrices with $d=a$ and $a+bc=1$: for $a=0$, $b=c=1$; for $a=1$, $bc=0$. So there are 4:

$$I_2,\quad \begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \begin{pmatrix}1&1\\0&1\end{pmatrix},\quad \begin{pmatrix}1&0\\1&1\end{pmatrix},$$

against 8 idempotents. The bijection of (c) fails because $2P-I_2=I_2$ for every P.

Every $A\in\mathbb{F}_2^{2\times 2}$ has $t,\Delta\in\{0,1\}$. If $\Delta=1$, A is invertible (Proposition 2.5.28(a)). If $\Delta=0$, then $A^2=tA$, which is 0 for $t=0$ and A for $t=1$. Over $\mathbb{F}_3$, $A=\left(\begin{smallmatrix}2&0\\0&0\end{smallmatrix}\right)$ has $\Delta=0$, so it is not invertible (Proposition 2.5.28(b)), and $A^2=\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$, which is neither 0 nor A.

Solution to Problem 2.9

> *Uses.* Proposition 2.4.14 (Algebra of matrices), Lemma 2.4.24 (Matrix units), Proposition 2.4.28 (Cyclicity of the trace)
> *Techniques.* Cyclicity of trace (Technique 2.4.29)
> *Hint 1.* Write out $(BA)^{k+1}=BABA\cdots BA$ and regroup by associativity, or induct on k.
> *Hint 2.* For (c), try matrix units: $U_{ij}U_{kl}=\delta_{jk}U_{il}$. For (d), apply cyclicity to $A\cdot\left(B(AB)^{k-1}\right)$.
> *Sketch.* (a) Induction: $(BA)^{k+2}=B(AB)^kABA=B(AB)^{k+1}A$. (b) If $(AB)^k=0$, then $(BA)^{k+1}=0$; swap the roles of A and B. (c) $A=U_{11}$, $B=U_{21}$: $AB=0$, $BA=U_{21}$. (d) $\operatorname{tr}\left((AB)^k\right)=\operatorname{tr}\left(A\left(B(AB)^{k-1}\right)\right)=\operatorname{tr}\left(B(AB)^{k-1}A\right)=\operatorname{tr}\left((BA)^k\right)$.

(a) For $k=0$ both sides are BA. If the identity holds for k, then by associativity

$$(BA)^{k+2}=(BA)^{k+1}(BA)=B(AB)^kABA=B(AB)^k(AB)A=B(AB)^{k+1}A.$$

(b) If $(AB)^k=0$, then $(BA)^{k+1}=B0A=0$ by (a), so BA is nilpotent. Exchanging the roles of A and B (the identity of (a) holds with A and B swapped, by the same proof), if BA is nilpotent, so is AB.

(c) Let $A=U_{11}$ and $B=U_{21}$, with $U_{ij}=\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}\in F^{2\times 2}$. By Lemma 2.4.24(c), $AB=U_{11}U_{21}=0$ and $BA=U_{21}U_{11}=U_{21}\neq 0$. Here $(AB)^1=0$ while $(BA)^1\neq 0$, and $(BA)^2=0$, as (b) predicts.

(d) For $k \geq 1$, by Proposition 2.4.28 applied to the product of A and $B(AB)^{k-1}$, and then (a),

$$\operatorname{tr}\Big((AB)^k\Big) = \operatorname{tr}\Big(A\Big(B(AB)^{k-1}\Big)\Big) = \operatorname{tr}\Big(B(AB)^{k-1}A\Big) = \operatorname{tr}\Big((BA)^k\Big).$$

So AB and BA cannot be told apart by the traces of their powers, although they may have different sizes and, as (c) shows, different nilpotency exponents.

Solution to Problem 2.10

Uses. Exercise 2.5.39, Proposition 2.5.17 (Algebra of inverses), Proposition 2.1.24 (Algebra of composition), Lemma 2.4.24 (Matrix units)
Techniques. Guess the inverse, then check both products (Technique 2.5.16), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (a) If $N_1^a = 0$ and $N_2^b = 0$, expand $(N_1 + N_2)^{a+b-1}$. (b) Try $U_{12} + U_{21}$.
Hint 2. (c) Factor $G + N = G\big(\mathrm{id} + G^{-1}N\big)$. Show that G^{-1} commutes with N, so that $G^{-1}N$ is nilpotent.
Sketch. (a) Every term $N_1^j N_2^{a+b-1-j}$ has $j \geq a$ or $a + b - 1 - j \geq b$. (b) $(U_{12} + U_{21})^2 = I_2$. (c) $G^{-1}N = NG^{-1}$, so $\big(G^{-1}N\big)^k = G^{-k}N^k = 0$, $\mathrm{id} + G^{-1}N$ is invertible by Exercise 2.5.39, and $(G+N)^{-1} = \sum_{j<k}(-1)^j G^{-j-1}N^j$. (d) $G = U_{12} + U_{21}$, $N = -U_{12}$: $G + N = U_{21}$.

(a) Let $N_1^a = 0$ and $N_2^b = 0$, with $N_1N_2 = N_2N_1$. By the binomial theorem,

$$(N_1 + N_2)^{a+b-1} = \sum_{j=0}^{a+b-1} \binom{a+b-1}{j} N_1^j N_2^{a+b-1-j}.$$

In each term, either $j \geq a$, and then $N_1^j = 0$, or $j \leq a-1$, and then $a + b - 1 - j \geq b$ and $N_2^{a+b-1-j} = 0$. So every term is 0.

(b) U_{12} and U_{21} satisfy $U_{12}^2 = U_{21}^2 = 0$ by Lemma 2.4.24(c), but $S = U_{12} + U_{21} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ satisfies $S^2 = U_{11} + U_{22} = I_2$, so it is invertible, and it is not nilpotent: S^k is I_2 or S. So the set of nilpotent matrices is not closed under addition. (Here $U_{12}U_{21} \neq U_{21}U_{12}$, so (a) does not apply.)

(c) Multiplying $GN = NG$ by G^{-1} on both sides gives $NG^{-1} = G^{-1}N$. So G^{-1} and N commute, and by induction $\big(G^{-1}N\big)^j = G^{-j}N^j$. If $N^k = 0$, then $N' = G^{-1}N$ satisfies $N'^k = 0$. By Exercise 2.5.39(a), applied to the nilpotent operator $-N'$, $\mathrm{id} + N'$ is invertible with inverse $\sum_{j=0}^{k-1}(-N')^j$. Now $G + N = G(\mathrm{id} + N')$ is a composite of invertible maps, so it is invertible (Proposition 2.5.17(b)), with

$$(G+N)^{-1} = (\mathrm{id} + N')^{-1}G^{-1} = \sum_{j=0}^{k-1}(-1)^j G^{-j}N^jG^{-1} = \sum_{j=0}^{k-1}(-1)^j G^{-j-1}N^j.$$

(d) Let $G = U_{12} + U_{21} \in F^{2\times 2}$, which is invertible by (b), and $N = -U_{12}$, which satisfies $N^2 = 0$. Then $G + N = U_{21}$ has a zero row, so it has no right inverse (Exercise 2.5.9(b)) and is not invertible. Indeed $GN = -U_{22}$ and $NG = -U_{11}$, so G and N do not commute.

In (c) the commutation is what lets the geometric series of Technique 2.5.16 run: without it, $\big(G^{-1}N\big)^j$ is not $G^{-j}N^j$, and $G^{-1}N$ need not be nilpotent at all.

Solution to Problem 2.11

Uses. Proposition 2.4.8 (Four views of the product), Proposition 2.3.13 (Linearity of the product), Proposition 2.4.14 (Algebra of matrices)
Hint 1. By the column view, column j of XN is $X\big(N\mathbf{e}_j\big)$. Which $N\mathbf{e}_j$ gives $\mathbf{x}_{j+1}$?
Hint 2. $N\mathbf{e}_j = \mathbf{e}_{j+1}$ for $j < n$ and $N\mathbf{e}_n = \mathbf{0}$. Compute $N^k\mathbf{e}_j$, and use $\big(XNX^{-1}\big)^k = XN^kX^{-1}$.
Sketch. (a) N has columns $\mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{0}$, so $XN\mathbf{e}_j = X\mathbf{e}_{j+1} = \mathbf{x}_{j+1}$. (b) $X^{-1}Y = N$, and $N^k\mathbf{e}_j = \mathbf{e}_{j+k}$ when $j + k \leq n$ and $\mathbf{0}$ otherwise, so $N^n = 0$ and $N^{n-1}\mathbf{e}_1 = \mathbf{e}_n$. Also $YX^{-1} = XNX^{-1}$ has k-th power XN^kX^{-1}, which is 0 exactly when $N^k = 0$.

(a) Let $N \in F^{n\times n}$ have columns $\mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{0}$, so $N\mathbf{e}_j = \mathbf{e}_{j+1}$ for $j < n$ and $N\mathbf{e}_n = \mathbf{0}$. By the column view (Proposition 2.4.8(b)) and Proposition 2.3.13(c), column j of XN is $X\big(N\mathbf{e}_j\big)$, which is $X\mathbf{e}_{j+1} = \mathbf{x}_{j+1}$ for $j < n$ and $X\mathbf{0} = \mathbf{0}$ for $j = n$. So $XN = Y$.

(b) $X^{-1}Y = N$. By induction on k, $N^k\mathbf{e}_j = \mathbf{e}_{j+k}$ if $j+k \le n$, and $N^k\mathbf{e}_j = \mathbf{0}$ otherwise. So $N^n\mathbf{e}_j = \mathbf{0}$ for every j, that is, $N^n = 0$, while $N^{n-1}\mathbf{e}_1 = \mathbf{e}_n \neq \mathbf{0}$, so $N^{n-1} \neq 0$.

$YX^{-1} = XNX^{-1}$. By induction, $\left(XNX^{-1}\right)^k = XN^kX^{-1}$: the inner factors $X^{-1}X$ cancel. So $\left(YX^{-1}\right)^n = X0X^{-1} = 0$. If $\left(YX^{-1}\right)^{n-1} = XN^{n-1}X^{-1}$ were 0, then $N^{n-1} = X^{-1}\left(XN^{n-1}X^{-1}\right)X = 0$, which is false.

The problem looks like a statement about X, but X only changes coordinates: both products are the shift N, seen in two different bases.

Solution to Problem 2.12

Uses. Lemma 2.4.24 (Matrix units), Proposition 1.4.14 (The span is the smallest subspace containing a list), Theorem 1.4.3 (Subspace criterion), Theorem 2.4.34 (Block multiplication), Proposition 2.6.8 (Kernels and images are subspaces), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. If $X \in W$ has a nonzero entry x_{jk}, multiply X on both sides by matrix units to isolate it: compute $U_{ij}XU_{kl}$.
Hint 2. For (c), show that $\ker\psi$ satisfies the hypothesis of (a).
Sketch. (a) $U_{ij}XU_{kl} = x_{jk}U_{il}$, so a nonzero entry x_{jk} puts every U_{il} in W, and the units span. (b) $A\left(\mathbf{x}\mathbf{e}_1^{\mathsf{T}}\right) = (A\mathbf{x})\mathbf{e}_1^{\mathsf{T}} \in W_1$, but $U_{11}U_{12} = U_{12} \notin W_1$. (c) $\ker\psi$ is a subspace with $\psi(AX) = \psi(A)\psi(X) = 0$ and $\psi(XA) = 0$ for $X \in \ker\psi$; it is not everything since $\psi \neq 0$, so it is $\{0\}$. Example: $\psi(X) = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}$.

(a) Suppose $W \neq \{0\}$, and let $X \in W$ have an entry $x_{jk} \neq 0$. For any i, l, the hypothesis applied twice gives $U_{ij}X \in W$ and then $U_{ij}XU_{kl} \in W$. By associativity and Lemma 2.4.24(b),

$$U_{ij}XU_{kl} = \mathbf{e}_i\left(\mathbf{e}_j^{\mathsf{T}}X\mathbf{e}_k\right)\mathbf{e}_l^{\mathsf{T}} = x_{jk}U_{il}.$$

Since W is a subspace, $U_{il} = x_{jk}^{-1}\left(x_{jk}U_{il}\right) \in W$. So W contains every matrix unit, and the matrix units span $F^{n\times n}$ (Lemma 2.4.24(a)), so $W = F^{n\times n}$ by Proposition 1.4.14(b).

(b) The matrices of W_1 are the $\mathbf{x}\mathbf{e}_1^{\mathsf{T}}$ with $\mathbf{x} \in F^n$, and they form a subspace (Theorem 1.4.3). For $A \in F^{n\times n}$, $A\left(\mathbf{x}\mathbf{e}_1^{\mathsf{T}}\right) = (A\mathbf{x})\mathbf{e}_1^{\mathsf{T}} \in W_1$. But $U_{11} \in W_1$ and $U_{11}U_{12} = U_{12}$, whose second column is $\mathbf{e}_1 \neq \mathbf{0}$, so $U_{12} \notin W_1$. So for one-sided closure there are more subspaces than $\{0\}$ and $F^{n\times n}$; in (a), the proof multiplied by matrix units on both sides, and closure on the right is exactly what W_1 lacks.

(c) $\ker\psi$ is a subspace of $F^{n\times n}$ (Proposition 2.6.8). If $X \in \ker\psi$ and $A \in F^{n\times n}$, then $\psi(AX) = \psi(A)\psi(X) = \psi(A)0 = 0$ and $\psi(XA) = \psi(X)\psi(A) = 0$, so $AX, XA \in \ker\psi$. By (a), $\ker\psi = \{0\}$ or $\ker\psi = F^{n\times n}$. The second would mean $\psi = 0$. So $\ker\psi = \{0\}$, and ψ is injective (Theorem 2.6.11).

For an example, let $\psi(X) = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} \in F^{2n\times 2n}$. It is linear, nonzero, and multiplicative, since block diagonal matrices multiply block by block (Theorem 2.4.34).

Solution to Problem 2.13

Uses. Theorem 2.4.34 (Block multiplication), Proposition 2.4.14 (Algebra of matrices), Proposition 1.2.19 (Restriction of scalars), Definition 2.3.1 (Matrix of a linear map), Proposition 2.5.20 (Inverses and matrices), Theorem 2.7.8 (Third Fundamental Theorem), Definition 2.4.26 (Trace)
Techniques. Block decomposition (Technique 2.4.35), Translate between maps and matrices (Technique 2.3.27)
Hint 1. Expand $(X+iY)(X'+iY')$ by distributivity, and compare with the block product $\rho(Z)\rho(W)$.
Hint 2. The $\mathcal{B}_m$-coordinates of $\mathbf{u}+i\mathbf{v}$, with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, are $(\mathbf{u}, \mathbf{v})$. For (c), use (b) and Proposition 2.5.20.
Sketch. (a) $ZW = (XX'-YY') + i(XY'+YX')$, and block multiplication gives $\rho(Z)\rho(W) = \begin{pmatrix} XX'-YY' & -(XY'+YX') \\ YX'+XY' & XX'-YY' \end{pmatrix}$. (b) $Z\mathbf{e}_j = X\mathbf{e}_j + iY\mathbf{e}_j$ has coordinates $\left(X\mathbf{e}_j, Y\mathbf{e}_j\right)$, column j of $\rho(Z)$; $Z\left(i\mathbf{e}_j\right) = -Y\mathbf{e}_j + iX\mathbf{e}_j$ gives column $n+j$. (c) If Z is invertible, $\rho(Z)\rho\left(Z^{-1}\right) = \rho(I_n) = I_{2n}$ in both orders. If $\rho(Z)$ is invertible, then T_Z is invertible as an $\mathbb{R}$-linear map, so it is bijective, and a bijective T_Z means Z is invertible. (d) $\operatorname{tr} X + \operatorname{tr} X$.

(a) Let $W = X' + iY'$. Matrix multiplication over $\mathbb{C}$ distributes, and real scalars and i pass through products (Proposition 2.4.14), so

$$ZW = (XX'-YY') + i(XY'+YX'),$$

with $XX'-YY'$ and $XY'+YX'$ real. By block multiplication (Theorem 2.4.34),

$$\rho(Z)\rho(W) = \begin{pmatrix} XX' - YY' & -XY' - YX' \\ YX' + XY' & -YY' + XX' \end{pmatrix} = \rho(ZW).$$

The real and imaginary parts of $Z + Z'$ and of cZ for real c are the sums and real multiples of those of Z and Z', so ρ is $\mathbb{R}$-linear. Finally $I_n = I_n + i0$, so $\rho(I_n) = \begin{pmatrix} I_n & 0 \\ 0 & I_n \end{pmatrix} = I_{2n}$.

(b) Every $\mathbf{z} \in \mathbb{C}^n$ can be written $\mathbf{z} = \mathbf{u} + i\mathbf{v}$ with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ in exactly one way (real and imaginary parts), and then

$$\mathbf{z} = u_1\mathbf{e}_1 + \cdots + u_n\mathbf{e}_n + v_1(i\mathbf{e}_1) + \cdots + v_n(i\mathbf{e}_n).$$

So $\mathcal{B}_n$ is a basis of $\mathbb{C}^n$ over $\mathbb{R}$, and the $\mathcal{B}_n$-coordinates of $\mathbf{u} + i\mathbf{v}$ are $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{2n}$. The map T_Z is $\mathbb{C}$-linear, so in particular $\mathbb{R}$-linear. Its matrix has as column j the coordinates of $Z\mathbf{e}_j = X\mathbf{e}_j + iY\mathbf{e}_j$, which are $\big(X\mathbf{e}_j, Y\mathbf{e}_j\big)$, column j of $\rho(Z)$; and as column $n + j$ the coordinates of $Z\big(i\mathbf{e}_j\big) = iX\mathbf{e}_j - Y\mathbf{e}_j$, which are $\big(-Y\mathbf{e}_j, X\mathbf{e}_j\big)$, column $n + j$ of $\rho(Z)$. So $[T_Z]_{\mathcal{B}_m\mathcal{B}_n} = \rho(Z)$.

(c) If Z is invertible, then by (a), $\rho(Z)\rho\big(Z^{-1}\big) = \rho(I_n) = I_{2n}$, and likewise in the other order, so $\rho(Z)$ is invertible with inverse $\rho\big(Z^{-1}\big)$.

Conversely, suppose $\rho(Z)$ is invertible. By (b) and Proposition 2.5.20(b), the $\mathbb{R}$-linear map $T_Z : \mathbb{C}^n \to \mathbb{C}^n$ is invertible, so it is a bijection. Bijectivity is a property of the function, not of the field of scalars, so T_Z is bijective as a $\mathbb{C}$-linear map too, and Z is invertible by Theorem 2.7.8, (I) ⇒ (V), and Proposition 2.5.20(c).

(d) The diagonal of $\rho(Z)$ consists of the diagonal of X twice, so $\operatorname{tr}\rho(Z) = 2\operatorname{tr} X = 2\operatorname{Re}(\operatorname{tr} Z)$.

For $n = 1$, $\rho(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ is the matrix of multiplication by $a + bi$ found in Exercise 2.3.35(a).

Solution to Problem 2.14

> *Uses.* Proposition 2.4.14 (Algebra of matrices), Proposition 2.6.15 (Kernels and images of composites), Example 2.6.4 (Injective, surjective, both, neither), Definition 2.4.26 (Trace), Lemma 2.4.24 (Matrix units), Definition 2.5.12 (Invertible)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18), Try small cases and small fields (Technique 1.2.38)
> *Hint 1.* (a) Expand, using $PQ = QP$ to collect terms; or note that $I_n - (P + Q - PQ) = (I_n - P)(I_n - Q)$. (b) and (c): look for counterexamples on $F[x]$ and over $\mathbb{F}_p$.
> *Hint 2.* (d) Try $P = U_{11}$ and an idempotent Q that shares its first column. (e) First show that S is injective, then cancel it.
> *Sketch.* (a) True: $I_n - (P + Q - PQ) = (I_n - P)(I_n - Q)$, a product of commuting idempotents. (b) False: multiplication by x on $F[x]$ has $\ker M = \ker M^2 = \{\mathbf{0}\}$ but $x \in \operatorname{im} M \setminus \operatorname{im} M^2$. (c) False over $\mathbb{F}_p$: I_p. (d) False: $P = U_{11}$ and $Q = U_{11} + U_{12}$ are idempotent, but $(P + Q)^2 = 4U_{11} + 2U_{12} \neq P + Q$. (e) True: TS injective makes S injective; $S(TS - \mathrm{id}) = (ST)S - S = 0$, so $TS = \mathrm{id}$.

(a) True. Write $R = P + Q - PQ$; then $I_n - R = (I_n - P)(I_n - Q)$ by distributivity. The matrices $I_n - P$ and $I_n - Q$ are idempotent ($(I_n - P)^2 = I_n - 2P + P^2 = I_n - P$) and commute, since P and Q do. A product EF of commuting idempotents is idempotent: $(EF)^2 = EFEF = EEFF = EF$. So $I_n - R$ is idempotent, and then so is $R = I_n - (I_n - R)$, by the same computation as for $I_n - P$.

(b) False. Let $V = F[x]$ and $T = M$, multiplication by x. It is injective (Example 2.6.4(e)), so M^2 is injective as well, and $\ker M = \ker M^2 = \{\mathbf{0}\}$. But $x = M(1) \in \operatorname{im} M$, while every element x^2q of $\operatorname{im} M^2$ has coefficient 0 at x, so $x \notin \operatorname{im} M^2$. (When V has a finite basis the statement is true; it follows from rank–nullity, Section 4.2.)

(c) False. Let p be a prime and $A = I_p \in \mathbb{F}_p^{p\times p}$. Then $\operatorname{tr}\big(A^k\big) = \operatorname{tr} I_p = p * 1 = 0$ for every k, but $A^k = I_p \neq 0$. Over $\mathbb{C}$ the statement is true, but its proof needs eigenvalues; the counterexample shows that some property of the field must enter.

(d) False. Let $P = U_{11}$ and $Q = U_{11} + U_{12}$, padded with zeros to size n. By Lemma 2.4.24(c), $P^2 = P$ and $Q^2 = U_{11} + U_{12} = Q$, since $U_{12}U_{11} = U_{12}U_{12} = 0$. But $P + Q = 2U_{11} + U_{12}$ and

$$(P + Q)^2 = 4U_{11} + 2U_{12},$$

whose entry in position $(1,2)$ is 2, while that of $P+Q$ is 1. Since $2 \neq 1$ in every field, $P+Q$ is not idempotent. When $PQ = QP = 0$, on the other hand, $(P+Q)^2 = P^2 + Q^2 = P + Q$; part (a) is another repair.

(e) True. Since TS is injective, so is S (Proposition 2.6.15(c)). Using $ST = \mathrm{id}_V$,

$$S(TS - \mathrm{id}_W) = (ST)S - S = S - S = 0.$$

So for every $\mathbf{w} \in W$, $S((TS - \mathrm{id}_W)(\mathbf{w})) = \mathbf{0}$, and injectivity of S gives $(TS - \mathrm{id}_W)(\mathbf{w}) = \mathbf{0}$. Hence $TS = \mathrm{id}_W$, and together with $ST = \mathrm{id}_V$, S is a two-sided inverse of T (Definition 2.5.12). Here S is cancelled on the left, which is legitimate because S is injective; compare Exercise 2.8.15, where the cancellation was not justified.

Solution to Problem 2.15

> *Uses.* Corollary 2.2.18 (Linear maps correspond to lists of vectors), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Exercise 1.5.43, Exercise 2.6.44, Proposition 1.5.14 (Appending a vector), Problem 1.13, Proposition 1.5.15 (Short lists, zeros, and repeats)
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4), Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
> *Hint 1.* $\mathbf{u}, \mathbf{e}_2, \mathbf{e}_3$ is a basis of $\mathbb{F}_p^3$. A linear map is the same thing as a list of values on it (Corollary 2.2.18).
> *Hint 2.* In (b), $\ker T = \ell$ exactly when $T(\mathbf{e}_2), T(\mathbf{e}_3)$ is independent. In (c), $|\ker T| = p^3/p^2$ by Exercise 2.6.44.
> *Sketch.* (a) T is determined by $T(\mathbf{u}) = \mathbf{0}$ and free values $T(\mathbf{e}_2), T(\mathbf{e}_3) \in \mathbb{F}_p^2$: p^4 maps. (b) $a\mathbf{u} + b\mathbf{e}_2 + c\mathbf{e}_3 \in \ker T$ iff $bT(\mathbf{e}_2) + cT(\mathbf{e}_3) = \mathbf{0}$; so $\ker T = \ell$ iff the two values are independent, and there are $(p^2-1)(p^2-p)$ independent pairs; two independent vectors span $\mathbb{F}_p^2$, so T is onto. (c) $|\ker T| = p$, and a nonzero $\mathbf{v} \in \ker T$ has $|\mathrm{span}(\mathbf{v})| = p$. Each of the $p^2 + p + 1$ lines is the kernel of $(p^2-1)(p^2-p)$ maps, by (b) after completing $\mathbf{v}$ to a basis. Total $(p^2+p+1)(p^2-1)(p^2-p) = (p^3-1)(p^3-p)$.

(a) The list $\mathbf{u}, \mathbf{e}_2, \mathbf{e}_3$ is a basis of $\mathbb{F}_p^3$: the equation $(x,y,z) = a\mathbf{u} + b\mathbf{e}_2 + c\mathbf{e}_3 = (a, a+b, a+c)$ has exactly one solution, $a = x, b = y - x, c = z - x$. By Corollary 2.2.18, a linear map $T : \mathbb{F}_p^3 \to \mathbb{F}_p^2$ is determined by the list $T(\mathbf{u}), T(\mathbf{e}_2), T(\mathbf{e}_3)$, and every list occurs. The condition fixes $T(\mathbf{u}) = \mathbf{0}$ and leaves the other two values free, so there are $(p^2)^2 = p^4$ such maps.

(b) Let $T(\mathbf{u}) = \mathbf{0}$, and write $\mathbf{w}_2 = T(\mathbf{e}_2)$, $\mathbf{w}_3 = T(\mathbf{e}_3)$. For $\mathbf{v} = a\mathbf{u} + b\mathbf{e}_2 + c\mathbf{e}_3$, $T(\mathbf{v}) = b\mathbf{w}_2 + c\mathbf{w}_3$.

If $\mathbf{w}_2, \mathbf{w}_3$ is independent, then $T(\mathbf{v}) = \mathbf{0}$ forces $b = c = 0$, so $\ker T \subseteq \ell$, and $\ell \subseteq \ker T$ since $T(\mathbf{u}) = \mathbf{0}$. If the pair is dependent, a nontrivial relation $b\mathbf{w}_2 + c\mathbf{w}_3 = \mathbf{0}$ gives $\mathbf{v} = b\mathbf{e}_2 + c\mathbf{e}_3 = (0, b, c) \in \ker T$. This vector is nonzero and not in ℓ, since the only multiple (a,a,a) of $\mathbf{u}$ with first entry 0 is $\mathbf{0}$. So $\ker T = \ell$ if and only if $\mathbf{w}_2, \mathbf{w}_3$ is independent. By Exercise 1.5.43(b), $\mathbb{F}_p^2$ contains exactly $(p^2-1)(p^2-p)$ independent lists of length 2, and each gives one T by (a).

For such a T, the span of the independent list $\mathbf{w}_2, \mathbf{w}_3$ has p^2 elements (Exercise 1.5.43(a)), so it is all of $\mathbb{F}_p^2$, and $\mathrm{im}\, T \supseteq \mathrm{span}(\mathbf{w}_2, \mathbf{w}_3) = \mathbb{F}_p^2$ (Proposition 2.1.16). So T is surjective.

(c) Let $T : \mathbb{F}_p^3 \to \mathbb{F}_p^2$ be surjective. By Exercise 2.6.44(a), $p^3 = |\ker T| \cdot |\mathrm{im}\, T| = |\ker T| \cdot p^2$, so $|\ker T| = p$. In particular $\ker T$ contains some $\mathbf{v} \neq \mathbf{0}$. The p multiples $a\mathbf{v}$ are distinct (if $a\mathbf{v} = a'\mathbf{v}$ then $(a - a')\mathbf{v} = \mathbf{0}$, so $a = a'$) and lie in $\ker T$, so $\ker T = \mathrm{span}(\mathbf{v})$, a line.

Conversely, fix a line $\mathrm{span}(\mathbf{v})$. By Proposition 1.5.14, appending to the list $\mathbf{v}$ each standard basis vector that is not in the span of the list so far gives an independent list whose span contains every $\mathbf{e}_i$, hence a basis of $\mathbb{F}_p^3$. It has exactly 3 vectors, since an independent list of length k spans p^k vectors (Exercise 1.5.43(a)). The argument of (a) and (b), with this basis in place of $\mathbf{u}, \mathbf{e}_2, \mathbf{e}_3$, shows that exactly $(p^2-1)(p^2-p)$ maps have kernel $\mathrm{span}(\mathbf{v})$, and that they are surjective.

So the surjective maps are sorted by their kernels, which are exactly the lines. By Problem 1.13(b), $\mathbb{F}_p^3$ has $(p^3-1)/(p-1) = p^2 + p + 1$ lines, so the number of surjective maps is

$$(p^2+p+1)(p^2-1)(p^2-p) = \frac{p^3-1}{p-1} \cdot (p^2-1) \cdot p(p-1) = (p^3-1)(p^3-p),$$

in agreement with Exercise 2.6.45, which counted them through their columns instead.

Solution to Problem 2.16

Uses. Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear)
Techniques. Cyclicity of trace (Technique 2.4.29), Block decomposition (Technique 2.4.35)
Hint 1. By Theorem 3.5.27, $P = G^{-1}N_kH^{-1}$ with G, H invertible. Set $M = H^{-1}G^{-1}$ and rewrite $P^2 = P$ as an equation involving N_kMN_k.
Hint 2. Partition M with a $k \times k$ block M_{11} in the corner, and compute N_kMN_k by blocks. Then $\operatorname{tr} P = \operatorname{tr}(N_kM)$ by cyclicity.
Sketch. (a) $P^2 = G^{-1}N_kMN_kH^{-1}$, so $P^2 = P$ means $N_kMN_k = N_k$. By blocks, $N_kMN_k = \left(\begin{smallmatrix} M_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right)$, so $M_{11} = I_k$, and $\operatorname{tr} P = \operatorname{tr}\left(G^{-1}N_kH^{-1}\right) = \operatorname{tr}(N_kM) = \operatorname{tr} M_{11} = k * 1$. (b) Same, with $N_kMN_k = 0$, so $M_{11} = 0$. (c) Take traces: $\sum_i \operatorname{rank} P_i = \operatorname{tr} I_n = n$ in $\mathbb{Q}$. Over $\mathbb{F}_2$, $I_1 + I_1 + I_1 = I_1$.

(a) By Theorem 3.5.27 there are invertible $G, H \in F^{n\times n}$ with $GPH = N_k$, so $P = G^{-1}N_kH^{-1}$. If $k = 0$, then $N_0 = 0$, so $P = 0$ and $\operatorname{tr} P = 0$. Let $k \geq 1$. Let $M = H^{-1}G^{-1}$. Then

$$P^2 = G^{-1}N_kH^{-1}G^{-1}N_kH^{-1} = G^{-1}N_kMN_kH^{-1},$$

and multiplying $P^2 = P$ by G on the left and H on the right gives $N_kMN_k = N_k$. Partition $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$ with $M_{11} \in F^{k\times k}$, matching $N_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$. By block multiplication,

$$N_kM = \begin{pmatrix} M_{11} & M_{12} \\ 0 & 0 \end{pmatrix}, \quad N_kMN_k = \begin{pmatrix} M_{11} & 0 \\ 0 & 0 \end{pmatrix}.$$

So $M_{11} = I_k$. By cyclicity (Proposition 2.4.28), applied to G^{-1} and N_kH^{-1},

$$\operatorname{tr} P = \operatorname{tr}\left(N_kH^{-1}G^{-1}\right) = \operatorname{tr}(N_kM) = \operatorname{tr} M_{11} = \operatorname{tr} I_k = k * 1.$$

(When $k = n$ the blocks other than M_{11} are empty and the same computation applies.)

(b) Let $k = \operatorname{rank} A$. If $k = 0$, then $A = G^{-1}N_0H^{-1} = 0$. Otherwise write $A = G^{-1}N_kH^{-1}$ and $M = H^{-1}G^{-1}$ as in (a). Now $A^2 = 0$ gives $N_kMN_k = 0$, so $M_{11} = 0$, and $\operatorname{tr} A = \operatorname{tr}(N_kM) = \operatorname{tr} M_{11} = 0$.

(c) Let $k_i = \operatorname{rank} P_i$. By Proposition 2.4.27 and (a),

$$n = \operatorname{tr} I_n = \operatorname{tr} P_1 + \cdots + \operatorname{tr} P_r = k_1 + \cdots + k_r$$

in $\mathbb{Q}$, where the k_i are integers; so $k_1 + \cdots + k_r = n$. Over $\mathbb{F}_2$, take $n = 1$, $r = 3$, and $P_1 = P_2 = P_3 = I_1$: they are idempotent, their sum is $(1 + 1 + 1)I_1 = I_1$, and their ranks add up to $3 \neq 1$. Part (a) still holds there, but the trace only records $3 * 1 = 1$.

The trace of an idempotent counts the dimension of the space it projects onto, but only modulo the characteristic. In particular, over $\mathbb{Q}$ an idempotent with trace 0 is 0, and over $\mathbb{F}_2$, I_2 is a nonzero idempotent with trace 0.

Solution to Problem 2.17

Uses. Proposition 2.4.14 (Algebra of matrices), Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Lemma 2.4.24 (Matrix units)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Expand $(A + B)^k$ as a sum of 2^k products $W_1W_2\cdots W_k$ with each $W_t \in \{A, B\}$. The products with only A or only B give A^k and B^k. Show that every other product has trace 0.
Hint 2. A product that contains both letters has, reading cyclically (so that W_k is followed by W_1), some A immediately followed by a B. Rotate the product with Proposition 2.4.28 so that this pair becomes adjacent.
Sketch. (a) In a mixed word some A is cyclically followed by B: otherwise every letter after an A is an A, and going around the cycle all letters are A. If the pair is inside the word, the product contains $AB = 0$. If it is $W_k = A$, $W_1 = B$, then $\operatorname{tr}(W_1\cdots W_k) = \operatorname{tr}(W_kW_1\cdots W_{k-1})$, which contains AB. (b) $A = U_{11}$, $B = U_{21}$: $(A + B)^2 = U_{11} + U_{21}$, but $A^2 + B^2 = U_{11}$.

(a) By distributivity (Proposition 2.4.14), $(A + B)^k$ is the sum, over all 2^k choices of $W_1, \ldots, W_k \in \{A, B\}$, of the products $W_1W_2\cdots W_k$. The choice with every $W_t = A$ gives A^k, and the choice with every $W_t = B$ gives B^k. By Proposition 2.4.27 it suffices to show that every other product has trace 0.

Consider a product in which both A and B occur, and read the indices cyclically, so that position k is followed by position 1. Some A is followed by a B: if every A were followed by an A, then starting from any position holding A and moving forward around the cycle, every position would hold A, contradicting the presence of B. So there is a t with $W_t = A$ and $W_{t+1} = B$, where W_{k+1} means W_1.

- If $t < k$, the product contains the adjacent factors $W_t W_{t+1} = AB = 0$, so it is 0 by associativity, and its trace is 0.
- If $t = k$, then $W_k = A$ and $W_1 = B$. By Proposition 2.4.28, applied to $X = W_1 \cdots W_{k-1}$ and $Y = W_k$,

$$\operatorname{tr}(W_1 \cdots W_k) = \operatorname{tr}(W_k W_1 \cdots W_{k-1}),$$

  and the product on the right contains $W_k W_1 = AB = 0$. Its trace is 0.

So $\operatorname{tr}\big((A+B)^k\big) = \operatorname{tr}\big(A^k\big) + \operatorname{tr}\big(B^k\big)$.

(b) Let $A = U_{11}$ and $B = U_{21}$, with $U_{ij} = \mathbf{e}_i \mathbf{e}_j^{\mathsf{T}} \in F^{2\times 2}$. By Lemma 2.4.24(c), $AB = 0$, $BA = U_{21}$, $A^2 = U_{11}$, and $B^2 = 0$. So

$$(A+B)^2 = A^2 + AB + BA + B^2 = U_{11} + U_{21} \neq U_{11} = A^2 + B^2.$$

The extra term BA has trace 0, as (a) requires. The trace cannot see a cyclic rotation of the factors, and one rotation is all it takes to bring an A next to a B.

Solution to Problem 2.18

Uses. Proposition 2.4.8 (Four views of the product), Definition 2.4.21 (Transpose), Proposition 2.5.4 (Left inverses give uniqueness), Exercise 2.6.44
Techniques. Test injectivity at **0** (Technique 2.6.12)
Hint 1. Record club j as the vector $\mathbf{v}_j \in \mathbb{F}_2^n$ with 1 in position r exactly when resident r is a member. Then $\mathbf{v}_i^{\mathsf{T}} \mathbf{v}_j$ is the number of common members, modulo 2.
Hint 2. Let $M \in \mathbb{F}_2^{n\times k}$ have columns $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Compute $M^{\mathsf{T}} M$. What does a left inverse say about $T_M : \mathbb{F}_2^k \to \mathbb{F}_2^n$?
Sketch. (a) The entries of $M^{\mathsf{T}} M$ are $\mathbf{v}_i^{\mathsf{T}} \mathbf{v}_j$, which are 1 on the diagonal (odd sizes) and 0 off it (even intersections), so $M^{\mathsf{T}} M = I_k$. Then T_M is injective, and an injective linear map $\mathbb{F}_2^k \to \mathbb{F}_2^n$ has $k \le n$ (Exercise 2.6.44). (b) The n one-member clubs. (c) Pair the residents as $\{1,2\}, \{3,4\}, \{5,6\}$, and take the 7 nonempty unions of pairs; the Gram matrix is now 0, and the argument collapses.

(a) Number the residents $1, \ldots, n$, and for club j let $\mathbf{v}_j \in \mathbb{F}_2^n$ have r-th entry 1 if resident r belongs to club j and 0 otherwise. For any clubs i and j,

$$\mathbf{v}_i^{\mathsf{T}} \mathbf{v}_j = \sum_{r=1}^{n} (\mathbf{v}_i)_r (\mathbf{v}_j)_r$$

adds a 1 for each common member, so it is the number of common members, reduced modulo 2. By hypothesis this is 1 for $i = j$ and 0 for $i \neq j$.

Let $M \in \mathbb{F}_2^{n\times k}$ have columns $\mathbf{v}_1, \ldots, \mathbf{v}_k$. The (i,j) entry of $M^{\mathsf{T}} M$ is row i of M^{T}, which is $\mathbf{v}_i^{\mathsf{T}}$, times column j of M (Proposition 2.4.8(a)), so $M^{\mathsf{T}} M = I_k$. Thus M^{T} is a left inverse of M, and $T_M : \mathbb{F}_2^k \to \mathbb{F}_2^n$ is injective (Proposition 2.5.4). By Exercise 2.6.44(b), $k \le n$.

(b) Let club j consist of resident j alone, for $j = 1, \ldots, n$. Each club has one member and two different clubs share none.

(c) Split six residents into the pairs $\{1,2\}, \{3,4\}, \{5,6\}$, and let the clubs be the 7 nonempty unions of these pairs:

$$\{1,2\}, \quad \{3,4\}, \quad \{5,6\}, \quad \{1,2,3,4\}, \quad \{1,2,5,6\}, \quad \{3,4,5,6\}, \quad \{1,\ldots,6\}.$$

Each is a union of pairs, so it has an even number of members, and the intersection of two of them is again a union of pairs (those they share), so it is even as well. That is $7 > 6$ clubs. In the vector language every

$\mathbf{v}_i^{\mathsf{T}}\mathbf{v}_j$ is now 0, so $M^{\mathsf{T}}M = 0$ and there is no left inverse to be had: the oddness of the clubs is what made the vectors independent.

This is the Oddtown theorem. For the even version, one can show that at most $2^{\lfloor n/2 \rfloor}$ clubs are possible, counting the empty club, but the proof needs more than this chapter.

Solution to Problem 2.19

Uses. Proposition 2.4.14 (Algebra of matrices), Proposition 2.5.4 (Left inverses give uniqueness), Lemma 2.4.24 (Matrix units)
Techniques. Test injectivity at **0** (Technique 2.6.12)
Hint 1. Multiply $A^2 + B^2$ by a nonzero matrix built from A and B, and use both hypotheses.
Hint 2. Expand $(A^2 + B^2)(A - B)$. For (b), use block diagonal matrices whose first block is the same and whose second blocks are different matrices with square 0.
Sketch. (a) No. $(A^2 + B^2)(A - B) = (A^3 - B^3) - (A^2B - B^2A) = 0$ and $A - B \neq 0$; a nonzero column $\mathbf{x}$ of $A - B$ has $(A^2 + B^2)\mathbf{x} = \mathbf{0}$, so $A^2 + B^2$ is not injective. (b) $A = U_{11} + U_{23}$, $B = U_{11} + U_{32}$ in $\mathbb{R}^{3\times 3}$: $A^2 = B^2 = A^3 = B^3 = U_{11}$, $A^2B = B^2A = U_{11}$, and $A^2 + B^2 = 2U_{11}$.

(a) No. By distributivity,

$$(A^2 + B^2)(A - B) = A^3 - A^2B + B^2A - B^3 = (A^3 - B^3) - (A^2B - B^2A) = 0.$$

Since $A \neq B$, some column $\mathbf{x}$ of $A - B$ is nonzero, and by the column view $(A^2 + B^2)\mathbf{x}$ is the corresponding column of 0. So $T_{A^2+B^2}$ is not injective, and $A^2 + B^2$ has no left inverse (Proposition 2.5.4), let alone an inverse. The argument works over every field.

(b) Let $A = U_{11} + U_{23}$ and $B = U_{11} + U_{32}$ in $\mathbb{R}^{3\times 3}$. By Lemma 2.4.24(c), U_{11} kills U_{23} and U_{32} on both sides, and $U_{23}^2 = U_{32}^2 = 0$. So $A^2 = B^2 = U_{11}$, hence $A^3 = U_{11}A = U_{11} = B^3$, and $A^2B = U_{11}B = U_{11} = U_{11}A = B^2A$. Also $A \neq B$. But $A^2 + B^2 = 2U_{11} \neq 0$.

This is Problem A2 of the 1991 Putnam Competition. The idea is to find a product with $A^2 + B^2$ that the hypotheses kill, which is the zero-divisor phenomenon of Example 2.4.19 put to work.

Solution to Problem 2.20

Uses. Problem 2.9, Proposition 2.5.28 (Inverse of a 2×2 matrix), Exercise 2.5.46, Exercise 2.5.32, Lemma 2.4.24 (Matrix units), Theorem 2.4.34 (Block multiplication)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Try small cases and small fields (Technique 1.2.38)
Hint 1. (a) By Problem 2.9(a), $(BA)^3 = B(AB)^2A = 0$. So $N = BA$ is a 2×2 matrix with $N^3 = 0$. What does Exercise 2.5.46(a) say about such an N?
Hint 2. N is not invertible, so $ad - bc = 0$ and $N^2 = (\operatorname{tr} N)N$. For (b), look for 3×3 examples among sums of two matrix units, with AB a single matrix unit of square 0 and BA a path of length 2.
Sketch. (a) $N = BA$ has $N^3 = 0$, so N is not invertible and its $ad - bc$ is 0; then $N^2 = tN$ with $t = \operatorname{tr} N$, so $0 = N^3 = t^2N$, and either $t = 0$ or $N = 0$. In both cases $N^2 = tN = 0$. (b) $A = U_{11} + U_{22}$, $B = U_{12} + U_{31}$: $AB = U_{12}$, $(AB)^2 = 0$, but $BA = U_{12} + U_{31}$ and $(BA)^2 = U_{32}$. Pad with zero blocks for $n > 3$.

(a) Let $N = BA$. By Problem 2.9(a) with $k = 2$, $N^3 = B(AB)^2A = B0A = 0$. So N is not invertible (Exercise 2.5.32(b)), and by Proposition 2.5.28(a), $\Delta = ad - bc = 0$ for its entries. By Exercise 2.5.46(a), $N^2 = tN - \Delta I_2 = tN$, where $t = \operatorname{tr} N$. Then

$$0 = N^3 = tN^2 = t^2N.$$

If $t \neq 0$, this gives $N = 0$; otherwise $t = 0$. Either way $N^2 = tN = 0$, that is, $BABA = 0$.

(b) Let $n = 3$, $A = U_{11} + U_{22}$, and $B = U_{12} + U_{31}$, where $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$. By Lemma 2.4.24(c),

$$AB = U_{11}U_{12} = U_{12}, \quad (AB)^2 = U_{12}U_{12} = 0,$$

since the other three products of units vanish. On the other side,

$$BA = U_{12}U_{22} + U_{31}U_{11} = U_{12} + U_{31}, \quad (BA)^2 = U_{31}U_{12} = U_{32} \neq 0.$$

Explicitly,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

For $n > 3$, use the block diagonal matrices $\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$ in $F^{n\times n}$; by Theorem 2.4.34 their products are computed in the top left 3×3 block, so the same identities hold there.

In size 2 a matrix with $N^3 = 0$ already has $N^2 = 0$, and that is what (a) used. In size 3, BA is still nilpotent by Problem 2.9, but it can take three steps to die, as $U_{12} + U_{31}$ does: $\mathbf{e}_2 \mapsto \mathbf{e}_1 \mapsto \mathbf{e}_3 \mapsto \mathbf{0}$. This is Problem A5 of the 1990 Putnam Competition.

Section 3.1

Solution to Exercise 3.1.7

Uses. Proposition 3.1.5 (Three faces of a linear system), Definition 3.1.4 (Coefficient matrix; matrix and vector forms)
Hint 1. Row i of A lists the coefficients of x_1, x_2, x_3 in equation i, with 0 for a missing unknown.
Hint 2. To test a vector, substitute it into each equation, or compute $A\mathbf{x}$ and compare with $\mathbf{y}$.
Sketch. A has rows $(1, 0, -1)$, $(0, 1, 2)$, $(1, 1, 0)$ and $\mathbf{y} = (2, -1, 1)$. Then $A(2, -1, 0) = (2, -1, 1)$, a solution, and $A(3, -3, 1) = (2, -1, 0)$, not a solution.

The matrix and vector forms are

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad x_1\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}.$$

For $\mathbf{x} = (2, -1, 0)$, the left sides are $2 - 0 = 2$, $-1 + 0 = -1$, and $2 - 1 = 1$, so $(2, -1, 0)$ is a solution. For $\mathbf{x} = (3, -3, 1)$, the first two equations hold ($3 - 1 = 2$ and $-3 + 2 = -1$), but $3 + (-3) = 0 \neq 1$, so $(3, -3, 1)$ is not a solution.

Solution to Exercise 3.1.11

Uses. Technique 3.1.9 (Undetermined coefficients), Definition 1.2.13 (Polynomials)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. Expand the right side and collect the coefficients of x^2, x, and 1.
Hint 2. Two polynomials are equal exactly when their coefficients agree, which gives three equations.
Sketch. The right side is $(c_1 + c_2)x^2 + (2c_1 - 2c_2)x + (c_1 + c_2 + c_3)$, so the system is $c_1 + c_2 = 1$, $2c_1 - 2c_2 = 0$, $c_1 + c_2 + c_3 = 0$, with unique solution $(1/2, 1/2, -1)$.

Expanding, $c_1(x + 1)^2 + c_2(x - 1)^2 + c_3 = (c_1 + c_2)x^2 + (2c_1 - 2c_2)x + (c_1 + c_2 + c_3)$. Comparing coefficients with $x^2 = 1x^2 + 0x + 0$ gives the system

$$c_1 + c_2 = 1, \quad 2c_1 - 2c_2 = 0, \quad c_1 + c_2 + c_3 = 0.$$

The second equation gives $c_1 = c_2$, and then the first gives $c_1 = c_2 = 1/2$; the third gives $c_3 = -1$. Conversely, $(1/2, 1/2, -1)$ satisfies all three equations. So the system has exactly one solution, and

$$x^2 = \frac{1}{2}(x + 1)^2 + \frac{1}{2}(x - 1)^2 - 1.$$

Solution to Exercise 3.1.21

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 3.1.16 (Homogeneous solution sets are subspaces)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. The difference of two solutions of the same system lies in the kernel.
Hint 2. Use linearity of T_A for the second question, and Proposition 3.1.19 for the third.
Sketch. $\mathbf{z} = (1, 2, 0) - (0, 1, 1) = (1, 1, -1) \in \ker T_A$, and it is nonzero. $A(2, 4, 0) = 2(3, 1) = (6, 2)$. The vectors $(1, 2, 0) + t(1, 1, -1)$, $t \in \mathbb{R}$, all solve $A\mathbf{x} = (3, 1)$.

Put $\mathbf{z} = (1,2,0) - (0,1,1) = (1,1,-1)$. Then $A\mathbf{z} = (3,1) - (3,1) = \mathbf{0}$, so $\mathbf{z}$ is a nonzero vector of $\ker T_A$. Next, $A(2,4,0) = 2A(1,2,0) = 2(3,1) = (6,2)$, so $(2,4,0)$ solves $A\mathbf{x} = (6,2)$. Finally, $\ker T_A$ is a subspace (Corollary 3.1.16), so $t\mathbf{z} \in \ker T_A$ for every $t \in \mathbb{R}$, and by Proposition 3.1.19 every vector $(1,2,0) + t(1,1,-1) = (1+t, 2+t, -t)$ is a solution of $A\mathbf{x} = (3,1)$. These vectors are distinct for distinct t (compare third entries), so there are infinitely many.

Solution to Exercise 3.1.28

Uses. Definition 3.1.1 (Linear system), Definition 3.1.4 (Coefficient matrix; matrix and vector forms), Definition 3.1.12 (Augmented matrix), Corollary 3.1.23 (Lines through solutions)
Hint 1. Move every term involving an unknown to the left and every constant to the right. Is each term then a scalar times a single unknown?
Hint 2. Coefficients may be any real numbers, including $\sin 1$ and π. A product of two unknowns is not allowed.
Sketch. (a), (c), and (d) are linear; rearranged they read $x_1 - 2x_2 + x_3 = 1$, $-4x_1 + x_3 = 0$; then $(\sin 1)x_1 + ex_2 - \pi x_3 = 0$; then $3x_1 - x_2 = -3$, $x_2 + x_3 = \sqrt{2}$. (b) contains the product x_1x_2.

Adding the same expression to both sides of an equation does not change its solutions, so we may move terms freely.

(a) Linear. Rearranged, the system is $x_1 - 2x_2 + x_3 = 1$, $-4x_1 + x_3 = 0$, so

$$A = \begin{pmatrix} 1 & -2 & 1 \\ -4 & 0 & 1 \end{pmatrix}, \quad (A \mid \mathbf{y}) = \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ -4 & 0 & 1 & 0 \end{array}\right).$$

(b) Not linear: the term x_1x_2 is a product of two unknowns. In fact its solution set, the points with $x_3 = 1$ and $x_1 = 0$ or $x_2 = 0$, is not the solution set of any linear system: it contains $(1,0,1)$ and $(0,1,1)$ but not $\left(\frac{1}{2}, \frac{1}{2}, 1\right)$, which lies on the line through them, contrary to Corollary 3.1.23.

(c) Linear: one equation $(\sin 1)x_1 + ex_2 - \pi x_3 = 0$, so

$$A = (\sin 1 \;\; e \;\; -\pi), \quad (A \mid \mathbf{y}) = (\sin 1 \;\; e \;\; -\pi \mid 0).$$

(d) Linear. Rearranged, the system is $3x_1 - x_2 = -3$, $x_2 + x_3 = \sqrt{2}$, so

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad (A \mid \mathbf{y}) = \left(\begin{array}{ccc|c} 3 & -1 & 0 & -3 \\ 0 & 1 & 1 & \sqrt{2} \end{array}\right).$$

Solution to Exercise 3.1.29

Uses. Technique 3.1.9 (Undetermined coefficients), Proposition 3.1.5 (Three faces of a linear system), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Particular plus homogeneous (Technique 3.1.18), Undetermined coefficients (Technique 3.1.9)
Hint 1. The combinations $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = \mathbf{y}$ are the solutions of a linear system with augmented matrix $(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3 \mid \mathbf{y})$. Eliminate.
Hint 2. Notice $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$. For (b), look for a linear condition that all three columns satisfy.
Sketch. Elimination reduces (a) to $c_1 + c_3 = 3$, $c_2 + c_3 = -1$, so the combinations are $(3-t)\mathbf{a}_1 + (-1-t)\mathbf{a}_2 + t\mathbf{a}_3$. In (b) elimination produces $0 = 1$; equivalently, every column satisfies $y_1 - 2y_2 + y_3 = 0$ and $(1,0,0)$ does not.

(a) We want all (c_1, c_2, c_3) with $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = (1,5,9)$. The augmented matrix is

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 5 \\ 3 & 0 & 3 & 9 \end{array}\right).$$

Subtracting 2 times row 1 from row 2 and 3 times row 1 from row 3 gives rows $(0,-3,-3 \mid 3)$ and $(0,-6,-6 \mid 6)$. Subtracting 2 times the new row 2 from the new row 3 gives the zero row, and scaling row 2 by $-1/3$ gives $(0,1,1 \mid -1)$. Finally, subtracting 2 times this row from row 1 gives $(1,0,1 \mid 3)$. So the system is equivalent to $c_1 + c_3 = 3$, $c_2 + c_3 = -1$, and its solutions are $(c_1, c_2, c_3) = (3-t, -1-t, t)$ for $t \in \mathbb{R}$. That is,

$$(1,5,9) = (3-t)\mathbf{a}_1 + (-1-t)\mathbf{a}_2 + t\mathbf{a}_3 \quad (t \in \mathbb{R}),$$

and these are all the ways. This matches Proposition 3.1.19: $(3,-1,0)$ is one solution, and the homogeneous solutions are the multiples of $(-1,-1,1)$, reflecting $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$.

(b) The same row operations applied to $(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3 \mid (1,0,0))$ give the rows $(1,2,3 \mid 1)$, $(0,-3,-3 \mid -2)$, $(0,-6,-6 \mid -3)$, and then the last row becomes $(0,0,0 \mid 1)$. So every solution would satisfy $0 = 1$, and there is none. Here is a second argument, from the column picture. Each column satisfies $y_1 - 2y_2 + y_3 = 0$: $1 - 4 + 3 = 0$, $2 - 2 + 0 = 0$, and $3 - 6 + 3 = 0$. This condition is preserved by sums and scalar multiples, so every linear combination of the columns satisfies it. But $(1,0,0)$ gives $1 \neq 0$.

Solution to Exercise 3.1.30

Uses. Corollary 3.1.25 (Counting solutions over a finite field), Corollary 3.1.24 (No, one, or infinitely many solutions)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. Solve the second equation for x_1. In $\mathbb{F}_3$, $-2 = 1$.
Hint 2. Then the first equation determines x_3 in terms of x_2, and x_2 is free. For the count, also solve the homogeneous system.
Sketch. Over $\mathbb{F}_3$: $x_1 = x_2$ and $x_3 = 1 + x_2$, giving $(0,0,1)$, $(1,1,2)$, $(2,2,0)$. The homogeneous solutions are (t,t,t), also three. Over $\mathbb{R}$: $(-2s, s, 1+s)$, infinitely many.

Over $\mathbb{F}_3$, the second equation says $x_1 = -2x_2 = x_2$. Substituting into the first, $x_3 = 1 - 2x_2 = 1 + x_2$. Conversely, for each $t \in \mathbb{F}_3$ the vector $(t, t, 1+t)$ satisfies both equations: $t + t + 1 + t = 1 + 3t = 1$ and $t + 2t = 3t = 0$. So the solutions are

$$(0,0,1), \quad (1,1,2), \quad (2,2,0).$$

The homogeneous system $x_1 + x_2 + x_3 = 0$, $x_1 + 2x_2 = 0$ gives $x_1 = x_2$ and $x_3 = -2x_2 = x_2$ in the same way, so its solutions are $(0,0,0)$, $(1,1,1)$, $(2,2,2)$: three, as Corollary 3.1.25 predicts.

Over $\mathbb{R}$, the same steps give $x_1 = -2x_2$ and $x_3 = 1 - x_1 - x_2 = 1 + x_2$, so the solutions are $(-2s, s, 1+s)$ for $s \in \mathbb{R}$. There are infinitely many, in line with Corollary 3.1.24.

Solution to Exercise 3.1.31

Uses. Definition 3.1.1 (Linear system)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. Once $x_1, \ldots, x_{n-1}$ are chosen, how many choices of x_n work?
Hint 2. For (b), encode a subset $X \subseteq \{1, \ldots, n\}$ as the vector $\mathbf{x} \in \mathbb{F}_2^n$ with $x_i = 1$ exactly when $i \in X$. What is $x_1 + \cdots + x_n$ in $\mathbb{F}_2$?
Sketch. The map $\mathbf{x} \mapsto (x_1, \ldots, x_{n-1})$ is a bijection from the solution set onto $\mathbb{F}_p^{n-1}$, since $x_n = c - x_1 - \cdots - x_{n-1}$ is forced. For (b), $x_1 + \cdots + x_n$ in $\mathbb{F}_2$ is the parity of $|X|$, so the even subsets correspond to the solutions of $x_1 + \cdots + x_n = 0$ over $\mathbb{F}_2$.

(a) If $\mathbf{x}$ is a solution, then $x_n = c - x_1 - \cdots - x_{n-1}$. Conversely, for any $x_1, \ldots, x_{n-1} \in \mathbb{F}_p$, this choice of x_n gives a solution. So $\mathbf{x} \mapsto (x_1, \ldots, x_{n-1})$ is a bijection from the solution set onto $\mathbb{F}_p^{n-1}$, which has p^{n-1} elements.

(b) Encode $X \subseteq \{1, \ldots, n\}$ as $\mathbf{x} \in \mathbb{F}_2^n$ with $x_i = 1$ if $i \in X$ and $x_i = 0$ otherwise; this is a bijection between subsets and vectors. The sum $x_1 + \cdots + x_n$ computed in $\mathbb{F}_2$ is the remainder of $|X|$ on division by 2 (Lemma 1.1.21), so $|X|$ is even exactly when $x_1 + \cdots + x_n = 0$ in $\mathbb{F}_2$. By (a) with $p = 2$ and $c = 0$, there are 2^{n-1} such vectors.

Solution to Exercise 3.1.31

Uses. Corollary 3.1.25 (Counting solutions over a finite field)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. Every $\mathbf{x} \in \mathbb{F}_p^n$ solves exactly one of the p equations $x_1 + \cdots + x_n = c$, $c \in \mathbb{F}_p$.
Hint 2. Each of these equations is consistent. What does Corollary 3.1.25 say about the sizes of their solution sets?
Sketch. The p solution sets partition $\mathbb{F}_p^n$, and each has the size of the homogeneous solution set, since each equation is consistent ($c\mathbf{e}_1$ is a solution). So each has $p^n/p = p^{n-1}$ elements.

(a) Every $\mathbf{x} \in \mathbb{F}_p^n$ solves exactly one of the equations $x_1 + \cdots + x_n = c$ with $c \in \mathbb{F}_p$, namely the one with $c = x_1 + \cdots + x_n$. So the p solution sets S_c partition $\mathbb{F}_p^n$. Each equation is consistent, since $c\mathbf{e}_1 = (c, 0, \ldots, 0)$ is a solution, so by Corollary 3.1.25 every S_c has as many elements as S_0. Hence $p\,|S_0| = p^n$, and $|S_c| = |S_0| = p^{n-1}$ for every c.

(b) As in the first approach, the even subsets correspond to S_0 over $\mathbb{F}_2$, which has 2^{n-1} elements. In this language the argument says something familiar: toggling whether $1 \in X$ (that is, adding $\mathbf{e}_1$) is a bijection between even and odd subsets.

Solution to Exercise 3.1.32

Uses. Proposition 1.1.17 (No zero divisors), Corollary 3.1.24 (No, one, or infinitely many solutions)
Hint 1. Subtract the second equation from the first and factor: $(1-k)(x_1-x_2)=0$.
Hint 2. Split into the cases $k=1$ and $k\neq 1$. In the second case $x_1=x_2$; substitute back. Which value of k is then special?
Sketch. Any solution satisfies $(1-k)(x_1-x_2)=0$. If $k=1$ both equations are $x_1+x_2=1$. If $k\neq 1$, then $x_1=x_2$ and $(1+k)x_1=1$: no solution if $k=-1$, else $x_1=x_2=1/(1+k)$. The argument works in any field; over $\mathbb{F}_3$, $k=1$ gives 3 solutions, $k=2=-1$ none, $k=0$ one.

The algebra below is valid over any field, so we do it once. Suppose (x_1,x_2) is a solution. Subtracting the equations gives $(1-k)x_1-(1-k)x_2=0$, that is, $(1-k)(x_1-x_2)=0$.

Case $k=1$. Both equations read $x_1+x_2=1$, so the solution set is $\{(1-t,t)\mid t\in F\}$, one solution for each $t\in F$.

Case $k\neq 1$. Then $1-k\neq 0$, so $x_1=x_2$ by Proposition 1.1.17, and either equation becomes $(1+k)x_1=1$. If $k=-1$, this says $0=1$, so there is no solution. If $k\neq -1$, then $x_1=x_2=1/(1+k)$. Conversely, this pair is a solution, since $(1+k)/(1+k)=1$ in both equations. So there is exactly one solution.

(a) Over $\mathbb{R}$: infinitely many solutions for $k=1$, none for $k=-1$, and exactly one, $(1/(1+k),1/(1+k))$, otherwise.

(b) Over $\mathbb{F}_3$: for $k=1$ there are exactly 3 solutions, $(1,0)$, $(0,1)$, $(2,2)$; for $k=2=-1$ there are none; for $k=0$ there is exactly one, $(1,1)$.

Solution to Exercise 3.1.33

Uses. Proposition 3.1.8 (Systems and the map T_A), Definition 2.3.15 (The linear map of a matrix), Proposition 2.6.8 (Kernels and images are subspaces)
Hint 1. For (a), add solutions and scale them.
Hint 2. For (b), the set of $\mathbf{y}$ for which the system is consistent is $\operatorname{im} T_A$, a subspace. Is its complement closed under addition? Try $\mathbf{y}_2=-\mathbf{y}_1$.
Sketch. (a) True: if $A\mathbf{u}=\mathbf{y}_1$ and $A\mathbf{v}=\mathbf{y}_2$, then $A(\mathbf{u}+\mathbf{v})=\mathbf{y}_1+\mathbf{y}_2$ and $A(c\mathbf{u})=c\mathbf{y}_1$. (b) False: with $A=\begin{pmatrix}1\\0\end{pmatrix}$, both $(0,1)$ and $(0,-1)$ are outside $\operatorname{im} T_A$, but their sum $\mathbf{0}$ is not.

(a) True. Let $A\mathbf{u}=\mathbf{y}_1$ and $A\mathbf{v}=\mathbf{y}_2$. Since T_A is linear, $A(\mathbf{u}+\mathbf{v})=A\mathbf{u}+A\mathbf{v}=\mathbf{y}_1+\mathbf{y}_2$ and $A(c\mathbf{u})=cA\mathbf{u}=c\mathbf{y}_1$. In other words, by Proposition 3.1.8 the consistent right-hand sides form $\operatorname{im} T_A$, which is a subspace (Proposition 2.6.8).

(b) False. Over any field, let $A=\begin{pmatrix}1\\0\end{pmatrix}\in F^{2\times 1}$, $\mathbf{y}_1=(0,1)$, and $\mathbf{y}_2=(0,-1)$. For every $x\in F$, $Ax=(x,0)$ has second entry 0, so both systems are inconsistent. But $\mathbf{y}_1+\mathbf{y}_2=\mathbf{0}$, and $A\mathbf{x}=\mathbf{0}$ is consistent.

Solution to Exercise 3.1.34

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 3.1.22 (Uniqueness depends only on A)
Hint 1. Reread the hypothesis of Proposition 3.1.19. What does it assume about $\mathbf{x}_0$?
Hint 2. Look for a matrix with more rows than columns.
Sketch. The proposition needs a solution $\mathbf{x}_0$ to exist, and the argument never shows one does. For $A=\begin{pmatrix}1\\1\end{pmatrix}$, $Ax=0$ forces $x=0$, but $Ax=(1,0)$ has no solution.

Proposition 3.1.19 describes the solution set *given* a solution $\mathbf{x}_0$. The argument uses $\mathbf{x}_0$ without showing that the system is consistent, so it only proves that $A\mathbf{x}=\mathbf{y}$ has *at most* one solution, which is Corollary 3.1.22.

The claim itself is false. Let $A=\begin{pmatrix}1\\1\end{pmatrix}\in\mathbb{R}^{2\times 1}$. Then $Ax=(x,x)$, so $Ax=\mathbf{0}$ only for $x=0$. But $Ax=(1,0)$ would need $x=1$ and $x=0$, so it has no solution.

Solution to Exercise 3.1.35

Uses. Definition 3.1.1 (Linear system)
Hint 1. Is $(2, 1)$ a solution of the original system?
Hint 2. Each new equation is a combination of the old ones, so no solution is lost. Can the old equations be recovered from the two new ones?
Sketch. The two replacements were made at the same time, and the second used the old first equation, which no longer appears in the new system. The new equations are negatives of each other and only say $x_2 = 1$; the original system cannot be recovered from them. The original solution set is $\{(1, 1)\}$, so the points $(t, 1)$ with $t \neq 1$ were gained and nothing was lost.

The original system has the unique solution $(1, 1)$: adding the equations gives $2x_1 = 2$, so $x_1 = 1$, and then $x_2 = 1$. The student's answer contains $(2, 1)$, which is not a solution, since $2 - 1 \neq 0$.

Every solution of the old system satisfies both new equations, because each new equation is a combination of the old ones. So no solution is lost, and the student's set does contain $(1, 1)$. The problem is the other direction. The two new equations, $2x_2 = 2$ and $-2x_2 = -2$, are negatives of each other and together say only $x_2 = 1$. The equation $x_1 + x_2 = 2$ cannot be recovered from them, so solutions are gained: all $(t, 1)$ with $t \neq 1$.

The error is performing two replacements simultaneously when the second one uses the equation that the first one replaces. Replacing an equation E_i by $E_i - E_j$ is reversible only while E_j is still present: then $E_i = \big(E_i - E_j\big) + E_j$. Done one at a time, the operations are harmless. Replacing the first equation by the first minus the second gives $2x_2 = 2$, $x_1 - x_2 = 0$, and then subtracting the new first equation from the second gives $2x_2 = 2$, $x_1 - 3x_2 = -2$, whose only solution is $(1, 1)$.

Solution to Exercise 3.1.36

Uses. Definition 3.1.1 (Linear system)
Hint 1. A plane can contain lines in many directions. Must a plane that contains the direction of L be parallel to one of the planes through L?
Hint 2. Try three planes that all contain lines parallel to the x_3-axis, for example planes whose equations do not involve x_3.
Sketch. The step "parallel to neither plane, hence not parallel to L" is unjustified. Counterexample: $x_1 = 0$, $x_2 = 0$, $x_1 + x_2 = 1$. No two coefficient vectors are proportional, but a common point would have $0 + 0 = 1$.

The unjustified step is "the third plane is parallel to neither of the first two, so it is not parallel to L". Being parallel to a line is much weaker than being parallel to a plane: a plane is parallel to L when it contains lines in the direction of L, and many non-parallel planes do. If the third plane is parallel to L and does not contain it, it misses L entirely.

Counterexample: the planes $x_1 = 0$, $x_2 = 0$, and $x_1 + x_2 = 1$. Their coefficient vectors $(1, 0, 0)$, $(0, 1, 0)$, $(1, 1, 0)$ are pairwise not proportional, so no two of the planes are parallel. The first two meet in the x_3-axis L, and the third contains the lines $\{(1, 0, t) \mid t \in \mathbb{R}\}$ and $\{(0, 1, t) \mid t \in \mathbb{R}\}$, both parallel to L. A common point would satisfy $x_1 = 0$, $x_2 = 0$, and $x_1 + x_2 = 1$, which gives $0 = 1$. So there is none. This is the configuration of Figure 3.1.3.

Solution to Exercise 3.1.37

Uses. Technique 3.1.9 (Undetermined coefficients), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Undetermined coefficients (Technique 3.1.9), Particular plus homogeneous (Technique 3.1.18)
Hint 1. Write $p = a + bx + cx^2$. Each condition $p(t) = v$ is a linear equation in a, b, c.
Hint 2. For (b), the answer to (a) is one particular solution. Find the cubics q with $q(-1) = q(0) = q(1) = 0$.
Sketch. (a) $a - b + c = 2$, $a = 1$, $a + b + c = 4$ give $p = 1 + x + 2x^2$, uniquely. (b) The homogeneous conditions force $a = c = 0$ and $d = -b$, so the solutions are $1 + x + 2x^2 + t\big(x^3 - x\big)$.

(a) Write $p = a + bx + cx^2$. The conditions are the linear system

$$a - b + c = 2, \quad a = 1, \quad a + b + c = 4.$$

With $a = 1$, the first and third equations become $-b + c = 1$ and $b + c = 3$; adding and subtracting gives $c = 2$ and $b = 1$. Conversely $p = 1 + x + 2x^2$ satisfies $p(-1) = 2$, $p(0) = 1$, $p(1) = 4$. So $p = 1 + x + 2x^2$ is the only such polynomial.

(b) Write $q = a + bx + cx^2 + dx^3$. The conditions are linear in (a, b, c, d), and $q = 1 + x + 2x^2$ is one solution by (a). The homogeneous conditions $q(-1) = q(0) = q(1) = 0$ read $a - b + c - d = 0$, $a = 0$, $a + b + c + d = 0$. Then $-b + c - d = 0$ and $b + c + d = 0$; adding gives $c = 0$, and then $d = -b$. So the homogeneous solutions are $bx - bx^3 = -b\big(x^3 - x\big)$, the multiples of $x^3 - x = (x + 1)x(x - 1)$. By Proposition 3.1.19, applied to the linear map $q \mapsto (q(-1), q(0), q(1))$ from $\mathbb{R}[x]_{\leq 3}$ to $\mathbb{R}^3$, the solutions are exactly

$$q = 1 + x + 2x^2 + t\big(x^3 - x\big), \quad t \in \mathbb{R}.$$

The kernel consists of the cubics vanishing at the three points. Section 3.4 shows that through $n + 1$ points with distinct first coordinates there is always exactly one polynomial of degree at most n.

Solution to Exercise 3.1.38

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Definition 3.1.14 (Homogeneous system)
Techniques. Undetermined coefficients (Technique 3.1.9), Particular plus homogeneous (Technique 3.1.18)
Hint 1. At A the arrivals are $300 + x_4$ and the departures are x_1. Write the other three balances similarly.
Hint 2. Solve for x_1, x_2, x_3 in terms of x_4. For (b), add the four equations written as "arrivals minus departures $= 0$".
Hint 3. The homogeneous solutions are the flows (t, t, t, t): cars circling the loop. For (c), which t are allowed?
Sketch. The balances give $x_1 = 300 + x_4$, $x_2 = x_1 - 200$, $x_3 = x_2 + 100$, $x_3 = x_4 + 200$, so the solutions are $(300 + t, 100 + t, 200 + t, t)$. The four balances sum to $0 = 0$ because each road leaves one intersection and enters another and the external flows cancel. Nonnegativity means $t \geq 0$, so the minima $300, 100, 200, 0$ all occur at $t = 0$.

(a) Balancing arrivals and departures at each intersection:

$$\begin{aligned} A: &\quad 300 + x_4 = x_1, \\ B: &\quad x_1 = x_2 + 200, \\ C: &\quad x_2 + 100 = x_3, \\ D: &\quad x_3 = x_4 + 200. \end{aligned}$$

Suppose (x_1, x_2, x_3, x_4) is a solution and put $t = x_4$. Then A gives $x_1 = 300 + t$, B gives $x_2 = 100 + t$, and C gives $x_3 = 200 + t$. Conversely, for every $t \in \mathbb{R}$ the vector $(300 + t, 100 + t, 200 + t, t)$ satisfies A, B, C by construction, and also D, since $200 + t = t + 200$. So the solution set is

$$(300, 100, 200, 0) + \operatorname{span}((1, 1, 1, 1)).$$

In the language of Proposition 3.1.19, the homogeneous solutions (t, t, t, t) are traffic circling the loop, which changes no balance.

(b) Write each balance as "arrivals minus departures $= 0$". Each road variable x_i appears once with a plus sign, at the intersection where the road ends, and once with a minus sign, where it starts. The external terms are $300 - 200 + 100 - 200 = 0$. So the sum of the four equations is $0 = 0$, and any one of them is the negative of the sum of the other three. (So the system is really three equations in four unknowns, and the computation in (a) shows that its homogeneous solutions form the line $\operatorname{span}((1, 1, 1, 1))$.)

(c) The flows are nonnegative exactly when $t \geq 0$, as then $300 + t$, $100 + t$, and $200 + t$ are positive too. So the smallest possible flows are $x_1 = 300$, $x_2 = 100$, $x_3 = 200$, $x_4 = 0$, and all four occur together at $t = 0$, when no traffic uses the road from D to A.

Solution to Exercise 3.1.39

Uses. Technique 3.1.9 (Undetermined coefficients), Corollary 3.1.16 (Homogeneous solution sets are subspaces)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. Write one equation per element: for copper in (a), $a = c$. Count oxygen carefully: $Cu(NO_3)_2$ contains six oxygen atoms.

Hint 2. In (a), express everything in terms of a. In (b) there are two equations in four unknowns; express a and b in terms of c and d.
Sketch. (a) Cu: $a = c$; H: $b = 2e$; N: $b = 2c + d$; O: $3b = 6c + d + e$. The solutions are the multiples of $(3, 8, 3, 2, 4)$. (b) H: $a = c + d$; O: $2b = c + 2d$; so $(a, b, c, d) = (c/2)(2, 1, 2, 0) + d(1, 1, 0, 1)$.

(a) Counting atoms of each element on both sides gives the homogeneous system

$$\text{Cu} : a = c, \quad \text{H} : b = 2e, \quad \text{N} : b = 2c + d, \quad \text{O} : 3b = 6c + d + e.$$

Suppose (a, b, c, d, e) is a solution. Then $c = a$, and the nitrogen equation gives $d = b - 2a = 2e - 2a$. The oxygen equation becomes $6e = 6a + (2e - 2a) + e$, that is, $3e = 4a$. So $e = 4a/3$, $b = 8a/3$, and $d = 2a/3$, which means

$$(a, b, c, d, e) = \frac{a}{3}(3, 8, 3, 2, 4).$$

Conversely, $(3, 8, 3, 2, 4)$ satisfies all four equations: $3 = 3$, $8 = 8$, $8 = 6 + 2$, and $24 = 18 + 2 + 4$. Since the solution set is a subspace (Corollary 3.1.16), all multiples are solutions too. So the balancings are the multiples of $(3, 8, 3, 2, 4)$, and the smallest positive integer one is

$$3\ \text{Cu} + 8\text{HNO}_3 \to 3\text{Cu}(\text{NO}_3)_2 + 2\ \text{NO} + 4\text{H}_2\text{O}.$$

A balancing with positive integer entries is $(a/3)(3, 8, 3, 2, 4)$ with a a positive integer. Its fourth entry $2a/3$ is an integer and $\gcd(2, 3) = 1$, so $3 \mid a$ and $a \geq 3$. Hence $a = 3$ gives the smallest one.

(b) The system is $2a = 2c + 2d$ (hydrogen) and $2b = c + 2d$ (oxygen). So any solution has $a = c + d$ and $b = c/2 + d$, and

$$(a, b, c, d) = \left(c + d, \frac{c}{2} + d, c, d\right) = \frac{c}{2}(2, 1, 2, 0) + d(1, 1, 0, 1).$$

Conversely, $(2, 1, 2, 0)$ and $(1, 1, 0, 1)$ both satisfy the two equations, so all their combinations do. Chemically, every balanced version of this equation is a mixture of the two independent reactions $2\text{H}_2 + \text{O}_2 \to 2\text{H}_2\text{O}$ and $\text{H}_2 + \text{O}_2 \to \text{H}_2\text{O}_2$. The solution set is a plane rather than a line, so balancing alone cannot determine the proportions.

Solution to Exercise 3.1.40

Uses. Proposition 3.1.5 (Three faces of a linear system), Definition 2.4.1 (Matrix product)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. Pressing a light twice does nothing, and the order of presses does not matter. Why? So only the parity of the number of presses of each light matters.
Hint 2. Column j of A records which lights pressing j toggles. For (b), entry (i, j) of A^2 counts the lights toggled both by pressing i and by pressing j (A is symmetric), modulo 2.
Hint 3. If $A^2 = I$, then $A\mathbf{x} = \mathbf{y}$ implies $\mathbf{x} = A\mathbf{y}$.
Sketch. Light i ends in state $y_i + \sum_j a_{ij}x_j$, so all lights end off exactly when $A\mathbf{x} = \mathbf{y}$ (over $\mathbb{F}_2$, $-\mathbf{y} = \mathbf{y}$). Here A has rows $(1, 1, 1, 0)$, $(1, 1, 0, 1)$, $(1, 0, 1, 1)$, $(0, 1, 1, 1)$; each row has three ones and any two rows share two, so $A^2 = I$. Then $A\mathbf{x} = \mathbf{y}$ if and only if $\mathbf{x} = A\mathbf{y}$. For light 1 alone: press 1, 2, 3.

(a) The final state of a light depends only on how many times it is toggled, and only on that number's parity. So the order of presses is irrelevant, pressing a light twice is the same as not pressing it, and a strategy is the vector $\mathbf{x} \in \mathbb{F}_2^4$ with $x_j = 1$ exactly when light j is pressed (an odd number of times). Let $a_{ij} = 1$ if pressing light j toggles light i, and $a_{ij} = 0$ otherwise. Starting from the pattern $\mathbf{y}$, light i ends in state $y_i + a_{i1}x_1 + \cdots + a_{i4}x_4$ in $\mathbb{F}_2$. All lights end off exactly when $A\mathbf{x} + \mathbf{y} = \mathbf{0}$, that is, $A\mathbf{x} = -\mathbf{y} = \mathbf{y}$, as $-1 = 1$ in $\mathbb{F}_2$. Light 1 toggles lights 1, 2, 3, light 2 toggles 1, 2, 4, light 3 toggles 1, 3, 4, and light 4 toggles 2, 3, 4, so

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

(b) A is symmetric, so entry (i,j) of A^2 is $\sum_k a_{ik}a_{kj} = \sum_k a_{ik}a_{jk}$, the number of positions where rows i and j both have a 1, reduced mod 2. Each row has three 1s, so the diagonal entries are $3 = 1$. Any two different rows have exactly two 1s in common (for instance, rows 1 and 2 share positions 1 and 2), so the off-diagonal entries are $2 = 0$. Hence $A^2 = I_4$.

Now let $\mathbf{y}$ be any pattern. If $A\mathbf{x} = \mathbf{y}$, then $\mathbf{x} = A^2\mathbf{x} = A\mathbf{y}$. Conversely, $\mathbf{x} = A\mathbf{y}$ satisfies $A\mathbf{x} = A^2\mathbf{y} = \mathbf{y}$. So $A\mathbf{x} = \mathbf{y}$ has exactly one solution, $\mathbf{x} = A\mathbf{y}$: every pattern can be switched off, in exactly one way.

(c) For light 1 alone, $\mathbf{y} = \mathbf{e}_1$ and $\mathbf{x} = A\mathbf{e}_1 = (1,1,1,0)$: press lights 1, 2, and 3. Light 1 is toggled three times and each other light twice. By symmetry, to switch off a single lit corner, press it and its two neighbors.

Solution to Exercise 3.1.41

Uses. Technique 3.1.9 (Undetermined coefficients), Corollary 3.1.16 (Homogeneous solution sets are subspaces), Proposition 1.4.14 (The span is the smallest subspace containing a list), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Undetermined coefficients (Technique 3.1.9), Double containment (Technique 1.3.14)
Hint 1. For (a), look for an equation $a_1x_1 + a_2x_2 + a_3x_3 = 0$ satisfied by both spanning vectors. That is a linear system in a_1, a_2, a_3.
Hint 2. Having found a candidate, prove that its solution set equals the span by double containment.
Hint 3. For (b), the line is $(1,0,2) + \operatorname{span}((1,1,-1))$. Find two equations whose left sides vanish at $(1,1,-1)$, and choose the right sides so that $(1,0,2)$ is a solution.
Sketch. (a) $a_1 + 2a_2 = 0$ and $a_2 + a_3 = 0$ give $(2,-1,1)$; the solutions of $2x_1 - x_2 + x_3 = 0$ are $x_1(1,2,0) + x_3(0,1,1)$. (b) $x_1 - x_2 = 1$, $x_2 + x_3 = 2$; its solutions are $(1,0,2) + s(1,1,-1)$.

(a) An equation $a_1x_1 + a_2x_2 + a_3x_3 = 0$ is satisfied by $(1,2,0)$ and $(0,1,1)$ exactly when $a_1 + 2a_2 = 0$ and $a_2 + a_3 = 0$. One solution of this system is $(a_1,a_2,a_3) = (2,-1,1)$. We claim that the solution set S of

$$2x_1 - x_2 + x_3 = 0$$

is $U = \operatorname{span}((1,2,0),(0,1,1))$.

$U \subseteq S$: both spanning vectors are solutions, since $2 - 2 + 0 = 0$ and $0 - 1 + 1 = 0$, and S is a subspace (Corollary 3.1.16), so S contains their span (Proposition 1.4.14).

$S \subseteq U$: if $2x_1 - x_2 + x_3 = 0$, then $x_2 = 2x_1 + x_3$, so

$$(x_1,x_2,x_3) = (x_1, 2x_1 + x_3, x_3) = x_1(1,2,0) + x_3(0,1,1) \in U.$$

(b) The line is $(1,0,2) + \operatorname{span}((1,1,-1))$, since $(2,1,1) - (1,0,2) = (1,1,-1)$. By Proposition 3.1.19, it suffices to find a system that has $(1,0,2)$ as a solution and whose homogeneous solution set is $\operatorname{span}((1,1,-1))$. The left sides $x_1 - x_2$ and $x_2 + x_3$ both vanish at $(1,1,-1)$, and at $(1,0,2)$ they equal 1 and 2. So consider

$$x_1 - x_2 = 1, \quad x_2 + x_3 = 2.$$

Any solution has $x_2 = x_1 - 1$ and $x_3 = 2 - x_2 = 3 - x_1$, so

$$(x_1,x_2,x_3) = (x_1, x_1 - 1, 3 - x_1) = (1,0,2) + (x_1 - 1)(1,1,-1),$$

a point of the line. Conversely, $(1,0,2) + s(1,1,-1) = (1+s, s, 2-s)$ satisfies $(1+s) - s = 1$ and $s + (2-s) = 2$. So the solution set is exactly the line.

Solution to Exercise 3.1.42

Uses. Corollary 3.1.23 (Lines through solutions)
Hint 1. By Corollary 3.1.23, a solution set that contains $\mathbf{u}$ and $\mathbf{v}$ contains $\mathbf{u} + t(\mathbf{v} - \mathbf{u})$ for every $t \in F$. Find two points of S_F.
Hint 2. In $\mathbb{F}_2$, what is x^2 for each x? In $\mathbb{F}_3$, list S_F by trying all nine points.
Sketch. (a) False: $(0,0),(1,0) \in S_{\mathbb{R}}$, but $(2,0) \notin S_{\mathbb{R}}$. (b) True: $x^2 = x$ for all $x \in \mathbb{F}_2$, so $S_{\mathbb{F}_2} = \mathbb{F}_2^2$, the solution set of $0x_1 + 0x_2 = 0$. (c) False: $S_{\mathbb{F}_3} = \{0,1\}^2$ contains $(0,0)$ and $(1,0)$ but not $(2,0)$.

In all three cases $(0,0)$ and $(1,0)$ lie in S_F, since $0 = 0$ and $1 = 1$. If S_F were the solution set of a linear system, then by Corollary 3.1.23 it would contain $(0,0) + t((1,0) - (0,0)) = (t,0)$ for every $t \in F$.

(a) False. $(2,0) \notin S_{\mathbb{R}}$, since $4 \neq 2$.

(b) True. In $\mathbb{F}_2$, $0^2 = 0$ and $1^2 = 1$, so $x^2 = x$ for every x, and every point of $\mathbb{F}_2^2$ satisfies the equation. So $S_{\mathbb{F}_2} = \mathbb{F}_2^2$, which is the solution set of the linear equation $0x_1 + 0x_2 = 0$.

(c) False. In $\mathbb{F}_3$, $x^2 - x$ equals $0, 0, 2$ for $x = 0, 1, 2$. So $(2,0) \notin S_{\mathbb{F}_3}$, since $\left(2^2 - 2\right) + \left(0^2 - 0\right) = 2 \neq 0$. (In fact $S_{\mathbb{F}_3} = \{(0,0),(0,1),(1,0),(1,1)\}$, as $2 + 2 = 4 = 1 \neq 0$ rules out $(2,2)$.)

Solution to Exercise 3.1.43

Uses. Definition 2.3.15 (The linear map of a matrix), Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product), Definition 2.5.2 (Right inverse), Definition 1.6.8 (Standard bases)
Hint 1. Write $\mathbf{y} = y_1\mathbf{e}_1 + \cdots + y_m\mathbf{e}_m$ and use linearity of T_A.
Hint 2. The jth column of AB is $A\mathbf{b}_j$.
Sketch. $A(y_1\mathbf{b}_1 + \cdots + y_m\mathbf{b}_m) = y_1\mathbf{e}_1 + \cdots + y_m\mathbf{e}_m = \mathbf{y}$. The columns of AB are $A\mathbf{b}_j = \mathbf{e}_j$, so $AB = I_m$, and $T_A \circ T_B = \mathrm{id}$ since $A(B\mathbf{y}) = \mathbf{y}$.

(a) Let $\mathbf{y} = (y_1, \ldots, y_m) = y_1\mathbf{e}_1 + \cdots + y_m\mathbf{e}_m$. Since T_A is linear,

$$A(y_1\mathbf{b}_1 + \cdots + y_m\mathbf{b}_m) = y_1A\mathbf{b}_1 + \cdots + y_mA\mathbf{b}_m = y_1\mathbf{e}_1 + \cdots + y_m\mathbf{e}_m = \mathbf{y}.$$

So $\mathbf{x} = y_1\mathbf{b}_1 + \cdots + y_m\mathbf{b}_m$ is a solution.

(b) The jth column of AB is $A\mathbf{b}_j = \mathbf{e}_j$ (Definition 2.4.1), which is the jth column of I_m. So $AB = I_m$. The solution in (a) is $\mathbf{x} = B\mathbf{y}$, by Definition 2.3.8, so $T_A(T_B(\mathbf{y})) = A(B\mathbf{y}) = \mathbf{y}$ for every $\mathbf{y} \in F^m$. That is, $T_A \circ T_B = \mathrm{id}_{F^m}$, and the linear map T_B is a right inverse of T_A. This is one direction of the equivalence between solvability for every $\mathbf{y}$ and right invertibility in Theorem 2.7.4, with an explicit right inverse: it suffices to solve the m systems $A\mathbf{x} = \mathbf{e}_i$.

Solution to Exercise 3.1.44

Uses. Corollary 3.1.23 (Lines through solutions), Proposition 1.2.35 (Zero products)
Hint 1. Let $\mathbf{u} \neq \mathbf{v}$ be solutions. Corollary 3.1.23 gives a solution for each $t \in F$.
Hint 2. Show that $t \mapsto \mathbf{u} + t(\mathbf{v} - \mathbf{u})$ is injective. Then compare sizes.
Sketch. If $\mathbf{u} \neq \mathbf{v}$ are solutions, $t \mapsto \mathbf{u} + t(\mathbf{v} - \mathbf{u})$ is an injective map from F into the solution set, so the solution set has at least $|F|$ elements. With exactly two solutions this gives $|F| \leq 2$, and $|F| \geq 2$ since $0 \neq 1$.

(b) Let $\mathbf{u} \neq \mathbf{v}$ be solutions, and define $\varphi : F \to S$, where S is the solution set, by $\varphi(t) = \mathbf{u} + t(\mathbf{v} - \mathbf{u})$. This lands in S by Corollary 3.1.23. It is injective: if $\varphi(t) = \varphi(t')$, then $(t - t')(\mathbf{v} - \mathbf{u}) = \mathbf{0}$, and $\mathbf{v} - \mathbf{u} \neq \mathbf{0}$ gives $t = t'$ by Proposition 1.2.35. So $|S| \geq |F|$.

(a) If S has exactly two elements, then F is finite, since $\varphi : F \to S$ is injective, and $|F| \leq |S| = 2$ by the argument in (b). Every field has at least two elements, 0 and 1, so $|F| = 2$.

Solution to Exercise 3.1.45

Uses. Notation 3.1.17 (Translates), Definition 1.4.1 (Subspace), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Double containment (Technique 1.3.14)
Hint 1. Since $\mathbf{0} \in U'$, the vector $\mathbf{v}'$ lies in $\mathbf{v} + U$. What does that say about $\mathbf{v}' - \mathbf{v}$?
Hint 2. Take $\mathbf{u}' \in U'$. Then $\mathbf{v}' + \mathbf{u}' \in \mathbf{v} + U$; subtract $\mathbf{v}$ and use that U is closed under subtraction. Then swap the roles of the two translates.
Sketch. $\mathbf{v}' = \mathbf{v}' + \mathbf{0} \in \mathbf{v} + U$, so $\mathbf{w} = \mathbf{v}' - \mathbf{v} \in U$. For $\mathbf{u}' \in U'$, $\mathbf{v}' + \mathbf{u}' = \mathbf{v} + \mathbf{u}$ with $\mathbf{u} \in U$, so $\mathbf{u}' = \mathbf{u} - \mathbf{w} \in U$. By symmetry $U \subseteq U'$. For (b), write the common solution set as $\mathbf{x}_0 + \ker T_A = \mathbf{x}_0 + \ker T_B$.

(a) Since $\mathbf{0} \in U'$, we have $\mathbf{v}' = \mathbf{v}' + \mathbf{0} \in \mathbf{v}' + U' = \mathbf{v} + U$, so $\mathbf{v}' = \mathbf{v} + \mathbf{w}$ for some $\mathbf{w} \in U$; that is, $\mathbf{v}' - \mathbf{v} = \mathbf{w} \in U$. By the symmetric argument, $\mathbf{v} - \mathbf{v}' \in U'$.

$U' \subseteq U$: let $\mathbf{u}' \in U'$. Then $\mathbf{v}' + \mathbf{u}' \in \mathbf{v}' + U' = \mathbf{v} + U$, so $\mathbf{v}' + \mathbf{u}' = \mathbf{v} + \mathbf{u}$ for some $\mathbf{u} \in U$. Hence $\mathbf{u}' = \mathbf{u} - (\mathbf{v}' - \mathbf{v}) = \mathbf{u} - \mathbf{w}$, which lies in U because U is a subspace.

$U \subseteq U'$: the same argument with the roles of $(\mathbf{v}, U)$ and $(\mathbf{v}', U')$ exchanged, using $\mathbf{v} - \mathbf{v}' \in U'$.

(b) Let S be the common solution set and $\mathbf{x}_0 \in S$. By Proposition 3.1.19, $\mathbf{x}_0 + \ker T_A = S = \mathbf{x}_0 + \ker T_B$, and $\ker T_A$, $\ker T_B$ are subspaces of F^n. By (a), $\ker T_A = \ker T_B$.

(c) Take $A = I_2$, $\mathbf{y} = (1,1)$, and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\mathbf{z} = (1,2)$. The first system says $x_1 = 1$, $x_2 = 1$. The second says $x_1 = 1$ and $x_1 + x_2 = 2$, which also forces $(x_1, x_2) = (1,1)$, and $(1,1)$ satisfies it. Both solution sets are $\{(1,1)\}$, and both kernels are $\{\mathbf{0}\}$.

Solution to Exercise 3.1.46

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Proposition 3.1.8 (Systems and the map T_A)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. $(1, 1, \ldots, 1)$ is a particular solution. Find the homogeneous solutions.
Hint 2. In the homogeneous system, $x_{i+1} = -x_i$ for each i, so $x_{k+1} = (-1)^k x_1$. Go once around the cycle: what does $x_n + x_1 = 0$ say?
Sketch. The homogeneous solutions have $x_{k+1} = (-1)^k x_1$, and the last equation gives $\big((-1)^{n-1} + 1\big)x_1 = 0$. For n odd this forces $x_1 = 0$, so $(1, \ldots, 1)$ is the unique solution; for n even the kernel is spanned by $(1, -1, 1, \ldots, -1)$. Over $\mathbb{F}_2$, $2 = 0$ and the equations say $x_1 = x_2 = \cdots = x_n$: two solutions.

Over $\mathbb{R}$, $(1, 1, \ldots, 1)$ is a solution. By Proposition 3.1.19, we need the solutions of the homogeneous system $x_i + x_{i+1} = 0$ $(1 \le i \le n-1)$, $x_n + x_1 = 0$. In such a solution, $x_{i+1} = -x_i$ for $i \le n-1$, so by induction $x_{k+1} = (-1)^k x_1$ for $0 \le k \le n-1$. The last equation then reads $(-1)^{n-1} x_1 + x_1 = 0$. Conversely, a vector with $x_{k+1} = (-1)^k x_1$ satisfies the first $n-1$ equations, and it satisfies the last one exactly when $\big((-1)^{n-1} + 1\big)x_1 = 0$.

If n is odd, this says $2x_1 = 0$, so $x_1 = 0$ and the kernel is $\{\mathbf{0}\}$. The unique solution is $(1, 1, \ldots, 1)$.

If n is even, it says $0 = 0$, so the kernel is spanned by $(1, -1, 1, -1, \ldots, 1, -1)$, and the solutions are

$$(1 + t, 1 - t, 1 + t, 1 - t, \ldots, 1 + t, 1 - t), \quad t \in \mathbb{R}.$$

Over $\mathbb{F}_2$, the right-hand sides are $2 = 0$ and $-x_i = x_i$, so the system says $x_{i+1} = x_i$ for all i (and $x_1 = x_n$). The solutions are $(0, \ldots, 0)$ and $(1, \ldots, 1)$, for every n. Over $\mathbb{R}$ the parity of n mattered only through $2x_1 = 0 \Rightarrow x_1 = 0$, which is exactly the step that fails in characteristic 2.

Solution to Exercise 3.1.47

Uses. Corollary 3.1.22 (Uniqueness depends only on A), Proposition 3.1.8 (Systems and the map T_A)
Techniques. Particular plus homogeneous (Technique 3.1.18)
Hint 1. By Corollary 3.1.22, it suffices to show that the homogeneous system, with all $b_q = 0$, has only the trivial solution.
Hint 2. In a homogeneous solution, let M be the largest u_p, and suppose $M > 0$. At a point where $u_p = M$, the average of four numbers that are all at most M equals M. What does that force?
Hint 3. Walk from a maximizing point to the right until you leave G. The first point outside G is a boundary point.
Sketch. Take a homogeneous solution, extended by $u_q = 0$ on ∂G, so each u_p is the average of its four neighbors. If $M = \max u_p > 0$ is attained at p, all four neighbors of p equal M. Walking right from p, every point of G along the way has value M, and the first point outside G is a boundary point adjacent to a point with value M, so its value 0 equals M, a contradiction. So $\max \le 0$; applying this to $-u$ gives $\min \ge 0$. Hence $u = \mathbf{0}$, and Corollary 3.1.22 finishes.

By Corollary 3.1.22, it suffices to show that the homogeneous system, the one with every $b_q = 0$, has only the trivial solution. Let $\big(u_p\big)_{p \in G}$ be a solution of the homogeneous system, and set $u_q = 0$ for $q \in \partial G$. Every neighbor of a point of G lies in G or in ∂G, so the equations say

$$u_p = \frac{1}{4} \sum_{q \sim p} u_q \quad \text{for every } p \in G,$$

the sum running over all four neighbors of p.

Suppose, for contradiction, that $M = \max_{p \in G} u_p > 0$ (the maximum exists because G is finite and, being the index set of the unknowns, nonempty). Then every u_q with $q \in G \cup \partial G$ is at most M. We first show: *if $p \in G$ and $u_p = M$, then all four neighbors of p have value M.* Indeed, $4M = 4u_p = \sum_{q \sim p} u_q$ is a sum of four numbers each at most M, so if one of them were less than M, the sum would be less than $4M$.

Pick $p_0 \in G$ with $u_{p_0} = M$, and consider the points $p_k = p_0 + (k, 0)$ for $k = 0, 1, 2, \ldots$. Since G is finite, some p_k lies outside G; let k be the smallest such index, so $k \geq 1$ and $p_0, \ldots, p_{k-1} \in G$. By the claim and induction, $u_{p_0} = u_{p_1} = \cdots = u_{p_{k-1}} = M$, and then $u_{p_k} = M$ too, since p_k is a neighbor of p_{k-1}. But $p_k \notin G$ is adjacent to $p_{k-1} \in G$, so $p_k \in \partial G$ and $u_{p_k} = 0 < M$. This contradiction shows $u_p \leq 0$ for all $p \in G$.

The vector $\left(-u_p\right)_{p \in G}$ also solves the homogeneous system, so the same argument gives $-u_p \leq 0$ for all p. Hence $u_p = 0$ for all $p \in G$: the homogeneous system has only the trivial solution, and the original system has at most one solution.

The system has as many equations as unknowns. For such square systems, Section 3.5 shows that "at most one solution for every right-hand side" implies "exactly one", so the equilibrium temperatures always exist and are unique.

Solution to Exercise 3.1.48

Uses. Definition 1.4.1 (Subspace), Theorem 1.4.3 (Subspace criterion), Notation 3.1.17 (Translates), Definition 1.1.28 (Characteristic)
Hint 1. Fix $\mathbf{v} \in S$ and let $U = \{\mathbf{s} - \mathbf{v} \mid \mathbf{s} \in S\}$, so $S = \mathbf{v} + U$. Check the subspace conditions for U.
Hint 2. Scalar multiples: $\mathbf{v} + t(\mathbf{s} - \mathbf{v}) \in S$. Sums are harder: first show that the midpoint $\frac{1}{2}\mathbf{s} + \frac{1}{2}\mathbf{s}'$ of two points of S lies in S; this step needs $2 \neq 0$.
Hint 3. For (b), over $\mathbb{F}_2$ the only values of t are 0 and 1. For the second part, a translate of a subspace by $\mathbf{v}$ contains $\mathbf{v}$; subtract it and count.
Sketch. With $U = S - \mathbf{v}$: $\mathbf{0} \in U$; $t(\mathbf{s} - \mathbf{v}) = (\mathbf{v} + t(\mathbf{s} - \mathbf{v})) - \mathbf{v} \in U$; the midpoint $\mathbf{m} = \mathbf{s} + \frac{1}{2}(\mathbf{s}' - \mathbf{s}) \in S$, and $(\mathbf{s} - \mathbf{v}) + (\mathbf{s}' - \mathbf{v}) = 2(\mathbf{m} - \mathbf{v}) \in U$. In (b), the line through two points of $\mathbb{F}_2^2$ is just those two points; a translate $\mathbf{v} + U$ with three elements would make U a subspace of $\mathbb{F}_2^2$ with three elements, but $\mathbf{u}, \mathbf{u}' \in U$ distinct and nonzero force a fourth element $\mathbf{u} + \mathbf{u}'$.

(a) Fix $\mathbf{v} \in S$ and put $U = \{\mathbf{s} - \mathbf{v} \mid \mathbf{s} \in S\}$, so that $S = \mathbf{v} + U$. We check that U is a subspace.

$\mathbf{0} \in U$, as $\mathbf{0} = \mathbf{v} - \mathbf{v}$.

Scalar multiples. Let $\mathbf{s} \in S$ and $t \in F$. By line-closedness with $\mathbf{u} = \mathbf{v}$, the point $\mathbf{v} + t(\mathbf{s} - \mathbf{v})$ lies in S, so $t(\mathbf{s} - \mathbf{v}) \in U$.

Sums. Let $\mathbf{s}, \mathbf{s}' \in S$. Since $2 = 1 + 1 \neq 0$, the scalar $\frac{1}{2}$ exists, and $\mathbf{m} = \mathbf{s} + \frac{1}{2}(\mathbf{s}' - \mathbf{s})$ lies in S. Then $\mathbf{m} - \mathbf{v} \in U$, and by the previous step $2(\mathbf{m} - \mathbf{v}) \in U$. But

$$2(\mathbf{m} - \mathbf{v}) = 2\mathbf{s} + (\mathbf{s}' - \mathbf{s}) - 2\mathbf{v} = (\mathbf{s} - \mathbf{v}) + (\mathbf{s}' - \mathbf{v}).$$

So U is closed under addition, and U is a subspace by Theorem 1.4.3. Hence $S = \mathbf{v} + U$ is a translate of a subspace.

(b) Over $\mathbb{F}_2$, t is 0 or 1, and $\mathbf{u} + t(\mathbf{v} - \mathbf{u})$ is then $\mathbf{u}$ or $\mathbf{v}$. So every subset of $\mathbb{F}_2^n$ is line-closed, in particular $S = \{(0, 0), (1, 0), (0, 1)\}$.

Suppose $S = \mathbf{v} + U$ for a subspace U. Then $U = \{\mathbf{s} - \mathbf{v} \mid \mathbf{s} \in S\}$ has exactly three elements, since $\mathbf{s} \mapsto \mathbf{s} - \mathbf{v}$ is injective. One of them is $\mathbf{0}$; call the other two $\mathbf{u} \neq \mathbf{u}'$, both nonzero. Since U is a subspace, $\mathbf{u} + \mathbf{u}' \in U$. But $\mathbf{u} + \mathbf{u}' \neq \mathbf{0}$, since otherwise $\mathbf{u}' = -\mathbf{u} = \mathbf{u}$ in characteristic 2; and $\mathbf{u} + \mathbf{u}' \neq \mathbf{u}$ and $\neq \mathbf{u}'$, since $\mathbf{u}'$ and $\mathbf{u}$ are nonzero. So U has a fourth element, a contradiction.

Part (b) shows that line-closedness is too weak a condition in characteristic 2. The right notion uses combinations of more than two points; see Section 4.7.

Solution to Exercise 3.1.49

Uses. Corollary 3.1.16 (Homogeneous solution sets are subspaces), Proposition 3.1.5 (Three faces of a linear system)
Hint 1. Fix the masses $\mathbf{w} = (w_1, \ldots, w_{13})$ and, for each i, one valid split. Encode the splits as a 13×13 matrix A with $A\mathbf{w} = \mathbf{0}$. What else lies in $\ker T_A$?
Hint 2. $(1, 1, \ldots, 1) \in \ker T_A$, and $\ker T_A$ is a subspace. Show that all entries of an integer vector in $\ker T_A$ have the same parity.

Hint 3. Subtract the smallest mass from all masses, then keep dividing by 2. Why must this stop, and what does that force?
Sketch. Row i of A has 0 in position i, $+1$ on one group and -1 on the other, so $A\mathbf{w} = \mathbf{0}$ and $A\mathbf{1} = \mathbf{0}$. For an integer $\mathbf{v} \in \ker T_A$, row i gives $\sum_{j\neq i} v_j \equiv 0 \bmod 2$, so every v_i has the parity of $\sum_j v_j$. Among nonzero integer vectors in $\ker T_A$ with nonnegative entries and a zero entry, take one with the smallest maximum entry; its entries are all even, and halving it gives a smaller one. So $\mathbf{w} - (\min_j w_j)\mathbf{1} = \mathbf{0}$.

Let $\mathbf{w} = (w_1, \ldots, w_{13}) \in \mathbb{Z}^{13}$ be the masses. For each i, fix a split of the other twelve weights into two groups of six with equal total mass, and let $A \in \mathbb{Q}^{13\times 13}$ have (i, j) entry 0 if $j = i$, $+1$ if weight j is in the first group of the split for i, and -1 if it is in the second. Then row i of $A\mathbf{w} = \mathbf{0}$ says that the two groups balance, so $\mathbf{w} \in \ker T_A$. Each row of A has six entries $+1$ and six entries -1, so $\mathbf{1} = (1, \ldots, 1) \in \ker T_A$ as well. By Corollary 3.1.16, $\ker T_A$ is a subspace of $\mathbb{Q}^{13}$.

Parity. Let $\mathbf{v} \in \ker T_A$ have integer entries, and let $N = v_1 + \cdots + v_{13}$. Row i says $\sum_{j\neq i} \pm v_j = 0$. Since $-v_j \equiv v_j$ modulo 2, this gives $\sum_{j\neq i} v_j \equiv 0$, that is, $N - v_i \equiv 0 \bmod 2$. So every v_i has the same parity as N, and all entries of $\mathbf{v}$ have the same parity.

Descent. Let K be the set of nonzero vectors $\mathbf{v} \in \ker T_A$ with integer entries, all entries ≥ 0, and at least one entry equal to 0. Suppose K is nonempty, and choose $\mathbf{v} \in K$ whose largest entry is as small as possible. Since $\mathbf{v}$ has a zero entry, all its entries are even by the parity step, so $\frac{1}{2}\mathbf{v}$ has integer entries. It lies in $\ker T_A$ (a subspace), is nonzero, has nonnegative entries and a zero entry, so $\frac{1}{2}\mathbf{v} \in K$. But its largest entry is half that of $\mathbf{v}$, which is positive, so it is smaller. This contradicts the choice of $\mathbf{v}$, so K is empty.

Conclusion. Let $c = \min_j w_j$ and $\mathbf{v} = \mathbf{w} - c\mathbf{1}$. Then $\mathbf{v} \in \ker T_A$, its entries are nonnegative integers, and at least one is 0. As K is empty, $\mathbf{v} = \mathbf{0}$, so $w_1 = \cdots = w_{13} = c$.

The same holds for rational masses (multiply by a common denominator) and in fact for real masses, but the real case needs more: one shows that the kernel of T_A over $\mathbb{R}$ is spanned by vectors with rational entries, which follows from Gauss–Jordan elimination never leaving the field of the entries (Section 3.4).

Section 3.2

Solution to Exercise 3.2.9

Uses. Proposition 3.2.7 (Row operations are left multiplications), Definition 3.2.2 (Elementary matrices)
Hint 1. Which rows of A changed, and how is each new row built from the old rows?
Hint 2. Row 3 of EA is $(5, 6) - 2(1, 2)$. Once you know the operation, perform it on I_3.
Sketch. EA is A with -2 times row 1 added to row 3, so $E = E_{1,-2,3}$. $E'A$ is A with rows 1 and 3 swapped, so $E' = E_{13}$.

Only row 3 changed in EA, and $(3, 2) = (5, 6) - 2(1, 2)$. So EA is A with -2 times row 1 added to row 3, and by Proposition 3.2.7 we can take $E = E_{1,-2,3}$. In $E'A$ rows 1 and 3 are swapped, so $E' = E_{13}$. Performing these operations on I_3,

$$E = E_{1,-2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, \quad E' = E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Solution to Exercise 3.2.13

Uses. Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.10 (Inverses of elementary matrices)
Hint 1. Invert each factor and reverse the order.
Hint 2. In $\mathbb{F}_5$, $3^{-1} = 2$ and $-4 = 1$.
Sketch. The inverse is $E_{2,-4,1}E_{1,3^{-1}}E_{12} = E_{2,1,1}E_{1,2}E_{12} = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 0 \end{smallmatrix}\right)$.

By Corollary 3.2.11 and Proposition 3.2.10,

$$\left(E_{12}E_{1,3}E_{2,4,1}\right)^{-1} = E_{2,4,1}^{-1}E_{1,3}^{-1}E_{12}^{-1} = E_{2,-4,1}E_{1,3^{-1}}E_{12} = E_{2,1,1}E_{1,2}E_{12},$$

since $-4 = 1$ and $3 \cdot 2 = 6 = 1$ in $\mathbb{F}_5$. Performing the operations of the factors on I_2 from right to left: swapping the rows gives $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, scaling row 1 by 2 gives $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, and adding row 2 to row 1 gives $\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$. As a check, $E_{12}E_{1,3}E_{2,4,1} = \begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}$, and

$$\begin{pmatrix} 0 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{in } \mathbb{F}_5.$$

Solution to Exercise 3.2.20

Uses. Corollary 3.2.17 (Row operations preserve relations among columns), Proposition 3.2.7 (Row operations are left multiplications), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. Clear the first column below the top entry, then work on the second column.
Hint 2. In B, the third column is $\mathbf{e}_1 + 2\mathbf{e}_2$ and the first two columns are $\mathbf{e}_1, \mathbf{e}_2$.
Sketch. Add -2 times row 1 to row 2 and -1 times row 1 to row 3, add row 2 to row 3, scale row 2 by $-1/2$, and add -1 times row 2 to row 1. Since $\mathbf{b}_3 = \mathbf{b}_1 + 2\mathbf{b}_2$ and $\mathbf{b}_1, \mathbf{b}_2$ is independent, $\mathbf{a}_3 = \mathbf{a}_1 + 2\mathbf{a}_2$ and $\mathbf{a}_1, \mathbf{a}_2$ is independent.

Adding -2 times row 1 to row 2 and -1 times row 1 to row 3 gives rows $(1, 1, 3)$, $(0, -2, -4)$, $(0, 2, 4)$. Adding row 2 to row 3 makes row 3 zero. Scaling row 2 by $-1/2$ gives $(0, 1, 2)$, and adding -1 times this row to row 1 gives $(1, 0, 1)$. The result is B. So $B = GA$, where G is the product of the five elementary matrices, which is invertible by Corollary 3.2.11.

The columns of B are $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_1 + 2\mathbf{e}_2$. By Corollary 3.2.17(b), the third column of A is 1 times the first plus 2 times the second; and by Corollary 3.2.17(a), the first two columns of A are linearly independent, since $\mathbf{e}_1, \mathbf{e}_2$ is. Directly: $(1, 2, 1) + 2(1, 0, 3) = (3, 2, 7)$, and if $s(1, 2, 1) + t(1, 0, 3) = \mathbf{0}$, then the second entry gives $s = 0$ and then the first gives $t = 0$.

Solution to Exercise 3.2.26

Uses. Definition 3.2.21 (Row equivalence), Corollary 3.2.23 (Row equivalent matrices have the same kernel), Proposition 3.2.22 (Row equivalence is an equivalence relation)
Hint 1. Try to reach each matrix from A by row operations. When that seems impossible, compare kernels.
Hint 2. $\ker T_A = \operatorname{span}((2, -1))$. What is the kernel of (ii)?
Sketch. (i) Add -2 times row 1 to row 2. (iii) From (i), swap the rows and scale row 2 by 3. (ii) is not: its kernel is $\operatorname{span}((1, -2))$, while $\ker T_A = \operatorname{span}((2, -1))$.

(i) Yes: adding -2 times row 1 to row 2 of A gives $\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$.

(iii) Yes: from $\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, swapping the rows gives $\begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix}$, and scaling row 2 by 3 gives $\begin{pmatrix} 0 & 0 \\ 3 & 6 \end{pmatrix}$. By transitivity (Proposition 3.2.22), (iii) is row equivalent to A.

(ii) No. $A(x_1, x_2) = (x_1 + 2x_2, 2(x_1 + 2x_2))$, so $\ker T_A = \{\mathbf{x} \mid x_1 + 2x_2 = 0\} = \operatorname{span}((2, -1))$. For the matrix in (ii), $\mathbf{x} \mapsto (2x_1 + x_2, 2(2x_1 + x_2))$, so its kernel is $\operatorname{span}((1, -2))$. These are different, since $(2, -1)$ is not a multiple of $(1, -2)$, so the matrices are not row equivalent by Corollary 3.2.23.

Solution to Exercise 3.2.32

Uses. Proposition 3.2.28 (Column operations are right multiplications), Proposition 3.2.7 (Row operations are left multiplications)
Hint 1. Which column changed, and by a multiple of which other column?
Hint 2. Adding c times column j to column i is right multiplication by E_{icj}.
Sketch. Column 2 became $(2, 5) - 2(1, 4)$, so $E = E_{2,-2,1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. On the left, $E_{2,-2,1}B$ is B with -2 times row 2 added to row 1.

Only column 2 changed, and $(0, -3) = (2, 5) - 2(1, 4)$. So AE is A with -2 times column 1 added to column 2, which by Proposition 3.2.28(c) (with $i = 2$, $j = 1$) is $AE_{2,-2,1}$, where

$$E_{2,-2,1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By Proposition 3.2.7, $E_{2,-2,1}B$ is B with -2 times row 2 added to row 1: the indices play the opposite roles.

Solution to Exercise 3.2.36

Uses. Definition 3.2.2 (Elementary matrices), Proposition 3.2.10 (Inverses of elementary matrices)
Hint 1. Compare each matrix with the identity. In how many entries can an elementary matrix differ from I_m, and what do those entries look like?
Hint 2. A scaling differs from I_m in one diagonal entry, which is nonzero; a replacement in one off-diagonal entry; a swap in exactly four entries, which form the pattern $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ in rows and columns i, j.
Sketch. Elementary: (a) I_2; (b) $E_{1,-3,2}$, inverse $E_{1,3,2}$; (c) E_{12}, its own inverse; (f) $E_{3,4,1}$, inverse $E_{3,-4,1}$; (h) $E_{1,2}$, its own inverse since $2 \cdot 2 = 1$ in $\mathbb{F}_3$; (i) E_{23}, its own inverse. Not elementary: (d), (e), (g).

From Definition 3.2.2: E_{ic} differs from I_m only in entry (i,i), where it has $c \neq 0$; E_{icj} differs only in entry (j,i), where it has c; and E_{ij} differs in exactly the entries $(i,i),(j,j),(i,j),(j,i)$, which are $0,0,1,1$. So a matrix is elementary exactly when it equals I_m, or differs from I_m in a single entry that is either off the diagonal or a nonzero diagonal entry, or differs from I_m in exactly four entries forming a swap pattern. Inverses come from Proposition 3.2.10.

(a) Elementary: $I_2 = E_{1,1}$ (scaling row 1 by 1, or adding 0 times row 1 to row 2). Its inverse is I_2.

(b) Elementary: the single extra entry -3 is in position $(2,1)$, so this is $E_{1,-3,2}$, which adds -3 times row 1 to row 2. Its inverse is $E_{1,3,2} = \left(\begin{smallmatrix}1&0\\3&1\end{smallmatrix}\right)$.

(c) Elementary: E_{12} swaps the two rows, and it is its own inverse.

(d) Not elementary. It differs from I_2 in all four entries, but it has the entry 2, which a swap does not. (It is $E_{1,2}E_{12}$; see Non-example 3.2.5.)

(e) Not elementary. It differs from I_3 only in the diagonal entry $(3,3)$, which is 0, and scaling by 0 is not allowed. (This matrix is not invertible: it sends $\mathbf{e}_3$ to $\mathbf{0}$.)

(f) Elementary: the extra entry 4 is in position $(1,3)$, that is, row $j = 1$ and column $i = 3$. So this is $E_{3,4,1}$, which adds 4 times row 3 to row 1. Its inverse is $E_{3,-4,1}$.

(g) Not elementary: it differs from I_2 in exactly two entries.

(h) Elementary: $E_{1,2}$ scales row 1 by 2. In $\mathbb{F}_3$, $2 \cdot 2 = 4 = 1$, so $2^{-1} = 2$ and $E_{1,2}$ is its own inverse.

(i) Elementary: E_{23} swaps rows 2 and 3, and it is its own inverse.

Solution to Exercise 3.2.37

Uses. Proposition 3.2.7 (Row operations are left multiplications), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. Make the $(1,1)$ entry 1, clear the entry below it, make the $(2,2)$ entry 1, and clear the entry above it.
Hint 2. In $\mathbb{F}_5$, $2^{-1} = 3$ and $4^{-1} = 4$. To compute G, perform the same operations on I_2.
Sketch. (a) Scale row 1 by 3, add -1 times row 1 to row 2, scale row 2 by 4, add -2 times row 2 to row 1; $G = \left(\begin{smallmatrix}2&2\\3&4\end{smallmatrix}\right)$. (b) Add $-i$ times row 1 to row 2, scale row 2 by $1/3$, add $-i$ times row 2 to row 1; $G = \frac{1}{3}\left(\begin{smallmatrix}2&-i\\-i&1\end{smallmatrix}\right)$.

(a) In $\mathbb{F}_5$: scaling row 1 by $2^{-1} = 3$ gives $\left(\begin{smallmatrix}1&12\\1&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&2\\1&1\end{smallmatrix}\right)$. Adding -1 times row 1 to row 2 gives $\left(\begin{smallmatrix}1&2\\0&-1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&2\\0&4\end{smallmatrix}\right)$. Scaling row 2 by $4^{-1} = 4$ gives $\left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$, and adding -2 times row 2 to row 1 gives I_2. So

$$G = E_{2,-2,1}E_{2,4}E_{1,-1,2}E_{1,3}.$$

Performing the same four operations on I_2: $\left(\begin{smallmatrix}3&0\\0&1\end{smallmatrix}\right)$, then $\left(\begin{smallmatrix}3&0\\-3&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}3&0\\2&1\end{smallmatrix}\right)$, then $\left(\begin{smallmatrix}3&0\\8&4\end{smallmatrix}\right) = \left(\begin{smallmatrix}3&0\\3&4\end{smallmatrix}\right)$, then $\left(\begin{smallmatrix}3-6&-8\\3&4\end{smallmatrix}\right) = \left(\begin{smallmatrix}2&2\\3&4\end{smallmatrix}\right)$. So $G = \left(\begin{smallmatrix}2&2\\3&4\end{smallmatrix}\right)$. Check:

$$AG = \begin{pmatrix}2&4\\1&1\end{pmatrix}\begin{pmatrix}2&2\\3&4\end{pmatrix} = \begin{pmatrix}16&20\\5&6\end{pmatrix} = \begin{pmatrix}1&0\\0&1\end{pmatrix} \quad \text{in } \mathbb{F}_5.$$

(b) Adding $-i$ times row 1 to row 2 gives $\begin{pmatrix}1 & i\\ 0 & 2-i^2\end{pmatrix} = \begin{pmatrix}1 & i\\ 0 & 3\end{pmatrix}$. Scaling row 2 by $1/3$ gives $\begin{pmatrix}1 & i\\ 0 & 1\end{pmatrix}$, and adding $-i$ times row 2 to row 1 gives I_2. So $G = E_{2,-i,1}E_{2,1/3}E_{1,-i,2}$. Performing these operations on I_2: $\begin{pmatrix}1 & 0\\ -i & 1\end{pmatrix}$, then $\begin{pmatrix}1 & 0\\ -i/3 & 1/3\end{pmatrix}$, then

$$G = \begin{pmatrix}1-i(-i/3) & -i/3\\ -i/3 & 1/3\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2 & -i\\ -i & 1\end{pmatrix},$$

using $1 + i^2/3 = 2/3$. Check:

$$AG = \frac{1}{3}\begin{pmatrix}1 & i\\ i & 2\end{pmatrix}\begin{pmatrix}2 & -i\\ -i & 1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2-i^2 & -i+i\\ 2i-2i & -i^2+2\end{pmatrix} = \frac{1}{3}\begin{pmatrix}3 & 0\\ 0 & 3\end{pmatrix} = I_2.$$

In both cases $GA = I_2 = AG$, so $G = A^{-1}$. Computing inverses this way is the subject of Algorithm 3.3.24.

Solution to Exercise 3.2.38

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.28 (Column operations are right multiplications)
Hint 1. On the left, $E_{1,2,3}$ adds 2 times row 1 to row 3. On the right, it adds 2 times column 3 to column 1.
Hint 2. For the last part, read the subscript right to left: adding c times column j to column i is right multiplication by E_{icj}.
Sketch. $E_{1,2,3}A = \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 9 & 12 & 15\end{pmatrix}$ and $AE_{1,2,3} = \begin{pmatrix}7 & 2 & 3\\ 16 & 5 & 6\\ 25 & 8 & 9\end{pmatrix}$. The last part needs $E = E_{3,2,1}$.

By Proposition 3.2.7, $E_{1,2,3}A$ is A with 2 times row 1 added to row 3:

$$E_{1,2,3}A = \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7+2 & 8+4 & 9+6\end{pmatrix} = \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 9 & 12 & 15\end{pmatrix}.$$

By Proposition 3.2.28(c) with $i = 1$, $j = 3$, $AE_{1,2,3}$ is A with 2 times column 3 added to column 1:

$$AE_{1,2,3} = \begin{pmatrix}1+6 & 2 & 3\\ 4+12 & 5 & 6\\ 7+18 & 8 & 9\end{pmatrix} = \begin{pmatrix}7 & 2 & 3\\ 16 & 5 & 6\\ 25 & 8 & 9\end{pmatrix}.$$

Adding 2 times column 1 to column 3 is Proposition 3.2.28(c) with $i = 3$, $j = 1$, so $E = E_{3,2,1}$. Indeed $AE_{3,2,1} = \begin{pmatrix}1 & 2 & 5\\ 4 & 5 & 14\\ 7 & 8 & 23\end{pmatrix}$.

Solution to Exercise 3.2.39

Uses. Proposition 3.2.7 (Row operations are left multiplications), Definition 3.2.2 (Elementary matrices), induction on k
Hint 1. By Proposition 3.2.7, $E_{icj}E_{idj}$ is E_{idj} with c times row i added to row j. What are rows i and j of E_{idj}?
Hint 2. For (b), induct on k. In $\mathbb{F}_p$, $pc = 0$.
Sketch. Row i of E_{idj} is row i of I_m, and row j has d in column i; adding c times row i to row j turns d into $c + d$. Row i of E_{id} is d times row i of I_m, and scaling it by c gives cd. Then induction gives $E_{icj}^k = E_{i(kc)j}$, and $E_{i(pc)j} = E_{i,0,j} = I_m$ in $\mathbb{F}_p$.

(a) By Proposition 3.2.7, $E_{icj}E_{idj}$ is E_{idj} with c times row i added to row j. As $i \neq j$, row i of E_{idj} is row i of I_m, with a 1 in column i, and row j has 1 in column j and d in column i. After the operation, row j has 1 in column j and $d + c$ in column i, and the other rows are unchanged. That is the matrix $E_{i(c+d)j}$.

Similarly, $E_{ic}E_{id}$ is E_{id} with row i scaled by c. Row i of E_{id} has d in column i and zeros elsewhere, so the result has cd there, and it is $E_{i(cd)}$ (note $cd \neq 0$).

(b) For $k = 1$ the claim is clear. If $E_{icj}^k = E_{i(kc)j}$, then by (a),

$$E_{icj}^{k+1} = E_{icj}E_{icj}^k = E_{icj}E_{i(kc)j} = E_{i(c+kc)j} = E_{i((k+1)c)j}.$$

In $\mathbb{F}_p$ we have $pc = 0$, so $E_{icj}^p = E_{i,0,j} = I_m$. For example, over $\mathbb{F}_2$ every type 3 elementary matrix is its own inverse.

Solution to Exercise 3.2.40

Uses. Proposition 3.2.7 (Row operations are left multiplications), Definition 3.2.2 (Elementary matrices)
Hint 1. Compute each product by performing the operations on I_3, the right factor first.
Hint 2. In $E_{2d3}E_{1c2}$, row 2 is changed first, and the changed row 2 is then added to row 3.
Hint 3. For (c), track what each product does to the rows $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ of an arbitrary A. Does either operation change a row that the other one adds?
Sketch. $E_{2d3}E_{1c2} = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ cd & d & 1 \end{pmatrix}$ and $E_{1c2}E_{2d3} = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & d & 1 \end{pmatrix}$: adding d times the new row 2 also adds cd times row 1 to row 3. For (b) and (c), compare the effects on rows $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ of any A.

Write $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ for the rows of a $3 \times n$ matrix A. By Proposition 3.2.7, a product $E'EA$ performs the operation of E first and then that of E'.

(a) $E_{2d3}E_{1c2}A$: first row 2 becomes $\mathbf{r}_2 + c\mathbf{r}_1$, then d times this new row 2 is added to row 3. The rows become

$$\mathbf{r}_1, \quad \mathbf{r}_2 + c\mathbf{r}_1, \quad \mathbf{r}_3 + d\mathbf{r}_2 + cd\mathbf{r}_1.$$

$E_{1c2}E_{2d3}A$: first row 3 becomes $\mathbf{r}_3 + d\mathbf{r}_2$, then row 2 becomes $\mathbf{r}_2 + c\mathbf{r}_1$. The rows become $\mathbf{r}_1$, $\mathbf{r}_2 + c\mathbf{r}_1$, $\mathbf{r}_3 + d\mathbf{r}_2$. Taking $A = I_3$,

$$E_{2d3}E_{1c2} = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ cd & d & 1 \end{pmatrix}, \quad E_{1c2}E_{2d3} = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & d & 1 \end{pmatrix}.$$

The products differ when $cd \neq 0$. The difference is the extra $cd\mathbf{r}_1$ in row 3: in the first order, the row that is added to row 3 has already absorbed $c\mathbf{r}_1$.

(b) $E_{1(cd)3}E_{1c2}E_{2d3}A$ performs, in order: $\mathbf{r}_3 \to \mathbf{r}_3 + d\mathbf{r}_2$; $\mathbf{r}_2 \to \mathbf{r}_2 + c\mathbf{r}_1$; then adds $cd\mathbf{r}_1$ to row 3. The rows become $\mathbf{r}_1$, $\mathbf{r}_2 + c\mathbf{r}_1$, $\mathbf{r}_3 + d\mathbf{r}_2 + cd\mathbf{r}_1$, the same as in the first computation of (a). Taking $A = I_3$ gives the identity.

(c) E_{1c2} and E_{1d3} both read only row 1, which neither changes, and they change different rows. In either order the rows of A become $\mathbf{r}_1$, $\mathbf{r}_2 + c\mathbf{r}_1$, $\mathbf{r}_3 + d\mathbf{r}_1$. Likewise E_{1c2} and E_{3d2} both change only row 2 and read rows 1 and 3, which neither changes; in either order row 2 becomes $\mathbf{r}_2 + c\mathbf{r}_1 + d\mathbf{r}_3$. Taking $A = I_3$, the matrices commute.

Solution to Exercise 3.2.41

Uses. Proposition 3.2.7 (Row operations are left multiplications), Definition 3.2.2 (Elementary matrices)
Hint 1. Follow only rows i and j; call them $\mathbf{u}$ and $\mathbf{v}$. The rightmost factor acts first.
Hint 2. The sequence is $(\mathbf{u}, \mathbf{v}) \to (\mathbf{u}, \mathbf{u}+\mathbf{v}) \to (-\mathbf{v}, \mathbf{u}+\mathbf{v}) \to \cdots$. In $\mathbb{F}_2$, what is $E_{i,-1}$?
Sketch. Rows i, j go $(\mathbf{u}, \mathbf{v}) \to (\mathbf{u}, \mathbf{u}+\mathbf{v}) \to (-\mathbf{v}, \mathbf{u}+\mathbf{v}) \to (-\mathbf{v}, \mathbf{u}) \to (\mathbf{v}, \mathbf{u})$. Over $\mathbb{F}_2$, $-1 = 1$, so $E_{i,-1} = I_m$ and $E_{ij} = E_{i,1,j}E_{j,1,i}E_{i,1,j}$. For (c), replace every swap by the product in (a).

(a) Let A be any matrix with m rows, and let $\mathbf{u}$ and $\mathbf{v}$ be its rows i and j. By Proposition 3.2.7, the product on the right performs four operations on A, starting with the rightmost factor. Only rows i and j change, as follows:

$$(\mathbf{u}, \mathbf{v}) \overset{E_{i,1,j}}{\to} (\mathbf{u}, \mathbf{u}+\mathbf{v}) \overset{E_{j,-1,i}}{\to} (\mathbf{u} - (\mathbf{u}+\mathbf{v}), \mathbf{u}+\mathbf{v}) = (-\mathbf{v}, \mathbf{u}+\mathbf{v})$$
$$\overset{E_{i,1,j}}{\to} (-\mathbf{v}, -\mathbf{v}+\mathbf{u}+\mathbf{v}) = (-\mathbf{v}, \mathbf{u}) \overset{E_{i,-1}}{\to} (\mathbf{v}, \mathbf{u}).$$

So the product performs the swap of rows i and j on every A; taking $A = I_m$ shows that it equals E_{ij}.

(b) In $\mathbb{F}_2$, $-1 = 1$, so $E_{i,-1} = E_{i,1} = I_m$ and $E_{j,-1,i} = E_{j,1,i}$. By (a), $E_{ij} = E_{i,1,j}E_{j,1,i}E_{i,1,j}$.

(c) In a product of elementary matrices, replace each factor E_{ij} by the product in (a). The result is the same matrix, written with factors of types 2 and 3 only.

Solution to Exercise 3.2.42

Uses. Proposition 3.2.7 (Row operations are left multiplications), Exercise 3.2.41
Hint 1. For (a), write out EA for each type and expand D.
Hint 2. For (b), compute D of a product $E_k \cdots E_1$ of type 3 matrices by writing it as $E_k \cdots E_1 I_2$.
Sketch. Adding t times row 1 to row 2 gives $a(d+tb) - b(c+ta) = ad - bc$, and similarly for row 2 into row 1; swapping gives $cb - da$; scaling a row by s multiplies D by s. So a product P of type 3 matrices has $D(P) = D(I_2) = 1$, while $D(E_{12}) = -1 \neq 1$ when $1 + 1 \neq 0$, and $D\big(E_{1,2}\big) = 2$.

(a) By Proposition 3.2.7, the type 3 matrices act as follows. Adding t times row 1 to row 2 gives $\begin{pmatrix} a & b \\ c+ta & d+tb \end{pmatrix}$, and

$$a(d+tb) - b(c+ta) = ad - bc.$$

Adding t times row 2 to row 1 gives $\begin{pmatrix} a+tc & b+td \\ c & d \end{pmatrix}$, and $(a+tc)d - (b+td)c = ad - bc$. Swapping the rows gives $\begin{pmatrix} c & d \\ a & b \end{pmatrix}$, with $D = cb - da = -D(A)$. Scaling row 1 by s gives $D\left(\begin{pmatrix} sa & sb \\ c & d \end{pmatrix}\right) = sad - sbc = sD(A)$, and scaling row 2 by s gives $a(sd) - b(sc) = sD(A)$.

(b) Let $P = E_k \cdots E_1$ with each E_l of type 3. Then $P = E_k \cdots E_1 I_2$, and applying (a) k times, $D(P) = D(I_2) = 1$. But $D(E_{12}) = 0 \cdot 0 - 1 \cdot 1 = -1$, and $-1 \neq 1$ because $1 + 1 \neq 0$. So E_{12} is not such a product. Over $\mathbb{F}_2$, $-1 = 1$ and the invariant D gives no obstruction, in line with Exercise 3.2.41(b), where E_{12} was written as a product of three type 3 matrices.

(c) No: $D\big(E_{1,2}\big) = 2 \neq 1$.

The number $D(A)$ is the determinant of A (Theorem 5.1.13), and (a) is the 2×2 case of the rules computed in Section 5.4.

Solution to Exercise 3.2.43

Uses. Proposition 3.2.14 (Invertible matrices preserve solution sets), Non-example 3.2.19 (Row operations move the column space)
Hint 1. Reread Proposition 3.2.14. Which right-hand side goes with GA?
Hint 2. Try $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and a single row operation.
Sketch. The proposition pairs $A\mathbf{x} = \mathbf{y}$ with $(GA)\mathbf{x} = G\mathbf{y}$, not with $(GA)\mathbf{x} = \mathbf{y}$. What it shows is $\mathbf{y} \in \operatorname{im} T_A \Leftrightarrow G\mathbf{y} \in \operatorname{im} T_{GA}$. Counterexample: $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $G = E_{1,-1,2}$, $GA = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Proposition 3.2.14 compares $A\mathbf{x} = \mathbf{y}$ with $(GA)\mathbf{x} = G\mathbf{y}$: the right-hand side is multiplied by G too. The argument silently replaces $G\mathbf{y}$ by $\mathbf{y}$. What the proposition does give is $\mathbf{y} \in \operatorname{im} T_A \Leftrightarrow G\mathbf{y} \in \operatorname{im} T_{GA}$, so $\operatorname{im} T_{GA} = \{G\mathbf{y} \mid \mathbf{y} \in \operatorname{im} T_A\}$, which is usually a different subspace.

Counterexample: let $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^{2\times 1}$ and $G = E_{1,-1,2}$, so $GA = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Then $\operatorname{im} T_A = \operatorname{span}((1,1))$ and $\operatorname{im} T_{GA} = \operatorname{span}((1,0))$, and $(1,1)$ is not a multiple of $(1,0)$.

Solution to Exercise 3.2.44

Uses. Definition 3.2.21 (Row equivalence), Corollary 3.2.23 (Row equivalent matrices have the same kernel), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. Write $B = GA$ with G a product of elementary matrices. Which statements follow from associativity alone?
Hint 2. To disprove, use kernels (Corollary 3.2.23). For (a) and (c), try $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ or $A = I_2$ with B obtained by a swap.
Sketch. (b) and (e) are true: $BC = G(AC)$, and $B = GA$ is a product of invertible matrices. (a), (c), (d) are false: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ but the transposes have kernels $\operatorname{span}(\mathbf{e}_2) \neq \operatorname{span}(\mathbf{e}_1)$; with $C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, CI_2 and CE_{12} have different kernels; $I_2 + I_2$ and $E_{12} + I_2$ have different kernels.

Throughout, $B = GA$ where $G = E_k \cdots E_1$ is a product of elementary matrices.

(a) False. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = E_{12}A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, so $A \sim B$. Then $A^{\mathsf{T}} = A$ has kernel $\operatorname{span}(\mathbf{e}_2)$, while $B^{\mathsf{T}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ sends $(x_1, x_2) \mapsto (x_2, 0)$ and has kernel $\operatorname{span}(\mathbf{e}_1)$. By Corollary 3.2.23, A^{T} and B^{T} are not row equivalent.

(b) True. $BC = (GA)C = G(AC) = E_k \cdots E_1(AC)$, so $AC \sim BC$.

(c) False. Let $A = I_2$, $B = E_{12}$ (so $A \sim B$), and $C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then $CA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $CB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, whose kernels are $\operatorname{span}(\mathbf{e}_2)$ and $\operatorname{span}(\mathbf{e}_1)$, so CA and CB are not row equivalent.

(d) False. Let $A = C = D = I_2$ and $B = E_{12}$. Then $A + C = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $B + D = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. The kernel of $B + D$ is $\operatorname{span}((1, -1))$. The kernel of $A + C$ is $\{\mathbf{0}\}$ if $1 + 1 \neq 0$ and all of F^2 if $1 + 1 = 0$; either way it is not $\operatorname{span}((1, -1))$. So $A + C$ and $B + D$ are not row equivalent.

(e) True. G is invertible by Corollary 3.2.11, so $B = GA$ is a product of invertible matrices, which is invertible with inverse $A^{-1}G^{-1}$.

Solution to Exercise 3.2.45

Uses. Proposition 3.2.14 (Invertible matrices preserve solution sets), Definition 2.5.1 (Left inverse), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Theorem 2.7.2 (First Fundamental Theorem)
Hint 1. Rerun the proof of Proposition 3.2.14. Where exactly was G^{-1} used?
Hint 2. For (b), let the third row of G produce the sum of the two equations.
Hint 3. For (c), suppose $G\mathbf{z} = \mathbf{0}$. Choose $A = 0$ and $\mathbf{y} = \mathbf{z}$ and compare the two solution sets.
Sketch. (a) $\Rightarrow$ is multiplication by G; $\Leftarrow$ multiplies by H: $A\mathbf{x} = HGA\mathbf{x} = HG\mathbf{y} = \mathbf{y}$. (b) $G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ appends the sum of the two equations; $H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. (c) If $G\mathbf{z} = \mathbf{0}$, then with $A = 0$, $\mathbf{y} = \mathbf{z}$, every x solves the second system, hence the first, so $\mathbf{z} = \mathbf{0}$.

(a) If $A\mathbf{x} = \mathbf{y}$, then $(GA)\mathbf{x} = G(A\mathbf{x}) = G\mathbf{y}$. Conversely, if $(GA)\mathbf{x} = G\mathbf{y}$, multiplying on the left by H gives

$$A\mathbf{x} = (HG)A\mathbf{x} = H((GA)\mathbf{x}) = H(G\mathbf{y}) = (HG)\mathbf{y} = \mathbf{y}.$$

Only the left inverse was used.

(b) Let $G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ and $H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Then $HG = I_2$. For a system of two equations with augmented matrix $(A \mid \mathbf{y})$, the matrix $G(A \mid \mathbf{y}) = (GA \mid G\mathbf{y})$ consists of the two original equations followed by their sum. By (a), appending this redundant equation does not change the solution set. (G is not square, so it is not invertible.)

(c) By Theorem 2.6.11 it suffices to show $\ker T_G = \{\mathbf{0}\}$. Let $\mathbf{z} \in F^m$ with $G\mathbf{z} = \mathbf{0}$. Take $A = 0 \in F^{m \times 1}$ and $\mathbf{y} = \mathbf{z}$. For every $x \in F = F^1$, $(GA)x = \mathbf{0} = G\mathbf{z} = G\mathbf{y}$, so every x solves the second system. By hypothesis every x then solves the first system, so $\mathbf{0} = Ax = \mathbf{y} = \mathbf{z}$. Hence $\ker T_G = \{\mathbf{0}\}$, and T_G is injective.

By Theorem 2.7.2, an injective T_G has a left inverse. So (a) and (c) together say: multiplying systems by G always preserves solution sets exactly when G has a left inverse.

Solution to Exercise 3.2.46

Uses. Proposition 3.2.30 (Column operations preserve the image), Proposition 2.4.23 (Transpose of a product), Corollary 3.2.11 (Products of elementary matrices), Definition 2.5.12 (Invertible)
Techniques. Transpose to trade rows for columns (Technique 3.2.27), Double containment (Technique 1.3.14)
Hint 1. If $B = GA$, then $B^{\mathsf{T}} = A^{\mathsf{T}}G^{\mathsf{T}}$. Which result about right multiplication applies?
Hint 2. You need G^{T} to be invertible. Transpose the equations $GG^{-1} = G^{-1}G = I_m$.
Sketch. $B^{\mathsf{T}} = A^{\mathsf{T}}G^{\mathsf{T}}$, and G^{T} is invertible with inverse $\left(G^{-1}\right)^{\mathsf{T}}$. By Proposition 3.2.30, $\operatorname{im} T_{B^{\mathsf{T}}} = \operatorname{im} T_{A^{\mathsf{T}}}$.

Let $A \sim B$, so $B = GA$ with G a product of elementary matrices, invertible by Corollary 3.2.11. By Proposition 2.4.23, $B^{\mathsf{T}} = A^{\mathsf{T}}G^{\mathsf{T}}$. Transposing $GG^{-1} = I_m$ and $G^{-1}G = I_m$ gives $\left(G^{-1}\right)^{\mathsf{T}}G^{\mathsf{T}} = I_m$ and $G^{\mathsf{T}}\left(G^{-1}\right)^{\mathsf{T}} = I_m$ (as $I_m^{\mathsf{T}} = I_m$), so G^{T} is invertible. By Proposition 3.2.30, $\operatorname{im} T_{B^{\mathsf{T}}} = \operatorname{im} T_{A^{\mathsf{T}}G^{\mathsf{T}}} = \operatorname{im} T_{A^{\mathsf{T}}}$.

Solution to Exercise 3.2.46

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.22 (Row equivalence is an equivalence relation), Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Transpose to trade rows for columns (Technique 3.2.27), Double containment (Technique 1.3.14)
Hint 1. It suffices to treat a single row operation, then use symmetry.
Hint 2. Each row of EA is a combination of rows of A, so it lies in the row space of A.
Sketch. Every row of EA is a row of A or a combination of two rows of A, so the row space of EA is contained in that of A. Since A is obtained back from EA by an elementary operation, the reverse containment holds as well; then induct on the number of operations.

Write $R(A)$ for the row space. First let E be elementary. By Proposition 3.2.7, every row of EA is a row of A, a nonzero multiple of a row of A, or a row of A plus a multiple of another row of A. So every row of EA lies in $R(A)$, and since $R(A)$ is a subspace, $R(EA) \subseteq R(A)$ (Proposition 1.4.14). Also $A = E^{-1}(EA)$ with E^{-1} elementary, so the same argument gives $R(A) \subseteq R(EA)$. Hence $R(EA) = R(A)$. If $B = E_k \cdots E_1 A$, applying this k times gives $R(B) = R(A)$.

Solution to Exercise 3.2.47

Uses. Definition 3.2.21 (Row equivalence), Definition 3.2.1 (Elementary row operations), Definition 2.4.1 (Matrix product), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. With only one row, which elementary row operations are available?
Hint 2. For (b), if $B = GA$, then the columns of B are G times the columns of A. What do you learn about the two columns of B?
Sketch. (a) A matrix with one row admits only scalings, and a sequence of them multiplies by a nonzero scalar. (b) $B = G\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ has both columns equal to $G\mathbf{e}_1 \neq \mathbf{0}$, so $B = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$ with $(a, b) \neq (0, 0)$; conversely each such B is reached by a scaling and a replacement, or a swap and a scaling.

(a) Swaps and replacements need two different rows, so the only elementary row operations on a $1 \times n$ matrix are scalings by nonzero scalars. A sequence of them multiplies the matrix by the product of the scalars, which is nonzero. So $A \sim B$ implies $B = sA$ with $s \neq 0$; conversely, if $B = sA$ with $s \neq 0$, scaling the row by s turns A into B.

(b) Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. We claim that $B \sim A$ if and only if $B = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$ with $(a, b) \neq (0, 0)$.

Suppose $B \sim A$, and write $B = GA$ with G a product of elementary matrices. By Definition 2.4.1, both columns of B equal $G\mathbf{e}_1$. This vector is nonzero, because G is invertible (Corollary 3.2.11) and so $G\mathbf{e}_1 = \mathbf{0}$ would give $\mathbf{e}_1 = G^{-1}\mathbf{0} = \mathbf{0}$. So $B = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$ with $(a, b) = G\mathbf{e}_1 \neq (0, 0)$.

Conversely, let $(a, b) \neq (0, 0)$. If $a \neq 0$, scale row 1 of A by a to get $\begin{pmatrix} a & a \\ 0 & 0 \end{pmatrix}$, then add b/a times row 1 to row 2 to get $\begin{pmatrix} a & a \\ b & b \end{pmatrix}$. If $a = 0$, then $b \neq 0$; swap the rows of A to get $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, then scale row 2 by b.

Solution to Exercise 3.2.48

Uses. Definition 3.2.2 (Elementary matrices)
Hint 1. Count each type, but watch for coincidences. Which choices of c give I_m? Can a matrix of one type equal a matrix of another type?
Hint 2. Compare the diagonal entries: a swap has two diagonal entries 0, a scaling other than I_m is diagonal, and a replacement other than I_m has a nonzero off-diagonal entry and all diagonal entries 1.
Sketch. Apart from I_m, the types are disjoint. There are $\binom{m}{2}$ swaps (as $E_{ij} = E_{ji}$), $m(p-2)$ scalings with $c \notin \{0, 1\}$, and $m(m-1)(p-1)$ replacements with $c \neq 0$. Total $1 + \binom{m}{2} + m(p-2) + m(m-1)(p-1)$; for $p = m = 2$ this is 4: I_2, E_{12}, $E_{1,1,2}$, $E_{2,1,1}$.

The identity is elementary ($I_m = E_{1,1}$). We count the elementary matrices other than I_m.

Swaps. E_{ij} has diagonal entries 0 in positions i and j. Since $E_{ij} = E_{ji}$ and different pairs $\{i, j\}$ give different matrices, there are $\binom{m}{2} = m(m-1)/2$ of them.

Scalings other than I_m. E_{ic} with $c \notin \{0, 1\}$ is diagonal and differs from I_m exactly in position (i, i), where it has c. Different pairs (i, c) give different matrices, so there are $m(p-2)$ of them.

Replacements other than I_m. E_{icj} with $c \neq 0$ has all diagonal entries 1 and a single nonzero off-diagonal entry c in position (j,i). Different triples (i,c,j) give different matrices, so there are $m(m-1)(p-1)$ of them.

No matrix lies in two of these groups: swaps have zero diagonal entries, scalings other than I_m are diagonal with a diagonal entry different from 1, and replacements other than I_m are not diagonal but have all diagonal entries 1. So the number of elementary matrices in $\mathbb{F}_p^{m\times m}$ is

$$1 + m\frac{m-1}{2} + m(p-2) + m(m-1)(p-1).$$

For $p=2$, $m=2$ this is $1+1+0+2=4$:

$$I_2, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_{1,1,2} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad E_{2,1,1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Solution to Exercise 3.2.49

Uses. Example 3.2.34 (A rotation built from shears)
Hint 1. Multiply out with $t = \tan(\theta/2)$ and $s = \sin\theta$, and simplify ts with the double-angle formulas.
Hint 2. $ts = 2\sin^2(\theta/2) = 1 - \cos\theta$. For (b), use $R_\pi = R_{\pi/2}^2$ and Example 3.2.34.
Sketch. With $t = \tan(\theta/2)$, $s = \sin\theta$, the product is $\begin{pmatrix} 1-ts & -t(2-ts) \\ s & 1-ts \end{pmatrix}$. Since $ts = 1-\cos\theta$, the diagonal entries are $\cos\theta$ and the top right entry is $-t(1+\cos\theta) = -\sin\theta$. For (b), $(S(-1)L(1)S(-1))^2 = S(-1)L(1)S(-2)L(1)S(-1)$.

(a) Let $t = \tan(\theta/2)$, which is defined because $\theta/2$ is not an odd multiple of $\pi/2$, and $s = \sin\theta$. Then

$$S(-t)L(s) = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} = \begin{pmatrix} 1-ts & -t \\ s & 1 \end{pmatrix}, \quad S(-t)L(s)S(-t) = \begin{pmatrix} 1-ts & -t(1-ts)-t \\ s & 1-ts \end{pmatrix}.$$

By the double-angle formulas, $ts = (\sin(\theta/2)/\cos(\theta/2)) \cdot 2\sin(\theta/2)\cos(\theta/2) = 2\sin^2(\theta/2) = 1-\cos\theta$. So $1-ts = \cos\theta$, and the top right entry is

$$-t(\cos\theta + 1) = -\frac{\sin(\theta/2)}{\cos(\theta/2)} \cdot 2\cos^2(\theta/2) = -2\sin(\theta/2)\cos(\theta/2) = -\sin\theta.$$

Hence the product is R_θ, as desired. For $\theta = \pi/2$ this is the factorization $R = SLS$ of Example 3.2.34.

(b) $R_\pi = R_{\pi/2}R_{\pi/2}$, and by Example 3.2.34 $R_{\pi/2} = S(-1)L(1)S(-1)$. So

$$-I_2 = S(-1)L(1)S(-1)S(-1)L(1)S(-1) = S(-1)L(1)S(-2)L(1)S(-1),$$

using $S(-1)S(-1) = S(-2)$ (Exercise 3.2.39). (The formula in (a) is how images are rotated in software that can only shift pixels along rows and columns: each shear moves every row, or every column, by a fixed amount.)

Solution to Exercise 3.2.50

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.10 (Inverses of elementary matrices), Corollary 3.2.11 (Products of elementary matrices)
Hint 1. If $a \neq 0$: scale row 1 by a^{-1}, clear the $(2,1)$ entry, and check that the new $(2,2)$ entry is $(ad-bc)/a$. If $a = 0$, swap first.
Hint 2. From $E_k\cdots E_1 A = I_2$, multiply on the left by $E_1^{-1}\cdots E_k^{-1}$.
Sketch. If $a \neq 0$: $E_{1a^{-1}}$, then $E_{1(-c)2}$, giving $\begin{pmatrix} 1 & b/a \\ 0 & (ad-bc)/a \end{pmatrix}$; then scale row 2 by $a/(ad-bc)$ and add $-b/a$ times row 2 to row 1. If $a = 0$ then $c \neq 0$; swap first. Then $A = E_1^{-1}\cdots E_k^{-1}$. For $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$: $A = E_{12}E_{1,-2}E_{2,-3/2,1}$.

(a) *Case $a \neq 0$.* Scaling row 1 by a^{-1} gives $\begin{pmatrix} 1 & b/a \\ c & d \end{pmatrix}$. Adding $-c$ times row 1 to row 2 gives

$$\begin{pmatrix} 1 & b/a \\ 0 & d - cb/a \end{pmatrix} = \begin{pmatrix} 1 & b/a \\ 0 & (ad-bc)/a \end{pmatrix}.$$

The $(2,2)$ entry $\delta = (ad-bc)/a$ is nonzero. Scaling row 2 by δ^{-1} gives $\begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix}$, and adding $-b/a$ times row 2 to row 1 gives I_2. With

$$E_1 = E_{1a^{-1}}, \quad E_2 = E_{1(-c)2}, \quad E_3 = E_{2\delta^{-1}}, \quad E_4 = E_{2(-b/a)1},$$

Proposition 3.2.7 gives $E_4E_3E_2E_1A = I_2$.

Case $a = 0$. Then $-bc = ad - bc \neq 0$, so $c \neq 0$. Swapping the rows gives $\begin{pmatrix} c & d \\ 0 & b \end{pmatrix}$, whose $(1,1)$ entry is nonzero and whose value of "$ad - bc$" is $cb - d \cdot 0 = cb \neq 0$. Apply the first case to it.

In either case $E_k \cdots E_1 A = I_2$. Multiplying on the left by $E_1^{-1} \cdots E_k^{-1}$, which is the inverse of $E_k \cdots E_1$ (Corollary 3.2.11), gives $A = E_1^{-1} \cdots E_k^{-1}$, and each E_l^{-1} is elementary by Proposition 3.2.10.

(b) Here $a = 0$. Swapping the rows gives $\begin{pmatrix} -2 & 3 \\ 0 & 1 \end{pmatrix}$. Scaling row 1 by $-1/2$ gives $\begin{pmatrix} 1 & -3/2 \\ 0 & 1 \end{pmatrix}$, and adding $3/2$ times row 2 to row 1 gives I_2. So $E_{2,3/2,1}E_{1,-1/2}E_{12}A = I_2$, and

$$A = E_{12}E_{1,-2}E_{2,-3/2,1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -3/2 \\ 0 & 1 \end{pmatrix}.$$

Indeed, $\begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -3/2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 0 & 1 \end{pmatrix}$, and swapping its rows gives A. So T_A is: a horizontal shear by $-3/2$, then a horizontal stretch by 2 combined with the reflection across the x_2-axis, then the reflection across the line $x_1 = x_2$.

Solution to Exercise 3.2.51

Uses. Proposition 3.2.28 (Column operations are right multiplications), Corollary 3.2.11 (Products of elementary matrices), Theorem 2.4.2 (The matrix of a composite), Remark 3.2.29 (One matrix, two operations)
Hint 1. Use column 1 to clear the rest of row 1, then fix row 2. Compute H by performing the same column operations on I_3.
Hint 2. For (b), write $\mathbf{x} = H\mathbf{z}$. Then $A\mathbf{x} = (AH)\mathbf{z} = (z_1, z_2)$.
Sketch. Add -2 times column 1 to column 2 and -3 times column 1 to column 3, swap columns 2 and 3, add -2 times column 2 to column 1. On I_3 this gives $H = \begin{pmatrix} 7 & -3 & -2 \\ 0 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$. Since H is invertible, every $\mathbf{x}$ is $H\mathbf{z}$, and $A\mathbf{x} = (z_1, z_2)$ vanishes exactly when $\mathbf{z} = t\mathbf{e}_3$, that is, $\mathbf{x} = t\mathbf{h}_3 = t(-2, 1, 0)$.

(a) Adding -2 times column 1 to column 2 gives $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 0 & 7 \end{pmatrix}$; adding -3 times column 1 to column 3 gives $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \end{pmatrix}$; swapping columns 2 and 3 gives $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \end{pmatrix}$; and adding -2 times column 2 to column 1 gives $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. By Proposition 3.2.28, and since the first column operation is the leftmost factor (Remark 3.2.29),

$$H = E_{2,-2,1}E_{3,-3,1}E_{23}E_{1,-2,2}.$$

Performing the four column operations on I_3 (because $I_3H = H$):

$$I_3 \to \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -3 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 7 & -3 & -2 \\ 0 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} = H.$$

(b) H is invertible (Corollary 3.2.11), so every $\mathbf{x} \in \mathbb{R}^3$ can be written as $\mathbf{x} = H\mathbf{z}$ with $\mathbf{z} = H^{-1}\mathbf{x}$. Then

$$A\mathbf{x} = (AH)\mathbf{z} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

So $\mathbf{x} \in \ker T_A$ if and only if $z_1 = z_2 = 0$, that is, $\mathbf{z} = t\mathbf{e}_3$ for some $t \in \mathbb{R}$, that is, $\mathbf{x} = tH\mathbf{e}_3 = t\mathbf{h}_3$. Hence $\ker T_A = \operatorname{span}(\mathbf{h}_3) = \operatorname{span}((-2, 1, 0))$. Check: $A(-2, 1, 0) = (-2 + 2, -4 + 4) = \mathbf{0}$.

Solution to Exercise 3.2.52

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.28 (Column operations are right multiplications)
Hint 1. It suffices to use the matrices $E_{i,1,j}$. Describe $ME_{i,1,j}$ and $E_{i,1,j}M$ as column and row operations on M.

Hint 2. $ME_{i,1,j}$ is M with column j added to column i, and $E_{i,1,j}M$ is M with row i added to row j. Compare the entries in column i.
Sketch. $ME_{i,1,j}$ differs from M only in column i, by m_{kj} in row k; $E_{i,1,j}M$ differs from M only in row j, by m_{il} in column l. Comparing entry (k,i) for $k \neq j$ gives $m_{kj} = 0$; comparing entry (j,i) gives $m_{jj} = m_{ii}$. So M is diagonal with equal diagonal entries.

If $M = \lambda I_m$, then $ME = \lambda E = EM$ for every E.

Conversely, suppose M commutes with every type 3 elementary matrix, and let $i \neq j$. By Proposition 3.2.28, $ME_{i,1,j}$ is M with column j added to column i: it agrees with M outside column i, and its (k,i) entry is $m_{ki} + m_{kj}$. By Proposition 3.2.7, $E_{i,1,j}M$ is M with row i added to row j: it agrees with M outside row j, and its (j,l) entry is $m_{jl} + m_{il}$.

Compare the (k,i) entries for $k \neq j$: on the left $m_{ki} + m_{kj}$, on the right m_{ki} (row k is unchanged). So $m_{kj} = 0$ for all $k \neq j$. Compare the (j,i) entries: $m_{ji} + m_{jj} = m_{ji} + m_{ii}$, so $m_{jj} = m_{ii}$.

Since $m \geq 2$, every j has some $i \neq j$. So every off-diagonal entry of M is 0 and all diagonal entries are equal: $M = \lambda I_m$ with $\lambda = m_{11}$.

Solution to Exercise 3.2.53

Uses. Exercise 3.2.42, Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.10 (Inverses of elementary matrices)
Hint 1. It suffices to reduce A to I_2 with at most four row replacements, because inverses of type 3 matrices are of type 3. Type 3 operations do not change $ad - bc$ (Exercise 3.2.42).
Hint 2. First make sure the $(2,1)$ entry is nonzero. Then one replacement makes the $(1,1)$ entry equal to 1.
Hint 3. With $(1,1)$ entry 1, clear the $(2,1)$ entry. What does $ad - bc = 1$ say about the new $(2,2)$ entry?
Sketch. If $c = 0$, then $a \neq 0$; add row 1 to row 2 to make $c \neq 0$. Add $(1-a)/c$ times row 2 to row 1, making $a = 1$; add $-c$ times row 1 to row 2, making $c = 0$, and then $d = 1$ since $ad - bc$ is still 1; add $-b$ times row 2 to row 1. That is at most four replacements, and A is the product of their inverses in reverse order.

By Exercise 3.2.42(a), type 3 row operations do not change $D(A) = ad - bc$, so every matrix below has $D = 1$. We find type 3 matrices $E_1, \ldots, E_k$, with $k \leq 4$, such that $E_k \cdots E_1 A = I_2$. Then $A = E_1^{-1} \cdots E_k^{-1}$, and each E_l^{-1} is of type 3 by Proposition 3.2.10.

Step 1. If $c = 0$, then $ad = 1$, so $a \neq 0$. Adding row 1 to row 2 changes the $(2,1)$ entry to $a \neq 0$. So after at most one operation the $(2,1)$ entry, which we still call c, is nonzero.

Step 2. Add $(1-a)/c$ times row 2 to row 1. The new $(1,1)$ entry is $a + ((1-a)/c)c = 1$.

Step 3. Add $-c$ times row 1 to row 2. The new $(2,1)$ entry is $c - c \cdot 1 = 0$, so the matrix is $\begin{pmatrix} 1 & b' \\ 0 & d' \end{pmatrix}$ for some b', d'. Its value of D is d', and it is still 1, so $d' = 1$.

Step 4. Add $-b'$ times row 2 to row 1. This gives I_2.

We used three operations if $c \neq 0$ at the start and four if $c = 0$. So A is a product of at most four shears. For example, over $\mathbb{F}_2$ the swap E_{12} has $D = -1 = 1$, and Exercise 3.2.41(b) wrote it as three shears.

Solution to Exercise 3.2.54

Uses. Proposition 3.2.10 (Inverses of elementary matrices), Proposition 3.2.7 (Row operations are left multiplications), the division algorithm, strong induction on $|b|$
Hint 1. Operations by integer elementary matrices are invertible over $\mathbb{Z}$. What does that say about the common divisors of the two entries?
Hint 2. Imitate the Euclidean algorithm. If $b \neq 0$, write $a = qb + r$ with $0 \leq r < |b|$. Which operation turns (a,b) into (r,b)?
Hint 3. For (c), look at the first row of $G = E_k \cdots E_1$.
Sketch. The operations preserve the common divisors of the two entries, so they preserve the gcd. If $b = 0$, scale by ± 1. Otherwise add $-q$ times row 2 to row 1 and swap, which replaces (a,b) by (b,r) with $|r| < |b|$; induct on $|b|$. Then $G\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$ with G an integer matrix, whose first row (u,v) satisfies $ua + vb = g$. For $(42, 30)$: $G = \begin{pmatrix} -2 & 3 \\ 5 & -7 \end{pmatrix}$, and $-2 \cdot 42 + 3 \cdot 30 = 6$.

(a) The matrices E_{12}, $E_{i,-1}$, and E_{icj} with $c \in \mathbb{Z}$ have entries in $\{0, 1, -1, c\}$, so they are integer matrices. By Proposition 3.2.10 their inverses are E_{12}, $E_{i,-1}$, and $E_{i(-c)j}$, which are again integer elementary matrices.

(b) *Common divisors are preserved.* Let E be an integer elementary matrix and $\mathbf{w} \in \mathbb{Z}^2$. Each entry of $E\mathbf{w}$ is an integer combination of the entries of $\mathbf{w}$, so every common divisor of the entries of $\mathbf{w}$ divides both entries of $E\mathbf{w}$. Since $\mathbf{w} = E^{-1}(E\mathbf{w})$ and E^{-1} is also an integer elementary matrix by (a), the converse holds too. So $\mathbf{w}$ and $E\mathbf{w}$ have the same common divisors, and in particular the same gcd.

The algorithm. We prove the claim by strong induction on $|b|$. If $b = 0$, then $g = |a|$, and we are done after applying $E_{1,-1}$ if $a < 0$. If $b \neq 0$, divide with remainder: $a = qb + r$ with $q, r \in \mathbb{Z}$ and $0 \leq r < |b|$. Adding $-q$ times row 2 to row 1 (the matrix $E_{2,-q,1}$) turns $\begin{pmatrix} a \\ b \end{pmatrix}$ into $\begin{pmatrix} r \\ b \end{pmatrix}$, and swapping turns it into $\begin{pmatrix} b \\ r \end{pmatrix}$. These have the same gcd g, and $|r| < |b|$, so by induction some product of integer elementary matrices turns $\begin{pmatrix} b \\ r \end{pmatrix}$ into $\begin{pmatrix} g \\ 0 \end{pmatrix}$. Composing gives the claim.

(c) Let $G = E_k \cdots E_1$. It has integer entries, being a product of integer matrices. If its first row is (u, v), then the first entry of $G\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$ says $ua + vb = g$.

(d) Adding -1 times row 2 to row 1 gives $\begin{pmatrix} 12 \\ 30 \end{pmatrix}$; adding -2 times row 1 to row 2 gives $\begin{pmatrix} 12 \\ 6 \end{pmatrix}$; adding -2 times row 2 to row 1 gives $\begin{pmatrix} 0 \\ 6 \end{pmatrix}$; swapping gives $\begin{pmatrix} 6 \\ 0 \end{pmatrix}$. So $G = E_{12}E_{2,-2,1}E_{1,-2,2}E_{2,-1,1}$, with the first operation as the rightmost factor. Performing the four operations on I_2: $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, then $\begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}$, then $\begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix}$, then

$$G = \begin{pmatrix} -2 & 3 \\ 5 & -7 \end{pmatrix}.$$

(This takes a shortcut through the algorithm of (b): instead of swapping after every division, it reduces whichever entry is larger.) So $u = -2$, $v = 3$: indeed $-2 \cdot 42 + 3 \cdot 30 = -84 + 90 = 6 = \gcd(42, 30)$. The second row says $5 \cdot 42 - 7 \cdot 30 = 0$.

Solution to Exercise 3.2.55

Uses. Exercise 3.2.42, Proposition 3.2.28 (Column operations are right multiplications)
Hint 1. For $ad - bc$, the row moves were handled in Exercise 3.2.42. Do the same computation for columns, or transpose.
Hint 2. For the gcd, every new entry is an integer combination of old entries, and every move can be undone by a move.
Hint 3. Compute both invariants for M and for each target. For a target that passes both tests, find the moves.
Sketch. Row moves keep $ad - bc$ by Exercise 3.2.42; a column move adding t times column 1 to column 2 gives $a(d + tc) - (b + ta)c = ad - bc$, and similarly for the other column move. Moves preserve the set of common divisors of the entries because they are integer combinations and reversible. M has $ad - bc = -8$ and gcd 2. The first target has gcd 1, the second has $ad - bc = 8$, and the third is reached by adding -3 times row 1 to row 2 and then -2 times column 1 to column 2.

(a) *The quantity $ad - bc$.* Row moves do not change it, by Exercise 3.2.42(a). Adding t times column 1 to column 2 gives $\begin{pmatrix} a & b+ta \\ c & d+tc \end{pmatrix}$, and $a(d + tc) - (b + ta)c = ad - bc$. Adding t times column 2 to column 1 gives $\begin{pmatrix} a+tb & b \\ c+td & d \end{pmatrix}$, and $(a + tb)d - b(c + td) = ad - bc$.

The gcd. After a move, each entry is an old entry plus an integer multiple of another old entry. So every integer dividing all four old entries divides all four new ones. The move is undone by the move with $-t$ in place of t, so conversely every common divisor of the new entries divides the old ones. The sets of common divisors agree, and so do the gcds.

(b) For M, $ad - bc = 16 - 24 = -8$ and the gcd of $2, 4, 6, 8$ is 2.

$\begin{pmatrix} 1 & 0 \\ 0 & -8 \end{pmatrix}$ has $ad - bc = -8$ but gcd 1, so it cannot be reached.

$\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ has gcd 2 but $ad - bc = 8 \neq -8$, so it cannot be reached.

$\begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$ can be reached. Adding -3 times row 1 to row 2 gives $\begin{pmatrix} 2 & 4 \\ 0 & -4 \end{pmatrix}$, and adding -2 times column 1 to column 2 gives $\begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$.

In matrix terms, a move multiplies on the left or on the right by an integer elementary matrix E_{itj} (Proposition 3.2.7, Proposition 3.2.28). The two invariants here are the start of the theory of the Smith normal form, which classifies integer matrices up to such operations.

Solution to Exercise 3.2.56

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Exercise 3.2.41, Exercise 3.2.53, Proposition 3.2.7 (Row operations are left multiplications)
Hint 1. By Exercise 3.2.41(c), A is a product of matrices of types 2 and 3. Try to move all the scalings to the right end.
Hint 2. Compute $DE_{icj}D^{-1}$ for a diagonal D with nonzero diagonal entries, by tracking rows. It is again of type 3.
Hint 3. A diagonal matrix with entries t and t^{-1} in positions $k, k+1$ and 1 elsewhere is a product of type 3 matrices, by Exercise 3.2.53 applied in rows $k, k+1$. Use such matrices to push all of a diagonal matrix's entries into the last position.
Sketch. Write A as a product of type 2 and type 3 matrices. For diagonal invertible D, $DE_{icj}D^{-1} = E_{i(cd_j/d_i)j}$, so the scalings can be moved to the right: $A = P_1 D$ with P_1 a product of type 3 matrices and $D = \operatorname{diag}(d_1, \ldots, d_n)$. With $p_k = d_1 \cdots d_k$, $D = D_1 \cdots D_{n-1} E_{np_n}$, where D_k has p_k, p_k^{-1} in positions $k, k+1$; each D_k is a product of shears by the 2×2 result.

Write $\operatorname{diag}(d_1, \ldots, d_n)$ for the diagonal matrix with diagonal entries $d_1, \ldots, d_n$. A product of type 2 matrices is diagonal with nonzero diagonal entries, and conversely $\operatorname{diag}(d_1, \ldots, d_n) = E_{1d_1} \cdots E_{nd_n}$ when all $d_k \neq 0$ (each factor scales one row of I_n).

Step 1: only types 2 and 3. By Theorem 3.5.19, A is a product of elementary matrices, and by Exercise 3.2.41(c) we may assume that every factor has type 2 or 3.

Step 2: moving scalings to the right. Let $D = \operatorname{diag}(d_1, \ldots, d_n)$ with all $d_k \neq 0$, and let $i \neq j$. We claim

$$DE_{icj} = E_{ic'j}D, \quad c' = cd_j/d_i.$$

Indeed, apply $DE_{icj}D^{-1}$ to a matrix with rows $\mathbf{r}_1, \ldots, \mathbf{r}_n$, using Proposition 3.2.7 three times. $D^{-1} = \operatorname{diag}\left(d_1^{-1}, \ldots, d_n^{-1}\right)$ gives rows $\mathbf{r}_k/d_k$; then E_{icj} changes row j to $\mathbf{r}_j/d_j + c\mathbf{r}_i/d_i$; then D multiplies row k by d_k. The result has rows $\mathbf{r}_k$ for $k \neq j$ and $\mathbf{r}_j + c(d_j/d_i)\mathbf{r}_i$ in row j, which is what $E_{ic'j}$ does. So $DE_{icj}D^{-1} = E_{ic'j}$, which gives the claim. Using it repeatedly, and merging adjacent diagonal factors (a product of diagonal matrices is diagonal), we can move every type 2 factor in the product of Step 1 to the right end. So $A = P_1 D$ with P_1 a product of type 3 matrices and $D = \operatorname{diag}(d_1, \ldots, d_n)$, all $d_k \neq 0$.

Step 3: the diagonal part. For $1 \leq k < n$ and $t \neq 0$, let $D_k(t)$ be the diagonal matrix with t in position k, t^{-1} in position $k+1$, and 1 elsewhere. The 2×2 matrix $\operatorname{diag}\left(t, t^{-1}\right)$ has $ad - bc = 1$, so by Exercise 3.2.53 it is a product of 2×2 shears E_{1s2} and E_{2s1}. Replace each of these by the $n \times n$ matrix $E_{k,s,k+1}$, respectively $E_{k+1,s,k}$. By Proposition 3.2.7, these $n \times n$ matrices perform on rows $k, k+1$ of I_n exactly the operations that the 2×2 ones perform on rows $1, 2$ of I_2, and they leave the other rows alone. So their product is $D_k(t)$, which is therefore a product of type 3 matrices.

Now let $p_k = d_1 d_2 \cdots d_k$, and consider $D_1(p_1)D_2(p_2)\cdots D_{n-1}(p_{n-1})E_{np_n}$. All factors are diagonal. Its first diagonal entry is $p_1 = d_1$; for $2 \leq k \leq n-1$ its kth entry is $p_{k-1}^{-1}p_k = d_k$; and its nth entry is $p_{n-1}^{-1}p_n = d_n$. So it equals D.

Therefore $A = P_1 D_1(p_1)\cdots D_{n-1}(p_{n-1})E_{nd}$ with $d = p_n = d_1 \cdots d_n \neq 0$, and $P = P_1 D_1(p_1)\cdots D_{n-1}(p_{n-1})$ is a product of type 3 matrices.

Section 3.3

Solution to Exercise 3.3.4

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Proposition 3.2.7 (Row operations are left multiplications)
Hint 1. Pass 1 picks column 1. The $(1,1)$ entry is 0, so the first operation is a swap.

Hint 2. In pass 2 the pivot is in position $(2,2)$ and has to be scaled. Remember to clear above it as well as below it.
Sketch. The operations are E_{12}, $E_{1,-2,3}$, $E_{2,1/2}$, $E_{2,-1,1}$, $E_{2,-1,3}$, and $\operatorname{rref}(B)$ has rows $(1,0,-1)$, $(0,1,2)$, $(0,0,0)$. The pivot columns are 1 and 2.

Pass 1. Column 1 has nonzero entries, so $j=1$. The $(1,1)$ entry is 0 and row 2 is the topmost row with a nonzero entry in column 1, so we swap rows 1 and 2 (E_{12}). The new $(1,1)$ entry is 1, so no scaling is needed; $p_1=1$. Row 3 has entry 2 in column 1, so we add -2 times row 1 to row 3 ($E_{1,-2,3}$):

$$B \xrightarrow{E_{12}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 2 & 3 & 4 \end{pmatrix} \xrightarrow{E_{1,-2,3}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \end{pmatrix}.$$

Pass 2. Column 2 has the nonzero entries $2, 1$ in the unpivoted rows $2, 3$, so $j=2$, and the $(2,2)$ entry is nonzero, so there is no swap. Scaling row 2 by $1/2$ ($E_{2,1/2}$) gives the row $(0,1,2)$, and $p_2=2$. Rows 1 and 3 have entry 1 in column 2, so we add -1 times row 2 to each ($E_{2,-1,1}$, $E_{2,-1,3}$):

$$\xrightarrow{E_{2,1/2}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow{E_{2,-1,1}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow{E_{2,-1,3}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

End. The only unpivoted row is zero, so the next scan finds no column and the algorithm halts. The pivot columns are 1 and 2.

Solution to Exercise 3.3.15

Uses. Theorem 3.3.12 (Matrices in RREF)
Hint 1. Locate the leading entry of each row, then check (R1)–(R4) one at a time.
Hint 2. For (b), sort by the number of nonzero rows, and then by the position of the leading entries.
Sketch. (a) In RREF: the second and fourth. The first fails (R3), the third fails (R1), the fifth fails (R4). (b) 0, $\left(\begin{smallmatrix} 1 & a \\ 0 & 0 \end{smallmatrix}\right)$ for $a \in F$, $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, and I_2; over $\mathbb{F}_p$ there are $p+3$.

(a) By Theorem 3.3.12 we check (R1)–(R4).

- The first matrix fails (R3): the leading entry of row 3 (column 2) lies to the left of that of row 2 (column 3).
- The second is in RREF: the leading entries are 1s in columns 1 and 3, and those columns are $\mathbf{e}_1$ and $\mathbf{e}_2$.
- The third fails (R1): the zero row 2 lies above the nonzero row 3.
- The fourth is in RREF: leading 1s in columns 1 and 2, which are $\mathbf{e}_1$ and $\mathbf{e}_2$, and the zero row is last.
- The fifth fails (R4): column 3 contains the leading entry of row 2 and also a nonzero entry in row 3. (It also fails (R3).)

(b) Let $R \in F^{2\times 2}$ be in RREF. If R has no nonzero row, $R=0$. If it has one, it is row 1 by (R1). Its leading 1 is in column 1 or column 2: in the first case $R = \left(\begin{smallmatrix} 1 & a \\ 0 & 0 \end{smallmatrix}\right)$ with $a \in F$ arbitrary (column 2 contains no leading entry, so (R4) says nothing about it), and in the second case $R = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. If both rows are nonzero, their leading 1s are in columns 1 and 2 by (R3), and by (R4) $R = I_2$. All these matrices satisfy (R1)–(R4). Over $\mathbb{F}_p$ there are $1 + p + 1 + 1 = p + 3$ of them.

Solution to Exercise 3.3.21

Uses. Corollary 3.3.18 (RREF and row equivalence), Algorithm 3.3.1 (Gauss–Jordan elimination)
Hint 1. By Corollary 3.3.18(b), compare reduced row echelon forms.
Hint 2. C has a row that is a multiple of the other, so it has one pivot.
Sketch. $\operatorname{rref}(A) = \operatorname{rref}(B) = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ and $\operatorname{rref}(C) = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{smallmatrix}\right)$, so $A \sim B$, and C is row equivalent to neither.

For A: adding -2 times row 1 to row 2 gives $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, and adding -3 times row 2 to row 1 gives $\left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. For B: swapping the rows gives $\left(\begin{smallmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, and adding -2 times row 2 to row 1 gives $\left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. For C: adding -3 times row 1 to row 2 gives $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{smallmatrix}\right)$. Each result satisfies (R1)–(R4), so by Corollary 3.3.18(c) these are the reduced row echelon forms. By Corollary 3.3.18(b), $A \sim B$, while C is row equivalent to neither.

Solution to Exercise 3.3.27

Uses. Algorithm 3.3.24 (Inverting a matrix by row reduction), Theorem 3.3.23 (Invertibility and RREF)
Hint 1. The two computations agree until some pivot has to be divided by. Which number is it?
Hint 2. Over $\mathbb{Q}$, pass 3 divides row 3 by 3. In $\mathbb{F}_3$, $3 = 0$.
Sketch. After two passes, valid over any field, the third row of $(A \mid I_3)$ is $(0, 0, 3 \mid -1, 2, 1)$. Over $\mathbb{Q}$, 3 is a pivot, and $3A^{-1}$ has rows $(1, -2, 2)$, $(1, 1, -1)$, $(-1, 2, 1)$. Over $\mathbb{F}_3$ the third row of the left block is zero, so A is not invertible.

Pass 1 adds -1 times row 1 to row 3; pass 2 adds -2 times row 2 to row 1 and 2 times row 2 to row 3. This is the same over any field:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & -2 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & -1 & 2 & 1 \end{array}\right).$$

Over $\mathbb{Q}$. Pass 3 scales row 3 by $1/3$ and then adds 2 times row 3 to row 1 and -1 times row 3 to row 2:

$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & -2/3 & 2/3 \\ 0 & 1 & 0 & 1/3 & 1/3 & -1/3 \\ 0 & 0 & 1 & -1/3 & 2/3 & 1/3 \end{array}\right).$$

So A is invertible over $\mathbb{Q}$ and

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 2 & 1 \end{pmatrix}.$$

Check: the first row of A times the columns of $3A^{-1}$ gives $(1 + 2, -2 + 2, 2 - 2) = (3, 0, 0)$, and similarly for the others.

Over $\mathbb{F}_3$. Here $3 = 0$, so the last row of the left block above is zero, and the algorithm stops: A is not invertible over $\mathbb{F}_3$. (Indeed, in the computation above, the last row of the right block gives -1 times row 1 of A plus 2 times row 2 plus row 3, which is $(-1 + 0 + 1, -2 + 2 + 0, 0 + 2 + 1) = (0, 0, 3)$, and $3 = 0$ in $\mathbb{F}_3$.)

Solution to Exercise 3.3.39

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Proposition 3.3.11 (Carrying extra columns along), Corollary 3.3.19 (Pivot columns, intrinsically)
Hint 1. No swaps are needed. Carry I_3 along to get G.
Hint 2. For (b), the entries of a non-pivot column of $\operatorname{rref}(A)$ are the coefficients (Corollary 3.3.19).
Sketch. $\operatorname{rref}(A)$ has rows $(1, 2, 0, 0, 5)$, $(0, 0, 1, 0, -4)$, $(0, 0, 0, 1, 2)$, so the pivot columns are $1, 3, 4$, and G has rows $(4, -2, 1)$, $(-3, 2, -1)$, $(1, -1, 1)$. Then $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_5 = 5\mathbf{a}_1 - 4\mathbf{a}_3 + 2\mathbf{a}_4$.

(a) We row reduce $(A \mid I_3)$, scanning only the first five columns (Proposition 3.3.11). Pass 1: the pivot is the $(1, 1)$ entry; add -2 times row 1 to row 2 and -1 times row 1 to row 3. Pass 2: column 2 now vanishes in rows 2 and 3, so the pivot is the $(2, 3)$ entry, which is 1; add -1 times row 2 to rows 1 and 3. Pass 3: the pivot is the $(3, 4)$ entry, which is 1; add row 3 to row 1 and -1 times row 3 to row 2. The stages after each pass are

$$\left(\begin{array}{ccccc|ccc} 1 & 2 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & -1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccccc|ccc} 1 & 2 & 0 & -1 & 3 & 3 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & -1 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|ccc} 1 & 2 & 0 & 0 & 5 & 4 & -2 & 1 \\ 0 & 0 & 1 & 0 & -4 & -3 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 & 1 & -1 & 1 \end{array}\right).$$

So

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 2 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}, \quad G = \begin{pmatrix} 4 & -2 & 1 \\ -3 & 2 & -1 \\ 1 & -1 & 1 \end{pmatrix},$$

with pivot columns $1, 3, 4$. As a check, row 1 of GA is

$$4(1,2,1,0,1) - 2(2,4,3,1,0) + (1,2,2,2,1) = (1,2,0,0,5).$$

(b) By Corollary 3.3.19(b), $\mathbf{a}_2 = 2\mathbf{a}_1$ and $\mathbf{a}_5 = 5\mathbf{a}_1 - 4\mathbf{a}_3 + 2\mathbf{a}_4$. Indeed,

$$5(1,2,1) - 4(1,3,2) + 2(0,1,2) = (1,0,1) = \mathbf{a}_5.$$

Solution to Exercise 3.3.40

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Corollary 3.3.19 (Pivot columns, intrinsically)
Hint 1. The first two passes involve only the scalars 1 and -1, so they read the same over both fields (remember $-1 = 1$ in $\mathbb{F}_2$).
Hint 2. After two passes the third row is $(0,0,2,0)$. Can 2 be a pivot?
Sketch. After two passes both computations reach the rows $(1,0,-1,0)$, $(0,1,1,1)$, $(0,0,2,0)$. Over $\mathbb{Q}$, 2 is a pivot and $\operatorname{rref}(A)$ has rows $(1,0,0,0)$, $(0,1,0,1)$, $(0,0,1,0)$. Over $\mathbb{F}_2$, $2 = 0$ and $-1 = 1$, so $\operatorname{rref}(A)$ has rows $(1,0,1,0)$, $(0,1,1,1)$, $(0,0,0,0)$.

Pass 1 adds -1 times row 1 to row 3, which becomes $(0,-1,1,-1)$. Pass 2, with pivot $(2,2)$, adds -1 times row 2 to row 1 and row 2 to row 3:

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 \end{pmatrix}.$$

So far every operation makes sense over any field.

Over $\mathbb{Q}$. The third row has the entry $2 \neq 0$ in column 3, which becomes the pivot of pass 3. Scaling row 3 by $1/2$ and adding row 3 to row 1 and -1 times row 3 to row 2 gives

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Over $\mathbb{F}_2$. Here $2 = 0$ and $-1 = 1$, so after pass 2 the matrix is

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Its unpivoted row is zero, so the algorithm halts, and this is $\operatorname{rref}(A)$ over $\mathbb{F}_2$.

The computations part ways at the entry 2 in pass 3, which is a valid pivot over $\mathbb{Q}$ and is 0 over $\mathbb{F}_2$. In terms of Corollary 3.3.19: over $\mathbb{F}_2$ the third column of A is the sum of the first two, $(1,0,1) + (1,1,0) = (0,1,1)$, so column 3 is not a pivot column. Over $\mathbb{Q}$, a combination $x(1,0,1) + y(1,1,0)$ equal to $(0,1,1)$ would need $x = 1$ and $y = 1$ (from the third and second entries), but then the first entry is $2 \neq 0$; so the third column is a pivot column.

Solution to Exercise 3.3.41

Uses. Algorithm 3.3.24 (Inverting a matrix by row reduction), Theorem 3.3.23 (Invertibility and RREF)

Hint 1. In pass 1 for A, add $-i$ times row 1 to row 3, and remember $i^2 = -1$.
Hint 2. Pass 3 for A divides by $1 - i$, whose inverse is $(1 + i)/2$. For B, the left block gets a zero row in pass 2; read the relation off the right block.
Sketch. For A, the passes use $E_{1,-i,3}$, then $E_{2,-i,1}$ and $E_{2,-1,3}$, then a scaling by $(1 + i)/2$ and two replacements; the inverse has entries $(\pm 1 \pm i)/2$. For B, two passes leave the left row $(0, 0, 0)$ with right row $(-i, -1, 1)$, so B is not invertible, and $-i$ times row 1, minus row 2, plus row 3 is zero.

The matrix A. Pass 1 adds $-i$ times row 1 to row 3, which becomes $\big(0, -i^2, 1 \mid -i, 0, 1\big) = (0, 1, 1 \mid -i, 0, 1)$. Pass 2 (pivot $(2, 2)$) adds $-i$ times row 2 to row 1 and -1 times row 2 to row 3:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -i & 0 \\ 0 & 1 & i & 0 & 1 & 0 \\ 0 & 0 & 1-i & -i & -1 & 1 \end{array}\right).$$

Pass 3 scales row 3 by $(1 - i)^{-1} = (1 + i)/2$, which turns the right part of row 3 into $\frac{1}{2}(1 - i, -1 - i, 1 + i)$, and then adds -1 times row 3 to row 1 and $-i$ times row 3 to row 2. The result is

$$A^{-1} = \frac{1}{2}\begin{pmatrix} 1+i & 1-i & -1-i \\ -1-i & 1+i & 1-i \\ 1-i & -1-i & 1+i \end{pmatrix}.$$

Check: the first row of A times the first column of $2A^{-1}$ is $(1 + i) + i(-1 - i) = 1 + i - i + 1 = 2$, times the second column it is $(1 - i) + i(1 + i) = 0$, and times the third $(-1 - i) + i(1 - i) = 0$; the other rows are similar.

The matrix B. Pass 1 adds $-i$ times row 1 to row 3: the left part becomes $(i, 1, 0) - i(1, 0, i) = (0, 1, 1)$ and the right part $(-i, 0, 1)$. Pass 2 adds -1 times row 2 to row 3, which becomes $(0, 0, 0 \mid -i, -1, 1)$. The left block now has a zero row, so B is not invertible. The right part of the zero row says that $-i$ times row 1 of B, minus row 2, plus row 3, is zero:

$$-i(1, 0, i) - (0, 1, 1) + (i, 1, 0) = (0, 0, 1 - 1) = \mathbf{0}.$$

Solution to Exercise 3.3.42

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Theorem 3.3.23 (Invertibility and RREF)
Hint 1. Pass 1 works for every c. Write down the resulting rows 2 and 3.
Hint 2. They are $(0, c - 4, 0)$ and $(0, 0, c - 1)$. Split into cases according to which of $c - 4$ and $c - 1$ vanish. In $\mathbb{F}_3$, $4 = 1$.
Sketch. After pass 1 the rows are $(1, 2, 1)$, $(0, c - 4, 0)$, $(0, 0, c - 1)$. Over $\mathbb{R}$: $\operatorname{rref}(A_c) = I_3$ for $c \neq 1, 4$; $\operatorname{rref}(A_4)$ has rows $(1, 2, 0)$, $(0, 0, 1)$, $\mathbf{0}$; $\operatorname{rref}(A_1)$ has rows $(1, 0, 1)$, $(0, 1, 0)$, $\mathbf{0}$. Over $\mathbb{F}_3$, $c - 4 = c - 1$: $\operatorname{rref}(A_1)$ has rows $(1, 2, 1)$, $\mathbf{0}$, $\mathbf{0}$, and $\operatorname{rref}(A_c) = I_3$ for $c = 0, 2$.

Pass 1 adds -2 times row 1 to row 2 and -1 times row 1 to row 3, giving

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & c-4 & 0 \\ 0 & 0 & c-1 \end{pmatrix}.$$

(a) *If $c \neq 4$ and $c \neq 1$:* pass 2 scales row 2 by $(c - 4)^{-1}$, making it $(0, 1, 0)$, and adds -2 times it to row 1, which becomes $(1, 0, 1)$. Pass 3 scales row 3 to $(0, 0, 1)$ and subtracts it from row 1. So $\operatorname{rref}(A_c) = I_3$.

If $c = 4$: the rows are $(1, 2, 1)$, $(0, 0, 0)$, $(0, 0, 3)$. In pass 2, column 2 vanishes in rows 2 and 3, so the scan moves to column 3; the $(2, 3)$ entry is 0, so rows 2 and 3 are swapped. Scaling by $1/3$ and subtracting the new row 2 from row 1 gives

$$\operatorname{rref}(A_4) = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

If $c = 1$: the rows are $(1, 2, 1)$, $(0, -3, 0)$, $(0, 0, 0)$. Pass 2 scales row 2 to $(0, 1, 0)$ and adds -2 times it to row 1, so

$$\operatorname{rref}(A_1) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

By Theorem 3.3.23, A_c is invertible exactly when $c \neq 1$ and $c \neq 4$.

(b) In $\mathbb{F}_3$, $c - 4 = c - 1$, since $3 = 0$. If $c = 1$, rows 2 and 3 are both zero after pass 1, and

$$\operatorname{rref}(A_1) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If $c = 0$ or $c = 2$, then $c - 4 = c - 1 \neq 0$, and the computation of the first case of (a) applies verbatim: $\operatorname{rref}(A_c) = I_3$. So over $\mathbb{F}_3$, A_c is invertible exactly when $c \neq 1$.

Solution to Exercise 3.3.43

Uses. Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.12 (Matrices in RREF), Proposition 3.3.11 (Carrying extra columns along)
Hint 1. In each part, find a matrix in RREF that is row equivalent to the given one, and use Corollary 3.3.18(c).
Hint 2. For (b), $G(A \mid A) = (R \mid R)$. Check (R1)–(R4) for $(R \mid R)$: where are its leading entries?
Hint 3. For (c), first subtract each of the top rows from the corresponding bottom row.
Sketch. (a) Scaling every row by c turns A into cA, so $cA \sim A$. (b) $(A \mid A) \sim (R \mid R)$, whose nonzero rows have their leading entries in the left block, and whose pivot columns are those of R; so it satisfies (R1)–(R4). (c) Subtracting rows gives $\begin{pmatrix} A \\ 0 \end{pmatrix}$, and the run of the algorithm on A, performed on the top m rows, gives $\begin{pmatrix} R \\ 0 \end{pmatrix}$, which satisfies (R1)–(R4).

In each part we exhibit a matrix that satisfies (R1)–(R4) and is row equivalent to the given matrix; by Corollary 3.3.18(c) it is the reduced row echelon form.

(a) Scaling each of the m rows of A by c turns A into cA, so $cA \sim A \sim R$, and R is in RREF. So $\operatorname{rref}(cA) = R$.

(b) By Proposition 3.3.11(a), $G(A \mid A) = (GA \mid GA) = (R \mid R)$, so $(A \mid A) \sim (R \mid R)$. The nonzero rows of $(R \mid R)$ are the nonzero rows of R, in the same positions, and since the left block is R, their leading entries are the leading entries of R. So (R1), (R2), and (R3) hold for $(R \mid R)$ because they hold for R. The columns containing leading entries are the pivot columns of R, in the left block, and each of them is a standard basis vector; this is (R4). So $\operatorname{rref}((A \mid A)) = (R \mid R)$.

(c) For $i = 1, \ldots, m$, add -1 times row i to row $m + i$. This turns $\begin{pmatrix} A \\ A \end{pmatrix}$ into $\begin{pmatrix} A \\ 0 \end{pmatrix}$. Then perform the row operations of the run of Gauss–Jordan elimination on A, applied to rows $1, \ldots, m$ only; they turn $\begin{pmatrix} A \\ 0 \end{pmatrix}$ into $\begin{pmatrix} R \\ 0 \end{pmatrix}$. This matrix satisfies (R1)–(R4): its nonzero rows are those of R, at the top, and it is R with m zero rows appended. So $\operatorname{rref}\begin{pmatrix} A \\ A \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix}$.

Solution to Exercise 3.3.44

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.11 (Carrying extra columns along), Theorem 3.3.12 (Matrices in RREF), Corollary 3.3.18 (RREF and row equivalence)
Hint 1. Write $\operatorname{rref}((A \mid B)) = (S \mid T) = H(A \mid B)$ with H invertible. Then $S = HA$. It suffices to show that S is in RREF. Why?
Hint 2. Some nonzero rows of $(S \mid T)$ may have their leading entry in the right block. What do these rows look like in S, and where are they?
Hint 3. For (b), try $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
Sketch. $(S \mid T) = (HA \mid HB)$. The rows of $(S \mid T)$ whose leading entries lie in the left block come first, by (R3), and the remaining rows vanish in the left block; so S satisfies (R1)–(R4), with pivot columns the pivot columns of $(S \mid T)$ among $1, \ldots, n$. Then $S = \operatorname{rref}(HA) = \operatorname{rref}(A)$. For $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$: $\operatorname{rref}((A \mid B)) = I_2$, while $GB = (1, 1)$.

(a) By Theorem 3.3.10 and Proposition 3.3.11(a), $\operatorname{rref}((A \mid B)) = H(A \mid B) = (HA \mid HB)$ for some invertible H. Write $S = HA$ and $T = HB$. We show that S is in RREF; then $S = \operatorname{rref}(S) = \operatorname{rref}(HA) = \operatorname{rref}(A)$ by Corollary 3.3.18(a).

Let rows $1, \dots, r$ be the nonzero rows of $(S \mid T)$, with leading entries in columns $c_1 < \cdots < c_r$, and let t be the number of c_s that are at most n, so that $c_1, \dots, c_t \le n < c_{t+1}, \dots, c_r$. For $s \le t$, row s of S is nonzero, with the same leading entry as row s of $(S \mid T)$. For $s > t$, the leading entry of row s of $(S \mid T)$ is in the right block, so row s of S is zero; and rows $r+1, \dots, m$ of S are zero as well. So the nonzero rows of S are rows $1, \dots, t$, which is (R1), and (R2) and (R3) for S follow from (R2) and (R3) for $(S \mid T)$. Finally, for $s \le t$, column c_s of $(S \mid T)$ is $\mathbf{e}_s$ by (R2) and (R4), and it is a column of S. This is (R4) for S. By Theorem 3.3.12, S is in RREF.

(b) Let $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. The run on A is the single operation $E_{1,-1,2}$, so $G = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ and $GB = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. But $(A \mid B) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ reduces to $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and then to I_2, whose last column is $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \neq GB$. The full algorithm also pivots in the column of B, and then clears above that pivot.

Solution to Exercise 3.3.45

Uses. Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 3.3.12 (Matrices in RREF)
Hint 1. Only one of the statements is true. Look for 2×2 counterexamples, or even 1×2 ones, with many zeros.
Hint 2. For (e), write $\operatorname{rref}(A) = GA$ with G invertible.
Sketch. (a) False for $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. (b) False for $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. (c) False for $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. (d) False for $\begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \end{pmatrix}$. (e) True: $\operatorname{rref}(A)B = G(AB)$, so Corollary 3.3.18(a) applies.

(a) False. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. It satisfies (R1)–(R4), so $\operatorname{rref}(A) = A$ and $\operatorname{rref}(A)^{\mathsf{T}} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, which violates (R1) and so is not even in RREF. In fact $\operatorname{rref}\left(A^{\mathsf{T}}\right) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(b) False. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Then $\operatorname{rref}(A) = A$, $\operatorname{rref}(B) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, and $\operatorname{rref}(A)\operatorname{rref}(B) = 0$. But $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, which is its own reduced row echelon form.

(c) False. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = E_{12}A$. Then $\operatorname{rref}(A) = \operatorname{rref}(B) = A$, but $\operatorname{im} T_A = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} T_B = \operatorname{span}(\mathbf{e}_2)$.

(d) False. Let $A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \end{pmatrix}$ in $F^{1\times 2}$. Both images are all of F, but both matrices are in RREF and they are different.

(e) True. By Theorem 3.3.10, $\operatorname{rref}(A) = GA$ with G invertible, so $\operatorname{rref}(A)B = G(AB)$, and Corollary 3.3.18(a) gives $\operatorname{rref}(G(AB)) = \operatorname{rref}(AB)$.

Solution to Exercise 3.3.46

Uses. Theorem 3.3.17 (Uniqueness of the reduced row echelon form), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.4.33 (Block matrix), Definition 2.5.12 (Invertible)
Hint 1. Reread the statement proved by induction in the first proof of Theorem 3.3.17. For which s does it say anything about $\mathbf{g}_s$?
Hint 2. Try $R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
Hint 3. For (b), show that $GR = R$ exactly when $G\mathbf{e}_s = \mathbf{e}_s$ for $s \le r$, and write G in block form.
Sketch. The induction gives $\mathbf{g}_s = \mathbf{e}_s$ only when $p_s \le n$, that is, for $s \le r$. Counterexample: $R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. In general $GR = R$ if and only if $G = \begin{pmatrix} I_r & X \\ 0 & Y \end{pmatrix}$, and such a G is invertible exactly when Y is.

(a) The induction in the first proof of Theorem 3.3.17 shows $p_s = q_s$ for all s, but $\mathbf{g}_s = \mathbf{e}_s$ only when $p_s \le n$, that is, for $s \le r$. It says nothing about the columns $\mathbf{g}_{r+1}, \dots, \mathbf{g}_m$, so the claim does not follow. It is false: for $R = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, G is invertible (with inverse $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$) and $GR = R$, but $G \neq I_2$.

(b) Let $p_1 < \cdots < p_r$ be the pivot columns of R. If $GR = R$, then $G\mathbf{e}_s = G\mathbf{r}_{p_s} = \mathbf{r}_{p_s} = \mathbf{e}_s$ for $s \le r$ (Lemma 3.3.16(b)). Conversely, if $G\mathbf{e}_s = \mathbf{e}_s$ for $s \le r$, then by Lemma 3.3.16(a) (with $s = r$) every column of

R is a combination $\sum_{t\le r} c_t \mathbf{e}_t$, and G fixes it; so $GR = R$. Hence $GR = R$ if and only if the first r columns of G are $\mathbf{e}_1, \ldots, \mathbf{e}_r$, that is,

$$G = \begin{pmatrix} I_r & X \\ 0 & Y \end{pmatrix}$$

for some $X \in F^{r\times(m-r)}$ and $Y \in F^{(m-r)\times(m-r)}$.

Such a G is invertible if and only if Y is. If Y is invertible, block multiplication (Definition 2.4.33) shows that $\begin{pmatrix} I_r & -XY^{-1} \\ 0 & Y^{-1} \end{pmatrix}$ is a two-sided inverse of G. Conversely, let $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$ be an inverse of G, in the same block sizes. The first block column of HG is $\begin{pmatrix} H_{11} \\ H_{21} \end{pmatrix}$, which must be $\begin{pmatrix} I_r \\ 0 \end{pmatrix}$, so $H_{21} = 0$. Then the lower right blocks of $HG = I_m$ and $GH = I_m$ say $H_{22}Y = I_{m-r}$ and $YH_{22} = I_{m-r}$. So Y is invertible.

So the invertible G with $GR = R$ are exactly the matrices $\begin{pmatrix} I_r & X \\ 0 & Y \end{pmatrix}$ with X arbitrary and Y invertible. Only when $r = m$ is $G = I_m$ forced.

Solution to Exercise 3.3.47

Uses. Lemma 3.3.7 (The loop invariant), Proposition 3.3.8 (Properties of the output)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Find the place in the proof of Lemma 3.3.7 where the choice of the *leftmost* column was used.
Hint 2. Run the modified algorithm on $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, choosing column 2 in the first pass.
Sketch. The proof used that rows $k+1, \ldots, m$ vanish to the left of the chosen column, which is what the leftmost choice guarantees; parts (b), (d), (e) of the invariant depend on it. Choosing column 2 first in $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ gives the output $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which violates (R3).

The proof of Lemma 3.3.7 began with the observation that rows $k+1, \ldots, m$ vanish in the columns before the chosen column j. That is exactly what choosing the *leftmost* column guarantees, and parts (b), (d), and (e) were derived from it. With an arbitrary choice, none of them survives.

Counterexample: run the modified algorithm on $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, choosing column 2 in pass 1. The $(1,2)$ entry is already 1, so it becomes a pivot with $p_1 = 2$, and clearing column 2 adds -1 times row 1 to row 2, giving $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$. In pass 2, the only unpivoted row is $(-1, 0)$, so $j = 1$; scaling by -1 gives $(1, 0)$, and clearing adds -1 times row 2 to row 1. The output is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which violates (R3), so it is not in RREF. (Here $p_1 = 2 > 1 = p_2$, so (b) of the invariant fails.)

What does survive is (a) and (c): the output is still row equivalent to A, and each pivot column is a standard basis vector, because a column $\mathbf{e}_s$ has zeros in all unpivoted rows whichever column is chosen next. Only the staircase is lost.

Solution to Exercise 3.3.48

Uses. Theorem 3.3.12 (Matrices in RREF), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Corollary 3.3.18 (RREF and row equivalence)
Hint 1. Sort the matrices in RREF by their set of pivot columns. Which entries are then forced, and which are free?
Hint 2. By Lemma 3.3.16(a), a non-pivot column has zeros below the rows of the pivots to its left, and is arbitrary in those rows.
Hint 3. For (b), how many matrices in RREF does a row equivalence class contain?
Sketch. No pivots: 1 matrix. One pivot in column 1, 2, 3: p^2, p, 1 matrices. Two pivots in columns $(1,2)$, $(1,3)$, $(2,3)$: p^2, p, 1. Total $1 + 2(p^2 + p + 1)$. By Corollary 3.3.18(b), each class contains exactly one matrix in RREF; for $p = 2$ there are 15 classes.

(a) By Theorem 3.3.12 we count the matrices satisfying (R1)–(R4), sorted by their pivot columns. The pivot columns are $\mathbf{e}_1, \mathbf{e}_2, \ldots$ in order, and by Lemma 3.3.16(a) a non-pivot column has zeros in every row below the rows of the pivots to its left, while its entries in those rows are arbitrary.

- No pivots: only the zero matrix. That is 1 matrix.
- One pivot, in column 1: $\begin{pmatrix} 1 & a & b \\ 0 & 0 & 0 \end{pmatrix}$, so p^2 matrices. In column 2: $\begin{pmatrix} 0 & 1 & a \\ 0 & 0 & 0 \end{pmatrix}$, so p. In column 3: $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, so 1.
- Two pivots, in columns 1, 2: $\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix}$, so p^2. In columns 1, 3: $\begin{pmatrix} 1 & a & 0 \\ 0 & 0 & 1 \end{pmatrix}$, so p. In columns 2, 3: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, so 1.

Each displayed matrix satisfies (R1)–(R4) for every choice of a, b. The total is $1 + 2\big(p^2 + p + 1\big) = 2p^2 + 2p + 3$.

(b) Every class contains a matrix in RREF, namely $\operatorname{rref}(A)$ for any A in it, and by Corollary 3.3.18(b) two matrices in RREF are row equivalent only if they are equal (as $\operatorname{rref}(R) = R$). So $R \mapsto$ (the class of R) is a bijection from the matrices in RREF to the classes, and there are $2p^2 + 2p + 3$ classes. For $p = 2$ this is 15.

Solution to Exercise 3.3.49

> *Uses.* Corollary 3.3.20 (The kernel is a complete invariant), Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$)
> *Techniques.* Reduce to RREF (Technique 3.3.22)
> *Hint 1.* Find one matrix in RREF with this kernel.
> *Hint 2.* $R_0 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$ works. By Corollary 3.3.20, A has the same kernel as R_0 exactly when $\operatorname{rref}(A) = R_0$. Which matrices have this reduced row echelon form?
> *Sketch.* The matrix R_0 is in RREF and has kernel $\operatorname{span}((1,1,1))$. By Corollary 3.3.20, $\ker T_A = \ker T_{R_0}$ if and only if $\operatorname{rref}(A) = R_0$, if and only if $A = HR_0$ with H invertible. So $A = \begin{pmatrix} a & b & -a-b \\ c & d & -c-d \end{pmatrix}$ with $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertible.

Let $R_0 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$, which is in RREF. Its kernel consists of the $\mathbf{x}$ with $x_1 - x_3 = 0$ and $x_2 - x_3 = 0$, that is, $\mathbf{x} = x_3(1,1,1)$; so $\ker T_{R_0} = \operatorname{span}((1,1,1))$. By Corollary 3.3.20, $\ker T_A = \operatorname{span}((1,1,1))$ if and only if $\operatorname{rref}(A) = \operatorname{rref}(R_0) = R_0$.

If $\operatorname{rref}(A) = R_0$, then by Theorem 3.3.10, $GA = R_0$ with G invertible, so $A = G^{-1}R_0$. Conversely, if $A = HR_0$ with H invertible, then $\operatorname{rref}(A) = \operatorname{rref}(R_0) = R_0$ by Corollary 3.3.18(a). So the answer is the set of matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} a & b & -a-b \\ c & d & -c-d \end{pmatrix}$$

with $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertible.

Solution to Exercise 3.3.50

> *Uses.* Definition 2.3.8 (Matrix–vector product), Corollary 3.3.20 (The kernel is a complete invariant), Exercise 3.2.46, Definition 1.3.10 (Span)
> *Techniques.* Reduce to RREF (Technique 3.3.22)
> *Hint 1.* The ith entry of $A\mathbf{x}$ is (row i of A) $\cdot\, \mathbf{x}$. How does $\mathbf{v} \cdot \mathbf{x}$ behave when $\mathbf{v}$ is a linear combination?
> *Hint 2.* For (b), one direction is Exercise 3.2.46. For the other, use (a) and Corollary 3.3.20.
> *Sketch.* (a) $A\mathbf{x} = \mathbf{0}$ says that every row is orthogonal to $\mathbf{x}$ in this sense, and $\mathbf{v} \mapsto \mathbf{v} \cdot \mathbf{x}$ respects linear combinations. (b) Row equivalent matrices have the same row space by Exercise 3.2.46. If the row spaces agree, (a) shows that the kernels agree, and Corollary 3.3.20 gives $A \sim B$.

(a) Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be the rows of A. By Definition 2.3.8, the ith entry of $A\mathbf{x}$ is $a_{i1}x_1 + \cdots + a_{in}x_n = \mathbf{v}_i \cdot \mathbf{x}$, so $A\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{v}_i \cdot \mathbf{x} = 0$ for all i. By distributivity, $(c\mathbf{u} + c'\mathbf{u}') \cdot \mathbf{x} = c(\mathbf{u} \cdot \mathbf{x}) + c'(\mathbf{u}' \cdot \mathbf{x})$, so if $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$ and every $\mathbf{v}_i \cdot \mathbf{x} = 0$, then $\mathbf{v} \cdot \mathbf{x} = 0$. Conversely, each $\mathbf{v}_i$ lies in the row space. This proves (a).

(b) If $A \sim B$, they have the same row space by Exercise 3.2.46. Conversely, suppose A and B have the same row space. By (a), $\ker T_A$ and $\ker T_B$ are both the set of $\mathbf{x}$ with $\mathbf{v} \cdot \mathbf{x} = 0$ for all $\mathbf{v}$ in this common row space. So $\ker T_A = \ker T_B$, and $A \sim B$ by Corollary 3.3.20.

Solution to Exercise 3.3.51

Uses. Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.23 (Invertibility and RREF), Proposition 3.3.11 (Carrying extra columns along)
Hint 1. Work from the bottom up: for $i = n-1, n-2, \ldots, 1$, add $-a$ times row $i+1$ to row i. Any sequence of row operations that reaches $(I_n \mid C)$ will do, by Corollary 3.3.18(c).
Hint 2. Prove by downward induction on i what row i looks like after it has been processed.
Hint 3. For (b), U is the answer to (a) for $a = -1$.
Sketch. Processing rows from the bottom up, row i becomes $\left(\mathbf{e}_i^{\mathsf{T}} \mid \mathbf{c}_i\right)$ with $\mathbf{c}_i = \sum_{j \geq i} (-a)^{j-i} \mathbf{e}_j^{\mathsf{T}}$, because $\mathbf{e}_i^{\mathsf{T}} - a\mathbf{c}_{i+1} = \mathbf{c}_i$. So $H(B \mid I_n) = (I_n \mid C)$ for a product H of elementary matrices, and $B^{-1} = C$. For $a = -1$, (a) says that the matrix with 1 on the diagonal and -1 just above it has inverse U; so U^{-1} is that matrix.

(a) For $i = 1, \ldots, n$, let

$$\mathbf{c}_i = \sum_{j=i}^{n} (-a)^{j-i} \mathbf{e}_j^{\mathsf{T}},$$

the row vector we expect as row i of B^{-1}. For $i < n$, row i of $(B \mid I_n)$ has left part $\mathbf{e}_i^{\mathsf{T}} + a\mathbf{e}_{i+1}^{\mathsf{T}}$ and right part $\mathbf{e}_i^{\mathsf{T}}$. Row n has left part $\mathbf{e}_n^{\mathsf{T}}$ and right part $\mathbf{e}_n^{\mathsf{T}} = \mathbf{c}_n$.

For $i = n-1, n-2, \ldots, 1$ in this order, add $-a$ times row $i+1$ to row i. We claim that afterwards row i is $\left(\mathbf{e}_i^{\mathsf{T}} \mid \mathbf{c}_i\right)$. When row i is processed it has not been changed yet, and row $i+1$ is already $\left(\mathbf{e}_{i+1}^{\mathsf{T}} \mid \mathbf{c}_{i+1}\right)$ (for $i+1 = n$ this was noted above, and otherwise by induction). So row i becomes

$$\left(\mathbf{e}_i^{\mathsf{T}} + a\mathbf{e}_{i+1}^{\mathsf{T}} - a\mathbf{e}_{i+1}^{\mathsf{T}} \mid \mathbf{e}_i^{\mathsf{T}} - a\mathbf{c}_{i+1}\right) = \left(\mathbf{e}_i^{\mathsf{T}} \mid \mathbf{c}_i\right),$$

since

$$\mathbf{e}_i^{\mathsf{T}} - a \sum_{j \geq i+1} (-a)^{j-i-1} \mathbf{e}_j^{\mathsf{T}} = \mathbf{e}_i^{\mathsf{T}} + \sum_{j \geq i+1} (-a)^{j-i} \mathbf{e}_j^{\mathsf{T}} = \mathbf{c}_i.$$

At the end the matrix is $(I_n \mid C)$, where C has rows $\mathbf{c}_1, \ldots, \mathbf{c}_n$. Let H be the product of the elementary matrices used, so $H(B \mid I_n) = (HB \mid H) = (I_n \mid C)$. Since I_n is in RREF and $B \sim I_n$, Corollary 3.3.18(c) gives $\operatorname{rref}(B) = I_n$, and H is a product of elementary matrices with $HB = \operatorname{rref}(B)$. By Theorem 3.3.23, B is invertible and $B^{-1} = H = C$, whose (i,j) entry is $(-a)^{j-i}$ for $j \geq i$ and 0 otherwise.

(b) Take $a = -1$ in (a): the matrix B with 1 on the diagonal and -1 directly above it has inverse with entries $1^{j-i} = 1$ for $j \geq i$ and 0 below the diagonal, which is U. So $U = B^{-1}$ is invertible and $U^{-1} = B$: it has 1 on the diagonal, -1 in each position $(i, i+1)$, and 0 elsewhere.

Solution to Exercise 3.3.52

Uses. Definition 2.5.12 (Invertible), Definition 2.4.1 (Matrix product)
Hint 1. Row reduce for $n = 2$ and $n = 3$ over $\mathbb{Q}$ and look for a pattern.
Hint 2. Let J be the all-ones matrix, so $A = J - I_n$. Compute J^2, and look for an inverse of the form $\alpha J + \beta I_n$.
Hint 3. If $c = 0$, compute AJ.
Sketch. With J the all-ones matrix, $A = J - I_n$ and $J^2 = (n * 1)J$. If $c \neq 0$, then $c^{-1}J - I_n$ is a two-sided inverse of $J - I_n$, by direct multiplication. If $c = 0$, then $AJ = cJ = 0$ with $J \neq 0$, so A is not invertible. Over $\mathbb{F}_p$: invertible exactly when p does not divide $n-1$.

Let $J \in F^{n \times n}$ be the matrix all of whose entries are 1, so $A = J - I_n$. Each entry of J^2 is a sum of n products $1 \cdot 1$, so $J^2 = (n * 1)J = (c+1)J$.

Suppose $c \neq 0$, and let $C = c^{-1}J - I_n$. Then

$$AC = (J - I_n)\left(c^{-1}J - I_n\right) = c^{-1}J^2 - J - c^{-1}J + I_n = c^{-1}(c+1)J - J - c^{-1}J + I_n = I_n,$$

and the same computation, with the factors in the other order, gives $CA = I_n$ (all the matrices involved are combinations of J and I_n, which commute). So A is invertible and $A^{-1} = c^{-1}J - I_n$: its diagonal entries are $c^{-1} - 1$ and its other entries are c^{-1}. For instance, for $n = 3$ over $\mathbb{Q}$,

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Suppose $c = 0$. Then $AJ = J^2 - J = (c+1)J - J = cJ = 0$. If A were invertible, we would get $J = A^{-1}AJ = 0$, which is false. So A is not invertible. (This includes $n = 1$, where $A = (0)$.)

Over $\mathbb{F}_p$, $c = 0$ exactly when p divides $n - 1$. So A is invertible over $\mathbb{F}_p$ if and only if $n - 1$ is not a multiple of p; for example, over $\mathbb{F}_2$, if and only if n is even.

Solution to Exercise 3.3.53

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Theorem 3.3.23 (Invertibility and RREF), Definition 2.4.1 (Matrix product)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. For (a), show by induction on k that before pass k the current matrix is upper triangular, its first $k-1$ columns are $\mathbf{e}_1, \ldots, \mathbf{e}_{k-1}$, and rows $k, \ldots, n$ are still the rows of U.
Hint 2. In pass k the pivot is the (k,k) entry, and only rows above row k need clearing. Where is the off-diagonal entry of E_{kci} when $i < k$? Products of upper triangular matrices are upper triangular: check it entrywise.
Hint 3. For (b), take the first k with $u_{kk} = 0$, put $x_k = 1$ and $x_j = 0$ for $j > k$, and solve for $x_{k-1}, \ldots, x_1$ from the bottom up.
Sketch. (a) In pass k, rows $k, \ldots, n$ vanish in columns $1, \ldots, k-1$ and the (k,k) entry is $u_{kk} \neq 0$, so the pivot is (k,k) with no swap; the scaling is diagonal, and the replacements add multiples of row k to rows $i < k$, with elementary matrices E_{kci} whose extra entry (i,k) is above the diagonal. So $U^{-1} = G$ is a product of upper triangular matrices, and its diagonal is the product of their diagonals. (b) With k minimal and $u_{kk} = 0$, back substitution gives $\mathbf{x} \neq \mathbf{0}$ with $x_k = 1$ and $U\mathbf{x} = \mathbf{0}$.

(a) We show by induction on $k = 1, \ldots, n$ that before pass k, the current matrix M is upper triangular, its first $k-1$ columns are $\mathbf{e}_1, \ldots, \mathbf{e}_{k-1}$, and its rows $k, \ldots, n$ are those of U. For $k = 1$, $M = U$.

Assume this before pass k. Rows $k, \ldots, n$ of M vanish in columns $1, \ldots, k-1$ (as M is upper triangular), and the (k,k) entry is $u_{kk} \neq 0$. So step 1 of Algorithm 3.3.1 picks column k, and there is no swap. step 3 of Algorithm 3.3.1 multiplies by the diagonal matrix $E_{k,u_{kk}^{-1}}$. In column k, the rows $i > k$ have entry $u_{ik} = 0$, so step 4 of Algorithm 3.3.1 only changes rows $i < k$, using the matrices E_{kci} with $i < k$; the single off-diagonal entry of such a matrix is in position (i,k), above the diagonal. Adding a multiple of row k, which vanishes in columns $1, \ldots, k-1$, to a row $i < k$ changes it only in columns $k, \ldots, n$, all to the right of the diagonal entry (i,i). So after the pass, M is still upper triangular, column k is $\mathbf{e}_k$, columns $1, \ldots, k-1$ are unchanged, and rows $k+1, \ldots, n$ are untouched. This completes the induction.

After pass n the matrix is I_n, so $\operatorname{rref}(U) = I_n$, and by Theorem 3.3.23 U is invertible with $U^{-1} = G$, the product of the elementary matrices used. Each of them is upper triangular. If V and W are upper triangular, then $(VW)_{ij} = \sum_l v_{il} w_{lj}$ and a term can be nonzero only if $i \le l \le j$; so $(VW)_{ij} = 0$ for $i > j$, and $(VW)_{ii} = v_{ii} w_{ii}$. Hence G is upper triangular, and its diagonal is the entrywise product of the diagonals of its factors. The scaling in pass k has diagonal entry u_{kk}^{-1} in position k and 1 elsewhere, and the replacements have only 1s on the diagonal, so $G_{kk} = u_{kk}^{-1}$.

(b) Let k be the smallest index with $u_{kk} = 0$, so $u_{ii} \neq 0$ for $i < k$. Put $x_j = 0$ for $j > k$ and $x_k = 1$, and for $i = k-1, k-2, \ldots, 1$ in turn, put

$$x_i = -u_{ii}^{-1} \sum_{j=i+1}^{k} u_{ij} x_j.$$

Since U is upper triangular, entry i of $U\mathbf{x}$ is $\sum_{j \ge i} u_{ij} x_j$. For $i > k$ this involves only x_j with $j > k$, which are 0. For $i = k$ it is $u_{kk} \cdot 1 = 0$. For $i < k$ it is $u_{ii} x_i + \sum_{j=i+1}^{k} u_{ij} x_j = 0$ by the choice of x_i. So $U\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \neq \mathbf{0}$. If U were invertible, then $\mathbf{x} = U^{-1}U\mathbf{x} = \mathbf{0}$. So U is not invertible.

Together, (a) and (b) show that an upper triangular matrix is invertible if and only if all its diagonal entries are nonzero.

Solution to Exercise 3.3.54

Uses. Definition 1.1.32 (Subfield), Algorithm 3.3.1 (Gauss–Jordan elimination), Theorem 3.3.23 (Invertibility and RREF), Corollary 3.3.18 (RREF and row equivalence)
Hint 1. Compare the two runs step by step. Which field operations does the algorithm use, and do they give the same results in F and in K?
Hint 2. For (b), run Algorithm 3.3.24 over F.
Hint 3. For (c), one direction is free. For the other, use (a) and Corollary 3.3.18(b) twice.
Sketch. The algorithm only tests entries for being 0, inverts nonzero entries, and adds and multiplies; a subfield is closed under these operations and computes them as K does. So by induction the runs over F and K perform the same operations. (b) Over K, $\operatorname{rref}(A) = I_n$; so the run over F gives $(I_n \mid G)$ with G over F, and $G = A^{-1}$. (c) $\operatorname{rref}_F(A) = \operatorname{rref}_K(A) = \operatorname{rref}_K(B) = \operatorname{rref}_F(B)$.

(a) By Definition 1.1.32, the sum and product of elements of F, computed in K, are their sum and product in F, and the inverse in K of a nonzero element of F lies in F and is its inverse there (inverses are unique). We show by induction on the number of operations performed that the runs of Algorithm 3.3.1 on A over F and over K perform the same operations, with the same current matrices, all of whose entries lie in F. At the start both current matrices are A. If the current matrices agree and have entries in F, then every decision of the next step (which column, whether an entry is 0, which row to swap) is the same, the scalars used (inverses and negatives of entries) are the same elements of F, and the operation produces the same matrix, with entries in F. So the outputs agree.

(b) Over K, A is invertible, so $\operatorname{rref}(A) = I_n$ by Theorem 3.3.23, and by (a) this also holds over F. Running Algorithm 3.3.24 over F produces $(I_n \mid G)$ with $G \in F^{n\times n}$ and $GA = I_n$, and by Theorem 3.3.23 G is the inverse of A, over F and hence over K (where inverses are unique).

(c) An elementary matrix over F is an elementary matrix over K, so row equivalence over F implies row equivalence over K. Conversely, if A and B are row equivalent over K, then by Corollary 3.3.18(b) over K and by (a),

$$\operatorname{rref}_F(A) = \operatorname{rref}_K(A) = \operatorname{rref}_K(B) = \operatorname{rref}_F(B),$$

where the subscript records the field used. By Corollary 3.3.18(b) over F, A and B are row equivalent over F.

Solution to Exercise 3.3.55

Uses. Proposition 3.3.11 (Carrying extra columns along), Corollary 3.3.18 (RREF and row equivalence), Theorem 3.3.23 (Invertibility and RREF)
Hint 1. Compute $\operatorname{rref}(A) = G_A A$ and $\operatorname{rref}(B) = G_B B$. When can $A \sim B$ hold, and how would you get from A to B?
Hint 2. If $G_A A = G_B B$, then $B = G_B^{-1} G_A A$. To compute $G_B^{-1} G_A$, you can row reduce $(G_B \mid G_A)$.
Sketch. Row reduce $(A \mid I_m)$ and $(B \mid I_m)$ to $(R_A \mid G_A)$ and $(R_B \mid G_B)$. If $R_A \neq R_B$, then A and B are not row equivalent. If $R_A = R_B$, then $G = G_B^{-1} G_A$ works. In (b), $G_A = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$, $G_B = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}$, and $G = \begin{pmatrix} -2 & 1 \\ -1 & 1 \end{pmatrix}$.

(a) Row reduce $(A \mid I_m)$ and $(B \mid I_m)$, scanning only the first n columns. By Proposition 3.3.11 this gives $(R_A \mid G_A)$ and $(R_B \mid G_B)$ with $R_A = \operatorname{rref}(A) = G_A A$ and $R_B = \operatorname{rref}(B) = G_B B$, where G_A, G_B are invertible. If $R_A \neq R_B$, then A and B are not row equivalent, by Corollary 3.3.18(b). If $R_A = R_B$, then $G_A A = G_B B$, so $B = GA$ with $G = G_B^{-1} G_A$, which is invertible. To compute it, row reduce $(G_B \mid G_A)$, scanning only the first m columns: since G_B is invertible, $\operatorname{rref}(G_B) = I_m$, and the run's product is G_B^{-1} (Theorem 3.3.23), so the result is $\left(I_m \mid G_B^{-1} G_A\right)$ by Proposition 3.3.11.

(b) For A: adding -2 times row 1 to row 2 gives $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, and adding -1 times row 2 to row 1 gives $R_A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The same operations on I_2 give $G_A = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$. For B: swapping the rows gives $\begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, and adding -2 times row 2 to row 1 gives $R_B = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_A$. The same operations on I_2 give $G_B = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}$. So $A \sim B$. Since $G_B = E_{2,-2,1} E_{12}$, its inverse is $E_{12} E_{2,2,1} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ (Corollary 3.2.11), and

$$G = G_B^{-1} G_A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -1 & 1 \end{pmatrix}.$$

Check:

$$GA = \begin{pmatrix} -2+2 & -4+4 & -2+3 \\ -1+2 & -2+4 & -1+3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix} = B.$$

Solution to Exercise 3.3.56

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Theorem 3.3.23 (Invertibility and RREF), Definition B.1.7 (Modulus)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Pass 1 is fine, since $a_{11} \neq 0$. Find an invariant that makes every later pass equally fine.
Hint 2. After pass 1, the entries in rows and columns $2, \dots, n$ are $a_{ij} - a_{i1}a_{1j}/a_{11}$. Show that this $(n-1) \times (n-1)$ block is again strictly diagonally dominant.
Hint 3. Bound $\sum_{j \geq 2, j \neq i} |a_{1j}|$ using row 1, and $\sum_{j \geq 2, j \neq i} |a_{ij}|$ using row i, and finish with $|x| - |y| \leq |x - y|$.
Sketch. Invariant before pass k: columns $1, \dots, k-1$ are $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$, rows $k, \dots, n$ vanish in those columns, and the block of rows and columns $k, \dots, n$ is strictly diagonally dominant. Then the (k,k) entry is a nonzero pivot, and the new block entries are $m_{ij} - m_{ik}m_{kj}/m_{kk}$. Bounding the off-diagonal entries of row i of the new block with the dominance of rows k and i and the triangle inequality shows that the new block is again strictly diagonally dominant. After n passes the matrix is I_n.

We show by induction on $k = 1, \dots, n$ that before pass k the current matrix M satisfies: (i) columns $1, \dots, k-1$ of M are $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$; (ii) rows $k, \dots, n$ of M vanish in columns $1, \dots, k-1$; (iii) the block $(m_{ij})_{k \leq i,j \leq n}$ is strictly diagonally dominant. For $k = 1$, $M = A$ and this is the hypothesis.

Assume (i)–(iii) before pass k. By (iii), $|m_{kk}| > \sum_{j>k, j \neq k} |m_{kj}| \geq 0$, so $m_{kk} \neq 0$. By (ii), the scan picks column k, and the (k,k) entry is nonzero, so there is no swap. The pass scales row k by m_{kk}^{-1} and then, for each $i \neq k$, adds $-m_{ik}$ times the new row k to row i. Row k vanishes in columns $1, \dots, k-1$ by (ii), so these columns do not change, and column k becomes $\mathbf{e}_k$. This gives (i) and (ii) for $k+1$. For $i, j > k$, the new entries are

$$m'_{ij} = m_{ij} - m_{ik}\frac{m_{kj}}{m_{kk}}.$$

To prove (iii) for $k+1$, fix $i > k$, and let $\sum'$ denote a sum over $j > k$ with $j \neq i$. The dominance of row k and of row i in the old block give

$$\sum{}' |m_{kj}| < |m_{kk}| - |m_{ki}|, \quad \sum{}' |m_{ij}| < |m_{ii}| - |m_{ik}|.$$

Using the triangle inequality and then these two bounds (the first multiplied by $|m_{ik}|/|m_{kk}| \geq 0$, so it may become an equality),

$$\begin{aligned} \sum{}' |m'_{ij}| &\leq \sum{}' |m_{ij}| + \frac{|m_{ik}|}{|m_{kk}|} \sum{}' |m_{kj}| \\ &< |m_{ii}| - |m_{ik}| + \frac{|m_{ik}|}{|m_{kk}|}(|m_{kk}| - |m_{ki}|) \\ &= |m_{ii}| - \left|\frac{m_{ik}m_{ki}}{m_{kk}}\right| \leq \left|m_{ii} - \frac{m_{ik}m_{ki}}{m_{kk}}\right| = |m'_{ii}|, \end{aligned}$$

where the last inequality is $|x| - |y| \leq |x - y|$. So the new block is strictly diagonally dominant, which completes the induction.

So every pass finds its pivot on the diagonal without a swap, and after pass n the matrix is I_n by (i). Hence $\operatorname{rref}(A) = I_n$, and A is invertible by Theorem 3.3.23.

Solution to Exercise 3.3.57

Uses. Definition 2.4.1 (Matrix product), Proposition 3.3.8 (Properties of the output), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.5.12 (Invertible)
Hint 1. The rows of G numbered $r+1, \dots, m$ are relations, since the corresponding rows of GA are zero.
Hint 2. Conversely, if $\mathbf{z}A = \mathbf{0}$, put $\mathbf{w} = \mathbf{z}G^{-1}$, so that $\mathbf{z} = \mathbf{w}G$ and $\mathbf{w} \operatorname{rref}(A) = \mathbf{0}$.

Hint 3. Look at the entries of $\mathbf{w}\,\mathrm{rref}(A)$ in the pivot columns.
Sketch. Row i of GA is (row i of G) A, so the last $m-r$ rows of G are relations, and so are their combinations. If $\mathbf{z}A=\mathbf{0}$, then $\mathbf{w}=\mathbf{z}G^{-1}$ satisfies $\mathbf{w}R=\mathbf{0}$ with $R=\mathrm{rref}(A)$; the entry of $\mathbf{w}R$ in column p_s is w_s, so $w_1=\cdots=w_r=0$, and $\mathbf{z}=\mathbf{w}G$ is a combination of rows $r+1,\ldots,m$ of G.

Let $R=\mathrm{rref}(A)=GA$ and let $\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_m$ be the rows of G. By Definition 2.4.1, row i of GA is $\boldsymbol{\gamma}_iA$, and $(c\mathbf{z}+c'\mathbf{z}')A=c\mathbf{z}A+c'\mathbf{z}'A$.

By Proposition 3.3.8(a), rows $r+1,\ldots,m$ of R are zero, so $\boldsymbol{\gamma}_iA=\mathbf{0}$ for $i>r$. Hence every linear combination of $\boldsymbol{\gamma}_{r+1},\ldots,\boldsymbol{\gamma}_m$ is a relation.

Conversely, let $\mathbf{z}A=\mathbf{0}$, and put $\mathbf{w}=\mathbf{z}G^{-1}$. Then $\mathbf{z}=\mathbf{w}G$, and

$$\mathbf{w}R=\mathbf{w}GA=\mathbf{z}A=\mathbf{0}.$$

Now $\mathbf{w}R$ is w_1 times row 1 of R, plus $\cdots$, plus w_m times row m of R. Let $p_1<\cdots<p_r$ be the pivot columns. Column p_s of R is $\mathbf{e}_s$ (Lemma 3.3.16(b)), so row i of R has entry 1 in column p_s if $i=s$ and 0 otherwise, and the entry of $\mathbf{w}R$ in column p_s is w_s. Hence $w_s=0$ for $s\le r$, and

$$\mathbf{z}=\mathbf{w}G=w_{r+1}\boldsymbol{\gamma}_{r+1}+\cdots+w_m\boldsymbol{\gamma}_m.$$

In the language of transposes, $\mathbf{z}A=\mathbf{0}$ says $A^{\mathsf{T}}\mathbf{z}^{\mathsf{T}}=\mathbf{0}$. So the transposes of the last $m-r$ rows of G span $\ker T_{A^{\mathsf{T}}}$, one of the four fundamental subspaces of Section 4.3.

Solution to Exercise 3.3.58

Uses. Theorem 3.3.12 (Matrices in RREF), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Corollary 3.3.18 (RREF and row equivalence)
Hint 1. Split according to whether the last column is a pivot column.
Hint 2. If column n is a pivot column, it must hold the last pivot. Delete the last row and the last column. If it is not, delete only the last column. In each case, what can the deleted column have been?
Hint 3. For (b), a class contains exactly one matrix in RREF, and a 2×4 matrix in RREF has 0, 1, or 2 nonzero rows, all at the top.
Sketch. If column n is a pivot column, it is the last one, so it is $\mathbf{e}_r$, row r is $(0,\ldots,0,1)$, and deleting row r and column n gives a bijection onto the $(r-1)\times(n-1)$ matrices in RREF without zero rows. Otherwise all r pivots lie in columns $1,\ldots,n-1$, those columns form an $r\times(n-1)$ matrix in RREF without zero rows, and column n is arbitrary: p^r choices. For $p=2$: $N(1,n)=2^n-1$, $N(2,3)=7$, $N(2,4)=4\cdot7+7=35$, and there are $1+15+35=51$ classes.

(a) Let R be an $r\times n$ matrix in RREF with no zero row. By Theorem 3.3.12, it has r pivots, one in each row, in columns $p_1<\cdots<p_r$.

Column n is a pivot column. Then $p_r=n$, column n is $\mathbf{e}_r$, and row r, whose leading entry is in column n, is $(0,\ldots,0,1)$. Deleting row r and column n leaves an $(r-1)\times(n-1)$ matrix R_0. Rows $1,\ldots,r-1$ of R have zero entries in column n, so they keep their leading entries, and R_0 satisfies (R1)–(R4) with no zero row. Conversely, for any such R_0, the matrix $\begin{pmatrix}R_0&0\\0&1\end{pmatrix}$ satisfies (R1)–(R4) with no zero row and has column n as a pivot column. So there are $N(r-1,n-1)$ matrices in this case.

Column n is not a pivot column. Then all r pivots lie in columns $1,\ldots,n-1$, and deleting column n leaves an $r\times(n-1)$ matrix in RREF without zero rows (the leading entries are unchanged). Conversely, given such a matrix, appending *any* column $\mathbf{c}\in\mathbb{F}_p^r$ gives a matrix satisfying (R1)–(R4): the leading entries are unchanged, and (R4) says nothing about a column without a leading entry. So there are $p^rN(r,n-1)$ matrices in this case. (When $r=n$ there are none, consistent with $N(n,n-1)=0$.)

Adding the two cases gives the recursion.

(b) With $p=2$: $N(1,1)=1$ and $N(1,n)=2N(1,n-1)+1$, so $N(1,2)=3$, $N(1,3)=7$, $N(1,4)=15$. Next $N(2,2)=4\cdot0+N(1,1)=1$, $N(2,3)=4\cdot1+3=7$, and

$$N(2,4)=4\cdot N(2,3)+N(1,3)=28+7=35.$$

A 2×4 matrix in RREF has $r \in \{0, 1, 2\}$ nonzero rows, which by (R1) are its top r rows, and they form an $r \times 4$ matrix in RREF without zero rows. So there are $N(0,4) + N(1,4) + N(2,4) = 1 + 15 + 35 = 51$ of them. As in Exercise 3.3.48, each row equivalence class contains exactly one matrix in RREF, so $\mathbb{F}_2^{2\times 4}$ has 51 classes.

The recursion is a p-analogue of Pascal's rule $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$. Using row spaces and dimension (Chapter 4), one can show that $N(r,n)$ is the number of r-dimensional subspaces of $\mathbb{F}_p^n$.

Solution to Exercise 3.3.59

Uses. Theorem 3.3.33 (PLU and LU factorizations), Theorem 3.3.37 (LDL^{T} factorization)
Hint 1. Run Gaussian elimination and record the multipliers; they are the entries of L below the diagonal.
Hint 2. For (b), the diagonal of U gives D. Here all its entries are equal.
Sketch. (a) The multipliers are $\ell_{21} = 3$, $\ell_{31} = -2$, $\ell_{32} = 3$, and U has rows $(1,2,0)$, $(0,1,2)$, $(0,0,-1)$. (b) The multipliers are $1/2$, $-1/2$, $1/2$, and U has diagonal $4,4,4$, so $D = 4I_3$ and $S = (2L)(2L)^{\mathsf{T}}$; the rows of $M = 2L$ are $(2,0,0)$, $(1,2,0)$, $(-1,1,2)$.

(a) Pass 1 uses the multipliers $\ell_{21} = 3$ and $\ell_{31} = -2$, leaving rows $(0,1,2)$ and $(0,3,5)$. Pass 2 uses $\ell_{32} = 3$, leaving $(0,0,-1)$. By Theorem 3.3.33(b),

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}.$$

Check the last row:

$$-2(1,2,0) + 3(0,1,2) + (0,0,-1) = (-2,-1,5).$$

(b) Pass 1 uses $\ell_{21} = 2/4 = 1/2$ and $\ell_{31} = -2/4 = -1/2$, leaving rows $(0,4,2)$ and $(0,2,5)$. Pass 2 uses $\ell_{32} = 1/2$, leaving $(0,0,4)$. So

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -1/2 & 1/2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & 2 & -2 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{pmatrix},$$

and $D = 4I_3$. By Theorem 3.3.37, $S = L(4I_3)L^{\mathsf{T}}$; indeed $4L^{\mathsf{T}} = U$. So $S = (2L)(2L)^{\mathsf{T}} = MM^{\mathsf{T}}$ with

$$M = 2L = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ -1 & 1 & 2 \end{pmatrix}.$$

Check: the rows of M have pairwise products $4, 2, -2$ (row 1 with rows $1,2,3$), $1 + 4 = 5$, $-1 + 2 = 1$, and $1 + 1 + 4 = 6$, which are the entries of S.

Solution to Exercise 3.3.60

Uses. Lemma 3.3.32 (Products and inverses of triangular matrices), Exercise 3.3.53, Definition 2.5.12 (Invertible)
Hint 1. Suppose $L_1U_1 = L_2U_2$. Why are U_1 and U_2 invertible, and why is U_1^{-1} upper triangular?
Hint 2. Rearrange to $L_2^{-1}L_1 = U_2U_1^{-1}$, and look at both sides.
Sketch. $U_i = L_i^{-1}A$ is invertible, so by Exercise 3.3.53 it has nonzero diagonal entries and an upper triangular inverse. Then $L_2^{-1}L_1 = U_2U_1^{-1}$ is unit lower triangular and upper triangular, hence I_n.

Suppose $A = L_1U_1 = L_2U_2$ with L_1, L_2 unit lower triangular and U_1, U_2 upper triangular. By Lemma 3.3.32, L_i is invertible, so $U_i = L_i^{-1}A$ is invertible, as a product of invertible matrices. By Exercise 3.3.53(b), the diagonal entries of U_1 are nonzero, and then by Exercise 3.3.53(a), U_1^{-1} is upper triangular. Multiplying $L_1U_1 = L_2U_2$ on the left by L_2^{-1} and on the right by U_1^{-1} gives

$$L_2^{-1}L_1 = U_2U_1^{-1}.$$

By Lemma 3.3.32, the left side is unit lower triangular and the right side is upper triangular. A matrix that is both is diagonal, with the diagonal entries of the left side, which are 1. So $L_2^{-1}L_1 = I_n = U_2U_1^{-1}$, that is, $L_1 = L_2$ and $U_1 = U_2$.

Solution to Exercise 3.3.61

Uses. Exercise 3.2.54, Proposition 3.2.10 (Inverses of elementary matrices), Definition 2.4.33 (Block matrix), induction on n

Hint 1. (ii) ⇒ (i) is quick. For the converse, induct on n and imitate the first pass of Gauss–Jordan elimination, using only integer operations.

Hint 2. Let $B = A^{-1}$. The $(1,1)$ entry of $BA = I_n$ shows that the entries of the first column of A have greatest common divisor 1.

Hint 3. Use Exercise 3.2.54 on rows 1 and i, for $i = 2, \ldots, n$ in turn, to make the first column $\mathbf{e}_1$. Then show the lower right $(n-1)\times(n-1)$ block has an integer inverse.

Sketch. (ii) ⇒ (i): inverses of integer elementary matrices are integer elementary. (i) ⇒ (ii): the first column of A has gcd 1, because $\sum_k b_{1k}a_{k1} = 1$. Euclid's algorithm on pairs of rows, done with integer elementary matrices, gives $HA = \begin{pmatrix} 1 & \mathbf{w} \\ 0 & A' \end{pmatrix}$. This matrix has an integer inverse, and comparing blocks shows that A' has one too. By induction A' reduces to I_{n-1} by integer operations on rows $2, \ldots, n$, and integer replacements clear $\mathbf{w}$. So $PA = I_n$ for a product P of integer elementary matrices, and $A = P^{-1}$.

(ii) ⇒ (i). Integer elementary matrices have integer entries, and by Proposition 3.2.10 their inverses E_{ij}, $E_{i,-1}$, $E_{i(-c)j}$ are again integer elementary. If $A = E_1 \cdots E_k$ with all E_l integer elementary, then A is invertible with $A^{-1} = E_k^{-1} \cdots E_1^{-1}$, a product of integer matrices, so it has integer entries.

(i) ⇒ (ii). We show by induction on n that there is a product P of integer elementary matrices with $PA = I_n$; then $A = P^{-1}$ is a product of integer elementary matrices by the first part. For $n = 1$, $A = (a)$ with a and $1/a$ integers, so $a = \pm 1$, and $P = (1)$ or $P = E_{1,-1}$ works.

Let $n \geq 2$ and $B = A^{-1} \in \mathbb{Z}^{n\times n}$. The $(1,1)$ entry of $BA = I_n$ is $\sum_k b_{1k}a_{k1} = 1$, so every common divisor of $a_{11}, \ldots, a_{n1}$ divides 1. For $i = 2, \ldots, n$ in turn, if the current entries in positions $(1,1)$ and $(i,1)$ are not both 0, apply Exercise 3.2.54(b) to them: it gives 2×2 integer elementary matrices that turn this pair into (their gcd, 0), and acting on rows 1 and i only, they are $n \times n$ integer elementary matrices. Row operations of this kind do not change the set of common divisors of the entries of the first column (each new entry is an integer combination of old ones, and the operations can be undone by integer elementary matrices), so at the end the first column is $(g, 0, \ldots, 0)$, where $g \geq 1$ is a common divisor of the original entries; hence $g = 1$. So we have a product H_1 of integer elementary matrices with

$$H_1 A = \begin{pmatrix} 1 & \mathbf{w} \\ 0 & A' \end{pmatrix}, \quad \mathbf{w} \in \mathbb{Z}^{1\times(n-1)}, \quad A' \in \mathbb{Z}^{(n-1)\times(n-1)}.$$

This matrix is invertible with integer inverse $C = A^{-1}H_1^{-1}$. Write $C = \begin{pmatrix} c & \mathbf{x} \\ \mathbf{y} & C' \end{pmatrix}$ in the same block sizes. Comparing first block columns in $C(H_1A) = I_n$ gives $c = 1$ and $\mathbf{y} = \mathbf{0}$, and then the lower right blocks of $C(H_1A) = I_n$ and $(H_1A)C = I_n$ give $C'A' = I_{n-1}$ and $A'C' = I_{n-1}$. So A' is invertible with integer inverse C', and by the induction hypothesis there is a product P' of $(n-1)\times(n-1)$ integer elementary matrices with $P'A' = I_{n-1}$. Each factor F of P' gives the $n \times n$ integer elementary matrix $\begin{pmatrix} 1 & 0 \\ 0 & F \end{pmatrix}$ (with indices shifted by one), so $H_2 = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}$ is a product of integer elementary matrices, and $H_2H_1A = \begin{pmatrix} 1 & \mathbf{w} \\ 0 & I_{n-1} \end{pmatrix}$. Finally, adding $-w_j$ times row $j+1$ to row 1, for $j = 1, \ldots, n-1$, clears $\mathbf{w}$ with integer replacements. This gives $PA = I_n$ with P a product of integer elementary matrices.

For the matrix of Example 3.3.25, which has an integer inverse, this explains why fractions could have been avoided there.

Section 3.4

Solution to Exercise 3.4.7

Uses. Proposition 3.4.1 (Solving with the reduced matrix), Corollary 3.4.4 (Consistency), Theorem 3.4.3 (Solving a system in reduced form)
Hint 1. Reduce $(A \mid \mathbf{y})$ with symbolic y_1, y_2, y_3. The first column is zero, so the first pivot is in column 2.
Hint 2. The first column is a free column. Its variable x_1 appears in no equation at all.
Sketch. A_r has rows $(0,1,2,0)$, $(0,0,0,1)$, $\mathbf{0}$, so x_2, x_4 are pivot variables and x_1, x_3 are free. The last entry of $G\mathbf{y}$ is $-3y_1 + y_2 + y_3$, so the system is consistent iff $y_2 + y_3 = 3y_1$. For $\mathbf{y} = (1,2,1)$: $x_2 = 1 - 2x_3$, $x_4 = 0$, with x_1, x_3 arbitrary.

We reduce $(A \mid \mathbf{y})$. Column 1 is zero, so pass 1 uses the pivot in position $(1,2)$: add -2 times row 1 to row 2 and -1 times row 1 to row 3. In pass 2, column 3 vanishes in rows 2 and 3, so the pivot is in position $(2,4)$; add -1 times row 2 to row 1 and row 2 to row 3:

$$\left(\begin{array}{cccc|c} 0 & 1 & 2 & 1 & y_1 \\ 0 & 2 & 4 & 3 & y_2 \\ 0 & 1 & 2 & 0 & y_3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 2 & 1 & y_1 \\ 0 & 0 & 0 & 1 & -2y_1 + y_2 \\ 0 & 0 & 0 & -1 & -y_1 + y_3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 0 & 1 & 2 & 0 & 3y_1 - y_2 \\ 0 & 0 & 0 & 1 & -2y_1 + y_2 \\ 0 & 0 & 0 & 0 & -3y_1 + y_2 + y_3 \end{array}\right).$$

So A_r has rows $(0,1,2,0)$, $(0,0,0,1)$, $(0,0,0,0)$, the pivot variables are x_2 and x_4, and the free variables are x_1 and x_3. By Corollary 3.4.4 the system is consistent if and only if $-3y_1 + y_2 + y_3 = 0$.

For $\mathbf{y} = (1,2,1)$ the condition holds $(-3 + 2 + 1 = 0)$, and the reduced system is $x_2 + 2x_3 = 1$, $x_4 = 0$. By Theorem 3.4.3(b), the solutions are the vectors $(x_1, 1 - 2x_3, x_3, 0)$ with $x_1, x_3 \in \mathbb{R}$ arbitrary. The unknown x_1 multiplies the zero column, so it occurs in no equation, and it is free for the most obvious reason. As a check, $A(0,1,0,0)$ is the second column $(1,2,1)$.

Solution to Exercise 3.4.18

Uses. Proposition 3.4.15 (A pivot in every column), Proposition 3.4.16 (A pivot in every row)
Hint 1. Locate the pivots. Is there a column without one? A row without one?
Hint 2. A column without a pivot is a free variable, which gives a nontrivial solution of $A\mathbf{x} = \mathbf{0}$. A row without a pivot is a zero row, which makes some $A\mathbf{x} = \mathbf{y}$ inconsistent.
Sketch. R_1: a pivot in every row but not in column 3, so always consistent, with nontrivial homogeneous solutions. R_2: a pivot in every column but not in row 3, so only the trivial solution, but not always consistent. R_3: neither. R_4: both.

By Proposition 3.4.15, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution exactly when every column of A_r has a pivot, and by Proposition 3.4.16, $A\mathbf{x} = \mathbf{y}$ is consistent for every $\mathbf{y}$ exactly when every row has a pivot.

- R_1 has pivots in columns $1, 2$ and in both rows. Column 3 is free, so $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution (for instance the basic solution $(-3, 1, 1)$), and $A\mathbf{x} = \mathbf{y}$ is consistent for every $\mathbf{y}$.
- R_2 has pivots in both columns but not in row 3. So $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, but $A\mathbf{x} = \mathbf{y}$ is inconsistent for some $\mathbf{y}$.
- R_3 has pivots in columns $1, 3$ and rows $1, 2$. Column 2 and row 3 have none, so both properties fail.
- $R_4 = I_3$ has a pivot in every row and every column, so both properties hold.

Solution to Exercise 3.4.29

Uses. Definition 3.4.26 (Lagrange polynomials), Proposition 3.4.27 (Lagrange's formula), Theorem 3.4.24 (Polynomial interpolation)
Hint 1. Each L_i is the solution of a 3×3 system. Alternatively, guess L_i (for instance from the closed form previewed in Example 3.4.28) and check its values at the three nodes.
Hint 2. For (b), check whether the answer to (a) still makes sense with coefficients in $\mathbb{F}_5$, and quote uniqueness.
Sketch. $L_0 = 1 - \frac{4}{3}x + \frac{1}{3}x^2$, $L_1 = \frac{3}{2}x - \frac{1}{2}x^2$, $L_2 = -\frac{1}{6}x + \frac{1}{6}x^2$, and $p = 2L_0 - L_1 + 5L_2 = 2 - 5x + 2x^2$. Over $\mathbb{F}_5$, $2 - 5x + 2x^2 = 2 + 2x^2$ takes the values $2, 4 = -1, 20 = 0 = 5$, so it is the answer by uniqueness.

(a) Expanding the closed form of Example 3.4.28 suggests the candidates

$$L_0 = 1 - \frac{4}{3}x + \frac{1}{3}x^2, \quad L_1 = \frac{3}{2}x - \frac{1}{2}x^2, \quad L_2 = -\frac{1}{6}x + \frac{1}{6}x^2.$$

We check the values directly. $L_0(0) = 1, L_0(1) = 1 - \frac{4}{3} + \frac{1}{3} = 0, L_0(3) = 1 - 4 + 3 = 0$; $L_1(0) = 0, L_1(1) = \frac{3}{2} - \frac{1}{2} = 1$, $L_1(3) = \frac{9}{2} - \frac{9}{2} = 0$; $L_2(0) = 0$, $L_2(1) = 0$, $L_2(3) = -\frac{1}{2} + \frac{3}{2} = 1$. Each has degree at most 2, so by the uniqueness in Theorem 3.4.24 these are the Lagrange polynomials. By Proposition 3.4.27,

$$p = 2L_0 - L_1 + 5L_2 = 2 + \left(-\frac{8}{3} - \frac{3}{2} - \frac{5}{6}\right)x + \left(\frac{2}{3} + \frac{1}{2} + \frac{5}{6}\right)x^2 = 2 - 5x + 2x^2.$$

Check: $p(0) = 2$, $p(1) = 2 - 5 + 2 = -1$, $p(3) = 2 - 15 + 18 = 5$.

(b) In $\mathbb{F}_5$ the nodes $0, 1, 3$ are still distinct, so there is exactly one such p by Theorem 3.4.24. The polynomial $2 - 5x + 2x^2 = 2 + 2x^2$ has coefficients in $\mathbb{F}_5$, and its values are 2, $4 = -1$, and $2 + 18 = 20 = 0 = 5$ in $\mathbb{F}_5$. So $p = 2 + 2x^2$. (Reducing the answer over $\mathbb{Q}$ worked because the computation only divided by 2, 3, and 6, which are invertible in $\mathbb{F}_5$.)

Solution to Exercise 3.4.30

Uses. Algorithm 3.4.13 (Solving a linear system), Corollary 3.4.4 (Consistency), Theorem 3.4.9 (All solutions)
Hint 1. Row reduce $(A \mid \mathbf{y})$ with symbolic y_1, y_2, y_3. No swaps or scalings are needed.
Hint 2. The pivot columns are 1 and 3. The basic solutions come from columns 2, 4, 5 of A_r.
Sketch. The reduction gives rows $(1, 2, 0, 1, 3 \mid y_1)$, $(0, 0, 1, 2, 1 \mid -2y_1 + y_2)$, $(\mathbf{0} \mid 3y_1 - y_2 + y_3)$. (a) Consistent iff $y_2 = 3y_1 + y_3$. (b) $\mathbf{x}_0 = (1, 0, 1, 0, 0)$, $\mathbf{s}_1 = (-2, 1, 0, 0, 0)$, $\mathbf{s}_2 = (-1, 0, -2, 1, 0)$, $\mathbf{s}_3 = (-3, 0, -1, 0, 1)$.

(a) Add -2 times row 1 to row 2 and row 1 to row 3. Then column 2 vanishes below row 1, so the next pivot is in position $(2, 3)$; add -1 times row 2 to row 3. Row 1 already has a 0 in column 3. The result is

$$\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 3 & y_1 \\ 0 & 0 & 1 & 2 & 1 & -2y_1 + y_2 \\ 0 & 0 & 0 & 0 & 0 & 3y_1 - y_2 + y_3 \end{array}\right).$$

So $k = 2$, and by Corollary 3.4.4 the system is consistent if and only if $3y_1 - y_2 + y_3 = 0$, that is, $y_2 = 3y_1 + y_3$.

(b) For $\mathbf{y} = (1, 3, 0)$ the last entry is $3 - 3 + 0 = 0$, and the last column becomes $(1, 1, 0)$. The pivot columns are $1, 3$ and the free columns $2, 4, 5$. Following Algorithm 3.4.13: $\mathbf{x}_0$ has $1, 1$ in positions $1, 3$, and the basic solutions come from columns 2, 4, 5 of A_r, which are $(2, 0, 0)$, $(1, 2, 0)$, $(3, 1, 0)$. So

$$\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{s}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} -3 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix},$$

and the solution set is $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3)$. As a check,

$$A\mathbf{x}_0 = \mathbf{a}_1 + \mathbf{a}_3 = (1, 2, -1) + (0, 1, 1) = (1, 3, 0),$$
$$A\mathbf{s}_3 = -3\mathbf{a}_1 - \mathbf{a}_3 + \mathbf{a}_5 = (-3, -6, 3) - (0, 1, 1) + (3, 7, -2) = \mathbf{0}.$$

Solution to Exercise 3.4.31

Uses. Algorithm 3.4.13 (Solving a linear system), Theorem 3.4.9 (All solutions), Corollary 3.4.14 (Counting solutions)
Hint 1. Reduce $(A \mid \mathbf{y})$; the first two passes are the same over both fields. Over $\mathbb{F}_2$, remember that $-1 = 1$ and $2 = 0$.
Hint 2. Over $\mathbb{Q}$ there are three pivots and one free variable; over $\mathbb{F}_2$ there are two of each.
Sketch. Over $\mathbb{Q}$, $(A \mid \mathbf{y})$ reduces to rows $(1, 0, 0, 0 \mid 1)$, $(0, 1, 0, 1 \mid 0)$, $(0, 0, 1, 0 \mid 0)$, so the solutions are $(1, 0, 0, 0) + t(0, -1, 0, 1)$. Over $\mathbb{F}_2$ it reduces to $(1, 0, 1, 0 \mid 1)$, $(0, 1, 1, 1 \mid 0)$, $(\mathbf{0} \mid 0)$, and there are $2^2 = 4$ solutions: $(1, 0, 0, 0)$, $(0, 1, 1, 0)$, $(1, 1, 0, 1)$, $(0, 0, 1, 1)$.

Pass 1 adds -1 times row 1 to row 3, and pass 2 adds -1 times row 2 to row 1 and row 2 to row 3:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{array}\right).$$

Over $\mathbb{Q}$. Pass 3 scales row 3 by $1/2$, then adds row 3 to row 1 and -1 times row 3 to row 2, giving

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

The pivot columns are $1, 2, 3$ and x_4 is free. So $\mathbf{x}_0 = (1, 0, 0, 0)$, $\mathbf{s}_1 = (0, -1, 0, 1)$, and the solutions are $(1, -t, 0, t)$ for $t \in \mathbb{Q}$: infinitely many.

Over $\mathbb{F}_2$. Here $-1 = 1$ and $2 = 0$, so the last matrix above is

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

which is in RREF. The last entry of the right side is 0, so the system is consistent. The pivot variables are x_1, x_2 and the free variables x_3, x_4, with $x_1 = 1 + x_3$ and $x_2 = x_3 + x_4$ (again since $-1 = 1$). The four choices of (x_3, x_4) give the solutions

$$(1, 0, 0, 0), \quad (0, 1, 1, 0), \quad (1, 1, 0, 1), \quad (0, 0, 1, 1).$$

That is 2^2 solutions, as Corollary 3.4.14(b) predicts with $q = 2$ and $\ell = 2$.

Solution to Exercise 3.4.32

Uses. Algorithm 3.4.13 (Solving a linear system), Corollary 3.4.4 (Consistency), Theorem 3.4.9 (All solutions)
Hint 1. Pass 1 adds $-i$ times row 1 to row 2. Remember $i^2 = -1$.
Hint 2. The coefficient matrix has pivot columns $1, 3$. For (c), when is $1 - it$ real for $t \in \mathbb{C}$ with t itself real?
Sketch. The reduction gives rows $(1, i, 0 \mid y_1)$, $(0, 0, 1 \mid -iy_1 + y_2)$, $(0, 0, 0 \mid iy_1 - y_2 + y_3)$. (a) $(1, 0, 2) + t(-i, 1, 0)$, $t \in \mathbb{C}$. (b) $iy_1 - y_2 + y_3 = 0$. (c) Only $(1, 0, 2)$.

Pass 1 adds $-i$ times row 1 to row 2: the new row 2 is

$$(i, -1, 1) - i(1, i, 0) = \left(0, -1 - i^2, 1\right) = (0, 0, 1),$$

with right side $y_2 - iy_1$. Then column 2 vanishes below row 1, so pass 2 uses the pivot in position $(2, 3)$ and adds -1 times row 2 to row 3:

$$\left(\begin{array}{ccc|c} 1 & i & 0 & y_1 \\ 0 & 0 & 1 & -iy_1 + y_2 \\ 0 & 0 & 0 & iy_1 - y_2 + y_3 \end{array}\right).$$

(a) For $\mathbf{y} = (1, 2 + i, 2)$ the right side is $(1, -i + 2 + i, i - 2 - i + 2) = (1, 2, 0)$. The pivot variables are x_1, x_3 and x_2 is free, with $x_1 = 1 - ix_2$ and $x_3 = 2$. So the solutions are

$$(1, 0, 2) + t(-i, 1, 0), \quad t \in \mathbb{C}.$$

Check: the left sides at $(-i, 1, 0)$ are $-i + i = 0$, $1 - 1 + 0 = 0$, and 0.

(b) By Corollary 3.4.4, exactly when $iy_1 - y_2 + y_3 = 0$.

(c) The solution with parameter t is $(1 - it, t, 2)$. If it is real, then $t = x_2$ is real, and then $1 - it$ is real only if $t = 0$. So $(1, 0, 2)$ is the only real solution. The complex solution set is a line in $\mathbb{C}^3$, and it meets $\mathbb{R}^3$ in a single point.

Solution to Exercise 3.4.33

Uses. Algorithm 3.4.13 (Solving a linear system), Corollary 3.4.4 (Consistency), Corollary 3.4.14 (Counting solutions)
Hint 1. Subtract row 1 from row 2 and a times row 1 from row 3. When can you use the $(2, 2)$ entry $a - 1$ as a pivot?
Hint 2. For $a \neq 1$, after clearing column 2 the last row is $(0, 0, (1 - a)(2 + a) \mid b - a)$. Treat $a = 1$ and $a = -2$ separately. In $\mathbb{F}_3$, $-2 = 1$.

Sketch. For $a \neq 1$ the reduction reaches rows $(1, 1, a \mid 1)$, $(0, 1, -1 \mid 0)$, $(0, 0, (1-a)(2+a) \mid b-a)$. Over $\mathbb{R}$: one solution if $a \neq 1, -2$; for $a = -2$, none if $b \neq -2$ and infinitely many if $b = -2$; for $a = 1$, none if $b \neq 1$ and infinitely many if $b = 1$. Over $\mathbb{F}_3$, $a = -2$ is $a = 1$: for $a = 1$ there are 9 solutions if $b = 1$ and none otherwise; for $a = 0, 2$ exactly one.

Adding -1 times row 1 to row 2 and $-a$ times row 1 to row 3 gives

$$\left(\begin{array}{ccc|c} 1 & 1 & a & 1 \\ 0 & a-1 & 1-a & 0 \\ 0 & 1-a & 1-a^2 & b-a \end{array}\right).$$

Case $a = 1$. Rows 2 and 3 become $(0, 0, 0 \mid 0)$ and $(0, 0, 0 \mid b-1)$, and the left block is in RREF with one pivot. If $b \neq 1$, the system is inconsistent. If $b = 1$, it is consistent with $\ell = 2$ free variables x_2, x_3.

Case $a \neq 1$. Scale row 2 by $(a-1)^{-1}$ to get $(0, 1, -1 \mid 0)$, and add $-(1-a)$ times it to row 3, which becomes

$$\big(0, 0, 1 - a^2 + (1-a) \mid b-a\big) = (0, 0, (1-a)(2+a) \mid b-a).$$

If $(1-a)(2+a) \neq 0$, this entry is a third pivot, and there are pivots in every column and every row; by Corollary 3.4.20 the system has exactly one solution. If $(1-a)(2+a) = 0$, then since $a \neq 1$ we have $2 + a = 0$, and the last row is $(0, 0, 0 \mid b-a)$: the system is inconsistent if $b \neq a$, and consistent with one free variable x_3 if $b = a$.

(a) Over $\mathbb{R}$, $2 + a = 0$ means $a = -2$. So the system has exactly one solution if $a \neq 1$ and $a \neq -2$; no solution if $a = 1, b \neq 1$ or $a = -2, b \neq -2$; and infinitely many (by Corollary 3.4.14(c)) if $(a, b) = (1, 1)$ or $(a, b) = (-2, -2)$.

(b) Over $\mathbb{F}_3$, $-2 = 1$, so the case $2 + a = 0$ with $a \neq 1$ never occurs. For $a = 1$: if $b = 1$ there are $3^2 = 9$ solutions by Corollary 3.4.14(b), and if $b \neq 1$ there are none. For $a = 0$ and $a = 2$, $(1-a)(2+a)$ equals 2 and $(-1)(4) = -1 = 2$, both nonzero, so there is exactly one solution for every b. The two exceptional values 1 and -2 of the real case collide in $\mathbb{F}_3$.

Solution to Exercise 3.4.34

Uses. Theorem 3.4.9 (All solutions), Remark 3.4.12 (Checking a description of the solutions), Proposition 3.1.19 (Solution sets are translates of the kernel)
Hint 1. Student 1 proved one containment. Which one? Compare the description with Example 3.4.10.
Hint 2. Every solution is $\mathbf{x}_0 + t_1\mathbf{s}_1 + t_2\mathbf{s}_2$. Is $\mathbf{x}_0 + \mathbf{s}_1$ in Student 1's set? For Student 2, express the new vectors through $\mathbf{x}_0, \mathbf{s}_1, \mathbf{s}_2$.
Sketch. (a) The check shows only that the line consists of solutions, not that it contains all of them; the solution set is a plane. The solution $(-3, 1, 2, 0) = \mathbf{x}_0 + \mathbf{s}_1$ is missing: its second and fourth coordinates would force $t = 1$ and $t = 0$. (b) Right: $(1, -1, 2, 0) = \mathbf{x}_0 - \mathbf{s}_1$ is a solution, $(0, 1, -2, 1) = \mathbf{s}_1 + \mathbf{s}_2$ and $(4, -1, -2, 1) = \mathbf{s}_2 - \mathbf{s}_1$ lie in $\ker T_A$, and $\mathbf{s}_1, \mathbf{s}_2$ are combinations of these two vectors.

By Example 3.4.10, the solution set is $\mathbf{x}_0 + \operatorname{span}(\mathbf{s}_1, \mathbf{s}_2)$ with $\mathbf{x}_0 = (-1, 0, 2, 0)$, $\mathbf{s}_1 = (-2, 1, 0, 0)$, and $\mathbf{s}_2 = (2, 0, -2, 1)$.

(a) Student 1 showed that every vector of the line $(-1, 0, 2, 0) + \operatorname{span}((0, 1, -2, 1))$ is a solution. That is one containment; the claim "the solution set is" needs the other one too, and it is false (Remark 3.4.12). For instance $\mathbf{x}_0 + \mathbf{s}_1 = (-3, 1, 2, 0)$ is a solution. If it were $(-1, 0, 2, 0) + t(0, 1, -2, 1)$, comparing second coordinates would give $t = 1$ and comparing fourth coordinates $t = 0$. So Student 1 has missed it. The solution set is a plane, with two free variables, and a line cannot be all of it.

(b) Student 2 is right. First, $(1, -1, 2, 0) = \mathbf{x}_0 - \mathbf{s}_1$ is a solution, and $(0, 1, -2, 1) = \mathbf{s}_1 + \mathbf{s}_2$ and $(4, -1, -2, 1) = \mathbf{s}_2 - \mathbf{s}_1$ lie in $\ker T_A$. Since $\ker T_A$ is a subspace, Student 2's set is contained in the solution set. Conversely, with $\mathbf{u}_1 = \mathbf{s}_1 + \mathbf{s}_2$ and $\mathbf{u}_2 = \mathbf{s}_2 - \mathbf{s}_1$, we have

$$\mathbf{s}_1 = \frac{1}{2}(\mathbf{u}_1 - \mathbf{u}_2), \quad \mathbf{s}_2 = \frac{1}{2}(\mathbf{u}_1 + \mathbf{u}_2),$$

so $\mathbf{s}_1, \mathbf{s}_2 \in \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$, and every solution

$$\mathbf{x}_0 + t_1\mathbf{s}_1 + t_2\mathbf{s}_2 = (\mathbf{x}_0 - \mathbf{s}_1) + (t_1 + 1)\mathbf{s}_1 + t_2\mathbf{s}_2$$

lies in $(1, -1, 2, 0) + \operatorname{span}(\mathbf{u}_1, \mathbf{u}_2)$. (This uses $1 + 1 \neq 0$, which holds in $\mathbb{Q}$. In characteristic 2 we would have $\mathbf{u}_1 = \mathbf{u}_2$, and Student 2's set would only be a line.)

Solution to Exercise 3.4.35

Uses. Corollary 3.1.22 (Uniqueness depends only on A), Proposition 3.4.15 (A pivot in every column), Proposition 3.4.16 (A pivot in every row), Definition 2.3.8 (Matrix–vector product)
Hint 1. (a): uniqueness for one right-hand side is a statement about $\ker T_A$ (Corollary 3.1.22).
Hint 2. For (c), translate both statements into homogeneous systems with Proposition 3.4.15, and pad a solution for B with zeros. For (b) and (d), look at 1×2 and 2×1 matrices.
Sketch. (a) True: by Corollary 3.1.22, $\ker T_A = \{\mathbf{0}\}$, then Proposition 3.4.15. (b) False: the 2×1 matrix with column $(1, 0)$, and $\mathbf{y} = \mathbf{e}_2$. (c) True: a nontrivial solution of $B\mathbf{x} = \mathbf{0}$, padded with zeros in the deleted positions, is a nontrivial solution of $A\mathbf{x} = \mathbf{0}$. (d) False: $A = \begin{pmatrix} 0 & 1 \end{pmatrix}$, delete column 2.

(a) True. By Corollary 3.1.22(a), applied to the consistent system with exactly one solution, $\ker T_A = \{\mathbf{0}\}$; that is, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. By Proposition 3.4.15, A_r has a pivot in every column.

(b) False. Let $A \in F^{2\times 1}$ be the matrix with the single column $(1, 0)$. Then $Ax = (x, 0)$ is $\mathbf{0}$ only for $x = 0$, but $Ax = (0, 1)$ has no solution.

(c) True. By Proposition 3.4.15, we must show: if $B\mathbf{x} = \mathbf{0}$ has a nontrivial solution, so does $A\mathbf{x} = \mathbf{0}$. Let B consist of the columns $\mathbf{a}_{l_1}, \ldots, \mathbf{a}_{l_r}$ of A, and let $B\mathbf{u} = \mathbf{0}$ with $\mathbf{u} \neq \mathbf{0}$. Define $\mathbf{x} \in F^n$ by $x_{l_q} = u_q$ for $q = 1, \ldots, r$ and $x_l = 0$ for the deleted columns l. By Definition 2.3.8,

$$A\mathbf{x} = \sum_{q=1}^{r} u_q \mathbf{a}_{l_q} = B\mathbf{u} = \mathbf{0},$$

and $\mathbf{x} \neq \mathbf{0}$. (In terms of columns: a sublist of a linearly independent list is linearly independent.)

(d) False. Let $A = \begin{pmatrix} 0 & 1 \end{pmatrix} \in F^{1\times 2}$, which is in RREF with a pivot in its only row. Deleting the second column leaves $B = \begin{pmatrix} 0 \end{pmatrix}$, which has no pivot.

Solution to Exercise 3.4.36

Uses. Corollary 3.4.4 (Consistency), Corollary 3.4.14 (Counting solutions), Definition 2.5.12 (Invertible)
Hint 1. By Corollary 3.4.4, $\mathbf{y}$ works exactly when $G\mathbf{y}$ has zeros in positions $k + 1, \ldots, m$. How many vectors $\mathbf{z}$ have this property, and how are they related to the good $\mathbf{y}$?
Hint 2. For (b), count the pairs $(\mathbf{x}, \mathbf{y})$ with $A\mathbf{x} = \mathbf{y}$ in two ways.
Sketch. (a) $\mathbf{y} \mapsto G\mathbf{y}$ is a bijection of F^m (inverse $\mathbf{z} \mapsto G^{-1}\mathbf{z}$), and it maps the good $\mathbf{y}$ onto the $\mathbf{z}$ with $z_{k+1} = \cdots = z_m = 0$, of which there are q^k. (b) Each $\mathbf{x} \in F^n$ determines $\mathbf{y} = A\mathbf{x}$, so the q^n vectors $\mathbf{x}$ are split among the q^k consistent $\mathbf{y}$, each receiving q^{n-k} of them.

(a) Let G be invertible with $GA = A_r$, and let $Z = \{\mathbf{z} \in F^m \mid z_{k+1} = \cdots = z_m = 0\}$. By Corollary 3.4.4, $A\mathbf{x} = \mathbf{y}$ is consistent if and only if $G\mathbf{y} \in Z$. The map $\mathbf{y} \mapsto G\mathbf{y}$ is a bijection $F^m \to F^m$, with inverse $\mathbf{z} \mapsto G^{-1}\mathbf{z}$, so it restricts to a bijection from the set of good $\mathbf{y}$ onto Z. A vector of Z is determined by its first k entries, which are arbitrary, so $|Z| = q^k$.

(b) By Corollary 3.4.14(b), each consistent system has $q^\ell = q^{n-k}$ solutions. Now count the pairs $(\mathbf{x}, \mathbf{y}) \in F^n \times F^m$ with $A\mathbf{x} = \mathbf{y}$. Each $\mathbf{x}$ belongs to exactly one pair, namely $(\mathbf{x}, A\mathbf{x})$, so there are q^n pairs. On the other hand, grouping the pairs by $\mathbf{y}$, only the q^k consistent $\mathbf{y}$ occur, each in q^{n-k} pairs. So $q^n = q^k \cdot q^{n-k}$, and the identity just says that the solution sets of the consistent systems partition F^n.

Solution to Exercise 3.4.37

Uses. Definition 3.4.8 (Basic solutions), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Corollary 3.3.19 (Pivot columns, intrinsically), Definition 2.3.8 (Matrix–vector product)
Hint 1. If $j = f_t$ is free, try the basic solution $\mathbf{s}_t$. Which of its coordinates beyond position j could be nonzero?
Hint 2. A pivot coordinate $(\mathbf{s}_t)_{p_i} = -r_{if_t}$ with $p_i > f_t$ vanishes by Lemma 3.3.16(a). For the converse, rewrite $A\mathbf{x} = \mathbf{0}$ as an expression for $\mathbf{a}_j$.

Sketch. If $j = f_t$, then $\mathbf{s}_t$ has 1 in position j, 0 in the other free positions, and $-r_{ij}$ in position p_i; for $p_i > j$, $r_{ij} = 0$ by Lemma 3.3.16(a). Conversely, such an $\mathbf{x}$ gives $\mathbf{a}_j = -\sum_{l<j} x_l \mathbf{a}_l$, so column j is not a pivot column by Corollary 3.3.19(a).

Suppose x_j is free, say $j = f_t$, and consider the basic solution $\mathbf{s}_t$, which solves $A\mathbf{x} = \mathbf{0}$ (Theorem 3.4.9). It has a 1 in position j, and 0 in every other free position. In a pivot position $p_i > j$ it has the entry $-r_{ij}$ of A_r, where r_{ij} is the entry in row i of column $j < p_i$; this is 0 by Lemma 3.3.16(a), since column j lies to the left of the ith pivot. So $(\mathbf{s}_t)_l = 0$ for all $l > j$.

Conversely, suppose $A\mathbf{x} = \mathbf{0}$ with $x_j = 1$ and $x_l = 0$ for $l > j$. By Definition 2.3.8,

$$\mathbf{0} = A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_{j-1}\mathbf{a}_{j-1} + \mathbf{a}_j,$$

so $\mathbf{a}_j = -x_1\mathbf{a}_1 - \cdots - x_{j-1}\mathbf{a}_{j-1}$ is a linear combination of the earlier columns. By Corollary 3.3.19(a), column j is not a pivot column, so x_j is free.

So the basic solutions are the "staircase" solutions: $\mathbf{s}_t$ ends with a 1 in position f_t, followed only by zeros.

Solution to Exercise 3.4.38

Uses. Exercise 3.3.54, Corollary 3.4.4 (Consistency), Theorem 3.4.9 (All solutions), Proposition 3.4.15 (A pivot in every column)
Hint 1. By Exercise 3.3.54, Gauss–Jordan elimination on $(A \mid \mathbf{y})$ performs the same operations over F and over K. What does that say about the reduced augmented matrix?
Hint 2. The particular and basic solutions are built from the entries of $(A_r \mid G\mathbf{y})$.
Sketch. Row reducing $(A \mid \mathbf{y})$ uses only field operations on entries of F, so $(A_r \mid G\mathbf{y})$, the pivots, $\mathbf{x}_0$, and the $\mathbf{s}_j$ are the same over F and over K, and have entries in F. (a) Consistency over K is consistency over F, and then $\mathbf{x}_0 \in F^n$. (b) If the solution over K is unique, it is $\mathbf{x}_0$. (c) A free column exists, and $\mathbf{s}_1 \in F^n$ is nonzero.

Run Gauss–Jordan elimination on $(A \mid \mathbf{y})$, scanning only the first n columns. Every step uses only additions, multiplications, and inverses of nonzero entries, all of which stay in F, and whether an entry is 0 does not depend on whether we regard it in F or in K. So the run is the same over F and over K (this is the argument of Exercise 3.3.54), and it produces the same matrix $(A_r \mid \mathbf{z})$, with entries in F, and the same pivot and free columns. Hence the particular solution $\mathbf{x}_0$ (when the system is consistent) and the basic solutions $\mathbf{s}_j$ are the same over both fields, and lie in F^n.

(a) By Corollary 3.4.4, consistency over either field means $z_i = 0$ for $i > k$, the same condition. If the system has a solution in K^n, this condition holds, so $\mathbf{x}_0 \in F^n$ is a solution.

(b) If the solution in K^n is unique, it is $\mathbf{x}_0$ (which is a solution over K), and $\mathbf{x}_0 \in F^n$.

(c) If $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution in K^n, then by Proposition 3.4.15 (over K) some column of A_r is free, so $\ell \geq 1$, and $\mathbf{s}_1 \in F^n$ is a nonzero solution.

Solution to Exercise 3.4.39

Uses. Lemma 3.4.23 (Too many roots), Theorem 3.4.24 (Polynomial interpolation), Proposition 3.4.15 (A pivot in every column), Proposition 3.4.16 (A pivot in every row), Corollary 3.4.19 (More unknowns than equations)
Hint 1. As in Theorem 3.4.24, $W\mathbf{a}$ lists the values of $a_0 + a_1x + \cdots + a_dx^d$ at $c_1, \ldots, c_m$.
Hint 2. For (a) with $m \geq d+1$, use Lemma 3.4.23 with $d+1$ of the nodes. For (b) with $m \leq d+1$, interpolate with a polynomial of degree at most $m-1$. The other directions are Corollary 3.4.19.
Sketch. $W\mathbf{a}$ lists the values of $p_\mathbf{a}$ at the nodes. (a) If $m \geq d+1$ and $W\mathbf{a} = \mathbf{0}$, then $p_\mathbf{a}$ vanishes at $d+1$ distinct points, so $\mathbf{a} = \mathbf{0}$. If $m < d+1$, there are more unknowns than equations. (b) If $m \leq d+1$, Theorem 3.4.24 gives p of degree at most $m-1 \leq d$ with any values. If $m > d+1$, there are more equations than unknowns. Interpretation: interpolating polynomials of degree at most d are unique when they exist iff $m \geq d+1$, and always exist iff $m \leq d+1$.

For $\mathbf{a} = (a_0, \ldots, a_d)$ let $p_\mathbf{a} = a_0 + a_1x + \cdots + a_dx^d$. As in the proof of Theorem 3.4.24, $W\mathbf{a} = (p_\mathbf{a}(c_1), \ldots, p_\mathbf{a}(c_m))$.

(a) Suppose $m \geq d+1$, and $W\mathbf{a} = \mathbf{0}$. Then $p_\mathbf{a}$ has degree at most d and vanishes at the $d+1$ distinct points $c_1, \ldots, c_{d+1}$, so $p_\mathbf{a} = \mathbf{0}$ by Lemma 3.4.23, and $\mathbf{a} = \mathbf{0}$. By Proposition 3.4.15, W_r has a pivot in every column. If $m < d+1$, then W has more columns than rows, and by Corollary 3.4.19(a) $W\mathbf{a} = \mathbf{0}$ has a nontrivial solution, so some column of W_r has no pivot.

(b) Suppose $m \leq d+1$, and let $\mathbf{v} \in F^m$. By Theorem 3.4.24 (with $n = m-1$ and the nodes $c_1, \ldots, c_m$) there is a polynomial p of degree at most $m-1 \leq d$ with $p(c_i) = v_i$ for all i. Its coefficients, padded with zeros up to a_d, form a solution of $W\mathbf{a} = \mathbf{v}$. So $W\mathbf{a} = \mathbf{v}$ is consistent for every $\mathbf{v}$, and by Proposition 3.4.16 W_r has a pivot in every row. If $m > d+1$, then W has more rows than columns, and by Corollary 3.4.19(b) some system $W\mathbf{a} = \mathbf{v}$ is inconsistent, so some row of W_r has no pivot.

In terms of polynomials: prescribing values at m distinct points, a polynomial of degree at most d with those values is unique (if it exists) exactly when $m \geq d+1$, and exists for all choices of values exactly when $m \leq d+1$. Both hold exactly when $m = d+1$, which is Theorem 3.4.24.

Solution to Exercise 3.4.40

Uses. Theorem 3.4.24 (Polynomial interpolation), Exercise 1.1.51, Definition 1.2.17 (Evaluation)
Hint 1. $\mathbb{F}_p$ has exactly p elements. Use them all as nodes.
Hint 2. For (b), Fermat's little theorem (Exercise 1.1.51) says $c^{p-1} = 1$ for $c \neq 0$.
Sketch. (a) The p elements $0, 1, \ldots, p-1$ of $\mathbb{F}_p$ are distinct nodes, so Theorem 3.4.24 with $n = p-1$ gives exactly one f of degree at most $p-1$ with $f(c) = g(c)$ for all c. Count: p^p functions and p^p polynomials (one for each list of p coefficients). (b) $f = 1 - x^{p-1}$, by Fermat.

(a) The elements $0, 1, \ldots, p-1$ of $\mathbb{F}_p$ are p distinct nodes. By Theorem 3.4.24 with $n = p-1$ and $d_c = g(c)$, there is exactly one $f \in \mathbb{F}_p[x]_{\leq p-1}$ with $f(c) = g(c)$ for all $c \in \mathbb{F}_p$, that is, whose polynomial function is g. Counting: there are p^p functions $\mathbb{F}_p \to \mathbb{F}_p$ (a value for each of p inputs), and p^p polynomials $a_0 + a_1 x + \cdots + a_{p-1}x^{p-1}$ (a coefficient for each of p positions). The map sending a polynomial to its function is therefore a bijection between two sets with p^p elements, consistent with what we proved.

(b) Let $f = 1 - x^{p-1}$, which has degree at most $p-1$. Then $f(0) = 1$, and for $c \neq 0$ Fermat's little theorem (Exercise 1.1.51) gives $c^{p-1} = 1$, so $f(c) = 0$. By the uniqueness in (a), f is the polynomial.

Polynomials of degree at least p are not needed to describe functions, and indeed $x^p - x$ has the zero function (Exercise 1.2.51); this does not contradict Lemma 3.4.23, since $x^p - x$ has degree p and only p roots are available.

Solution to Exercise 3.4.41

Uses. Theorem 3.4.24 (Polynomial interpolation), Definition 3.4.22 (Vandermonde matrix), Corollary 3.4.4 (Consistency)
Hint 1. The coefficient matrix is $V(1, 2, 3)$: the system asks for a polynomial of degree at most 2 with values $1, 2, 4$ at $1, 2, 3$. When are the nodes distinct in $\mathbb{F}_p$?
Hint 2. For $p = 2$ two nodes coincide. Compare the corresponding equations.
Sketch. For $p \geq 3$, the nodes $1, 2, 3$ are distinct in $\mathbb{F}_p$ ($3-1=2$, $3-2=1$, $2-1=1$ are nonzero), so there is exactly one solution by Theorem 3.4.24. For $p = 2$, the first and third equations read $a_0 + a_1 + a_2 = 1$ and $a_0 + a_1 + a_2 = 0$: no solution.

The coefficient matrix has rows $\left(1, c, c^2\right)$ for $c = 1, 2, 3$, so it is $V(1, 2, 3)$, and the system asks for $p = a_0 + a_1 x + a_2 x^2$ with $p(1) = 1$, $p(2) = 2$, $p(3) = 4$.

If $p \geq 3$, the differences $2-1 = 1$, $3-2 = 1$, and $3-1 = 2$ are nonzero in $\mathbb{F}_p$, so the nodes are distinct, and by Theorem 3.4.24 there is exactly one solution. (Over $\mathbb{Q}$ it is $(1, -1/2, 1/2)$, and the same computation works in $\mathbb{F}_p$ for $p \geq 3$, where 2 is invertible.)

If $p = 2$, then $3 = 1$ and $9 = 1$ in $\mathbb{F}_2$, so the third equation reads $a_0 + a_1 + a_2 = 4 = 0$, while the first reads $a_0 + a_1 + a_2 = 1$. Subtracting gives $0 = 1$, so there is no solution. (In the language of Corollary 3.4.4: the reduced augmented matrix has a row $(0, 0, 0 \mid 1)$.)

So the number of solutions is 1 for every odd prime, and 0 for $p = 2$.

Solution to Exercise 3.4.42

Uses. Corollary 3.4.20 (Square systems), Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations)
Techniques. Uniqueness gives existence (Technique 3.4.21)

Hint 1. The four conditions are linear in a_0, a_1, a_2, a_3, so this is a square system. By Technique 3.4.21 it is enough to treat $u_0 = v_0 = u_1 = v_1 = 0$.
Hint 2. For (b), solve each system directly; then use that the conditions are linear in p.
Sketch. The conditions are $a_0 = u_0$, $a_1 = v_0$, $a_0 + a_1 + a_2 + a_3 = u_1$, $a_1 + 2a_2 + 3a_3 = v_1$. With zero data, $a_0 = a_1 = 0$, $a_2 + a_3 = 0$, $2a_2 + 3a_3 = 0$ force $\mathbf{a} = \mathbf{0}$, so Corollary 3.4.20 applies. $h_1 = 1 - 3x^2 + 2x^3$, $h_2 = x - 2x^2 + x^3$, $h_3 = 3x^2 - 2x^3$, $h_4 = -x^2 + x^3$, and $p = u_0 h_1 + v_0 h_2 + u_1 h_3 + v_1 h_4$.

(a) By Definition 1.2.17, the four conditions are

$$a_0 = u_0, \quad a_1 = v_0, \quad a_0 + a_1 + a_2 + a_3 = u_1, \quad a_1 + 2a_2 + 3a_3 = v_1,$$

a square linear system in a_0, a_1, a_2, a_3. By Corollary 3.4.20 it suffices to show that the homogeneous system has only the trivial solution. If all data are 0, then $a_0 = a_1 = 0$, so $a_2 + a_3 = 0$ and $2a_2 + 3a_3 = 0$; subtracting twice the first from the second gives $a_3 = 0$, and then $a_2 = 0$. So for all data there is exactly one solution.

(b) Solving the four systems:

- $u_0 = 1$: $a_0 = 1, a_1 = 0, a_2 + a_3 = -1, 2a_2 + 3a_3 = 0$, so $a_3 = 2$, $a_2 = -3$: $h_1 = 1 - 3x^2 + 2x^3$.
- $v_0 = 1$: $a_0 = 0, a_1 = 1, a_2 + a_3 = -1, 2a_2 + 3a_3 = -1$, so $a_3 = 1$, $a_2 = -2$: $h_2 = x - 2x^2 + x^3$.
- $u_1 = 1$: $a_0 = a_1 = 0, a_2 + a_3 = 1, 2a_2 + 3a_3 = 0$, so $a_3 = -2$, $a_2 = 3$: $h_3 = 3x^2 - 2x^3$.
- $v_1 = 1$: $a_0 = a_1 = 0, a_2 + a_3 = 0, 2a_2 + 3a_3 = 1$, so $a_3 = 1$, $a_2 = -1$: $h_4 = -x^2 + x^3$.

The four conditions depend linearly on p: values by Proposition 1.2.18(c), and p' because its coefficients are linear in those of p. So $u_0 h_1 + v_0 h_2 + u_1 h_3 + v_1 h_4$ satisfies the conditions with data u_0, v_0, u_1, v_1, and by uniqueness it is p. (These are the cubic Hermite polynomials, used in computer graphics to join curve segments smoothly.)

Solution to Exercise 3.4.43

Uses. Proposition 1.2.18 (Degree and evaluation versus the operations), Corollary 3.4.20 (Square systems), Definition 1.2.16 (Degree)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. Both sides depend linearly on p. So it is enough to require the equation for $p = 1$, $p = x$, and $p = x^2$.
Hint 2. This gives a 3×3 system for w_0, w_1, w_2 whose coefficient matrix is the transpose of $V(-1, 0, 1)$.
Sketch. For $p = 1, x, x^2$ the conditions are $w_0 + w_1 + w_2 = 2$, $-w_0 + w_2 = 0$, $w_0 + w_2 = 2/3$, with the unique solution $(1/3, 4/3, 1/3)$. These three conditions suffice by linearity. For x^3 both sides are 0; for x^4 the integral is $2/5$ and the rule gives $2/3$.

(a) The integral and the right side both depend linearly on p: the right side by Proposition 1.2.18(c), and the integral because $\int (f + g) = \int f + \int g$ and $\int cf = c \int f$. So if the equation holds for $p = 1$, x, and x^2, it holds for every $a_0 + a_1 x + a_2 x^2$; conversely it must hold for these three. Since $\int_{-1}^{1} 1 = 2$, $\int_{-1}^{1} t \, dt = 0$, and $\int_{-1}^{1} t^2 \, dt = 2/3$, the conditions are

$$w_0 + w_1 + w_2 = 2, \quad -w_0 + w_2 = 0, \quad w_0 + w_2 = \frac{2}{3}.$$

The last two give $w_0 = w_2 = 1/3$, and the first gives $w_1 = 4/3$. So the only rule is

$$\int_{-1}^{1} p(t) \, dt = \frac{1}{3} p(-1) + \frac{4}{3} p(0) + \frac{1}{3} p(1),$$

Simpson's rule. (The coefficient matrix of the system is the transpose of $V(-1, 0, 1)$, which is invertible because $V(-1, 0, 1)$ is (Theorem 3.4.24, Proposition 2.4.23). This is why the weights exist and are unique for any three distinct nodes.)

(b) By linearity it is enough to check x^3: the integral is 0, and the rule gives $\frac{1}{3}(-1) + 0 + \frac{1}{3}(1) = 0$. For x^4 the integral is $2/5$, while the rule gives $\frac{1}{3} + \frac{1}{3} = 2/3$.

Solution to Exercise 3.4.44

Uses. Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 3.3.20 (The kernel is a complete invariant), Definition 3.2.21 (Row equivalence), Definition 2.5.12 (Invertible)
Hint 1. Two systems with the same nonempty solution set have the same homogeneous solution set. Why?
Hint 2. So $\ker T_A = \ker T_B$. What does Corollary 3.3.20 give? Then use a common solution $\mathbf{x}_0$ to compare $\mathbf{y}$ and $\mathbf{w}$.
Hint 3. For (b), two inconsistent systems have the same (empty) solution set.
Sketch. If S is the common solution set and $\mathbf{x}_0 \in S$, then $S = \mathbf{x}_0 + \ker T_A = \mathbf{x}_0 + \ker T_B$, so $\ker T_A = \ker T_B$, and Corollary 3.3.20 gives $B = HA$ with H a product of elementary matrices. Then $\mathbf{w} = B\mathbf{x}_0 = HA\mathbf{x}_0 = H\mathbf{y}$, so $(B \mid \mathbf{w}) = H(A \mid \mathbf{y})$. (b) The inconsistent systems $0x = 1, 0x = 0$ and $x = 0, x = 1$ have the same (empty) solution set, but their augmented matrices have different reduced forms.

(a) Let S be the common solution set and $\mathbf{x}_0 \in S$. By Proposition 3.1.19, $S = \mathbf{x}_0 + \ker T_A$ and $S = \mathbf{x}_0 + \ker T_B$. So $\mathbf{z} \in \ker T_A$ if and only if $\mathbf{x}_0 + \mathbf{z} \in S$ if and only if $\mathbf{z} \in \ker T_B$; that is, $\ker T_A = \ker T_B$. By Corollary 3.3.20, $A \sim B$, so $B = HA$ for a product H of elementary matrices (Definition 3.2.21). Then

$$\mathbf{w} = B\mathbf{x}_0 = HA\mathbf{x}_0 = H\mathbf{y},$$

and therefore $(B \mid \mathbf{w}) = (HA \mid H\mathbf{y}) = H(A \mid \mathbf{y})$, by the column-by-column description of HM (Proposition 3.2.14). Since H is a product of elementary matrices, this says $(A \mid \mathbf{y}) \sim (B \mid \mathbf{w})$.

(b) Take $m = 2$, $n = 1$, and the systems

$$0x = 1, \quad 0x = 0 \quad \text{and} \quad x = 0, \quad x = 1.$$

Both are inconsistent, so they have the same (empty) solution set. Their augmented matrices are $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The first is in RREF, and the second reduces to I_2, so their reduced row echelon forms differ and they are not row equivalent (Corollary 3.3.18(b)).

Solution to Exercise 3.4.45

Uses. Corollary 3.4.4 (Consistency), Proposition 3.1.8 (Systems and the map T_A), Proposition 3.1.19 (Solution sets are translates of the kernel), Proposition 3.3.11 (Carrying extra columns along), Definition 2.4.1 (Matrix product)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. Let U be the $n \times r$ matrix with columns $\mathbf{u}_1, \ldots, \mathbf{u}_r$, so that the span is $\operatorname{im} T_U$. By Corollary 3.4.4, $\operatorname{im} T_U$ is cut out by linear equations. Which ones?
Hint 2. If $GU = \operatorname{rref}(U)$ has k pivots, let B consist of the last $n - k$ rows of G. Then $\ker T_B = \operatorname{im} T_U$. Now choose the right-hand side so that $\mathbf{x}_0$ is a solution.
Sketch. Row reduce $(U \mid I_n)$ to $(\operatorname{rref}(U) \mid G)$, with k pivots, and let B be the last $n - k$ rows of G. By Corollary 3.4.4, $B\mathbf{y} = \mathbf{0}$ if and only if $\mathbf{y} \in \operatorname{im} T_U = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$. The system $B\mathbf{x} = B\mathbf{x}_0$ has the solution $\mathbf{x}_0$ and homogeneous solution set $\ker T_B$, so its solution set is $\mathbf{x}_0 + \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$. In (b), B has rows $(-1, 1, 0, 0)$ and $(0, 0, -1, 1)$, and the system is $-x_1 + x_2 = -1$, $-x_3 + x_4 = -2$.

(a) Let $U \in F^{n \times r}$ have columns $\mathbf{u}_1, \ldots, \mathbf{u}_r$, so that $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r) = \operatorname{im} T_U$ by Proposition 3.1.8(a). The procedure:

1. Row reduce $(U \mid I_n)$, scanning only the first r columns, to $(\operatorname{rref}(U) \mid G)$ (Proposition 3.3.11). Let k be the number of pivots.
2. If $k = n$, output the empty system (or $0x_1 = 0$): its solution set is F^n.
3. Otherwise, let $B \in F^{(n-k) \times n}$ consist of the last $n - k$ rows of G, and output the system $B\mathbf{x} = B\mathbf{x}_0$.

Correctness. The jth entry of $B\mathbf{y}$ is the $(k + j)$th entry of $G\mathbf{y}$ (Definition 2.4.1), so $B\mathbf{y} = \mathbf{0}$ if and only if $(G\mathbf{y})_i = 0$ for all $i > k$. By Corollary 3.4.4, applied to U, this holds if and only if $\mathbf{y} \in \operatorname{im} T_U$. So $\ker T_B = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$. (If $k = n$, the conditions are vacuous and $\operatorname{im} T_U = F^n$.) The system $B\mathbf{x} = B\mathbf{x}_0$ has the solution $\mathbf{x}_0$, so by Proposition 3.1.19 its solution set is $\mathbf{x}_0 + \ker T_B = \mathbf{x}_0 + \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_r)$.

(b) Here U has rows $(1, 0)$, $(1, 0)$, $(0, 1)$, $(0, 1)$. Pass 1 adds -1 times row 1 to row 2, which becomes $(0, 0 \mid -1, 1, 0, 0)$. In pass 2, column 2 has a 0 in row 2, so rows 2 and 3 are swapped, and then -1 times the new row 2, which is $(0, 1 \mid 0, 0, 1, 0)$, is added to row 4:

$$\left(\begin{array}{cc|cccc} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array}\right).$$

So $k = 2$, and B has rows $(-1, 1, 0, 0)$ and $(0, 0, -1, 1)$. Since $B\mathbf{x}_0 = (-1, -2)$, the system is

$$-x_1 + x_2 = -1, \quad -x_3 + x_4 = -2.$$

Check: $\mathbf{x}_0$ satisfies both equations, and $\mathbf{u}_1$ and $\mathbf{u}_2$ satisfy the homogeneous ones. Its solution set has two free variables, x_2 and x_4, matching the two parameters of $\mathbf{x}_0 + s\mathbf{u}_1 + t\mathbf{u}_2$. Part (a) of Exercise 3.1.41 found such a system by inspection; the procedure always works.

Solution to Exercise 3.4.46

Uses. Theorem 3.4.24 (Polynomial interpolation)
Hint 1. The nodes $1, \ldots, N$ are distinct in $\mathbb{F}_p$ because $N < p$, and none of them is 0.
Hint 2. For (a), $t + 1$ shares are the values of f at $t + 1$ distinct nodes, and f has degree at most t. For (b), add the condition $f(0) = s'$ to the t shares.
Sketch. (a) f is the unique polynomial of degree at most t with the given $t + 1$ values (Theorem 3.4.24), so the group can compute f and then $s = f(0)$. (b) The conditions $f(0) = s'$ and $f(i_q) = v_q$ prescribe values at the $t + 1$ distinct nodes $0, i_1, \ldots, i_t$, so exactly one f of degree at most t satisfies them, and its coefficients $a_1, \ldots, a_t$ are the unique choice. Every candidate secret is explained by exactly one choice of the random coefficients, so the observed shares are equally likely under every candidate secret, and they carry no information about s.

Since $N < p$, the elements $1, 2, \ldots, N$ of $\mathbb{F}_p$ are distinct and nonzero.

(a) Suppose people $i_0, \ldots, i_t$ pool their shares $f(i_0), \ldots, f(i_t)$. These are the values of f, a polynomial of degree at most t, at $t + 1$ distinct nodes. By Theorem 3.4.24 there is exactly one polynomial of degree at most t with these values, so the group can compute it (by solving the Vandermonde system, or with Proposition 3.4.27), and it is f. Then $s = f(0)$.

(b) Fix $s' \in \mathbb{F}_p$. A dealer with secret s' who chose $a_1, \ldots, a_t$ uses $f' = s' + a_1 x + \cdots + a_t x^t$, and produces the shares $v_1, \ldots, v_t$ exactly when $f'(i_q) = v_q$ for $q = 1, \ldots, t$. So the possible choices correspond to the polynomials f' of degree at most t with

$$f'(0) = s', \quad f'(i_1) = v_1, \quad \ldots, \quad f'(i_t) = v_t.$$

The $t + 1$ nodes $0, i_1, \ldots, i_t$ are distinct, so by Theorem 3.4.24 there is exactly one such f', and hence exactly one choice of $(a_1, \ldots, a_t)$.

If the dealer chooses $(a_1, \ldots, a_t)$ uniformly at random, each of the p^t choices has probability p^{-t}. By what we just proved, for each candidate secret s' exactly one of these choices produces the observed shares, so the observed shares are equally likely under every candidate secret. Seeing them gives no reason to prefer one value of s over another. (This is Shamir's secret sharing scheme.)

Solution to Exercise 3.4.47

Uses. Theorem 3.4.24 (Polynomial interpolation), Definition 1.2.17 (Evaluation), Definition 1.2.16 (Degree), induction on n
Hint 1. Compute p_0, p_1, p_2 and look for a pattern. Then consider the differences $p_n(x + 1) - p_n(x)$.
Hint 2. Show that there is a polynomial q of degree at most $n - 1$ with $q(t) = p_n(t + 1) - p_n(t)$ for all $t \in \mathbb{Q}$. What are its values at $0, 1, \ldots, n - 1$?
Hint 3. By uniqueness, $q = p_{n-1}$. Now induct.
Sketch. Induction on n; $p_0 = 1$ gives $p_0(1) = 1$. Expanding $(t + 1)^j$ by the binomial theorem, $p_n(t + 1) - p_n(t) = q(t)$ for a polynomial q of degree at most $n - 1$ (the t^n terms cancel). Then $q(k) = 2^{k+1} - 2^k = 2^k$ for $k = 0, \ldots, n - 1$, so $q = p_{n-1}$ by Theorem 3.4.24. Hence $p_n(n + 1) = 2^n + p_{n-1}(n)$, and induction gives $2^n + 2^n - 1$.

The polynomials p_n exist and are unique by Theorem 3.4.24, since $0, 1, \ldots, n$ are distinct in $\mathbb{Q}$. We induct on n. For $n = 0$, $p_0 = 1$ and $p_0(1) = 1 = 2^1 - 1$.

Let $n \geq 1$ and write $p_n = \sum_{j=0}^{n} a_j x^j$. For $t \in \mathbb{Q}$, the binomial theorem gives

$$p_n(t+1) - p_n(t) = \sum_{j=0}^{n} a_j\big((t+1)^j - t^j\big) = \sum_{j=1}^{n} a_j \sum_{i=0}^{j-1} \binom{j}{i} t^i = \sum_{i=0}^{n-1} b_i t^i,$$

where $b_i = \sum_{j=i+1}^{n} \binom{j}{i} a_j$. So with

$$q = b_0 + b_1 x + \cdots + b_{n-1} x^{n-1} \in \mathbb{Q}[x]_{\leq n-1}$$

we have $q(t) = p_n(t+1) - p_n(t)$ for every $t \in \mathbb{Q}$. In particular, for $k = 0, 1, \ldots, n-1$,

$$q(k) = p_n(k+1) - p_n(k) = 2^{k+1} - 2^k = 2^k.$$

By the uniqueness in Theorem 3.4.24 (with the n nodes $0, \ldots, n-1$), $q = p_{n-1}$. Therefore, using the induction hypothesis $p_{n-1}(n) = 2^n - 1$,

$$p_n(n+1) = p_n(n) + q(n) = 2^n + p_{n-1}(n) = 2^n + 2^n - 1 = 2^{n+1} - 1,$$

as desired. (For instance $p_2 = 1 + x/2 + x^2/2$ takes the values $1, 2, 4$ at $0, 1, 2$, and $p_2(3) = 1 + 3/2 + 9/2 = 7$.)

Section 3.5

Solution to Exercise 3.5.4

Uses. Definition 3.5.1 (Rank), Proposition 3.5.2 (First properties of rank), Theorem 3.4.24 (Polynomial interpolation), Theorem 3.3.23 (Invertibility and RREF)
Hint 1. For (a) and (c), use Proposition 3.5.2(d): count the columns that are not in the span of the earlier ones. Here every column is a multiple of the first.
Hint 2. For (b) over $\mathbb{Q}$, the nodes $1, 2, 3$ are distinct. What happens to them in $\mathbb{F}_2$?
Sketch. (a) Every column is a multiple of the first, which is nonzero: rank 1. (b) Over $\mathbb{Q}$, V is invertible by Theorem 3.4.24, so its reduced form is I_3 and the rank is 3. Over $\mathbb{F}_2$, rows 1 and 3 coincide and the rank is 2. (c) Rank 1.

(a) The first column $\mathbf{a}_1 = (1, -2, 3)$ is nonzero, and $\mathbf{a}_2 = -\mathbf{a}_1$, $\mathbf{a}_3 = 2\mathbf{a}_1$. By Proposition 3.5.2(d), only $j = 1$ counts, so the rank is 1.

(b) Over $\mathbb{Q}$ the nodes $1, 2, 3$ are distinct, so $V(1, 2, 3)$ is invertible by Theorem 3.4.24. By Theorem 3.3.23 its reduced form is I_3, which has 3 pivots. Over $\mathbb{F}_2$ we have $3 = 1$, $4 = 0$, and $9 = 1$, so the matrix is $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Adding row 1 to rows 2 and 3 gives the rows $(1, 1, 1)$, $(0, 1, 1)$, $(0, 0, 0)$, and adding row 2 to row 1 gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which is in RREF with 2 pivots. So the rank is 2 over $\mathbb{F}_2$: the nodes 1 and 3 coincide in $\mathbb{F}_2$, so two rows of V coincide.

(c) Every column equals the first, which is nonzero. By Proposition 3.5.2(d), the rank is 1.

Solution to Exercise 3.5.5

Uses. Proposition 3.5.2 (First properties of rank), Definition 1.3.10 (Span)
Hint 1. Write each column as a combination of two fixed vectors: $\mathbf{u} = (1, 2, \ldots, n)$ and the all-ones vector $\mathbf{1}$.
Hint 2. Column j is $\mathbf{u} + j\mathbf{1}$. Express column j through columns 1 and 2, and use Proposition 3.5.2(d).
Hint 3. To see that column 2 is not a multiple of column 1 in every field, look at the first two entries of both columns.
Sketch. Column j is $\mathbf{c}_j = \mathbf{u} + j\mathbf{1}$, and $\mathbf{c}_j = (2 - j)\mathbf{c}_1 + (j - 1)\mathbf{c}_2$. So only $j = 1, 2$ can count in Proposition 3.5.2(d). Both do: $\mathbf{c}_1 \neq \mathbf{0}$ because its first two entries $2, 3$ differ by 1, and $\mathbf{c}_2 = t\mathbf{c}_1$ would give $3 = 2t$ and $4 = 3t$, hence $1 = t$ and then $3 = 2$, which is false in every field.

Let $\mathbf{u} = (1, 2, \ldots, n)$ and $\mathbf{1} = (1, 1, \ldots, 1)$ in F^n. Column j of A is

$$\mathbf{c}_j = (1+j, 2+j, \ldots, n+j) = \mathbf{u} + j\mathbf{1}.$$

For every j, in the combination $(2-j)\mathbf{c}_1 + (j-1)\mathbf{c}_2 = (2-j)(\mathbf{u}+\mathbf{1}) + (j-1)(\mathbf{u}+2\mathbf{1})$ the coefficient of $\mathbf{u}$ is $(2-j)+(j-1) = 1$ and that of $\mathbf{1}$ is $(2-j)+2(j-1) = j$. So

$$(2-j)\mathbf{c}_1 + (j-1)\mathbf{c}_2 = \mathbf{u} + j\mathbf{1} = \mathbf{c}_j,$$

and for $j \geq 3$ the column $\mathbf{c}_j$ lies in $\operatorname{span}(\mathbf{c}_1, \mathbf{c}_2)$, hence in $\operatorname{span}\big(\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}\big)$. By Proposition 3.5.2(d), $\operatorname{rank} A \leq 2$, and it remains to see that $j = 1$ and $j = 2$ both count.

The first two entries of $\mathbf{c}_1$ are 2 and 3, and $3 - 2 = 1 \neq 0$, so they are not both 0: $\mathbf{c}_1 \neq \mathbf{0}$. Suppose $\mathbf{c}_2 = t\mathbf{c}_1$ for some $t \in F$. Comparing the first two entries gives $3 = 2t$ and $4 = 3t$. Subtracting, $1 = t$, and then $3 = 2$, that is, $1 = 0$, which is false. So $\mathbf{c}_2 \notin \operatorname{span}(\mathbf{c}_1)$, and $\operatorname{rank} A = 2$.

Solution to Exercise 3.5.11

Uses. Proposition 3.5.8 (Right inverses from the pivots), Proposition 3.3.11 (Carrying extra columns along)
Hint 1. Reduce $(A \mid I_2)$. Only one row operation is needed.
Hint 2. The pivot columns are 1 and 2, so B has its 1s in positions $(1,1)$ and $(2,2)$: BG is G on top of two zero rows.
Sketch. Adding -1 times row 2 to row 1 gives $A_r = \left(\begin{smallmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{smallmatrix}\right)$ and $G = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$. So $\operatorname{rank} A = 2$, $p_1 = 1$, $p_2 = 2$, and $BG = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

Row 1 already has a leading 1 in column 1, and column 1 is clear below it. In pass 2, the entry in position $(2,2)$ is 1, and we add -1 times row 2 to row 1:

$$\left(\begin{array}{cccc|cc} 1 & 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|cc} 1 & 0 & -1 & 1 & 1 & -1 \\ 0 & 1 & 1 & 1 & 0 & 1 \end{array}\right).$$

So $\operatorname{rank} A = 2$, the pivot columns are 1 and 2, and $G = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$. The matrix $B \in \mathbb{Q}^{4\times 2}$ has 1s in positions $(1,1)$ and $(2,2)$, so

$$BG = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Check: $A(BG)$ is A times the columns $(1,0,0,0)$ and $(-1,1,0,0)$, that is, $\mathbf{a}_1 = (1,0)$ and $-\mathbf{a}_1 + \mathbf{a}_2 = (-1,0) + (1,1) = (0,1)$. So $A(BG) = I_2$.

Solution to Exercise 3.5.18

Uses. Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Theorem 3.5.14 (Third Fundamental Theorem, completed), Corollary 3.5.15 (Invertible matrices are square)
Hint 1. Each "moreover" clause is a constraint on the shape. Check it first.
Hint 2. For (d), compare with condition (VI) of Theorem 3.5.14.
Sketch. (a) None: injective needs $n \leq m$, but $5 > 3$. (b) $A = (I_3 \;\; 0)$, rank $3 = m$. (c) None: a right inverse needs $n \geq m$, but $3 < 5$. (d) None: rank $A = 4 = m = n$ makes T_A invertible.

(a) None. By Theorem 3.5.12, T_A injective forces $n \leq m$, and here $n = 5 > 3 = m$. (Concretely, $\operatorname{rank} A \leq 3 < 5$.)

(b) $A = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{smallmatrix}\right)$ is in RREF with 3 pivots, so $\operatorname{rank} A = 3 = m$ and T_A is surjective by Theorem 3.5.13.

(c) None. By Corollary 3.5.15(b), a matrix with a right inverse has $n \geq m$, and here $n = 3 < 5 = m$. The condition $\operatorname{rank} A = 3$ is possible, and it gives a left inverse instead.

(d) None. If $\operatorname{rank} A = 4$, then $\operatorname{rank} A = m = n$, so T_A is invertible by Theorem 3.5.14, and so is A (Remark 3.5.7).

Solution to Exercise 3.5.26

Uses. Corollary 3.5.15 (Invertible matrices are square), Theorem 3.5.23 (One-sided inverses of square matrices)
Hint 1. Cancelling A on the left from $AX = AY$ means multiplying on the left by a left inverse of A. Does A have one?
Hint 2. Try $A = \begin{pmatrix}1 & 0\end{pmatrix}$.
Sketch. Cancelling A on the left needs a left inverse of A, which a matrix with a right inverse has only when it is square. Counterexample: $A = \begin{pmatrix}1 & 0\end{pmatrix}$, $B = \begin{pmatrix}1\\0\end{pmatrix}$: $AB = I_1$, $BA = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \neq I_2$.

The step "cancelling A on the left" is the error. From $AX = AY$ we may conclude $X = Y$ if A has a left inverse C, by multiplying by C; in general we may not, since T_A need not be injective. Here A has a right inverse B, so $n \geq m$ by Corollary 3.5.15(b), and if $n > m$ then A has no left inverse at all, by Corollary 3.5.15(a).

The conclusion is false: for $A = \begin{pmatrix}1 & 0\end{pmatrix} \in F^{1\times 2}$ and $B = \begin{pmatrix}1\\0\end{pmatrix} \in F^{2\times 1}$, we have $AB = (1) = I_1$, but

$$BA = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \neq I_2.$$

When $m = n$, the conclusion is true, but its proof needs Theorem 3.5.23, not cancellation.

Solution to Exercise 3.5.30

Uses. Theorem 3.5.27 (Rank normal form), Proposition 3.2.7 (Row operations are left multiplications)
Hint 1. Compare row 2 with i times row 1.
Hint 2. One row operation reduces A. Then follow the first proof of Theorem 3.5.27: the only pivot column is column 1, so $P = I_3$.
Sketch. Row 2 is i times row 1, so $E_{1,-i,2}$ gives $A_r = \begin{pmatrix}1 & i & 1+i\\0 & 0 & 0\end{pmatrix}$ and $k = 1$, with $G = \begin{pmatrix}1 & 0\\-i & 1\end{pmatrix}$. Then $X = \begin{pmatrix}i & 1+i\end{pmatrix}$ and $H = U = \begin{pmatrix}1 & -i & -1-i\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}$.

Since $i(1, i, 1+i) = (i, -1, i-1)$, row 2 is i times row 1. Adding $-i$ times row 1 to row 2 gives

$$A_r = \begin{pmatrix}1 & i & 1+i\\0 & 0 & 0\end{pmatrix}, \quad G = E_{1,-i,2} = \begin{pmatrix}1 & 0\\-i & 1\end{pmatrix}.$$

So $k = \operatorname{rank} A = 1$, the pivot column is 1, and the free columns are $2, 3$. As in the first proof of Theorem 3.5.27, $P = I_3$, $X = \begin{pmatrix}i & 1+i\end{pmatrix}$, and

$$H = U = \begin{pmatrix}1 & -i & -1-i\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}.$$

Check: $A_r H$ has first column $(1, 0)$, second column $(-i+i, 0) = \mathbf{0}$, and third column $(-1-i+1+i, 0) = \mathbf{0}$. So $GAH = A_r H = \begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\end{pmatrix} = N_1$.

Solution to Exercise 3.5.31

Uses. Proposition 3.5.2 (First properties of rank), Definition 3.5.1 (Rank), Algorithm 3.3.1 (Gauss–Jordan elimination), Proposition 1.1.17 (No zero divisors)
Hint 1. Row operations do not change the rank (Proposition 3.5.2(c)), so you may reorder the rows first. Put the row $(1, 1, t)$ on top.
Hint 2. After clearing the first column, add the new second row to the third. The last entry becomes $2 - t - t^2 = -(t-1)(t+2)$.
Hint 3. Treat the cases $t = 1$, $t = -2$, and the rest separately. In $\mathbb{F}_3$, which of these cases coincide?
Sketch. Swapping rows 1 and 3 and eliminating gives rows $(1, 1, t)$, $(0, t-1, 1-t)$, $(0, 0, -(t-1)(t+2))$. (a) Rank 1 for $t = 1$, rank 2 for $t = -2$, rank 3 otherwise. (b) In $\mathbb{F}_3$, $-2 = 1$, so the two exceptional values merge: rank 1 for $t = 1$ and rank 3 for $t = 0, 2$. Rank 2 never occurs.

By Proposition 3.5.2(c), row operations do not change the rank, so we may use any convenient sequence of them; at the end we reduce to RREF and count pivots. Swap rows 1 and 3, add -1 times the new row 1 to row 2, and add $-t$ times it to row 3:

$$A_t \rightsquigarrow \begin{pmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & t \\ 0 & t-1 & 1-t \\ 0 & 1-t & 1-t^2 \end{pmatrix}.$$

Adding row 2 to row 3 makes the last entry $(1-t)+\left(1-t^2\right) = 2-t-t^2 = -(t-1)(t+2)$:

$$\rightsquigarrow \begin{pmatrix} 1 & 1 & t \\ 0 & t-1 & 1-t \\ 0 & 0 & -(t-1)(t+2) \end{pmatrix}.$$

These operations make sense over every field. Call this matrix M_t.

(a) $t = 1$. Rows 2 and 3 of M_1 are zero, and $M_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is in RREF. So rank $A_1 = 1$.

$t \neq 1$. Scale row 2 by $1/(t-1)$, making it $(0, 1, -1)$, and add -1 times it to row 1, which becomes $(1, 0, t+1)$. If $t = -2$, row 3 is zero, and the result $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ is in RREF with 2 pivots, so rank $A_{-2} = 2$. If $t \neq 1, -2$, the entry $-(t-1)(t+2)$ is nonzero, so scaling row 3 and clearing column 3 gives I_3, and rank $A_t = 3$.

(b) In $\mathbb{F}_3$ we have $-2 = 1$, so $t+2 = t-1$, and the last entry of M_t is $-(t-1)^2$. For $t = 1$ the computation in (a) gives rank 1. For $t = 0$ and $t = 2$, the entry $t-1$ is nonzero and so is $-(t-1)^2$, by Proposition 1.1.17, so the computation in (a) gives rank 3. Rank 2 never occurs: over $\mathbb{R}$ it requires $t = -2$ with $t \neq 1$, but in $\mathbb{F}_3$ the values -2 and 1 are the same. (Over an arbitrary field, rank $A_t = 2$ exactly when $t = -2 \neq 1$, that is, when $t = -2$ and $3 \neq 0$ in F.)

Solution to Exercise 3.5.32

> *Uses.* Proposition 3.5.9 (Left inverses), Proposition 3.3.11 (Carrying extra columns along), Exercise 3.5.33
> *Hint 1.* Reduce $(A \mid I_3)$ over $\mathbb{F}_5$, scanning only the first two columns. You will need $3^{-1} = 2$.
> *Hint 2.* The first two rows of the right block form a left inverse. The third row $\mathbf{z}$ of the right block satisfies $\mathbf{z}A = \mathbf{0}$; add it to a row of the left inverse.
> *Sketch.* Reducing $(A \mid I_3)$ gives $A_r = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and G with rows $(3, 1, 0)$, $(4, 2, 0)$, $(1, 3, 1)$. So rank $A = 2$ and $C = \begin{pmatrix} 3 & 1 & 0 \\ 4 & 2 & 0 \end{pmatrix}$ is a left inverse. Adding the third row of G to the first row of C gives another, $\begin{pmatrix} 4 & 4 & 1 \\ 4 & 2 & 0 \end{pmatrix}$. There are $5^{2\cdot 1} = 25$ left inverses.

We reduce $(A \mid I_3)$ over $\mathbb{F}_5$, scanning the first two columns. Pass 1: the pivot is $(1, 1)$, and we add $-3 = 2$ times row 1 to row 2. Pass 2: the entry in position $(2, 2)$ is now $4 + 4 = 8 = 3$; we scale row 2 by $3^{-1} = 2$, then add $-2 = 3$ times row 2 to row 1 and $-1 = 4$ times row 2 to row 3:

$$\left(\begin{array}{cc|ccc} 1 & 2 & 1 & 0 & 0 \\ 3 & 4 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|ccc} 1 & 2 & 1 & 0 & 0 \\ 0 & 3 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cc|ccc} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 4 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cc|ccc} 1 & 0 & 3 & 1 & 0 \\ 0 & 1 & 4 & 2 & 0 \\ 0 & 0 & 1 & 3 & 1 \end{array}\right).$$

So $A_r = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, rank $A = 2$, and by Proposition 3.3.11 the right block is G with $GA = A_r$. By the first approach to Proposition 3.5.9, the first two rows of G form a left inverse

$$C = \begin{pmatrix} 3 & 1 & 0 \\ 4 & 2 & 0 \end{pmatrix}.$$

Check: $3(1,2) + 1(3,4) = (6,10) = (1,0)$ and $4(1,2) + 2(3,4) = (10,16) = (0,1)$ in $\mathbb{F}_5$, so $CA = I_2$.

The third row $\mathbf{z} = (1,3,1)$ of G satisfies $\mathbf{z}A = \mathbf{0}$, since $\mathbf{z}A$ is row 3 of $GA = A_r$. So adding $\mathbf{z}$ to the first row of C gives another left inverse,

$$C' = \begin{pmatrix} 4 & 4 & 1 \\ 4 & 2 & 0 \end{pmatrix},$$

and indeed $4(1,2) + 4(3,4) + (0,1) = (16,25) = (1,0)$. By Exercise 3.5.33(c), with $q = 5$, $m = 3$, $n = 2$, there are $5^{2 \cdot 1} = 25$ left inverses: each row of C may be changed by any of the 5 multiples of $\mathbf{z}$.

Solution to Exercise 3.5.33

> *Uses.* Proposition 3.5.9 (Left inverses), Proposition 3.5.8 (Right inverses from the pivots), Corollary 3.4.14 (Counting solutions), Definition 2.4.1 (Matrix product), Corollary 3.5.15 (Invertible matrices are square), Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 3.5.2 (First properties of rank)
> *Hint 1.* For (a) and (b), subtract: $AX = I_m = AX_0$ if and only if $A(X - X_0) = 0$, and similarly on the left.
> *Hint 2.* For (b), a row vector $\mathbf{z}$ satisfies $\mathbf{z}A = \mathbf{0}$ if and only if $\mathbf{w}A_r = \mathbf{0}$, where $\mathbf{w} = \mathbf{z}G^{-1}$. What is $\mathbf{w}A_r$ when $A_r = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$?
> *Hint 3.* For (c), count the possible columns of K with Corollary 3.4.14(b), and the possible rows of $C - C_0$ using that the rows of G are linearly independent.
> *Sketch.* (a) $AX = I_m$ iff $A(X - X_0) = 0$. (b) $CA = I_n$ iff $ZA = 0$ for $Z = C - C_0$. A row $\mathbf{z}$ has $\mathbf{z}A = \mathbf{0}$ iff $\mathbf{w} = \mathbf{z}G^{-1}$ has $\mathbf{w}A_r = (w_1, \ldots, w_n) = \mathbf{0}$, iff $\mathbf{z} = \mathbf{w}G$ is a combination of rows $n+1, \ldots, m$ of G. (c) $\ker T_A$ has q^{n-m} elements, and there are m columns; the span of $m - n$ independent rows has q^{m-n} elements, and there are n rows. (d) If $n > m$, a nonzero kernel vector perturbs a right inverse; if $m > n$, row m of G perturbs a left inverse; if $m = n$, the inverse is unique.

(a) If $AX_0 = I_m$, then $AX = I_m$ if and only if $A(X - X_0) = AX - AX_0 = 0$. By Definition 2.4.1, column j of AK is $A\mathbf{k}_j$, so $AK = 0$ says exactly that each column $\mathbf{k}_j$ of K lies in $\ker T_A$.

(b) By the first approach to Proposition 3.5.9, $C_0 A = I_n$, and $A_r = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. So $CA = I_n$ if and only if $ZA = 0$, where $Z = C - C_0$, that is, if and only if every row $\mathbf{z}$ of Z satisfies $\mathbf{z}A = \mathbf{0}$. Let $\mathbf{z} \in F^{1 \times m}$ and $\mathbf{w} = \mathbf{z}G^{-1}$, so that $\mathbf{z} = \mathbf{w}G$. Since $A = G^{-1}A_r$,

$$\mathbf{z}A = \mathbf{z}G^{-1}A_r = \mathbf{w}A_r = \mathbf{w}\begin{pmatrix} I_n \\ 0 \end{pmatrix} = (w_1, \ldots, w_n).$$

So $\mathbf{z}A = \mathbf{0}$ if and only if $w_1 = \cdots = w_n = 0$, if and only if $\mathbf{z} = \mathbf{w}G = w_{n+1}\mathbf{g}_{n+1} + \cdots + w_m\mathbf{g}_m$, where $\mathbf{g}_i$ is row i of G (Definition 2.4.1). This is the claim.

(c) *Right inverses.* Fix a right inverse X_0 (one exists, Proposition 3.5.8). By (a), the right inverses correspond bijectively to the matrices K whose m columns lie in $\ker T_A$. The system $A\mathbf{x} = \mathbf{0}$ is consistent with $n - m$ free variables (Proposition 3.5.2(b)), so $\ker T_A$ has q^{n-m} elements by Corollary 3.4.14(b). Choosing the m columns independently gives $(q^{n-m})^m = q^{m(n-m)}$ matrices K.

Left inverses. The rows $\mathbf{g}_{n+1}, \ldots, \mathbf{g}_m$ of G are linearly independent: if $\mathbf{c}G = \mathbf{0}$ for a row vector $\mathbf{c}$, then $\mathbf{c} = \mathbf{c}GG^{-1} = \mathbf{0}$. So distinct coefficient lists $(w_{n+1}, \ldots, w_m)$ give distinct combinations, and their span has exactly q^{m-n} elements. By (b), a left inverse is determined by choosing each of the n rows of $C - C_0$ in this span, which gives $(q^{m-n})^n = q^{n(m-n)}$ left inverses.

(d) If A is square and has a one-sided inverse, it is invertible and every one-sided inverse equals A^{-1} (Theorem 3.5.23). Suppose A is not square. If A has a right inverse X_0, then $n > m$ (Corollary 3.5.15(b)), so $\operatorname{rank} A = m < n$ and A has a free column; the corresponding basic solution $\mathbf{s} \neq \mathbf{0}$ lies in $\ker T_A$, and by (a) the matrix $X_0 + \mathbf{s}\mathbf{e}_1^{\mathsf{T}}$, whose first column is changed by $\mathbf{s}$, is a second right inverse. If A has a left inverse, then $m > n$, so G has a row $\mathbf{g}_m$ beyond the first n, and $\mathbf{g}_m \neq \mathbf{0}$ as G is invertible. By (b), adding $\mathbf{g}_m$ to the first row of C_0 gives a second left inverse.

Solution to Exercise 3.5.34

Uses. Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.13 (Second Fundamental Theorem, completed), Theorem 2.4.2 (The matrix of a composite), Proposition 2.4.23 (Transpose of a product), Remark 3.5.7 (Inverses of T_A and inverses of A), Theorem 3.5.14 (Third Fundamental Theorem, completed)
Hint 1. Translate rank conditions into injectivity or surjectivity of T_A, T_B, and $T_{AB} = T_A \circ T_B$.
Hint 2. For (c), think of a matrix that kills its own image. For (d), transpose a left inverse. For (e), try $A = 0$.
Hint 3. For (f), which of T_A and T_B fails to be injective?
Sketch. (a) True: a composition of injective maps is injective. (b) True: a composition of surjective maps is surjective. (c) False: $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has rank 1 and $A^2 = 0$. (d) True: if $CA = I_n$, then $A^{\mathsf{T}}C^{\mathsf{T}} = I_n$, so $T_{A^{\mathsf{T}}}$ is surjective onto F^n. (e) False: for $A = 0 \in F^{2\times 2}$, every pair of invertible matrices works. (f) True: $T_B : F^m \to F^n$ is not injective, hence neither is $T_A \circ T_B$.

(a) True. By Theorem 3.5.12, T_A and T_B are injective. If $(T_A \circ T_B)(\mathbf{x}) = (T_A \circ T_B)(\mathbf{x}')$, then $T_B\mathbf{x} = T_B\mathbf{x}'$ by injectivity of T_A, and then $\mathbf{x} = \mathbf{x}'$. So $T_{AB} = T_A \circ T_B$ (Theorem 2.4.2(b)) is injective, and since $AB \in F^{m\times p}$, Theorem 3.5.12 gives $\operatorname{rank}(AB) = p$.

(b) True. By Theorem 3.5.13, $T_A : F^n \to F^m$ and $T_B : F^p \to F^n$ are surjective. Given $\mathbf{y} \in F^m$, choose $\mathbf{u} \in F^n$ with $A\mathbf{u} = \mathbf{y}$ and then $\mathbf{x} \in F^p$ with $B\mathbf{x} = \mathbf{u}$; then $AB\mathbf{x} = \mathbf{y}$. So T_{AB} is surjective, and $\operatorname{rank}(AB) = m$ by Theorem 3.5.13.

(c) False. For $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, which is in RREF with one pivot, $\operatorname{rank} A = 1$, but $A^2 = 0$ has rank 0.

(d) True. By Theorem 3.5.12 and Remark 3.5.7, A has a left inverse $C \in F^{n\times m}$, $CA = I_n$. By Proposition 2.4.23, $A^{\mathsf{T}}C^{\mathsf{T}} = (CA)^{\mathsf{T}} = I_n^{\mathsf{T}} = I_n$. So $A^{\mathsf{T}} \in F^{n\times m}$ has a right inverse, and by Theorem 3.5.13 (applied to A^{T}, which has n rows), $\operatorname{rank} A^{\mathsf{T}} = n$.

(e) False. Let $A = 0 \in F^{2\times 2}$, so $k = 0$ and $N_0 = 0$. Then $GAH = 0 = N_0$ for every pair of invertible G, H, for instance $G = H = I_2$ and $G = H = E_{12}$.

(f) True. Since $m > n$, the map $T_B : F^m \to F^n$ is not injective, by the "moreover" part of Theorem 3.5.12: there is $\mathbf{x} \neq \mathbf{0}$ with $B\mathbf{x} = \mathbf{0}$. Then $AB\mathbf{x} = \mathbf{0}$, so T_{AB} is not injective, and the $m \times m$ matrix AB is not invertible (Theorem 3.5.14).

Solution to Exercise 3.5.35

Uses. Corollary 3.5.15 (Invertible matrices are square), Theorem 3.5.23 (One-sided inverses of square matrices), Definition 2.5.12 (Invertible)
Hint 1. Which step assumes something that has not been shown?
Hint 2. For (b), take $n = 1$ and $m = 2$.
Hint 3. For (c), if $P = I_m$, then B is both a left and a right inverse of A. What does Corollary 3.5.15 say?
Sketch. (a) P^{-1} need not exist; the argument assumes the conclusion that P is invertible. (b) $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \end{pmatrix}$: $BA = I_1$ but $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. (c) $P^2 = P$ as computed. If $m = n$, then $AB = I_n$ by Theorem 3.5.23; if $AB = I_m$, then A is invertible, so $m = n$ by Corollary 3.5.15.

(a) The computation $P^2 = P$ is correct. The error is the last step: it multiplies by P^{-1}, but nothing shows that P is invertible. In fact, an invertible P with $P^2 = P$ is I_m, so assuming that P is invertible is assuming the conclusion.

(b) Let $A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in F^{2\times 1}$ and $B = \begin{pmatrix} 1 & 0 \end{pmatrix} \in F^{1\times 2}$. Then $BA = (1) = I_1$, but $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq I_2$.

(c) $P^2 = P$ holds whenever $BA = I_n$, by the computation in the "proof". If $m = n$, then $BA = I_n$ gives $AB = I_n$ by Theorem 3.5.23. Conversely, if $P = AB = I_m$, then $AB = I_m$ and $BA = I_n$, so A is invertible (Definition 2.5.12) and $m = n$ by Corollary 3.5.15(c). So for $m > n$ (the only other possibility, by Corollary 3.5.15(a)), P is an idempotent different from I_m; in (b) it is the projection $(x_1, x_2) \mapsto (x_1, 0)$.

Solution to Exercise 3.5.36

Uses. Proposition 3.5.2 (First properties of rank), Definition 2.4.33 (Block matrix), Definition 1.3.10 (Span)
Hint 1. Use Proposition 3.5.2(d): the rank counts the columns that are not combinations of earlier columns.

Hint 2. A combination of earlier columns of $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ has top block a combination of columns of A and bottom block a combination of columns of B. Treat the first n columns and the last q columns separately.
Hint 3. For (b), only the bottom block of a column $(\mathbf{c}_l; \mathbf{b}_l)$ is needed to show it is new. For strictness, try 1×1 blocks.
Sketch. (a) For $j \le n$, $(\mathbf{a}_j; \mathbf{0})$ is a combination of earlier columns iff $\mathbf{a}_j$ is one of $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$. For $(\mathbf{0}; \mathbf{b}_l)$, a combination $\sum c_i(\mathbf{a}_i; \mathbf{0}) + \sum_{i<l} d_i(\mathbf{0}; \mathbf{b}_i)$ equals it iff $\sum c_i \mathbf{a}_i = \mathbf{0}$ and $\sum d_i \mathbf{b}_i = \mathbf{b}_l$, which is possible iff $\mathbf{b}_l \in \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{l-1})$. Count with Proposition 3.5.2(d). (b) The same count works for the first n columns, and a column $(\mathbf{c}_l; \mathbf{b}_l)$ with $\mathbf{b}_l$ new is new, by looking at bottom blocks. Strict: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Write $\mathbf{a}_1, \ldots, \mathbf{a}_n$ for the columns of A, $\mathbf{b}_1, \ldots, \mathbf{b}_q$ for those of B, and $\mathbf{c}_1, \ldots, \mathbf{c}_q$ for those of C. For column vectors $\mathbf{u} \in F^m$ and $\mathbf{v} \in F^p$, write $(\mathbf{u}; \mathbf{v}) \in F^{m+p}$ for the column with $\mathbf{u}$ on top of $\mathbf{v}$. By Proposition 3.5.2(d), the rank of a matrix is the number of its *new* columns, those not in the span of the columns before them.

(a) The columns of $M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ are $(\mathbf{a}_1; \mathbf{0}), \ldots, (\mathbf{a}_n; \mathbf{0}), (\mathbf{0}; \mathbf{b}_1), \ldots, (\mathbf{0}; \mathbf{b}_q)$. For $j \le n$, the combinations of the earlier columns are the vectors $\left(\sum_{i<j} c_i \mathbf{a}_i; \mathbf{0}\right)$, so $(\mathbf{a}_j; \mathbf{0})$ is new in M if and only if $\mathbf{a}_j$ is new in A. For $1 \le l \le q$, the combinations of the columns before $(\mathbf{0}; \mathbf{b}_l)$ are the vectors

$$\left(\sum_{i=1}^{n} c_i \mathbf{a}_i; \sum_{i<l} d_i \mathbf{b}_i\right).$$

Such a vector equals $(\mathbf{0}; \mathbf{b}_l)$ for some choice of scalars if and only if $\mathbf{b}_l \in \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{l-1})$: if $\mathbf{b}_l = \sum_{i<l} d_i \mathbf{b}_i$, take all $c_i = 0$, and conversely the bottom block gives such an expression. So $(\mathbf{0}; \mathbf{b}_l)$ is new in M if and only if $\mathbf{b}_l$ is new in B. Counting, $\operatorname{rank} M = \operatorname{rank} A + \operatorname{rank} B$.

(b) For $M' = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, the first n columns are as in (a), so exactly $\operatorname{rank} A$ of them are new. Let $1 \le l \le q$ with $\mathbf{b}_l$ new in B. Every combination of the columns of M' before $(\mathbf{c}_l; \mathbf{b}_l)$ has bottom block $\sum_{i<l} d_i \mathbf{b}_i$, which is not $\mathbf{b}_l$. So $(\mathbf{c}_l; \mathbf{b}_l)$ is new in M'. Hence M' has at least $\operatorname{rank} A + \operatorname{rank} B$ new columns. For strict inequality, take 1×1 blocks $A = B = (0)$ and $C = (1)$: then $M' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has rank $1 > 0 + 0$.

Solution to Exercise 3.5.37

Uses. Corollary 3.4.4 (Consistency), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.8 (Properties of the output), Definition 2.4.1 (Matrix product)
Hint 1. If both held, compute $\mathbf{z}^{\mathsf{T}} A \mathbf{x}$ in two ways.
Hint 2. If (i) fails, Corollary 3.4.4 gives an $i > \operatorname{rank} A$ with $(G\mathbf{y})_i \neq 0$. Take $\mathbf{z}^{\mathsf{T}}$ to be row i of G.
Sketch. Both cannot hold: $0 = \left(\mathbf{z}^{\mathsf{T}} A\right)\mathbf{x} = \mathbf{z}^{\mathsf{T}}\mathbf{y} \neq 0$. If (i) fails, some $i > \operatorname{rank} A$ has $(G\mathbf{y})_i \neq 0$; with $\mathbf{z}^{\mathsf{T}}$ the ith row of G, $\mathbf{z}^{\mathsf{T}} A$ is row i of A_r, which is zero, and $\mathbf{z}^{\mathsf{T}}\mathbf{y} = (G\mathbf{y})_i$. For Example 3.4.6, $\mathbf{z} = (-1, -1/2, 1)$ and $\mathbf{z}^{\mathsf{T}}\mathbf{y} = 1$.

Not both. If $A\mathbf{x} = \mathbf{y}$ and $\mathbf{z}^{\mathsf{T}} A = \mathbf{0}$, then $\mathbf{z}^{\mathsf{T}}\mathbf{y} = \mathbf{z}^{\mathsf{T}}(A\mathbf{x}) = \left(\mathbf{z}^{\mathsf{T}} A\right)\mathbf{x} = 0$, so (ii) fails.

At least one. Suppose (i) fails. Let G be invertible with $GA = A_r$, and $k = \operatorname{rank} A$. By Corollary 3.4.4 there is an i with $k < i \le m$ and $(G\mathbf{y})_i \neq 0$. Let $\mathbf{z}^{\mathsf{T}}$ be row i of G. By Definition 2.4.1, $\mathbf{z}^{\mathsf{T}} A$ is row i of $GA = A_r$, which is zero since $i > k$ (Proposition 3.3.8(a)), and $\mathbf{z}^{\mathsf{T}}\mathbf{y}$ is the ith entry of $G\mathbf{y}$, which is nonzero. So (ii) holds.

The example. In Example 3.4.6, $k = 2$ and the last row of G is $(-1, -1/2, 1)$. So $\mathbf{z} = (-1, -1/2, 1)$, and indeed $\mathbf{z}^{\mathsf{T}} A = -(0, 0, 2, 4) - \frac{1}{2}(2, 4, 2, 0) + (1, 2, 3, 4) = \mathbf{0}$ while $\mathbf{z}^{\mathsf{T}}\mathbf{y} = -4 - 1 + 6 = 1$. In words: minus the first equation, minus half the second, plus the third, reads $0 = 1$. Such a $\mathbf{z}$ is a certificate of inconsistency that anyone can check without redoing the elimination.

Solution to Exercise 3.5.38

Uses. Theorem 3.5.23 (One-sided inverses of square matrices)
Hint 1. Group ABC as $A(BC)$, or as $(AB)C$, and apply Theorem 3.5.23.
Hint 2. For (c), take $C = (AB)^{-1}$ with A and B invertible and $AB \neq BA$.
Sketch. (a) $A(BC) = I_n$ gives $(BC)A = I_n$. (b) $(AB)C = I_n$ gives $C(AB) = I_n$. (c) False: with $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and $C = (AB)^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, we get $BAC = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix}$.

(a) True. $A(BC) = I_n$, so by Theorem 3.5.23, $(BC)A = I_n$.

(b) True. $(AB)C = I_n$, so by Theorem 3.5.23, $C(AB) = I_n$.

(c) False. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then $AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, whose inverse is $C = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ (check: $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$). So $ABC = I_2$. But $BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, and

$$BAC = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix} \neq I_2.$$

This works over every field, since the $(1,1)$ entry is $0 \neq 1$. In general $BAC = BA(AB)^{-1}$, which is I_n only if $BA = AB$. Cyclic shifts of a product equal to I_n are again I_n, but other rearrangements need not be.

Solution to Exercise 3.5.39

Uses. Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 2.4.14 (Algebra of matrices)
Techniques. Uniqueness gives existence (Technique 3.4.21)
Hint 1. Rewrite $AB - A - B = 0$ so that it becomes a product equal to I_n. Add I_n to both sides.
Hint 2. $(A - I_n)(B - I_n) = I_n$. Now use that a one-sided inverse of a square matrix is two-sided.
Sketch. $(A - I_n)(B - I_n) = AB - A - B + I_n = I_n$. By Theorem 3.5.23, $(B - I_n)(A - I_n) = I_n$ as well, that is, $BA - A - B + I_n = I_n$. So $BA = A + B = AB$.

Expanding with the distributive laws (Proposition 2.4.14),

$$\begin{aligned}(A - I_n)(B - I_n) &= AB - A - B + I_n \\ &= (A + B) - A - B + I_n = I_n.\end{aligned}$$

So $B - I_n$ is a right inverse of the square matrix $A - I_n$. By Theorem 3.5.23 it is also a left inverse:

$$I_n = (B - I_n)(A - I_n) = BA - B - A + I_n.$$

Hence $BA = A + B = AB$, as desired.

The proof computes nothing: the commutation comes entirely from Theorem 3.5.23, applied to the square matrix $A - I_n$.

Solution to Exercise 3.5.40

Uses. Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Proposition 1.5.14 (Appending a vector), Proposition 1.4.14 (The span is the smallest subspace containing a list), Definition 1.6.1 (Basis), Theorem 2.2.1 (Linear maps are determined by a basis)
Techniques. Check it on a basis (Technique 2.2.7), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. For (a), imitate the second approach to Proposition 3.5.9. First show that $T\mathbf{v}_1, \ldots, T\mathbf{v}_n$ is linearly independent.
Hint 2. Append to $T\mathbf{v}_1, \ldots, T\mathbf{v}_n$ each $\mathbf{w}_i$ that is not yet in the span. Then define S on the resulting basis of W.
Hint 3. For (b), choose a preimage of each $\mathbf{c}_i$ and define R on the basis $\mathbf{c}_1, \ldots, \mathbf{c}_m$.
Sketch. (a) $T\mathbf{v}_1, \ldots, T\mathbf{v}_n$ is independent since $\ker T = \{\mathbf{0}\}$. Appending each $\mathbf{w}_i$ not yet in the span keeps independence (Proposition 1.5.14) and ends with a list spanning W, hence a basis. Define $S(T\mathbf{v}_j) = \mathbf{v}_j$ and $S = \mathbf{0}$ on the appended vectors; $S \circ T$ fixes each $\mathbf{v}_j$, so it is id_V. (b) Pick $\mathbf{u}_i$ with $T\mathbf{u}_i = \mathbf{c}_i$ and let $R(\mathbf{c}_i) = \mathbf{u}_i$; $T \circ R$ fixes each $\mathbf{c}_i$, so it is id_W.

(a) *The images are independent.* If $c_1 T\mathbf{v}_1 + \cdots + c_n T\mathbf{v}_n = \mathbf{0}$, then $T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = \mathbf{0}$ by linearity, so $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$, because $\ker T = \{\mathbf{0}\}$ (Theorem 2.6.11), and all $c_j = 0$ because $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is independent.

Extending to a basis of W. Let $\mathcal{L}_0 = T\mathbf{v}_1, \ldots, T\mathbf{v}_n$, and for $i = 1, \ldots, p$ let $\mathcal{L}_i = \mathcal{L}_{i-1}$ if $\mathbf{w}_i \in \operatorname{span}(\mathcal{L}_{i-1})$, and $\mathcal{L}_i = \mathcal{L}_{i-1}, \mathbf{w}_i$ otherwise. Each $\mathcal{L}_i$ is independent by Proposition 1.5.14, and every $\mathbf{w}_i$ lies in $\operatorname{span}(\mathcal{L}_p)$. So $\operatorname{span}(\mathcal{L}_p)$ is a subspace containing a spanning list of W, hence equals W (Proposition 1.4.14), and $\mathcal{L}_p = T\mathbf{v}_1, \ldots, T\mathbf{v}_n, \mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_t}$ is a basis of W (Definition 1.6.1).

Defining S. By Theorem 2.2.1 there is a linear $S : W \to V$ with $S(T\mathbf{v}_j) = \mathbf{v}_j$ for each j and $S(\mathbf{w}_{i_s}) = \mathbf{0}$ for each s. Then $(S \circ T)(\mathbf{v}_j) = \mathbf{v}_j$ for every j, so the linear maps $S \circ T$ and id_V agree on a basis of V and are equal by the uniqueness in Theorem 2.2.1.

(b) Since T is surjective, for each i there is $\mathbf{u}_i \in V$ with $T\mathbf{u}_i = \mathbf{c}_i$; fix one. By Theorem 2.2.1 there is a linear $R : W \to V$ with $R(\mathbf{c}_i) = \mathbf{u}_i$ for each i. Then $(T \circ R)(\mathbf{c}_i) = \mathbf{c}_i$ for every i, so $T \circ R = \mathrm{id}_W$, again by uniqueness in Theorem 2.2.1.

In (a), some finiteness assumption on W is what lets the extension stop; Section 4.1 develops this. In (b) the choice of the $\mathbf{u}_i$ is where right inverses get their freedom, as in Exercise 3.5.33(a).

Solution to Exercise 3.5.41

Uses. Theorem 3.5.27 (Rank normal form), Proposition 3.5.2 (First properties of rank), Theorem 2.4.34 (Block multiplication), Theorem 3.5.13 (Second Fundamental Theorem, completed), Theorem 2.4.2 (The matrix of a composite), Theorem 3.3.12 (Matrices in RREF)
Hint 1. For (a), $N_kQ = \begin{pmatrix} Q_1 \\ 0 \end{pmatrix}$, where Q_1 consists of the first k rows of Q. Zero rows do not affect relations among the columns, so $\operatorname{rank}(N_kQ) = \operatorname{rank} Q_1$ by Proposition 3.5.2(d).
Hint 2. $Q_1 = (I_k \;\; 0)Q$, and both $T_{(I_k \;\; 0)}$ and T_Q are surjective. Apply Theorem 3.5.13 to Q_1.
Hint 3. For (b), compare a given normal form with the one from Theorem 3.5.27: $N_j = \left(GG_0^{-1}\right)N_k\left(H_0^{-1}H\right)$. For (c), find a normal form of AH.
Sketch. (a) By block multiplication $N_kQ = \begin{pmatrix} Q_1 \\ 0 \end{pmatrix}$, and its columns satisfy the same relations as those of Q_1, so $\operatorname{rank}(N_kQ) = \operatorname{rank} Q_1$. As $T_{Q_1} = T_{(I_k \;\; 0)} \circ T_Q$ is surjective onto F^k, $\operatorname{rank} Q_1 = k$. (b) With $G_0AH_0 = N_k$, $k = \operatorname{rank} A$: $N_j = \left(GG_0^{-1}\right)N_k\left(H_0^{-1}H\right)$, so $j = \operatorname{rank} N_j = \operatorname{rank}\left(N_kH_0^{-1}H\right) = k$ by row invariance and (a). (c) $G_0(AH)\left(H^{-1}H_0\right) = N_k$, so $\operatorname{rank}(AH) = k$ by (b). (d) Use (c) and row invariance for $\Rightarrow$; both are equivalent to N_k for $\Leftarrow$.

(a) Partition Q into its first k rows $Q_1 \in F^{k\times n}$ and the remaining rows Q_2, and N_k into blocks with row sizes $k, m-k$ and column sizes $k, n-k$. By Theorem 2.4.34,

$$N_kQ = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} = \begin{pmatrix} Q_1 \\ 0 \end{pmatrix}.$$

A list of scalars $x_1, \dots, x_n$ satisfies $\sum_l x_l\mathbf{q}'_l = \mathbf{0}$ for the columns $\mathbf{q}'_l$ of $\begin{pmatrix} Q_1 \\ 0 \end{pmatrix}$ if and only if it satisfies $\sum_l x_l\mathbf{q}_l = \mathbf{0}$ for the columns $\mathbf{q}_l$ of Q_1, since the extra entries are all 0. In particular, a column of $\begin{pmatrix} Q_1 \\ 0 \end{pmatrix}$ is a combination of the earlier ones exactly when the same holds in Q_1, so $\operatorname{rank}(N_kQ) = \operatorname{rank} Q_1$ by Proposition 3.5.2(d).

Now $Q_1 = (I_k \;\; 0)Q$, where $(I_k \;\; 0) \in F^{k\times n}$ (again by block multiplication). The map $T_{(I_k \;\; 0)}$ sends $(x_1, \dots, x_n)$ to $(x_1, \dots, x_k)$, so it is surjective onto F^k, and T_Q is surjective because Q is invertible. So $T_{Q_1} = T_{(I_k \;\; 0)} \circ T_Q$ (Theorem 2.4.2(b)) is surjective, and $\operatorname{rank} Q_1 = k$ by Theorem 3.5.13. Hence $\operatorname{rank}(N_kQ) = k$.

(b) By Theorem 3.5.27 there are invertible G_0, H_0 with $G_0AH_0 = N_k$, where $k = \operatorname{rank} A$. So $A = G_0^{-1}N_kH_0^{-1}$, and if $GAH = N_j$, then

$$N_j = \left(GG_0^{-1}\right)N_k\left(H_0^{-1}H\right).$$

The matrix N_j satisfies (R1)–(R4), so it is in RREF with j pivots (Theorem 3.3.12), and $\operatorname{rank} N_j = j$. On the other hand, GG_0^{-1} is invertible, so by Proposition 3.5.2(c) and (a) with $Q = H_0^{-1}H$,

$$\operatorname{rank} N_j = \operatorname{rank}\left(N_kH_0^{-1}H\right) = k.$$

So $j = k = \operatorname{rank} A$.

(c) With G_0, H_0, k as in (b), $G_0(AH)\left(H^{-1}H_0\right) = G_0AH_0 = N_k$, and $H^{-1}H_0$ is invertible. By (b), applied to the matrix AH, $\operatorname{rank}(AH) = k = \operatorname{rank} A$.

(d) If $B = GAH$, then $\operatorname{rank} B = \operatorname{rank}(G(AH)) = \operatorname{rank}(AH) = \operatorname{rank} A$ by Proposition 3.5.2(c) and (c). Conversely, if both have rank k, then $G_1AH_1 = N_k = G_2BH_2$ for invertible G_1, H_1, G_2, H_2, so

$$B = \left(G_2^{-1}G_1\right)A\left(H_1H_2^{-1}\right),$$

with both brackets invertible.

Solution to Exercise 3.5.42

Uses. Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Proposition 2.4.14 (Algebra of matrices)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. Start from $A = G^{-1}N_kH^{-1}$. Which square matrices are obviously idempotent?
Hint 2. $N_k^2 = N_k$. Conjugates of idempotents are idempotent: $\left(XN_kX^{-1}\right)^2 = XN_k^2X^{-1}$. Insert $H^{-1}H$ or GG^{-1} in the right place.
Sketch. $A = G^{-1}N_kH^{-1} = \left(G^{-1}H^{-1}\right)\left(HN_kH^{-1}\right)$, and $\left(HN_kH^{-1}\right)^2 = HN_k^2H^{-1} = HN_kH^{-1}$ since $N_k^2 = N_k$. Similarly $A = \left(G^{-1}N_kG\right)\left(G^{-1}H^{-1}\right)$.

By Theorem 3.5.27, $GAH = N_k$ for invertible G, H and $k = \operatorname{rank} A$, so $A = G^{-1}N_kH^{-1}$. Here $N_k \in F^{n\times n}$, and by block multiplication $N_k^2 = \begin{pmatrix} I_k^2 & 0 \\ 0 & 0 \end{pmatrix} = N_k$. Now

$$A = G^{-1}N_kH^{-1} = \left(G^{-1}H^{-1}\right)\left(HN_kH^{-1}\right).$$

The first factor $S = G^{-1}H^{-1}$ is invertible, with inverse HG. The second, $P = HN_kH^{-1}$, satisfies

$$P^2 = HN_k\left(H^{-1}H\right)N_kH^{-1} = HN_k^2H^{-1} = HN_kH^{-1} = P.$$

Similarly $A = \left(G^{-1}N_kG\right)\left(G^{-1}H^{-1}\right)$, and $P' = G^{-1}N_kG$ is idempotent by the same computation, while $S' = G^{-1}H^{-1}$ is invertible.

So up to an invertible factor, every square matrix is an idempotent. The idempotent $P = HN_kH^{-1}$ fixes the first k columns of H and kills the others; in the language of Corollary 3.5.29, it is the projection onto the span of $\mathbf{b}_1, \ldots, \mathbf{b}_k$ along the kernel of A.

Solution to Exercise 3.5.43

Uses. Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Definition 2.4.21 (Transpose)
Hint 1. Solve the problem for N_k first: which $n \times m$ matrix Y satisfies $N_kYN_k = N_k$ and $YN_kY = Y$?
Hint 2. $Y = N_k^{\mathsf{T}}$ works. Transport it with $A = G^{-1}N_kH^{-1}$: try $X = HN_k^{\mathsf{T}}G$.
Hint 3. For (c), write $\mathbf{y} = A\mathbf{x}_0$ and compute $AX\mathbf{y}$.
Sketch. (a) By block multiplication $N_kN_k^{\mathsf{T}}N_k = N_k$ and $N_k^{\mathsf{T}}N_kN_k^{\mathsf{T}} = N_k^{\mathsf{T}}$. With $A = G^{-1}N_kH^{-1}$ and $X = HN_k^{\mathsf{T}}G$, the factors $H^{-1}H$ and GG^{-1} cancel in AXA and XAX. (b) If $CA = I_n$, then $ACA = A$ and $CAC = C$; similarly for right inverses. (c) If $\mathbf{y} = A\mathbf{x}_0$, then $AX\mathbf{y} = AXA\mathbf{x}_0 = A\mathbf{x}_0 = \mathbf{y}$.

(a) First take $A = N_k$, and let $Y = N_k^{\mathsf{T}} \in F^{n\times m}$, which has 1 in positions $(1,1), \ldots, (k,k)$ and 0 elsewhere. By block multiplication, $N_kN_k^{\mathsf{T}} \in F^{m\times m}$ has the block form $\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, and then

$$N_kN_k^{\mathsf{T}}N_k = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} = N_k,$$

where the second factor is $N_k \in F^{m\times n}$. In the same way $N_k^{\mathsf{T}}N_kN_k^{\mathsf{T}} = N_k^{\mathsf{T}}$.

For general A, take G, H invertible with $GAH = N_k$ (Theorem 3.5.27), so $A = G^{-1}N_kH^{-1}$, and let $X = HN_k^{\mathsf{T}}G$. Then

$$\begin{aligned} AXA &= G^{-1}N_k\left(H^{-1}H\right)N_k^{\mathsf{T}}\left(GG^{-1}\right)N_kH^{-1} \\ &= G^{-1}N_kN_k^{\mathsf{T}}N_kH^{-1} = G^{-1}N_kH^{-1} = A, \\ XAX &= HN_k^{\mathsf{T}}\left(GG^{-1}\right)N_k\left(H^{-1}H\right)N_k^{\mathsf{T}}G \\ &= HN_k^{\mathsf{T}}N_kN_k^{\mathsf{T}}G = HN_k^{\mathsf{T}}G = X. \end{aligned}$$

(b) If $CA = I_n$, then $ACA = A(CA) = A$ and $CAC = (CA)C = C$. If $AC = I_m$, then $ACA = (AC)A = A$ and $CAC = C(AC) = C$.

(c) If $A\mathbf{x} = \mathbf{y}$ is consistent, say $\mathbf{y} = A\mathbf{x}_0$, then $A(X\mathbf{y}) = AXA\mathbf{x}_0 = A\mathbf{x}_0 = \mathbf{y}$.

So a generalized inverse solves every consistent system $A\mathbf{x} = \mathbf{y}$ by a single matrix multiplication, even when A has neither kind of one-sided inverse. Generalized inverses are not unique: by (b) and Exercise 3.5.33(d), a non-square matrix with a one-sided inverse has more than one.

Solution to Exercise 3.5.44

Uses. Theorem 3.5.27 (Rank normal form), Corollary 3.5.25 (The invertible matrix theorem), Definition 2.4.1 (Matrix product)
Hint 1. One direction is quick: if A is invertible and $AB = 0$, multiply by A^{-1}.
Hint 2. For the other, solve the problem for N_k with $k < n$ first. Which nonzero matrix U has $N_kU = UN_k = 0$?
Hint 3. Take U with a single 1 in position (n, n), and transport it: $B = HUG$.
Sketch. If A is invertible and $AB = 0$, then $B = A^{-1}AB = 0$. If not, $k = \operatorname{rank} A < n$ and $GAH = N_k$. The matrix U with a single 1 in position (n, n) has $N_kU = 0$ and $UN_k = 0$, because column n and row n of N_k are zero. Then $B = HUG \neq 0$ satisfies $AB = G^{-1}N_kUG = 0$ and $BA = HUN_kH^{-1} = 0$.

If A is invertible and $AB = 0$, then $B = A^{-1}(AB) = 0$. So an invertible A has no such B.

Conversely, suppose A is not invertible, so $k = \operatorname{rank} A < n$ (Corollary 3.5.25). By Theorem 3.5.27, $GAH = N_k$ for invertible G, H, so $A = G^{-1}N_kH^{-1}$. Let $U \in F^{n\times n}$ have entry 1 in position (n, n) and 0 elsewhere. Since $k < n$, column n and row n of N_k are zero. By Definition 2.4.1, N_kU has as its only possibly nonzero column $N_k\mathbf{e}_n$, which is column n of N_k, so $N_kU = 0$; and UN_k has as its only possibly nonzero row $\mathbf{e}_n^{\mathsf{T}}N_k$, which is row n of N_k, so $UN_k = 0$. Let $B = HUG$. Then $B \neq 0$, since otherwise $U = H^{-1}BG^{-1} = 0$, and

$$AB = G^{-1}N_k\big(H^{-1}H\big)UG = G^{-1}(N_kU)G = 0,$$
$$BA = HU\big(GG^{-1}\big)N_kH^{-1} = H(UN_k)H^{-1} = 0,$$

as desired. (In fact $B = HUG$ has rank 1, the rank of U, by Exercise 3.5.41(c) and Proposition 3.5.2(c).)

Solution to Exercise 3.5.45

Uses. Theorem 3.5.27 (Rank normal form), Corollary 3.5.24 (Products of square matrices)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. If $GAH = N_k$ and $N_k = D_1 + D_2$ with D_1, D_2 invertible, what can you say about A?
Hint 2. Choose $c \in F$ with $c \neq 0, 1$. Split each diagonal entry of N_k, which is 1 or 0, as a sum of two nonzero scalars.
Sketch. (a) $A = G^{-1}N_kH^{-1}$, so it suffices to split N_k. With $c \neq 0, 1$, let D_1 be diagonal with entries c (k times) and then 1, and D_2 diagonal with entries $1 - c$ (k times) and then -1. Both are invertible and $D_1 + D_2 = N_k$; then $A = G^{-1}D_1H^{-1} + G^{-1}D_2H^{-1}$. (b) The only invertible 1×1 matrix over $\mathbb{F}_2$ is (1), and $(1) + (1) = (0)$.

(a) By Theorem 3.5.27, $A = G^{-1}N_kH^{-1}$ with G, H invertible. Suppose $N_k = D_1 + D_2$ with D_1, D_2 invertible. Then

$$A = G^{-1}D_1H^{-1} + G^{-1}D_2H^{-1},$$

and each term is a product of invertible matrices, hence invertible (Corollary 3.5.24). So it suffices to split N_k.

Since F has at least three elements, there is $c \in F$ with $c \neq 0$ and $c \neq 1$. Let D_1 be the diagonal matrix with diagonal entries $c, \ldots, c$ (k times) followed by $1, \ldots, 1$ ($n - k$ times), and D_2 the diagonal matrix with entries $1 - c, \ldots, 1 - c$ (k times) followed by $-1, \ldots, -1$. Their sum is diagonal with entries $c + (1 - c) = 1$ (k times) and $1 + (-1) = 0$, that is, N_k. A diagonal matrix whose diagonal entries d_i are all nonzero is invertible, with inverse the diagonal matrix with entries d_i^{-1}; and c, 1, $1 - c$, -1 are all nonzero. So D_1 and D_2 are invertible.

(b) Over $\mathbb{F}_2$, a 1×1 matrix (a) is invertible if and only if $a \neq 0$, that is, $a = 1$. The only sum of two invertible 1×1 matrices is $(1) + (1) = (0)$, so (1) is not such a sum.

Solution to Exercise 3.5.46

Uses. Exercise 3.5.45, Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Theorem 3.3.23 (Invertibility and RREF), Theorem 3.5.23 (One-sided inverses of square matrices), Corollary 3.5.24 (Products of square matrices)
Techniques. Block decomposition (Technique 2.4.35)

Hint 1. As in Exercise 3.5.45, it suffices to write each $N_k \in \mathbb{F}_2^{n\times n}$ as a sum of two invertible matrices. Diagonal splittings are impossible now, since the only nonzero scalar is 1.
Hint 2. Work with small diagonal blocks. Split $(0) = (1) + (1)$, $I_2 = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&1\end{pmatrix}$, and $\begin{pmatrix}1&0\\0&0\end{pmatrix} = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&0\end{pmatrix}$. A block diagonal matrix with invertible blocks is invertible.
Hint 3. The blocks above handle every N_k except $N_n = I_n$ for odd n. For that you need a 3×3 block: find an invertible X such that $I_3 + X$ is also invertible.
Sketch. Reduce to N_k as in Exercise 3.5.45. Write N_k as block diagonal with blocks I_2, $\begin{pmatrix}1&0\\0&0\end{pmatrix}$, (0), and I_3, each split into two invertible blocks: $(0) = (1) + (1)$, $I_2 = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&1\end{pmatrix}$, $\begin{pmatrix}1&0\\0&0\end{pmatrix} = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&0\end{pmatrix}$, $I_3 = X + (I_3 + X)$ with $X = \begin{pmatrix}0&0&1\\1&0&1\\0&1&0\end{pmatrix}$. For k even use $k/2$ blocks I_2; for k odd and $k < n$ use one $\begin{pmatrix}1&0\\0&0\end{pmatrix}$; for $k = n$ odd use one I_3. The remaining diagonal positions are (0) blocks.

As in the proof of Exercise 3.5.45(a), $A = G^{-1}N_kH^{-1}$, and it suffices to write $N_k = X + Y$ with X, Y invertible. We use block diagonal matrices (Theorem 2.4.34): if $X = \operatorname{diag}(X_1, \ldots, X_r)$ and $Y = \operatorname{diag}(Y_1, \ldots, Y_r)$ are block diagonal with the same block sizes, then $X + Y = \operatorname{diag}(X_1 + Y_1, \ldots, X_r + Y_r)$, and if each X_i is invertible, then X is invertible, with inverse $\operatorname{diag}\left(X_1^{-1}, \ldots, X_r^{-1}\right)$, by block multiplication.

Building blocks. Over $\mathbb{F}_2$:

- $(0) = (1) + (1)$.
- $I_2 = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&1\end{pmatrix}$, and $\begin{pmatrix}1&0\\0&0\end{pmatrix} = \begin{pmatrix}1&1\\1&0\end{pmatrix} + \begin{pmatrix}0&1\\1&0\end{pmatrix}$. Here $\begin{pmatrix}1&1\\1&0\end{pmatrix}$ reduces to I_2 by adding row 1 to row 2 and then row 2 to row 1; $\begin{pmatrix}0&1\\1&1\end{pmatrix}$ reduces to I_2 by a swap and then adding row 2 to row 1; and $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ is a swap. So all three are invertible (Theorem 3.3.23).
- $I_3 = X + (I_3 + X)$ with
$$X = \begin{pmatrix}0&0&1\\1&0&1\\0&1&0\end{pmatrix}, \quad I_3 + X = \begin{pmatrix}1&0&1\\1&1&1\\0&1&1\end{pmatrix}.$$
Both are invertible: one checks that $X\begin{pmatrix}1&1&0\\0&0&1\\1&0&0\end{pmatrix} = I_3$ and $(I_3 + X)\begin{pmatrix}0&1&1\\1&1&0\\1&1&1\end{pmatrix} = I_3$, and a one-sided inverse of a square matrix is an inverse (Theorem 3.5.23).

Assembling N_k. The diagonal of N_k is k ones followed by $n - k$ zeros. Partition it into consecutive diagonal blocks as follows, putting (0) blocks on all remaining zero positions.

- If k is even, use $k/2$ blocks I_2.
- If k is odd and $k < n$, use $(k-1)/2$ blocks I_2 and then one block $\begin{pmatrix}1&0\\0&0\end{pmatrix}$, which covers the last one and the first zero.
- If $k = n$ is odd, then $n \geq 3$; use one block I_3 and $(n-3)/2$ blocks I_2.

In each case N_k is block diagonal with blocks from the list above, so splitting each block as shown writes N_k as a sum of two block diagonal matrices with invertible blocks, which are invertible.

Finally $A = G^{-1}XH^{-1} + G^{-1}YH^{-1}$, a sum of two invertible matrices (Corollary 3.5.24). The case $k = n = 1$ is exactly the one excluded by $n \geq 2$.

Section 3.6

Solution to Exercise 3.6.6

Uses. Corollary 3.3.19 (Pivot columns, intrinsically), Corollary 3.2.17 (Row operations preserve relations among columns), Proposition 3.1.8 (Systems and the map T_A), Corollary 3.4.4 (Consistency), Non-example 3.2.19 (Row operations move the column space)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. Every column of A has third entry equal to the sum of the first two. Does $(1, 0, 0)$?

Hint 2. Row operations preserve the relations among the columns, so they preserve *which* columns are pivot columns. Do they preserve the columns themselves?
Hint 3. Use Corollary 3.3.19(b): every column of A is a combination of the pivot columns *of* A.
Sketch. (a) The columns $(1,2,3)$, $(2,4,6)$, $(0,1,1)$ all satisfy $y_3 = y_1 + y_2$, so every vector of $\operatorname{im} T_A$ does, and $(1,0,0)$ does not. The relations transfer, but the columns of $\operatorname{rref}(A) = GA$ are $G\mathbf{a}_j$, and they span $\operatorname{im} T_{GA}$, not $\operatorname{im} T_A$. (b) The pivot columns of A, $(1,2,3)$ and $(0,1,1)$, span $\operatorname{im} T_A$, which is the plane $y_3 = y_1 + y_2$.

(a) The columns of A are $\mathbf{a}_1 = (1,2,3)$, $\mathbf{a}_2 = (2,4,6)$, and $\mathbf{a}_3 = (0,1,1)$, and each satisfies $y_3 = y_1 + y_2$. This condition is preserved by sums and scalar multiples, so every linear combination of the columns satisfies it, and by Proposition 3.1.8(a) every vector of $\operatorname{im} T_A$ does. The vector $(1,0,0)$ has $0 \neq 1 + 0$, so it is not in $\operatorname{im} T_A$.

The first sentence of the argument is correct (Corollary 3.2.17), and it does show that columns 1 and 3 of A are independent and that column 2 is a combination of them. The error is the step from relations to the vectors themselves. Writing $\operatorname{rref}(A) = GA$ with G invertible, the columns of $\operatorname{rref}(A)$ are $G\mathbf{a}_1, G\mathbf{a}_2, G\mathbf{a}_3$, so they span $\operatorname{im} T_{GA}$, the image of T_A moved by G. Row operations move the image (Non-example 3.2.19).

(b) By Corollary 3.3.19(b), the non-pivot column $\mathbf{a}_2$ is a combination of the pivot column $\mathbf{a}_1$ (indeed $\mathbf{a}_2 = 2\mathbf{a}_1$, read off the entry 2 of $\operatorname{rref}(A)$). So $\operatorname{im} T_A = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_3) = \operatorname{span}((1,2,3),(0,1,1))$. For an equation: we showed that every vector of $\operatorname{im} T_A$ satisfies $y_3 = y_1 + y_2$. Conversely, if $y_3 = y_1 + y_2$, then

$$y_1(1,2,3) + (y_2 - 2y_1)(0,1,1) = (y_1, y_2, y_1 + y_2) = \mathbf{y},$$

so $\mathbf{y} \in \operatorname{im} T_A$. Hence $\operatorname{im} T_A = \{\mathbf{y} \in \mathbb{R}^3 \mid y_3 = y_1 + y_2\}$. (The same equation comes out of Corollary 3.4.4: reducing A uses $E_{1,-2,2}$, $E_{1,-3,3}$, and $E_{2,-1,3}$, and the last row of G is $(-1,-1,1)$.)

Solution to Exercise 3.6.7

Uses. Corollary 3.3.19 (Pivot columns, intrinsically), Theorem 3.4.9 (All solutions), Definition 3.4.8 (Basic solutions), Proposition 3.1.8 (Systems and the map T_A)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. Columns 1 and 3 are the pivot columns. What does Corollary 3.3.19(b) say about columns 2 and 4?
Hint 2. The kernel depends only on $\operatorname{rref}(A)$: use the basic solutions. The image is spanned by the pivot columns of A.
Sketch. (a) $\mathbf{a}_2 = 2\mathbf{a}_1 = (2,2,4)$ and $\mathbf{a}_4 = -\mathbf{a}_1 + 3\mathbf{a}_3 = (-1,2,-5)$. (b) $\ker T_A = \operatorname{span}((-2,1,0,0),(1,0,-3,1))$, and $\operatorname{im} T_A = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_3)$, which is the plane $y_3 = 3y_1 - y_2$.

(a) The pivot columns of $\operatorname{rref}(A)$ are 1 and 3, so these are the pivot columns of A, and columns 2 and 4 are free. By Corollary 3.3.19(b), each free column of A is the combination of the pivot columns to its left whose coefficients are the entries of the corresponding column of $\operatorname{rref}(A)$:

$$\mathbf{a}_2 = 2\mathbf{a}_1 = (2,2,4), \quad \mathbf{a}_4 = -\mathbf{a}_1 + 3\mathbf{a}_3 = (-1, -1+3, -2-3) = (-1,2,-5).$$

So

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 1 & 2 & 1 & 2 \\ 2 & 4 & -1 & -5 \end{pmatrix}.$$

(b) By Theorem 3.4.9, $\ker T_A$ is spanned by the basic solutions (Definition 3.4.8), one for each free column, read off from $\operatorname{rref}(A)$: $\mathbf{s}_1 = (-2,1,0,0)$ and $\mathbf{s}_2 = (1,0,-3,1)$. As a check, $A\mathbf{s}_1 = -2\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$ and $A\mathbf{s}_2 = \mathbf{a}_1 - 3\mathbf{a}_3 + \mathbf{a}_4 = \mathbf{0}$.

By (a), every column of A is a combination of $\mathbf{a}_1$ and $\mathbf{a}_3$, so $\operatorname{im} T_A = \operatorname{span}(\mathbf{a}_1, \mathbf{a}_3)$ (Proposition 3.1.8(a)). A vector $\mathbf{y}$ lies in it if and only if $\mathbf{y} = s(1,1,2) + t(0,1,-1) = (s, s+t, 2s-t)$ for some s,t. Such s,t must be $s = y_1$ and $t = y_2 - y_1$, and then the third entry is $2y_1 - (y_2 - y_1) = 3y_1 - y_2$. So $\operatorname{im} T_A = \{\mathbf{y} \in \mathbb{R}^3 \mid y_3 = 3y_1 - y_2\}$.

Solution to Exercise 3.6.8

Uses. Proposition 3.2.30 (Column operations preserve the image), Non-example 3.2.31 (Column operations change the kernel), Corollary 3.2.23 (Row equivalent matrices have the same kernel), Proposition 2.4.23 (Transpose of a product), Exercise 3.5.41, Definition 2.5.1 (Left inverse)
Techniques. Transpose to trade rows for columns (Technique 3.2.27)
Hint 1. (a) and (b) are in Section 3.2.5. For (c), transpose: $\mathbf{z}A = \mathbf{0}$ says $A^\mathsf{T}\mathbf{z}^\mathsf{T} = \mathbf{0}$. What is $(AH)^\mathsf{T}$?
Hint 2. For (d), try $A = \begin{pmatrix}0 & 1\end{pmatrix}$ and a swap. For (f), if $CA = I_n$, find a left inverse of AH.
Sketch. (a) Same, Proposition 3.2.30. (b) Not always: $\begin{pmatrix}1 & 1\end{pmatrix}$ and $\begin{pmatrix}0 & 1\end{pmatrix}$. (c) Same: $(AH)^\mathsf{T} = H^\mathsf{T}A^\mathsf{T}$ is a row operation applied to A^T, which keeps the kernel; or directly, $\mathbf{z}AH = \mathbf{0} \Leftrightarrow \mathbf{z}A = \mathbf{0}$. (d) Not always: $\begin{pmatrix}0 & 1\end{pmatrix}$ and $\begin{pmatrix}1 & 0\end{pmatrix}$. (e) Same, by Exercise 3.5.41(c). (f) Same: $\left(H^{-1}C\right)(AH) = I_n$.

Since $A = (AH)H^{-1}$ and H^{-1} is invertible, a property that passes from A to AH for every invertible H also passes back from AH to A. So in each "same" case it suffices to show one direction.

(a) Always the same, by Proposition 3.2.30.

(b) Not always. For $A = \begin{pmatrix}1 & 1\end{pmatrix}$ and $H = E_{1,-1,2}$ (which adds -1 times column 2 to column 1), $AH = \begin{pmatrix}0 & 1\end{pmatrix}$, and the kernels are $\operatorname{span}((1,-1))$ and $\operatorname{span}((1,0))$ (Non-example 3.2.31).

(c) Always the same. Transposing, $\mathbf{z}A = \mathbf{0}$ says $A^\mathsf{T}\mathbf{z}^\mathsf{T} = \mathbf{0}$, that is, $\mathbf{z}^\mathsf{T} \in \ker T_{A^\mathsf{T}}$. By Proposition 2.4.23, $(AH)^\mathsf{T} = H^\mathsf{T}A^\mathsf{T}$, and H^T is invertible, with inverse $\left(H^{-1}\right)^\mathsf{T}$. So by Corollary 3.2.23, $\ker T_{(AH)^\mathsf{T}} = \ker T_{H^\mathsf{T}A^\mathsf{T}} = \ker T_{A^\mathsf{T}}$. (Directly: if $\mathbf{z}A = \mathbf{0}$ then $\mathbf{z}AH = \mathbf{0}$, and the converse follows by multiplying on the right by H^{-1}.) This is Corollary 3.2.23 seen in the mirror of Technique 3.2.27.

(d) Not always. $A = \begin{pmatrix}0 & 1\end{pmatrix}$ is in RREF, and with $H = E_{12}$, $AH = \begin{pmatrix}1 & 0\end{pmatrix}$ is also in RREF. They are different matrices, so their reduced forms differ (and so do their pivot columns).

(e) Always the same, by Exercise 3.5.41(c).

(f) Always the same. If $CA = I_n$, then $\left(H^{-1}C\right)(AH) = H^{-1}(CA)H = H^{-1}H = I_n$, so $H^{-1}C$ is a left inverse of AH.

Solution to Exercise 3.6.9

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Corollary 3.2.23 (Row equivalent matrices have the same kernel), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product)
Techniques. Reduce to RREF (Technique 3.3.22)
Hint 1. One direction is one line. For the other, first treat the case where $A = R$ is in RREF, then use $R = GA$.
Hint 2. If $A = R$ is in RREF with pivot columns $p_1, \ldots, p_r$, guess Y with $YR = B$ from the pivot columns: $R\mathbf{e}_{p_s} = \mathbf{e}_s$, so Y must send $\mathbf{e}_s$ to column p_s of B.
Hint 3. For a free column j, the relation $\mathbf{r}_j = \sum_s r_{sj}\mathbf{r}_{p_s}$ among the columns of R is a vector of $\ker T_R$. Transfer it to B.
Sketch. If $B = XA$ and $A\mathbf{x} = \mathbf{0}$, then $B\mathbf{x} = \mathbf{0}$. Conversely, let $R = GA = \operatorname{rref}(A)$, so $\ker T_R = \ker T_A \subseteq \ker T_B$. Let Y have column s equal to column p_s of B for $s \le r$ and $\mathbf{0}$ for $s > r$. Every column of R is $\mathbf{r}_j = \sum_{s\le r} r_{sj}\mathbf{e}_s$, and the relation $\mathbf{r}_j = \sum_s r_{sj}\mathbf{r}_{p_s}$ lies in $\ker T_R$, so it holds in B: $\mathbf{b}_j = \sum_s r_{sj}\mathbf{b}_{p_s} = Y\mathbf{r}_j$. Hence $YR = B$, and $X = YG$ works. (b) Row k of XA is the combination of the rows of A with coefficients from row k of X.

(a) Suppose $B = XA$. If $A\mathbf{x} = \mathbf{0}$, then $B\mathbf{x} = X(A\mathbf{x}) = \mathbf{0}$. So $\ker T_A \subseteq \ker T_B$.

Conversely, suppose $\ker T_A \subseteq \ker T_B$. By Theorem 3.3.10, $R = \operatorname{rref}(A) = GA$ for an invertible G, and $\ker T_R = \ker T_A$ (Corollary 3.2.23), so $\ker T_R \subseteq \ker T_B$. Let $p_1 < \cdots < p_r$ be the pivot columns, write $\mathbf{r}_j$ and $\mathbf{b}_j$ for the columns of R and B, and let $Y \in F^{p\times m}$ be the matrix whose column s is $\mathbf{b}_{p_s}$ for $s \le r$ and $\mathbf{0}$ for $s > r$. We show $YR = B$, column by column.

Fix j. By Lemma 3.3.16(a) (with $s = r$), $\mathbf{r}_j = \sum_{s=1}^{r} r_{sj}\mathbf{e}_s$, so by Definition 2.3.8,

$$Y\mathbf{r}_j = \sum_{s=1}^{r} r_{sj}Y\mathbf{e}_s = \sum_{s=1}^{r} r_{sj}\mathbf{b}_{p_s}.$$

On the other hand, $\mathbf{r}_{p_s} = \mathbf{e}_s$ (Lemma 3.3.16(b)), so $\mathbf{r}_j = \sum_{s=1}^{r} r_{sj}\mathbf{r}_{p_s}$. This says that the vector

$$\mathbf{x} = \mathbf{e}_j - \sum_{s=1}^{r} r_{sj}\mathbf{e}_{p_s}$$

satisfies $R\mathbf{x} = \mathbf{0}$. (If $j = p_t$ is itself a pivot column, then $r_{sj} = 0$ for $s \neq t$ and $r_{tj} = 1$, so $\mathbf{x} = \mathbf{0}$, and the argument still works.) Hence $\mathbf{x} \in \ker T_B$, that is, $\mathbf{b}_j = \sum_s r_{sj}\mathbf{b}_{p_s} = Y\mathbf{r}_j$. So the columns of YR are those of B, and $YR = B$ (Definition 2.4.1). With $X = YG$, we get $XA = Y(GA) = YR = B$, as desired.

(b) By the row description of a product (Definition 2.4.1, as in the second proof of Proposition 3.2.7), row k of XA is x_{k1} times row 1 of A, plus $\cdots$, plus x_{km} times row m of A. So "$B = XA$ for some X" says exactly that every row of B is a linear combination of the rows of A, and (b) is a restatement of (a).

If $\ker T_A = \ker T_B$, then (b) applies in both directions, so A and B have the same row space; this generalizes one half of Exercise 3.3.50(b) to matrices with different numbers of rows. The proof of (a) is the second proof of Theorem 3.3.17 once more: it used only relations among the columns.

Solution to Exercise 3.6.10

Uses. Exercise 3.2.42, Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 3.3.23 (Invertibility and RREF), Corollary 3.4.20 (Square systems), Proposition 3.3.8 (Properties of the output), Proposition 3.2.10 (Inverses of elementary matrices)
Techniques. Reduce to elementary matrices (Technique 5.4.6), Reduce to RREF (Technique 3.3.22)
Hint 1. By Exercise 3.2.42, $D(EY) = D(E)D(Y)$ for every elementary E and every Y: check that $D(E)$ is 1, -1, or s for the three types.
Hint 2. If $X = E_1 \cdots E_t$, apply the previous hint t times, once to XY and once to $X = XI_2$.
Hint 3. If X is not invertible, write $X = G^{-1}R$ with $R = \operatorname{rref}(X) \neq I_2$. What is the second row of R, and of RY?
Sketch. For elementary E, $D(EY) = D(E)D(Y)$: the factors $1, -1, s$ of Exercise 3.2.42 are $D(E)$. For $X = E_1 \cdots E_t$, induction gives $D(XY) = D(E_1)\cdots D(E_t)D(Y)$, and with $Y = I_2$, $D(X) = D(E_1)\cdots D(E_t)$. If X is not invertible, $X = PR$ with $P = G^{-1}$ a product of elementary matrices and $R = \operatorname{rref}(X)$ having zero second row; then $D(XY) = D(P)D(RY) = 0$ and $D(X) = D(P)D(R) = 0$.

Step 1: elementary matrices. By Exercise 3.2.42, for every $Y \in F^{2\times 2}$, $D(EY) = D(Y)$ if E has type 3, $D(E_{12}Y) = -D(Y)$, and $D(E_{is}Y) = sD(Y)$. Directly from the definition, $D(E) = 1$ for $E = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $D(E_{12}) = 0 \cdot 0 - 1 \cdot 1 = -1$, and $D(E_{1s}) = D(E_{2s}) = s$. So in all cases

$$D(EY) = D(E)D(Y) \quad \text{for every elementary } E \text{ and every } Y. \tag{C.1.6}$$

Step 2: invertible X. By Theorem 3.5.19, $X = E_1E_2\cdots E_t$ with each E_i elementary. Applying (C.1.6) t times,

$$D(XY) = D(E_1(E_2\cdots E_tY)) = D(E_1)D(E_2\cdots E_tY) = \cdots = D(E_1)\cdots D(E_t)D(Y).$$

With $Y = I_2$, and $D(I_2) = 1$, this gives $D(X) = D(E_1)\cdots D(E_t)$. Hence $D(XY) = D(X)D(Y)$.

Step 3: X not invertible. Let $R = \operatorname{rref}(X) = GX$ with G a product of elementary matrices (Theorem 3.3.10). Since X is not invertible, $R \neq I_2$ (Theorem 3.3.23). A 2×2 matrix in RREF with two pivots is I_2 (Corollary 3.4.20, (i) $\Leftrightarrow$ (iii)), so R has at most one pivot and its second row is zero (Proposition 3.3.8(a)). Then the second row of RY is zero as well (row 2 of RY is row 2 of R times Y), so $D(R) = 0$ and $D(RY) = 0$, since a matrix with zero second row has $c = d = 0$. Now $P = G^{-1}$ is a product of elementary matrices (Proposition 3.2.10), and $X = PR$, so Step 2 applied to P gives

$$D(XY) = D(P(RY)) = D(P)D(RY) = 0, \quad D(X) = D(PR) = D(P)D(R) = 0.$$

So $D(XY) = 0 = D(X)D(Y)$.

Step 3 is the part that is easy to forget, since a matrix that is not invertible is not a product of elementary matrices at all. In Section 5.4 the same strategy (elementary matrices, then invertible matrices, then a separate argument for the rest) proves $\det(AB) = \det A \det B$ for every n (Theorem 5.4.10).

Solution to Exercise 3.6.11

Uses. Proposition 3.2.7 (Row operations are left multiplications), Proposition 3.2.28 (Column operations are right multiplications)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Look for a quantity that every move increases. What happens to the sum of all entries when a row with sum $s < 0$ is negated?
Hint 2. An increasing quantity proves termination only if it can take finitely many values. Which matrices can occur? Find an invariant relating each entry to the starting entry.
Sketch. Invariant: every entry is $\pm$ the corresponding starting entry, so at most 2^{mn} matrices can occur. Monovariant: negating a row or column with sum $s < 0$ changes the total sum S by $-2s > 0$. So S strictly increases, no matrix occurs twice, and there are fewer than 2^{mn} moves. When no move is possible, all row and column sums are ≥ 0. Each move is left multiplication by $E_{i,-1}$ or right multiplication by $E_{j,-1}$.

(a) Let A be the starting matrix and M the current one.

Invariant. Every entry of M is $\pm$ the corresponding entry of A. This holds at the start, and a move multiplies some entries by -1, which keeps it true. So M is one of at most 2^{mn} matrices (two choices of sign for each entry).

Monovariant. Let $S(M)$ be the sum of all entries of M. If a move negates a row or column whose entries sum to $s < 0$, the entries of that line now sum to $-s$, and no other entry changes, so S changes by $-s - s = -2s > 0$. So S strictly increases with every move.

In a sequence of moves $M_0 = A, M_1, M_2, \ldots$, the values $S(M_0) < S(M_1) < \cdots$ are distinct, so the matrices $M_0, M_1, \ldots$ are distinct. By the invariant they come from a set of at most 2^{mn} matrices, so the sequence has at most 2^{mn} terms, that is, fewer than 2^{mn} moves.

(b) Make moves as long as one is possible. By (a) this stops, and it stops exactly when no row and no column has a negative sum.

(c) Negating row i is left multiplication by $E_{i,-1}$ (Proposition 3.2.7), and negating column j is right multiplication by the $n \times n$ matrix $E_{j,-1}$ (Proposition 3.2.28). Since $(EX)E' = E(XE')$, the final matrix is $D_1 A D_2$, where D_1 is the product of the row negations (the last one leftmost) and D_2 the product of the column negations (the first one leftmost). Both are diagonal with entries ± 1, and they are products of type 2 elementary matrices.

The proof follows the two halves of Technique 3.3.5. The monovariant alone would not do: an increasing real number can increase forever. The invariant supplies the finiteness.

Solution to Exercise 3.6.12

Uses. Definition 1.2.17 (Evaluation), Lemma 3.4.23 (Too many roots), Proposition 3.1.19 (Solution sets are translates of the kernel)
Techniques. Undetermined coefficients (Technique 3.1.9), Particular plus homogeneous (Technique 3.1.18)
Hint 1. Write $p = a_0 + a_1 x + \cdots + a_4 x^4$ and expand $(t+1)^k - t^k$ for $k = 1, \ldots, 4$. The difference $p(t+1) - p(t)$ is then a polynomial expression of degree at most 3 in t.
Hint 2. A polynomial of degree at most 3 that vanishes at every $t \in \mathbb{Q}$ is the zero polynomial (Lemma 3.4.23). So you may compare coefficients: four linear equations in $a_1, \ldots, a_4$.
Hint 3. The system is triangular. What does a_0 do? For (b), sum $p(k+1) - p(k)$ over $k = 0, \ldots, N-1$.
Sketch. The difference is $(a_1 + a_2 + a_3 + a_4) + (2a_2 + 3a_3 + 4a_4)t + (3a_3 + 6a_4)t^2 + 4a_4 t^3$. Comparing with t^3: $a_4 = 1/4$, $a_3 = -1/2$, $a_2 = 1/4$, $a_1 = 0$, and a_0 is free. So $p = (x(x-1)/2)^2 + a_0$: a particular solution plus the constants, which are the homogeneous solutions. Telescoping gives (b).

(a) Write $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$. By the binomial theorem,

$$(t+1) - t = 1, \quad (t+1)^2 - t^2 = 2t + 1,$$
$$(t+1)^3 - t^3 = 3t^2 + 3t + 1, \quad (t+1)^4 - t^4 = 4t^3 + 6t^2 + 4t + 1,$$

so for every $t \in \mathbb{Q}$ (Definition 1.2.17),

$$p(t+1)-p(t)=c_0+c_1t+c_2t^2+c_3t^3,$$

where

$$c_0=a_1+a_2+a_3+a_4,\quad c_1=2a_2+3a_3+4a_4,\quad c_2=3a_3+6a_4,\quad c_3=4a_4.$$

The condition is that $q=(c_0-0)+c_1x+c_2x^2+(c_3-1)x^3\in\mathbb{Q}[x]_{\le 3}$ satisfies $q(t)=0$ for every $t\in\mathbb{Q}$. If q vanishes at all of $\mathbb{Q}$, it vanishes at the four distinct points $0,1,2,3$, so $q=\mathbf{0}$ by Lemma 3.4.23; conversely, $q=\mathbf{0}$ certainly vanishes everywhere. So the condition is the linear system

$$a_1+a_2+a_3+a_4=0,\quad 2a_2+3a_3+4a_4=0,\quad 3a_3+6a_4=0,\quad 4a_4=1$$

in the unknowns $a_0,\dots,a_4$, in which a_0 does not occur. Solving from the bottom: $a_4=1/4$, $a_3=-2a_4=-1/2$, $2a_2=-3a_3-4a_4=3/2-1=1/2$, so $a_2=1/4$, and $a_1=-(a_2+a_3+a_4)=0$. So the solutions are

$$p=a_0+\frac{1}{4}x^2-\frac{1}{2}x^3+\frac{1}{4}x^4\quad(a_0\in\mathbb{Q}),$$

and $\frac{1}{4}x^2-\frac{1}{2}x^3+\frac{1}{4}x^4$ is the polynomial whose value at t is $(t(t-1)/2)^2$, since $t^2(t-1)^2=t^4-2t^3+t^2$. In the language of Proposition 3.1.19, applied to the linear map $p\mapsto(p(t+1)-p(t))_t$, this is a particular solution plus the homogeneous solutions, which are the constants: the same triangular system with right-hand side 0 forces $a_1=\cdots=a_4=0$.

(b) Let p be the solution with $a_0=0$, so $p(t)=(t(t-1)/2)^2$. Summing $p(k+1)-p(k)=k^3$ over $k=0,\dots,N-1$, the left side telescopes:

$$0^3+1^3+\cdots+(N-1)^3=p(N)-p(0)=(N(N-1)/2)^2.$$

The method needs a field in which comparing coefficients is allowed, which used four distinct points $0,1,2,3$, and in which 2, 3, 4 are nonzero, for the triangular solve. Over $\mathbb{F}_p$ with $p\ge 5$ both hold. Over $\mathbb{F}_2$ and $\mathbb{F}_3$ both fail: for instance x^3-x vanishes at every point of $\mathbb{F}_3$, so a polynomial identity on the field no longer determines the coefficients.

Solution to Exercise 3.6.13

> *Uses.* Corollary 3.4.20 (Square systems), Exercise 3.5.37, Proposition 2.4.23 (Transpose of a product)
> *Techniques.* Uniqueness gives existence (Technique 3.4.21), Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* The system in (a) is square. By Corollary 3.4.20 it suffices to show that $4x_i=x_{i-1}+x_{i+1}$ for all i forces $\mathbf{x}=\mathbf{0}$.
> *Hint 2.* Choose k with $|x_k|$ as large as possible, and compare $4|x_k|$ with $|x_{k-1}|+|x_{k+1}|$.
> *Hint 3.* In (b), the vector $(1,\dots,1)$ is a homogeneous solution. Show it spans the homogeneous solutions (look at a largest entry again), note that the matrix is symmetric, and apply Exercise 3.5.37.
> *Sketch.* (a) For a homogeneous solution, take k with $|x_k|$ maximal: $4|x_k|\le|x_{k-1}|+|x_{k+1}|\le 2|x_k|$, so $x_k=0$ and $\mathbf{x}=\mathbf{0}$; the system is square, so Corollary 3.4.20 gives existence and uniqueness. (b) Summing the equations gives $0=\sum y_i$, which is necessary. The homogeneous solutions are the constant vectors (at a largest entry x_k, both neighbors equal x_k; go around the cycle). The matrix L is symmetric, so $\mathbf{z}^{\mathsf{T}}L=\mathbf{0}$ iff $L\mathbf{z}=\mathbf{0}$ iff $\mathbf{z}$ is constant, and Exercise 3.5.37 says $L\mathbf{x}=\mathbf{y}$ is solvable iff $\mathbf{z}^{\mathsf{T}}\mathbf{y}=0$ for those $\mathbf{z}$, that is, iff $\sum y_i=0$.

(a) This is a system of n equations in n unknowns, so by Corollary 3.4.20 ((iv) ⇒ (vi)) it suffices to show that the homogeneous system has only the trivial solution. Let $\mathbf{x}$ satisfy $4x_i=x_{i-1}+x_{i+1}$ for all i, and choose k with $|x_k|\ge|x_i|$ for all i (an extremal choice). Then

$$4|x_k|=|x_{k-1}+x_{k+1}|\le|x_{k-1}|+|x_{k+1}|\le 2|x_k|,$$

so $2|x_k|\le 0$ and $x_k=0$. As $|x_k|$ is the largest of the $|x_i|$, every $x_i=0$.

(b) Let $L\in\mathbb{R}^{n\times n}$ be the coefficient matrix: 2 on the diagonal, -1 in the positions $(i,i-1)$ and $(i,i+1)$ (modulo n), and 0 elsewhere. It is symmetric, $L^{\mathsf{T}}=L$, because $j=i+1$ exactly when $i=j-1$. Let $\mathbf{1}=(1,\dots,1)$.

Necessity. Each x_j occurs in three equations, with total coefficient $2-1-1=0$ (in equation j with 2, and in equations $j-1$ and $j+1$ with -1; these are three different equations since $n \geq 3$). So summing all equations gives $0 = y_1 + \cdots + y_n$, and a solution can exist only if $\sum y_i = 0$.

The homogeneous solutions are the constant vectors. Constant vectors are solutions, since $2c - c - c = 0$. Conversely, let $L\mathbf{x} = \mathbf{0}$, and let x_k be a largest entry. Then $2x_k = x_{k-1} + x_{k+1}$ with $x_{k-1} \leq x_k$ and $x_{k+1} \leq x_k$, which forces $x_{k-1} = x_{k+1} = x_k$. So the neighbors of a largest entry are largest entries too, and going around the cycle, all entries equal x_k. Thus $L\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{x} \in \operatorname{span}(\mathbf{1})$.

Sufficiency. Suppose $\sum y_i = 0$ but $L\mathbf{x} = \mathbf{y}$ has no solution. By Exercise 3.5.37 there is $\mathbf{z}$ with $\mathbf{z}^{\mathsf{T}}L = \mathbf{0}$ and $\mathbf{z}^{\mathsf{T}}\mathbf{y} \neq 0$. Transposing (Proposition 2.4.23), $L^{\mathsf{T}}\mathbf{z} = L\mathbf{z} = \mathbf{0}$, so $\mathbf{z} = c\mathbf{1}$ for some c, and then $\mathbf{z}^{\mathsf{T}}\mathbf{y} = c(y_1 + \cdots + y_n) = 0$, a contradiction.

So with 4 the homogeneous system has only the trivial solution and every $\mathbf{y}$ is reached; with 2 there is a homogeneous solution $\mathbf{1}$, and in exchange one condition on $\mathbf{y}$. For a square system the two failures come together, as Corollary 3.4.20 says they must.

Solution to Exercise 3.6.14

> *Uses.* Corollary 3.4.20 (Square systems), Proposition 3.1.19 (Solution sets are translates of the kernel)
> *Techniques.* Uniqueness gives existence (Technique 3.4.21)
> *Hint 1.* Squareness only sets the technique up; its content is the homogeneous uniqueness. Check it: is there a nonzero p with $p(0) = p(1) = p'(1/2) = 0$?
> *Hint 2.* Compare $p'(1/2)$ with $p(1) - p(0)$ for a general p.
> *Sketch.* The argument never proves that the homogeneous system has only the trivial solution, and it does not: $x^2 - x$ is a nonzero solution. Since $p'(1/2) = b + c = p(1) - p(0)$ for every p, a solution exists only if $w = v - u$, and then the solutions are $u + (v-u)x + t\big(x^2 - x\big)$, infinitely many. So the claim fails for every (u, v, w).

The argument checks that the problem is linear and square, which is what makes Technique 3.4.21 *applicable*, and then skips the technique itself: the homogeneous problem must be shown to have only the trivial solution, and here it does not. For $p = a + bx + cx^2$, the homogeneous conditions are $a = 0$, $a + b + c = 0$, and $b + c = 0$, and $p = x^2 - x$ (that is, $a = 0$, $b = -1$, $c = 1$) satisfies all three.

In fact, for every p,

$$p(1) - p(0) = b + c = p'(1/2).$$

So a solution can exist only if $w = v - u$. If $w = v - u$, then $u + (v-u)x$ is a solution, and by Proposition 3.1.19 the solutions are $u + (v-u)x + t\big(x^2 - x\big)$ for $t \in \mathbb{R}$, since the homogeneous solutions are the multiples of $x^2 - x$ (the first two homogeneous equations give $a = 0$ and $c = -b$, and the third is then automatic). So the conclusion "exactly one" holds for no (u, v, w): there are no solutions when $w \neq v - u$ and infinitely many when $w = v - u$. By Corollary 3.4.20 this had to be so once uniqueness failed.

Problems for Chapter 3

Solution to Problem 3.1

> *Uses.* Definition 2.1.1 (Linear map), Corollary 3.1.22 (Uniqueness depends only on A), Proposition 3.1.8 (Systems and the map T_A), Definition 2.6.5 (Kernel), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
> *Techniques.* Particular plus homogeneous (Technique 3.1.18), Double containment (Technique 1.3.14)
> *Hint 1.* For (a), the statement is Proposition 3.1.19. Prove the two containments, and check that $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ is injective.
> *Hint 2.* For (b), split into the cases $\ker T_A = \{\mathbf{0}\}$ and $\ker T_A \neq \{\mathbf{0}\}$. In the second case, how large is the solution set of a consistent system?
> *Sketch.* (a) If $T(\mathbf{x}_0) = \mathbf{y}$, then $S = \mathbf{x}_0 + \ker T$: $T(\mathbf{x}_0 + \mathbf{z}) = \mathbf{y} + \mathbf{0}$, and $T(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}$ for $\mathbf{x} \in S$; translation by $\mathbf{x}_0$ is injective. (b) If $\ker T_A = \{\mathbf{0}\}$, every consistent system has exactly one solution, so the set is $\operatorname{im} T_A$. Otherwise each consistent system has a solution set in bijection with $\ker T_A$, which has at least two elements, and inconsistent ones have none, so the set is empty.

(a) *Statement* (Proposition 3.1.19). Let $T : V \to W$ be linear, $\mathbf{y} \in W$, and $S = \{\mathbf{x} \in V \mid T(\mathbf{x}) = \mathbf{y}\}$. If $\mathbf{x}_0 \in S$, then $S = \mathbf{x}_0 + \ker T$, and $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ is a bijection from $\ker T$ onto S.

Proof. We prove both containments. If $\mathbf{z} \in \ker T$, then $T(\mathbf{x}_0 + \mathbf{z}) = T(\mathbf{x}_0) + T(\mathbf{z}) = \mathbf{y} + \mathbf{0} = \mathbf{y}$, so $\mathbf{x}_0 + \mathbf{z} \in S$. Conversely, if $\mathbf{x} \in S$, put $\mathbf{z} = \mathbf{x} - \mathbf{x}_0$; then $T(\mathbf{z}) = T(\mathbf{x}) - T(\mathbf{x}_0) = \mathbf{y} - \mathbf{y} = \mathbf{0}$, so $\mathbf{z} \in \ker T$ and $\mathbf{x} = \mathbf{x}_0 + \mathbf{z} \in \mathbf{x}_0 + \ker T$. Hence $S = \mathbf{x}_0 + \ker T$, so the map $\mathbf{z} \mapsto \mathbf{x}_0 + \mathbf{z}$ sends $\ker T$ onto S. It is injective, since $\mathbf{x}_0 + \mathbf{z} = \mathbf{x}_0 + \mathbf{z}'$ gives $\mathbf{z} = \mathbf{z}'$ after adding $-\mathbf{x}_0$. So it is a bijection.

(b) Let U be the set of $\mathbf{y}$ for which $A\mathbf{x} = \mathbf{y}$ has exactly one solution. Every $\mathbf{y} \in U$ makes the system consistent, so $U \subseteq \operatorname{im} T_A$ (Proposition 3.1.8(a)).

Case $\ker T_A = \{\mathbf{0}\}$. Let $\mathbf{y} \in \operatorname{im} T_A$. The system $A\mathbf{x} = \mathbf{y}$ is consistent, so by Corollary 3.1.22(a) it has exactly one solution, and $\mathbf{y} \in U$. Hence $U = \operatorname{im} T_A$.

Case $\ker T_A \neq \{\mathbf{0}\}$. Then $\ker T_A$ contains $\mathbf{0}$ (Proposition 2.1.4) and some $\mathbf{z} \neq \mathbf{0}$, so it has at least two elements. If $A\mathbf{x} = \mathbf{y}$ is consistent, its solution set is in bijection with $\ker T_A$ by (a), applied to $T = T_A$, so it has at least two elements. If it is inconsistent, it has none. Either way $\mathbf{y} \notin U$, so $U = \varnothing$.

So the set is $\operatorname{im} T_A$ when $\ker T_A = \{\mathbf{0}\}$, and empty otherwise. Uniqueness is all or nothing: it holds for every consistent right-hand side or for none.

Solution to Problem 3.2

> *Uses.* Corollary 3.2.17 (Row operations preserve relations among columns), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Definition 2.4.1 (Matrix product), Theorem 3.3.12 (Matrices in RREF), Proposition 3.3.8 (Properties of the output)
> *Techniques.* Uniqueness via two candidates (Technique 1.1.8), Transpose to trade rows for columns (Technique 3.2.27)
> *Hint 1.* For (a), R and $R' = GR$ have the same relations among their columns (Corollary 3.2.17). Which column is a pivot column, and what a free column is, are statements about such relations (Lemma 3.3.16).
> *Hint 2.* For (b), the smallest example has one row and two columns.
> *Hint 3.* For (c), count nonzero rows. The nonzero rows of a matrix in RREF are linearly independent: look at the pivot columns. Right multiplication by an invertible H keeps rows independent.
> *Sketch.* (a) Column j is a combination of earlier columns, with given coefficients, in R exactly when it is in R'; so the pivot columns agree, the pivot columns of both are $\mathbf{e}_1, \ldots, \mathbf{e}_r$, and each free column has the same expression through the pivot columns in both. (b) $R = \begin{pmatrix} 0 & 1 \end{pmatrix}$, $H = E_{12}$, $R' = \begin{pmatrix} 1 & 0 \end{pmatrix}$. (c) If R has r nonzero rows, they are independent (entry p_t isolates the coefficient of row t). Row s of RH is (row s of R)H, so the zero rows stay zero and the others stay independent, hence nonzero. So RH has r nonzero rows, which are its pivot rows.

(a) *Statement* (Theorem 3.3.17). If $R, R' \in F^{m \times n}$ are in RREF and $R' = GR$ for an invertible $G \in F^{m \times m}$, then $R = R'$.

Proof. Write $\mathbf{r}_j$ and $\mathbf{r}'_j = G\mathbf{r}_j$ for the columns. By Corollary 3.2.17, for all scalars $x_1, \ldots, x_n$,

$$x_1\mathbf{r}_1 + \cdots + x_n\mathbf{r}_n = \mathbf{0} \quad \Leftrightarrow \quad x_1\mathbf{r}'_1 + \cdots + x_n\mathbf{r}'_n = \mathbf{0}.$$

An equation $\mathbf{r}_j = c_1\mathbf{r}_1 + \cdots + c_{j-1}\mathbf{r}_{j-1}$ is such a relation (with $x_j = -1$, $x_l = c_l$ for $l < j$, and $x_l = 0$ for $l > j$), so it holds in R if and only if the same equation holds in R'. Hence $\mathbf{r}_j$ is a combination of the earlier columns of R exactly when $\mathbf{r}'_j$ is a combination of the earlier columns of R', and by Lemma 3.3.16(c) the two matrices have the same pivot columns $p_1 < \cdots < p_r$. By Lemma 3.3.16(b), $\mathbf{r}_{p_s} = \mathbf{e}_s = \mathbf{r}'_{p_s}$. If j is a free column with pivot columns $p_1, \ldots, p_t$ to its left, Lemma 3.3.16(d) gives the relation $\mathbf{r}_j = \sum_{s \le t} r_{sj}\mathbf{r}_{p_s}$, which therefore also holds in R':

$$\mathbf{r}'_j = \sum_{s \le t} r_{sj}\mathbf{r}'_{p_s} = \sum_{s \le t} r_{sj}\mathbf{e}_s = \mathbf{r}_j.$$

So every column of R' equals the corresponding column of R.

(b) $R = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $R' = \begin{pmatrix} 1 & 0 \end{pmatrix}$ both satisfy (R1)–(R4), so they are in RREF (Theorem 3.3.12), and $R' = RE_{12}$ with E_{12} the 2×2 swap, which is invertible. The proof of (a) breaks because right multiplication does not preserve the relations among the columns: here $\mathbf{r}_1 = \mathbf{0}$ but $\mathbf{r}'_1 \neq \mathbf{0}$.

(c) Let R have pivot columns $p_1 < \cdots < p_r$. By Theorem 3.3.12 and Proposition 3.3.8, rows $1, \ldots, r$ of R are nonzero and the others are zero; write $\mathbf{u}_1, \ldots, \mathbf{u}_r \in F^{1\times n}$ for the nonzero rows. They are linearly independent: if $c_1\mathbf{u}_1 + \cdots + c_r\mathbf{u}_r = \mathbf{0}$, look at entry p_t. Column p_t of R is $\mathbf{e}_t$ (Lemma 3.3.16(b)), so $\mathbf{u}_s$ has entry 1 there if $s = t$ and 0 otherwise, and the entry p_t of the combination is c_t. So every $c_t = 0$.

By Definition 2.4.1, row s of RH is $\mathbf{u}_sH$ for $s \le r$, and $\mathbf{0}H = \mathbf{0}$ for $s > r$. The rows $\mathbf{u}_1H, \ldots, \mathbf{u}_rH$ are again independent: if $\sum_s c_s\mathbf{u}_sH = \mathbf{0}$, then $\big(\sum_s c_s\mathbf{u}_s\big)H = \mathbf{0}$, multiplying on the right by H^{-1} gives $\sum_s c_s\mathbf{u}_s = \mathbf{0}$, and so every $c_s = 0$. In particular each $\mathbf{u}_sH$ is nonzero (a list containing $\mathbf{0}$ is dependent). So RH has exactly r nonzero rows. Since RH is in RREF, its pivots are the leading entries of its nonzero rows (Theorem 3.3.12), so it has r pivots, as R does.

Part (c) is the transpose of a row statement in disguise (Technique 3.2.27): the rows of RH are the columns of $H^{\mathsf{T}}R^{\mathsf{T}}$, and left multiplication by an invertible matrix preserves independence of columns (Corollary 3.2.17).

Solution to Problem 3.3

> *Uses.* Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Proposition 3.3.11 (Carrying extra columns along), Proposition 3.3.8 (Properties of the output), Corollary 3.2.11 (Products of elementary matrices), Proposition 3.2.7 (Row operations are left multiplications), Definition 2.4.1 (Matrix product)
> *Techniques.* Reduce to RREF (Technique 3.3.22)
> *Hint 1.* For (a), if $GA = I_n$ with G invertible, then $A = G^{-1}$. If $\operatorname{rref}(A) \neq I_n$, its last row is zero, and a matrix with a zero row has no right inverse.
> *Hint 2.* For (b), at every stage the current matrix is $G'(A \mid I_n) = (G'A \mid G')$ for a product G' of elementary matrices. What is row i of $G'A$ in terms of row i of G'?
> *Hint 3.* If A were invertible, compute $\mathbf{z} = \mathbf{z}AA^{-1}$.
> *Sketch.* (a) A is invertible iff $\operatorname{rref}(A) = I_n$, and then $A^{-1} = G$; the algorithm computes $(\operatorname{rref}(A) \mid G)$ by Proposition 3.3.11. (b) The current matrix is $(G'A \mid G')$ with G' invertible, so $\mathbf{z} =$ row i of G' is nonzero and $\mathbf{z}A =$ row i of $G'A = \mathbf{0}$; if A were invertible, $\mathbf{z} = \mathbf{z}AA^{-1} = \mathbf{0}$. (c) Over $\mathbb{F}_5$, two passes give the left row $(0,0,0)$ with right row $(4,4,1)$, and indeed $-\mathbf{r}_1 - \mathbf{r}_2 + \mathbf{r}_3 = (0,0,5) = \mathbf{0}$ there. Over $\mathbb{Q}$ the same row is $(0,0,5)$, a third pivot, so A is invertible over $\mathbb{Q}$.

(a) *Statement* (Theorem 3.3.23). Let $A \in F^{n\times n}$, and let G be a product of elementary matrices with $GA = \operatorname{rref}(A)$. Then A is invertible if and only if $\operatorname{rref}(A) = I_n$, and in that case $A^{-1} = G$.

Proof. If $\operatorname{rref}(A) = I_n$, then $GA = I_n$, and G is invertible (Corollary 3.2.11), so $A = G^{-1}$ and $AG = G^{-1}G = I_n$; thus $A^{-1} = G$. Conversely, suppose $R = \operatorname{rref}(A) \neq I_n$. If R had n pivots, they would lie in columns $p_1 < \cdots < p_n$ among $1, \ldots, n$, forcing $p_s = s$, and then $R = I_n$ by Proposition 3.3.8(c). So R has fewer than n pivots, and its last row is zero (Proposition 3.3.8(a)). Then the last row of RX is zero for every X, so $RX \neq I_n$ and R is not invertible. If A were invertible, $R = GA$ would be invertible with inverse $A^{-1}G^{-1}$. So A is not invertible.

The algorithm. By Proposition 3.3.11, Gauss–Jordan elimination on $(A \mid I_n)$, scanning only the left block, ends at $(\operatorname{rref}(A) \mid G)$, where G is the product of the elementary matrices of the run on A. By the theorem, if the left block is I_n then $A^{-1} = G$, and otherwise A is not invertible.

(b) Each row operation is left multiplication by an elementary matrix (Proposition 3.2.7), so at every stage the current matrix is $G'(A \mid I_n) = (G'A \mid G')$ for some product G' of elementary matrices, which is invertible (Corollary 3.2.11). Let $\mathbf{z}$ be row i of G', the row i of the right block. By Definition 2.4.1, row i of $G'A$ is $\mathbf{z}A$, and this is the zero row by assumption. And $\mathbf{z} \neq \mathbf{0}$: otherwise row i of $G'X$ would be zero for every X, and G' would have no right inverse.

If A were invertible, then $\mathbf{z} = \mathbf{z}\big(AA^{-1}\big) = (\mathbf{z}A)A^{-1} = \mathbf{0}A^{-1} = \mathbf{0}$, a contradiction. So A is not invertible. The vector $\mathbf{z}$ is a *certificate*: anyone can check $\mathbf{z}A = \mathbf{0}$ by one multiplication, without redoing the elimination.

(c) Over $\mathbb{F}_5$ the entry 9 is 4. Pass 1 adds -2 times row 1 to row 2 and -3 times row 1 to row 3:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 4 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 3 & 1 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 \end{array}\right),$$

using $-4 = 1$, $-5 = 0$, $-2 = 3$, and $-3 = 2$ in $\mathbb{F}_5$. Pass 2 has its pivot at $(2,2)$; adding -1 times row 2 to row 3 makes row 3 of the left block zero, with right block row $(2-3, 0-1, 1) = (-1,-1,1) = (4,4,1)$. So $\mathbf{z} = (4,4,1)$, and indeed

$$\mathbf{z}A = -(1,2,3) - (2,0,1) + (3,2,9) = (0,0,5) = (0,0,0) \quad \text{in } \mathbb{F}_5.$$

So A is not invertible over $\mathbb{F}_5$: there, its third row is the sum of the first two.

Over $\mathbb{Q}$ the same computation gives $\mathbf{z}A = (0,0,5) \neq \mathbf{0}$, so the certificate is lost. Indeed, over $\mathbb{Q}$ pass 1 leaves the rows $(0,-4,-5)$ and $(0,-4,0)$ below the first, pass 2 scales the first of them to $(0,1,5/4)$ and adds 4 times it to the last, which becomes $(0,0,5)$. So a third pivot appears, $\operatorname{rref}(A) = I_3$, and A is invertible over $\mathbb{Q}$. Whether a matrix with integer entries is invertible can depend on the field in which the entries are read (compare Example 3.5.3).

Solution to Problem 3.4

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Lemma 3.3.16 (Reading the columns of a matrix in RREF), Theorem 2.4.34 (Block multiplication), Theorem 3.5.14 (Third Fundamental Theorem, completed), Lemma 2.4.24 (Matrix units), Theorem 3.3.12 (Matrices in RREF), Proposition 3.5.2 (First properties of rank), Exercise 3.5.41
Techniques. Reduce to RREF (Technique 3.3.22), Block decomposition (Technique 2.4.35)
Hint 1. For (a), start from $GA = A_r$. Permute the columns of A_r so that the pivot columns come first, and then clear the block next to I_k with one block column operation.
Hint 2. For (b), $N_k = \sum_{i=1}^{k} \mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$. Rank is unchanged by invertible factors on either side.
Hint 3. For (c), factor N_k as a product of an $m \times k$ and a $k \times n$ matrix.
Sketch. (a) $GA = A_r$; with P the permutation putting the pivot columns first, $A_rP = \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}$, and $U = \begin{pmatrix} I_k & -X \\ 0 & I \end{pmatrix}$ clears X: $GA(PU) = N_k$. (b) $A = G^{-1}N_kH^{-1} = \sum_{i \le k} G^{-1}\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}H^{-1}$, and each term has the rank of $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$, which is 1. (c) $N_k = \begin{pmatrix} I_k \\ 0 \end{pmatrix}\begin{pmatrix} I_k & 0 \end{pmatrix}$, so $A = CS$ with C the first k columns of G^{-1} and S the first k rows of H^{-1}.

(a) *Statement* (Theorem 3.5.27). Let $A \in F^{m\times n}$ have rank k, and let $N_k \in F^{m\times n}$ have 1 in positions $(1,1), \dots, (k,k)$ and 0 elsewhere. Then $GAH = N_k$ for some invertible $G \in F^{m\times m}$ and $H \in F^{n\times n}$.

Proof. Let G be invertible with $GA = A_r$ (Theorem 3.3.10), with pivot columns $p_1 < \cdots < p_k$ and free columns $f_1 < \cdots < f_\ell$, $\ell = n - k$. Let P have columns $\mathbf{e}_{p_1}, \dots, \mathbf{e}_{p_k}, \mathbf{e}_{f_1}, \dots, \mathbf{e}_{f_\ell}$; these are the standard basis vectors in another order, so they form a basis of F^n, and P is invertible (Theorem 3.5.14, Remark 3.5.7). Column j of A_rP is column p_j of A_r, which is $\mathbf{e}_j$, for $j \le k$ (Lemma 3.3.16(b)), and column f_i of A_r for $j = k + i$. All rows of A_r below row k are zero, so in block form $A_rP = \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}$ for some $X \in F^{k\times\ell}$. Let $U = \begin{pmatrix} I_k & -X \\ 0 & I_\ell \end{pmatrix}$. By block multiplication (Theorem 2.4.34), $U\begin{pmatrix} I_k & X \\ 0 & I_\ell \end{pmatrix} = I_n = \begin{pmatrix} I_k & X \\ 0 & I_\ell \end{pmatrix}U$, so U is invertible, and

$$A_rPU = \begin{pmatrix} I_k & X \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_k & -X \\ 0 & I_\ell \end{pmatrix} = \begin{pmatrix} I_k & -X + X \\ 0 & 0 \end{pmatrix} = N_k.$$

With $H = PU$, we get $GAH = A_rPU = N_k$. (If $k = 0$ or $\ell = 0$, the absent blocks are simply omitted.)

(b) By (a), $A = G^{-1}N_kH^{-1}$. By Lemma 2.4.24, $N_k = \sum_{i=1}^{k} \mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$, where $\mathbf{e}_i \in F^m$ on the left and $\mathbf{e}_i \in F^n$ on the right. So

$$A = \sum_{i=1}^{k} B_i, \quad B_i = G^{-1}\left(\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}\right)H^{-1}.$$

The matrix $\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ has a single nonzero entry, a 1 in position (i,i). Swapping row i to the top gives a matrix whose only nonzero row is $\mathbf{e}_i^{\mathsf{T}}$, which satisfies (R1)–(R4) and so is in RREF with one pivot (Theorem 3.3.12); so it has rank 1. Multiplying by invertible matrices on the left (Proposition 3.5.2(c)) and on the right (Exercise 3.5.41(c)) does not change the rank, so $\operatorname{rank} B_i = 1$.

(c) By block multiplication,

$$\begin{pmatrix} I_k \\ 0 \end{pmatrix}\begin{pmatrix} I_k & 0 \end{pmatrix} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} = N_k,$$

where the left factor is $m \times k$ and the right one is $k \times n$. So

$$A = G^{-1}N_kH^{-1} = \left(G^{-1}\begin{pmatrix} I_k \\ 0 \end{pmatrix}\right)\left(\begin{pmatrix} I_k & 0 \end{pmatrix}H^{-1}\right) = CS,$$

where C consists of the first k columns of G^{-1} and S of the first k rows of H^{-1}.

Problem 3.17 finds a factorization of this shape in which S is the nonzero part of $\operatorname{rref}(A)$, and shows that it is unique. Section 4.3 studies such factorizations in general.

Solution to Problem 3.5

Uses. Algorithm 3.3.1 (Gauss–Jordan elimination), Proposition 3.3.11 (Carrying extra columns along), Exercise 3.3.57, Corollary 3.3.19 (Pivot columns, intrinsically), Definition 3.4.8 (Basic solutions), Theorem 3.4.9 (All solutions), Corollary 3.4.4 (Consistency), Corollary 3.4.14 (Counting solutions), Proposition 3.5.8 (Right inverses from the pivots), Theorem 3.5.13 (Second Fundamental Theorem, completed)
Techniques. Reduce to RREF (Technique 3.3.22), Particular plus homogeneous (Technique 3.1.18)
Hint 1. Carry I_3 along. The two computations agree until a pivot must be divided by 5.
Hint 2. Over $\mathbb{F}_5$, read the relations off the free columns of $\operatorname{rref}(A)$, and the consistency condition off the last row of G.
Hint 3. For (d), use Proposition 3.5.8: place the rows of G in the pivot rows of a 5×3 matrix.
Sketch. After two passes, valid over any field, the rows are $(1,2,0,1,3)$, $(0,0,1,2,-5)$, $(0,0,0,0,5)$. Over $\mathbb{Q}$, $\operatorname{rank} A = 3$ with pivot columns $1,3,5$. Over $\mathbb{F}_5$, $5 = 0$: $\operatorname{rank} A = 2$, and $\mathbf{r}_1 - \mathbf{r}_2 + \mathbf{r}_3 = (0,0,0,0,5) = \mathbf{0}$. Over $\mathbb{F}_5$: $\mathbf{a}_2 = 2\mathbf{a}_1$, $\mathbf{a}_4 = \mathbf{a}_1 + 2\mathbf{a}_3$, $\mathbf{a}_5 = 3\mathbf{a}_1$; the condition is $y_1 - y_2 + y_3 = 0$ (25 vectors); the solutions of $A\mathbf{x} = (1,1,0)$ are $(1,0,4,0,0) + \operatorname{span}((3,1,0,0,0),(4,0,3,1,0),(2,0,0,0,1))$, 125 of them. (d) Over $\mathbb{Q}$ a right inverse has rows $(2/5, 3/5, -3/5)$, $\mathbf{0}$, $(-1,0,1)$, $\mathbf{0}$, $(1/5, -1/5, 1/5)$; over $\mathbb{F}_5$ there is none, since $\operatorname{rank} A = 2 < 3$.

(a) We reduce $(A \mid I_3)$. Pass 1 adds -2 times row 1 to row 2 and -1 times row 1 to row 3. In pass 2, column 2 vanishes in rows 2 and 3, so the pivot is at $(2,3)$, and we add -1 times row 2 to row 3. These steps are the same over any field:

$$\left(\begin{array}{ccccc|ccc} 1 & 2 & 0 & 1 & 3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 4 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 3 & 3 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccccc|ccc} 1 & 2 & 0 & 1 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & -5 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 & -1 & 1 \end{array}\right).$$

Over $\mathbb{Q}$. Pass 3 scales row 3 by $1/5$ and clears column 5, adding -3 times row 3 to row 1 and 5 times row 3 to row 2:

$$\rightsquigarrow \left(\begin{array}{ccccc|ccc} 1 & 2 & 0 & 1 & 0 & 2/5 & 3/5 & -3/5 \\ 0 & 0 & 1 & 2 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1/5 & -1/5 & 1/5 \end{array}\right).$$

So $\operatorname{rank} A = 3$, with pivot columns $1,3,5$.

Over $\mathbb{F}_5$. Here $5 = 0$ and $-5 = 0$, so after pass 2 the third row of the left block is zero and the algorithm halts:

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 2 & 0 & 1 & 3 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & 4 & 1 \end{pmatrix}.$$

So $\operatorname{rank} A = 2$, with pivot columns $1,3$. The last row of G is a relation among the rows $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ of A (Exercise 3.3.57):

$$\mathbf{r}_1 - \mathbf{r}_2 + \mathbf{r}_3 = (1-2+1, 2-4+2, 0-1+1, 1-4+3, 3-1+3) = (0,0,0,0,5).$$

This is $\mathbf{0}$ in $\mathbb{F}_5$ but not in $\mathbb{Q}$, and it is exactly why the third pivot disappears.

(b) Over $\mathbb{F}_5$ the free columns are $2,4,5$. By Corollary 3.3.19(b), with the entries of $\operatorname{rref}(A)$ as coefficients,

$$\mathbf{a}_2 = 2\mathbf{a}_1, \quad \mathbf{a}_4 = \mathbf{a}_1 + 2\mathbf{a}_3, \quad \mathbf{a}_5 = 3\mathbf{a}_1.$$

(Check: $\mathbf{a}_1 + 2\mathbf{a}_3 = (1,2,1) + (0,2,2) = (1,4,3) = \mathbf{a}_4$, and $3\mathbf{a}_1 = (3,6,3) = (3,1,3) = \mathbf{a}_5$.) The basic solutions (Definition 3.4.8) are

$$\mathbf{s}_1 = (-2,1,0,0,0) = (3,1,0,0,0), \quad \mathbf{s}_2 = (-1,0,-2,1,0) = (4,0,3,1,0), \quad \mathbf{s}_3 = (-3,0,0,0,1) = (2,0,0,0,1),$$

and $\ker T_A = \operatorname{span}(\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3)$ by Theorem 3.4.9(b).

(c) By Corollary 3.4.4, $A\mathbf{x} = \mathbf{y}$ is consistent if and only if the third entry of $G\mathbf{y}$ vanishes, that is, $y_1 - y_2 + y_3 = 0$. The vectors satisfying it are determined by $y_1, y_2 \in \mathbb{F}_5$ (then $y_3 = y_2 - y_1$), so there are 25 of them. For $\mathbf{y} = (1,1,0)$, $G\mathbf{y} = (1, -2+1, 1-1+0) = (1,4,0)$, so the system is consistent, with particular solution $\mathbf{x}_0 = (1,0,4,0,0)$ (pivot variables $x_1 = 1, x_3 = 4$, free variables 0). Check: $A\mathbf{x}_0 = \mathbf{a}_1 + 4\mathbf{a}_3 = (1, 2+4, 1+4) = (1,1,0)$. By Theorem 3.4.9 the solutions are

$$\mathbf{x}_0 + t_1\mathbf{s}_1 + t_2\mathbf{s}_2 + t_3\mathbf{s}_3 \quad (t_1, t_2, t_3 \in \mathbb{F}_5),$$

and there are $5^3 = 125$ of them (Corollary 3.4.14(b)).

(d) Over $\mathbb{Q}$, $\operatorname{rank} A = 3$ is the number of rows, so by Proposition 3.5.8 the matrix $X = BG$, which has row i of G in row p_i and zeros in the free rows, is a right inverse:

$$X = \begin{pmatrix} 2/5 & 3/5 & -3/5 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1/5 & -1/5 & 1/5 \end{pmatrix}.$$

Check: the first row of A times the columns of X gives $(2/5 + 3/5, 3/5 - 3/5, -3/5 + 3/5) = (1,0,0)$, and the other rows work the same way. Over $\mathbb{F}_5$, $\operatorname{rank} A = 2 < 3$, so T_A is not surjective and A has no right inverse (Theorem 3.5.13).

Solution to Problem 3.6

Uses. Theorem 3.3.23 (Invertibility and RREF), Theorem 3.4.9 (All solutions), Definition 1.1.29 (Complex numbers)
Techniques. Reduce to RREF (Technique 3.3.22), Try small cases and small fields (Technique 1.2.38)
Hint 1. Row reduce $A - \lambda I_2 = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}$ with λ as a symbol. Swap the rows first, so that you never divide by λ.
Hint 2. You should reach $\begin{pmatrix} 1 & -\lambda \\ 0 & -(1+\lambda^2) \end{pmatrix}$. So everything depends on whether $\lambda^2 = -1$ in the field.
Sketch. Swapping and adding λ times row 1 to row 2 gives $\begin{pmatrix} 1 & -\lambda \\ 0 & -(1+\lambda^2) \end{pmatrix}$, so $A - \lambda I_2$ is invertible iff $\lambda^2 \neq -1$, and otherwise its kernel is $\operatorname{span}((\lambda, 1))$. Solutions of $\lambda^2 = -1$: none in $\mathbb{R}$ and $\mathbb{F}_3$; $\pm i$ in $\mathbb{C}$; 1 in $\mathbb{F}_2$; 2, 3 in $\mathbb{F}_5$.

The equation $A\mathbf{x} = \lambda\mathbf{x}$ says $(A - \lambda I_2)\mathbf{x} = \mathbf{0}$. We row reduce $A - \lambda I_2 = \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix}$ over any field. Swapping the rows and then adding λ times row 1 to row 2 gives

$$\begin{pmatrix} 1 & -\lambda \\ -\lambda & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -\lambda \\ 0 & -1 - \lambda^2 \end{pmatrix}.$$

If $1 + \lambda^2 \neq 0$, the second row has a pivot, and clearing gives I_2; so $A - \lambda I_2$ is invertible (Theorem 3.3.23) and only $\mathbf{x} = \mathbf{0}$ satisfies $A\mathbf{x} = \lambda\mathbf{x}$. If $1 + \lambda^2 = 0$, the reduced form is $\begin{pmatrix} 1 & -\lambda \\ 0 & 0 \end{pmatrix}$, so $A - \lambda I_2$ is not invertible, and by Theorem 3.4.9 the solutions are the multiples of the basic solution $(\lambda, 1)$. (Check: $A(\lambda,1) = (-1, \lambda) = (\lambda^2, \lambda) = \lambda(\lambda,1)$, using $\lambda^2 = -1$.)

So we need the solutions of $\lambda^2 = -1$ in each field.

- $\mathbb{R}$: none, since $\lambda^2 \geq 0 > -1$. So $A - \lambda I_2$ is invertible for every real λ.
- $\mathbb{C}$: $\lambda = \pm i$, with $\mathbf{x} \in \operatorname{span}((i,1))$, respectively $\operatorname{span}((-i,1))$.
- $\mathbb{F}_2$: $-1 = 1$, and $\lambda = 1$ works ($0^2 = 0$ does not). The solutions are the multiples of $(1,1)$, that is, $(0,0)$ and $(1,1)$.

- $\mathbb{F}_3$: the squares are $0, 1, 1$, and $-1 = 2$ is not among them. No λ.
- $\mathbb{F}_5$: $-1 = 4 = 2^2 = 3^2$, so $\lambda = 2$ or 3, with $\mathbf{x} \in \operatorname{span}((2,1))$, respectively $\operatorname{span}((3,1))$.

Over $\mathbb{R}$, T_A rotates the plane by $90°$, so no nonzero vector is sent to a multiple of itself, and the computation agrees. Whether such vectors exist depends on the field, through the single equation $\lambda^2 = -1$; Section 7.2 takes this up.

Solution to Problem 3.7

Uses. Proposition 3.5.2 (First properties of rank), Proposition 3.3.8 (Properties of the output), Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Corollary 3.2.11 (Products of elementary matrices)
Techniques. Reduce to elementary matrices (Technique 5.4.6), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (a) Does a pass of the algorithm ever start on a row of zeros? Try the zero matrix.
Hint 2. (b) Which matrices are products of elementary matrices? Try $A = 0$.
Sketch. (a) A pass creates a pivot only in a column with a nonzero entry in an unpivoted row; zero rows never get pivots. $\operatorname{rank} A \le \min(m,n)$ bounds the rank from above only. Counterexample: $A = (0) \in F^{1\times 1}$, or $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. (b) Only invertible matrices are products of elementary matrices. Counterexample: $A = 0$, $B = I_m$.

(a) The step "every row receives a pivot" is wrong. A pass of Gauss–Jordan elimination creates a pivot only in a column that has a nonzero entry in an unpivoted row, and the algorithm halts as soon as the unpivoted rows are zero (step 1 of Algorithm 3.3.1). The count of rows and columns gives $\operatorname{rank} A \le \min(m,n)$ (Proposition 3.5.2(a)), an upper bound, not equality. For example, $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has $m = n = 2$ and $\operatorname{rref}(A) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, so $\operatorname{rank} A = 1 < 2$. It has no right inverse: if $AX = I_2$, then the two rows of AX would be equal (both rows of A are equal), but the rows of I_2 are not. The smallest counterexample is $A = (0) \in F^{1\times 1}$.

(b) The first step is wrong: Theorem 3.5.19 applies only to invertible matrices, and only those are products of elementary matrices (Corollary 3.2.11). For $A = 0$ and $B = I_m$, $\operatorname{rank}(AB) = \operatorname{rank} 0 = 0$, while $\operatorname{rank} B = m \ge 1$. The argument does prove the claim when A is invertible, which is Proposition 3.5.2(c). (In general only $\operatorname{rank}(AB) \le \operatorname{rank} B$ holds; see Section 4.3.)

Solution to Problem 3.8

Uses. Corollary 3.1.22 (Uniqueness depends only on A), Theorem 3.5.12 (First Fundamental Theorem, completed), Exercise 3.4.37, Proposition 3.5.2 (First properties of rank), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 3.3.12 (Matrices in RREF)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Reduce to RREF (Technique 3.3.22)
Hint 1. (a) A unique solution for one $\mathbf{y}$ says something about $\ker T_A$. (b) Take $\mathbf{y} = A\mathbf{x}$ for an arbitrary $\mathbf{x}$.
Hint 2. (c) Exercise 3.4.37 describes the free columns using only the kernel. (d) Column operations move pivots: try a 1×2 example.
Hint 3. (e) Column j of GA is $G\mathbf{a}_j$. (f) Add two matrix units.
Sketch. (a) True: $\ker T_A = \{\mathbf{0}\}$ by Corollary 3.1.22, so $\operatorname{rank} A = n \le m$. (b) True: $\mathbf{x}$ solves $A\mathbf{x} = A\mathbf{x}$, hence $B\mathbf{x} = A\mathbf{x}$ for every $\mathbf{x}$. (c) True: by Exercise 3.4.37 the free columns are determined by the kernel, and rank = n − (number of free columns). (d) False: $(0\ \ 1)$ and $(0\ \ 1)E_{12} = (1\ \ 0)$. (e) True: $G\mathbf{a}_j = \mathbf{0}$ with G invertible. (f) False: $\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}} + \mathbf{e}_2\mathbf{e}_2^{\mathsf{T}} = I_2$ has rank 2.

(a) True. If $A\mathbf{x} = \mathbf{y}$ has exactly one solution, then $\ker T_A = \{\mathbf{0}\}$ by Corollary 3.1.22(a). By Theorem 3.5.12, $\operatorname{rank} A = n$, and $n = \operatorname{rank} A \le m$.

(b) True. Let $\mathbf{x} \in F^n$ and $\mathbf{y} = A\mathbf{x}$. Then $\mathbf{x}$ solves $A\mathbf{x} = \mathbf{y}$, so by hypothesis it solves $B\mathbf{x} = \mathbf{y}$, that is, $B\mathbf{x} = A\mathbf{x}$. With $\mathbf{x} = \mathbf{e}_j$, column j of B equals column j of A, for every j.

(c) True. By Exercise 3.4.37, x_j is a free variable of A if and only if some $\mathbf{x} \in \ker T_A$ has $x_j = 1$ and $x_l = 0$ for all $l > j$. This condition involves only $\ker T_A$, so A and B have the same free columns, and by Proposition 3.5.2(b), $\operatorname{rank} A = n -$ (number of free columns) $= \operatorname{rank} B$. The number of rows plays no role.

(d) False. Let $A = (0\ \ 1)$ and $H = E_{12}$. Both A and $AH = (1\ \ 0)$ are in RREF (Theorem 3.3.12), with pivot columns 2 and 1 respectively.

(e) True. Write $\operatorname{rref}(A) = GA$ with G invertible (Theorem 3.3.10). Column j of GA is $G\mathbf{a}_j$, so if it is zero, then $\mathbf{a}_j = G^{-1}G\mathbf{a}_j = \mathbf{0}$.

(f) False. The matrix units $\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{e}_2\mathbf{e}_2^{\mathsf{T}} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ each have exactly one nonzero column, so by Proposition 3.5.2(d) each has rank 1. Their sum I_2 has rank 2. So the set is not closed under addition.

Solution to Problem 3.9

> *Uses.* Exercise 3.4.39, Proposition 3.4.16 (A pivot in every row), Proposition 3.5.2 (First properties of rank), Corollary 3.4.14 (Counting solutions), Proposition 3.4.15 (A pivot in every column), Corollary 3.1.22 (Uniqueness depends only on A), Definition 1.2.17 (Evaluation)
> *Techniques.* Undetermined coefficients (Technique 3.1.9), Particular plus homogeneous (Technique 3.1.18)
> *Hint 1.* Write $p = a_0 + a_1x + \cdots + a_nx^n$. The conditions are a linear system $W\mathbf{a} = \mathbf{d}$ whose coefficient matrix has rows $(1, c_i, \ldots, c_i^n)$.
> *Hint 2.* Exercise 3.4.39 says where W has its pivots. Count free columns.
> *Sketch.* The conditions are $W\mathbf{a} = \mathbf{d}$ with $W \in F^{k\times(n+1)}$, rows $(1, c_i, \ldots, c_i^n)$. (a) If $k \le n+1$, W has a pivot in every row, so $\operatorname{rank} W = k$: the system is consistent and has $n+1-k$ free variables, hence q^{n+1-k} solutions. (b) If $k > n+1$, W has a pivot in every column (at most one solution) but not in every row (some $\mathbf{d}$ inconsistent). (c) $7^2 = 49$ and $7^1 = 7$.

Write $p = a_0 + a_1x + \cdots + a_nx^n$, so that p corresponds to $\mathbf{a} = (a_0, \ldots, a_n) \in F^{n+1}$. By Definition 1.2.17, $p(c_i) = d_i$ reads $a_0 + a_1c_i + \cdots + a_nc_i^n = d_i$. So the conditions form the linear system $W\mathbf{a} = \mathbf{d}$, where $\mathbf{d} = (d_1, \ldots, d_k)$ and $W \in F^{k\times(n+1)}$ has rows $\left(1, c_i, c_i^2, \ldots, c_i^n\right)$, the matrix of Exercise 3.4.39 with $m = k$ and $d = n$.

(a) Since $k \le n+1$, Exercise 3.4.39(b) shows that W_r has a pivot in every row, so $\operatorname{rank} W = k$ and the system is consistent for every $\mathbf{d}$ (Proposition 3.4.16). It has $(n+1) - k$ free columns (Proposition 3.5.2(b)), so by Corollary 3.4.14(b) it has exactly q^{n+1-k} solutions. Each solution $\mathbf{a}$ is one polynomial.

(b) Since $k > n+1$, Exercise 3.4.39(a) shows that W_r has a pivot in every column, so $W\mathbf{a} = \mathbf{0}$ has only the trivial solution (Proposition 3.4.15), and every system $W\mathbf{a} = \mathbf{d}$ has at most one solution (Corollary 3.1.22). By Exercise 3.4.39(b), W_r does not have a pivot in every row, so some $W\mathbf{a} = \mathbf{d}$ is inconsistent (Proposition 3.4.16).

(c) Here $q = 7$, $n = 4$, and the nodes $1, 2, 4$ are distinct, so $k = 3 \le 5$ and there are $7^{5-3} = 49$ such polynomials. Adding the condition $p(0) = 0$ gives $k = 4$ distinct nodes $0, 1, 2, 4$, and $7^{5-4} = 7$ polynomials.

In words: each condition at a new node cuts the number of polynomials by a factor of q, until there are $n+1$ conditions and exactly one polynomial is left (Theorem 3.4.24).

Solution to Problem 3.10

> *Uses.* Corollary 3.3.19 (Pivot columns, intrinsically), Lemma 1.5.10 (Linear dependence lemma), Proposition 1.4.14 (The span is the smallest subspace containing a list), Definition 1.6.1 (Basis), Theorem 3.5.13 (Second Fundamental Theorem, completed)
> *Techniques.* Reduce to RREF (Technique 3.3.22)
> *Hint 1.* By Corollary 3.3.19(a), a column is a pivot column exactly when it is not a combination of the columns before it. Apply this to the columns of $(A \mid I_m)$.
> *Hint 2.* The pivot columns are independent: none is a combination of the earlier ones. They span: every column of $(A \mid I_m)$, in particular every $\mathbf{e}_i$, is a combination of them.
> *Hint 3.* For (c), compare the span of the list built so far with the span of all columns of $(A \mid I_m)$ before column $n+i$.
> *Sketch.* (a) For $j \le n$, $\mathbf{a}_j \notin \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1})$ by independence, so columns $1, \ldots, n$ are pivot columns. No pivot column is a combination of earlier pivot columns, so the pivot columns are independent (Lemma 1.5.10), and every column is a combination of them (Corollary 3.3.19(b)), so their span contains $\mathbf{e}_1, \ldots, \mathbf{e}_m$. The columns of $(A \mid I_m)$ span F^m, so its rank is m. (b) Pivot columns $1, 2, 3, 6$: the basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}_1, \mathbf{e}_4$, with $\mathbf{e}_2 = \mathbf{a}_1 - 2\mathbf{e}_1$, $\mathbf{e}_3 = -\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{e}_1$. (c) By induction, the list built before step i spans the same subspace as $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_1, \ldots, \mathbf{e}_{i-1}$.

(a) The columns of $M = (A \mid I_m)$ are $\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_1, \ldots, \mathbf{e}_m$. By Corollary 3.3.19(a), applied to M, column j is a pivot column if and only if it is not a linear combination of columns $1, \ldots, j-1$. For $j \le n$ these are $\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}$, and $\mathbf{a}_j$ is not a combination of them because the list $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is independent (a combination would give a nontrivial relation with coefficient -1 on $\mathbf{a}_j$). So columns $1, \ldots, n$ are pivot columns; let the others be $n + i_1 < \cdots < n + i_t$. Let $\mathcal{P} = \mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_t}$ be the list of pivot columns, in order.

$\mathcal{P}$ *is independent.* No pivot column is a combination of the columns before it, so in particular none is a combination of the pivot columns before it. If $\mathcal{P}$ were dependent, the linear dependence lemma (Lemma 1.5.10) would give an entry in the span of the earlier entries. So $\mathcal{P}$ is independent.

$\mathcal{P}$ *spans* F^m. By Corollary 3.3.19(b), every non-pivot column of M is a combination of the pivot columns to its left, so every column of M lies in $\operatorname{span}(\mathcal{P})$. In particular $\mathbf{e}_1, \ldots, \mathbf{e}_m \in \operatorname{span}(\mathcal{P})$, and since $\operatorname{span}(\mathcal{P})$ is a subspace (Proposition 1.4.14), it contains every combination of them, which is all of F^m.

So $\mathcal{P}$ is a basis of F^m (Definition 1.6.1). Finally, the columns of M span F^m, so $\operatorname{rank} M = m$ by Theorem 3.5.13 ((III) ⇔ (VI)), and $\operatorname{rank} M$ is the number of pivot columns, $n + t$.

(b) We apply the criterion of (a) to the columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ one at a time, without reducing.

- $\mathbf{a}_1 \neq \mathbf{0}$ and $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$ (third entries), as (a) predicts.
- $\mathbf{e}_1$: a combination $x\mathbf{a}_1 + y\mathbf{a}_2 = (2x + y, x + y, y, 0)$ equal to $\mathbf{e}_1$ needs $y = 0$, then $x = 0$, and then $2x + y = 0 \neq 1$. So column 3 is a pivot column.
- $\mathbf{e}_2 = \mathbf{a}_1 - 2\mathbf{e}_1$, so column 4 is not.
- $\mathbf{e}_3 = -\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{e}_1$, since $(-2 + 1 + 1, -1 + 1, 1, 0) = (0, 0, 1, 0)$; so column 5 is not.
- $\mathbf{e}_4$: every earlier column has fourth entry 0, and so does every combination of them. So column 6 is a pivot column.

The pivot columns are $1, 2, 3, 6$, and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{e}_1, \mathbf{e}_4$ is a basis of $\mathbb{Q}^4$. (Row reducing $(A \mid I_4)$ confirms this: the result has pivots in columns $1, 2, 3, 6$, and its columns 4 and 5 are $(1, 0, -2, 0)$ and $(-1, 1, 1, 0)$, the coefficients found above.)

(c) Let $\mathcal{L}_{i-1}$ be the list built before step i, and $V_{i-1} = \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n, \mathbf{e}_1, \ldots, \mathbf{e}_{i-1})$, the span of the columns of M before column $n + i$. We show by induction on i that $\operatorname{span}(\mathcal{L}_{i-1}) = V_{i-1}$. For $i = 1$ both are $\operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Suppose it holds for i. If $\mathbf{e}_i \in V_{i-1} = \operatorname{span}(\mathcal{L}_{i-1})$, then $\mathcal{L}_i = \mathcal{L}_{i-1}$ and $V_i = V_{i-1}$, because adding a vector of a span to the spanning list does not change the span. Otherwise $\mathbf{e}_i$ is appended to both lists, and the spans stay equal. This completes the induction.

So the procedure appends $\mathbf{e}_i$ exactly when $\mathbf{e}_i \notin V_{i-1}$, that is, by Corollary 3.3.19(a), exactly when column $n + i$ of M is a pivot column. The greedy extension and the pivot columns of $(A \mid I_m)$ are the same construction; the second one can be carried out by a single row reduction.

Solution to Problem 3.11

> *Uses.* Exercise 3.6.9, Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Definition 2.3.8 (Matrix–vector product), Definition 2.4.1 (Matrix product), Theorem 3.5.13 (Second Fundamental Theorem, completed), Proposition 3.3.8 (Properties of the output), Theorem 3.4.9 (All solutions), Exercise 3.3.57
> *Techniques.* Undetermined coefficients (Technique 3.1.9), Reduce to RREF (Technique 3.3.22)
> *Hint 1.* Let U and W be the matrices with columns $\mathbf{u}_j$ and $\mathbf{w}_j$, and write $T = T_X$. Then the conditions say $XU = W$, and the relations among the $\mathbf{u}_j$ are the vectors of $\ker T_U$.
> *Hint 2.* Now (a) is Exercise 3.6.9. For (b), two solutions differ by a Z with $ZU = 0$. If the $\mathbf{u}_j$ do not span, the last row of an invertible G with $GU = \operatorname{rref}(U)$ gives such a Z.
> *Hint 3.* For (c), every relation among the $\mathbf{u}_j$ is a multiple of one basic solution. Build X from G as in the proof of Exercise 3.6.9.
> *Sketch.* With $T = T_X$, the conditions are $XU = W$, and "relations transfer" is $\ker T_U \subseteq \ker T_W$; (a) is Exercise 3.6.9. (b) If the $\mathbf{u}_j$ span, $ZU = 0$ forces $Z = 0$. If not, $\operatorname{rank} U < n$, row n of $\operatorname{rref}(U) = GU$ is zero, and $\mathbf{z} =$ row n of G is nonzero with $\mathbf{z}U = \mathbf{0}$; adding $\mathbf{z}$ to the first row of X gives another solution. (c) The relations are the multiples of $(2, 1, -1)$, and $2\mathbf{w}_1 + \mathbf{w}_2 - \mathbf{w}_3 = \mathbf{0}$; the solutions are $X = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix}(-1 \;\; 1 \;\; 1)$. With $\mathbf{w}_3 = (2, 5)$ there is none.

Let $U \in F^{n \times k}$ and $W \in F^{m \times k}$ have columns $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_k$. Every linear map $T : F^n \to F^m$ is T_X for exactly one $X \in F^{m \times n}$ (Proposition 2.3.17), and column j of XU is $X\mathbf{u}_j$ (Definition 2.4.1). So $T_X(\mathbf{u}_j) = \mathbf{w}_j$ for all j if and only if $XU = W$. By Definition 2.3.8, a relation $\sum_j c_j \mathbf{u}_j = \mathbf{0}$ is a vector $\mathbf{c} \in \ker T_U$, and it holds among the $\mathbf{w}_j$ exactly when $\mathbf{c} \in \ker T_W$.

(a) So the condition on relations says $\ker T_U \subseteq \ker T_W$, and the existence of T says $W = XU$ for some X. These are equivalent by Exercise 3.6.9(a), applied with $A = U$ and $B = W$.

(b) Let $XU = W$. Another matrix X' satisfies $X'U = W$ if and only if $Z = X' - X$ satisfies $ZU = 0$. So T is unique if and only if $ZU = 0$ forces $Z = 0$.

Suppose the $\mathbf{u}_j$ span F^n, and $ZU = 0$. Then $Z\mathbf{u}_j = \mathbf{0}$ for every j, so $Z\mathbf{v} = \mathbf{0}$ for every combination $\mathbf{v}$ of the $\mathbf{u}_j$, which is every $\mathbf{v} \in F^n$. In particular every column $Z\mathbf{e}_i$ of Z is zero, and $Z = 0$.

Suppose the $\mathbf{u}_j$ do not span F^n. By Theorem 3.5.13 ((III) $\Leftrightarrow$ (VI)), $\operatorname{rank} U < n$, so row n of $\operatorname{rref}(U)$ is zero (Proposition 3.3.8(a)). Let G be invertible with $GU = \operatorname{rref}(U)$, and let $\mathbf{z}$ be row n of G. Then $\mathbf{z}U$ is row n of GU, which is zero, and $\mathbf{z} \neq \mathbf{0}$ since G is invertible. Let $Z \in F^{m\times n}$ have first row $\mathbf{z}$ and all other rows zero. Then $Z \neq 0$ and $ZU = 0$ (its first row is $\mathbf{z}U = \mathbf{0}$), so $X + Z$ is a second solution.

(c) Row reducing U (adding -1 times row 1 to row 3, then -1 times row 2 to row 1 and row 2 to row 3) gives

$$U = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}, \quad \operatorname{rref}(U) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

The basic solution is $(-2, -1, 1)$, so by Theorem 3.4.9 the relations among the $\mathbf{u}_j$ are exactly the multiples of $-2\mathbf{u}_1 - \mathbf{u}_2 + \mathbf{u}_3 = \mathbf{0}$. And $-2\mathbf{w}_1 - \mathbf{w}_2 + \mathbf{w}_3 = (-2, -2) - (0, 2) + (2, 4) = \mathbf{0}$, so by (a) a map exists. Following the proof of Exercise 3.6.9, let Y have columns $\mathbf{w}_1, \mathbf{w}_2$ (the $\mathbf{w}$'s at the pivot columns $1, 2$) and $\mathbf{0}$; then

$$X = YG = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Check: $X\mathbf{u}_1 = (1, 1)$, $X\mathbf{u}_2 = (1 - 1, 1 + 1) = (0, 2)$, $X\mathbf{u}_3 = (3 - 1, 3 + 1) = (2, 4)$. Since $\operatorname{rank} U = 2 < 3$, T is not unique. By (b) and its proof, the other solutions differ from X by matrices Z with $ZU = 0$, that is, whose rows $\mathbf{z}$ satisfy $\mathbf{z}U = \mathbf{0}$. By Exercise 3.3.57 these rows are the multiples of the last row $(-1, 1, 1)$ of G. So the maps are exactly

$$T(\mathbf{x}) = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}\mathbf{x} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix}(-x_1 + x_2 + x_3) \quad (\alpha, \beta \in \mathbb{Q}).$$

If $\mathbf{w}_3 = (2, 5)$, then $-2\mathbf{w}_1 - \mathbf{w}_2 + \mathbf{w}_3 = (0, 1) \neq \mathbf{0}$: the relation among the $\mathbf{u}_j$ fails for the $\mathbf{w}_j$, so by (a) there is no such map.

Solution to Problem 3.12

Uses. Lemma 2.4.24 (Matrix units), Definition 2.3.1 (Matrix of a linear map), Proposition 2.4.14 (Algebra of matrices), Proposition 3.5.2 (First properties of rank), Definition 1.3.10 (Span)
Techniques. Translate between maps and matrices (Technique 2.3.27), Reduce to RREF (Technique 3.3.22)
Hint 1. Column j of M is the coordinate vector of C_A applied to the jth matrix unit. Compute $A\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}} - \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}A$ entry by entry, or compute $C_A(X)$ for a general X and read off the coefficients.
Hint 2. For (b), use Proposition 3.5.2(d): count the columns of M that are not combinations of the earlier ones. Split into the cases $c \neq 0$; $c = 0, b \neq 0$; $b = c = 0$.
Sketch. $C_A(X)$ has entries $bx_{21} - cx_{12}$, $-bx_{11} + (a - d)x_{12} + bx_{22}$, $cx_{11} + (d - a)x_{21} - cx_{22}$, $cx_{12} - bx_{21}$, so M has columns $\mathbf{m}_1 = (0, -b, c, 0)$, $\mathbf{m}_2 = (-c, a - d, 0, c)$, $\mathbf{m}_3 = (b, 0, d - a, -b)$, $\mathbf{m}_4 = -\mathbf{m}_1$. Column 4 never counts. If $c \neq 0$: $\mathbf{m}_1, \mathbf{m}_2$ count, $\mathbf{m}_3 = ((d - a)/c)\mathbf{m}_1 - (b/c)\mathbf{m}_2$ does not. If $c = 0 \neq b$: $\mathbf{m}_1$ and $\mathbf{m}_3$ count. If $b = c = 0$: rank 2 when $a \neq d$, and $M = 0$ when $a = d$. (c) $C_A(X) = \left(\begin{smallmatrix} x_{21} & x_{22} - x_{11} \\ 0 & -x_{21} \end{smallmatrix}\right)$: kernel $\operatorname{span}(I_2, A)$, image the Y with $y_{21} = 0$ and $y_{11} + y_{22} = 0$.

(a) By the distributive laws for matrix multiplication (Proposition 2.4.14), $C_A(X + X') = AX + AX' - XA - X'A = C_A(X) + C_A(X')$ and $C_A(tX) = tC_A(X)$, so C_A is linear. For $X = \left(\begin{smallmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{smallmatrix}\right)$,

$$AX - XA = \begin{pmatrix} ax_{11} + bx_{21} & ax_{12} + bx_{22} \\ cx_{11} + dx_{21} & cx_{12} + dx_{22} \end{pmatrix} - \begin{pmatrix} ax_{11} + cx_{12} & bx_{11} + dx_{12} \\ ax_{21} + cx_{22} & bx_{21} + dx_{22} \end{pmatrix},$$

so

$$C_A(X) = \begin{pmatrix} bx_{21} - cx_{12} & -bx_{11} + (a-d)x_{12} + bx_{22} \\ cx_{11} + (d-a)x_{21} - cx_{22} & cx_{12} - bx_{21} \end{pmatrix}.$$

The coordinates of X in the basis of matrix units are $(x_{11}, x_{12}, x_{21}, x_{22})$ (Lemma 2.4.24(a)), so by Definition 2.3.1, M is the matrix taking these coordinates to those of $C_A(X)$:

$$M = \begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}.$$

Its columns are $\mathbf{m}_1 = (0, -b, c, 0)$, $\mathbf{m}_2 = (-c, a-d, 0, c)$, $\mathbf{m}_3 = (b, 0, d-a, -b)$, and $\mathbf{m}_4 = (0, b, -c, 0) = -\mathbf{m}_1$.

(b) By Proposition 3.5.2(d), $\operatorname{rank} M$ is the number of j with $\mathbf{m}_j \notin \operatorname{span}(\mathbf{m}_1, \ldots, \mathbf{m}_{j-1})$. Since $\mathbf{m}_4 = -\mathbf{m}_1$, $j = 4$ never counts.

Case $c \neq 0$. Then $\mathbf{m}_1 \neq \mathbf{0}$, so $j = 1$ counts; and $\mathbf{m}_2 \notin \operatorname{span}(\mathbf{m}_1)$, since its first entry is $-c \neq 0$ while that of $\mathbf{m}_1$ is 0. But $\mathbf{m}_3$ is a combination of $\mathbf{m}_1$ and $\mathbf{m}_2$:

$$\frac{d-a}{c}\mathbf{m}_1 - \frac{b}{c}\mathbf{m}_2 = (b, -(d-a)b/c - (a-d)b/c, d-a, -b) = (b, 0, d-a, -b) = \mathbf{m}_3.$$

So $\operatorname{rank} M = 2$.

Case $c = 0$, $b \neq 0$. Then $\mathbf{m}_1 = (0, -b, 0, 0) \neq \mathbf{0}$ counts, $\mathbf{m}_2 = (0, a-d, 0, 0)$ is a multiple of $\mathbf{m}_1$, and $\mathbf{m}_3 = (b, 0, d-a, -b)$ is not in $\operatorname{span}(\mathbf{m}_1, \mathbf{m}_2)$, whose vectors have first entry 0. So $\operatorname{rank} M = 2$.

Case $b = c = 0$. Then $\mathbf{m}_1 = \mathbf{0}$, $\mathbf{m}_2 = (0, a-d, 0, 0)$, $\mathbf{m}_3 = (0, 0, d-a, 0)$. If $a \neq d$, then $j = 2$ and $j = 3$ count, and $\operatorname{rank} M = 2$. If $a = d$, then $M = 0$ and $\operatorname{rank} M = 0$.

The case $b = c = 0$, $a = d$ is exactly $A = aI_2$. So $\operatorname{rank} M = 0$ for scalar A and $\operatorname{rank} M = 2$ otherwise.

(c) Here $a = b = d = 1$ and $c = 0$, so by (a)

$$C_A(X) = \begin{pmatrix} x_{21} & x_{22} - x_{11} \\ 0 & -x_{21} \end{pmatrix}.$$

So $AX = XA$ if and only if $x_{21} = 0$ and $x_{11} = x_{22}$, that is, $X = \begin{pmatrix} s & t \\ 0 & s \end{pmatrix} = (s-t)I_2 + tA$ for some s, t. The matrices commuting with A are exactly those in $\operatorname{span}(I_2, A)$. The image consists of matrices Y with $y_{21} = 0$ and $y_{22} = -y_{11}$, and every such Y is an image: take $x_{21} = y_{11}$, $x_{22} = y_{12}$, $x_{11} = x_{12} = 0$. So the image is $\{Y \mid y_{21} = 0, y_{11} + y_{22} = 0\}$. Both descriptions use two parameters, matching $\operatorname{rank} M = 2$ and the $4 - 2$ free columns.

Solution to Problem 3.13

Uses. Corollary 3.4.20 (Square systems), Definition 1.2.17 (Evaluation)
Techniques. Uniqueness gives existence (Technique 3.4.21), Undetermined coefficients (Technique 3.1.9)
Hint 1. Count: eight coefficients, and the conditions $p(0) = d_1 = q(0)$ count as two equations.
Hint 2. For (b), it suffices to solve the homogeneous system. The conditions at 0 say that p and q have the same coefficients of 1, x, and x^2.
Hint 3. For (c), if (p, q) is the spline for symmetric data, then so is the pair $(\tilde{q}, \tilde{p})$, where $\tilde{q}(t) = q(-t)$ and $\tilde{p}(t) = p(-t)$. Uniqueness then says that $p(t) = q(-t)$.
Sketch. (a) Eight linear equations ($p(-1)$, $p(0)$, $q(0)$, $q(1)$, two matching conditions, two end conditions) in eight coefficients. (b) Homogeneous: $a_0 = b_0 = 0$, $a_1 = b_1 = u$, $a_2 = b_2 = w$; the end conditions give $a_3 = w/3$, $b_3 = -w/3$; then $p(-1) = 0$ gives $u = 2w/3$ and $q(1) = 0$ gives $u = -2w/3$, so $u = w = 0$. By Corollary 3.4.20 every system has exactly one solution. (c) By symmetry and uniqueness $p(t) = q(-t)$, so $b_1 = 0$, and $q(1) = 0$, $q''(1) = 0$ give $q = 1 - \frac{3}{2}x^2 + \frac{1}{2}x^3$, $p = 1 - \frac{3}{2}x^2 - \frac{1}{2}x^3$.

Write $p = a_0 + a_1x + a_2x^2 + a_3x^3$ and $q = b_0 + b_1x + b_2x^2 + b_3x^3$.

(a) By Definition 1.2.17 and the formulas for p' and p'', the conditions read

$$a_0 - a_1 + a_2 - a_3 = d_0, \quad a_0 = d_1, \quad b_0 = d_1, \quad b_0 + b_1 + b_2 + b_3 = d_2,$$
$$a_1 = b_1, \quad 2a_2 = 2b_2, \quad 2a_2 - 6a_3 = 0, \quad 2b_2 + 6b_3 = 0.$$

These are eight linear equations in the eight unknowns $a_0, \dots, a_3, b_0, \dots, b_3$.

(b) By Corollary 3.4.20 ((iv) ⇒ (vi)) it suffices to show that the homogeneous system, with $d_0 = d_1 = d_2 = 0$, has only the trivial solution. Its equations give $a_0 = b_0 = 0$, $a_1 = b_1 =: u$, $a_2 = b_2 =: w$, $a_3 = w/3$, and $b_3 = -w/3$. The two remaining equations become

$$-u + w - \frac{w}{3} = 0 \quad \text{and} \quad u + w - \frac{w}{3} = 0,$$

that is, $u = 2w/3$ and $u = -2w/3$. So $4w/3 = 0$, hence $w = 0$ and $u = 0$, and all eight coefficients vanish. Therefore every system of this form, for every d_0, d_1, d_2, has exactly one solution.

(c) Let (p, q) be the spline for $(0, 1, 0)$. Define $\tilde{p} = \sum_k (-1)^k a_k x^k$ and $\tilde{q} = \sum_k (-1)^k b_k x^k$, so that $\tilde{p}(t) = p(-t)$ and $\tilde{q}(t) = q(-t)$ for all t. From the formulas, $\tilde{q}'(t) = -q'(-t)$ and $\tilde{q}''(t) = q''(-t)$, and likewise for $\tilde{p}$. We check that $(\tilde{q}, \tilde{p})$ is also a natural cubic spline for $(0, 1, 0)$:

- $\tilde{q}(-1) = q(1) = 0$, $\tilde{q}(0) = q(0) = 1 = p(0) = \tilde{p}(0)$, and $\tilde{p}(1) = p(-1) = 0$;
- $\tilde{q}'(0) = -q'(0) = -p'(0) = \tilde{p}'(0)$ and $\tilde{q}''(0) = q''(0) = p''(0) = \tilde{p}''(0)$;
- $\tilde{q}''(-1) = q''(1) = 0$ and $\tilde{p}''(1) = p''(-1) = 0$.

By the uniqueness in (b), $(\tilde{q}, \tilde{p}) = (p, q)$, so $a_k = (-1)^k b_k$ for every k.

So it remains to find q. From $a_1 = -b_1$ and $a_1 = b_1$ we get $b_1 = 0$; and $b_0 = 1$. The conditions $q(1) = 0$ and $q''(1) = 0$ give $1 + b_2 + b_3 = 0$ and $2b_2 + 6b_3 = 0$, so $b_2 = -3b_3$ and $1 - 2b_3 = 0$: $b_3 = 1/2$, $b_2 = -3/2$. Hence

$$q = 1 - \frac{3}{2}x^2 + \frac{1}{2}x^3, \quad p = 1 - \frac{3}{2}x^2 - \frac{1}{2}x^3.$$

Check the conditions not used: $p(-1) = 1 - 3/2 + 1/2 = 0$ and $p''(-1) = -3 - 3 \cdot (-1) = 0$. The spline is the function $t \mapsto 1 - \frac{3}{2}t^2 + \frac{1}{2}|t|^3$ on $[-1, 1]$.

The same holds for any nodes $t_0 < t_1 < \cdots < t_N$, with one cubic on each interval. Counting conditions again gives a square system, and uniqueness follows from an integration by parts: for a homogeneous spline s, $\int (s'')^2 = 0$, so s'' vanishes, s is linear on each interval and zero at the nodes. Solving the system by hand would be hopeless, but uniqueness is enough.

Solution to Problem 3.14

Uses. Theorem 2.4.34 (Block multiplication), Definition 2.5.12 (Invertible), Proposition 3.5.2 (First properties of rank), Exercise 3.5.41, Exercise 3.5.36, Corollary 3.5.25 (The invertible matrix theorem), Theorem 3.5.14 (Third Fundamental Theorem, completed)
Techniques. Block decomposition (Technique 2.4.35), Reduce to RREF (Technique 3.3.22)
Hint 1. Imitate one pass of elimination with blocks: subtract CA^{-1} times the first block row from the second. Then clear B with a block column operation.
Hint 2. $P = \begin{pmatrix} I_k & 0 \\ -CA^{-1} & I_\ell \end{pmatrix}$ and $Q = \begin{pmatrix} I_k & -A^{-1}B \\ 0 & I_\ell \end{pmatrix}$. Their inverses change the sign of the off-diagonal block.
Hint 3. Rank is unchanged by invertible factors on both sides, and Exercise 3.5.36 computes the rank of a block diagonal matrix.
Sketch. $PM = \begin{pmatrix} A & B \\ 0 & S \end{pmatrix}$ and $(PM)Q = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}$, with P, Q invertible (flip the sign of the off-diagonal block). So $\operatorname{rank} M = \operatorname{rank}\begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} = \operatorname{rank} A + \operatorname{rank} S = k + \operatorname{rank} S$. M is invertible iff $\operatorname{rank} M = k + \ell$ iff $\operatorname{rank} S = \ell$. (c) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = C = I_2$, $S = D - A^{-1} = \begin{pmatrix} t+2 & 0 \\ -1/2 & t+1/2 \end{pmatrix}$, which has rank 2 unless $t = -2$ or $t = -1/2$.

(a) Let

$$P = \begin{pmatrix} I_k & 0 \\ -CA^{-1} & I_\ell \end{pmatrix}, \quad Q = \begin{pmatrix} I_k & -A^{-1}B \\ 0 & I_\ell \end{pmatrix}.$$

By block multiplication (Theorem 2.4.34),

$$PM = \begin{pmatrix} A & B \\ -CA^{-1}A + C & -CA^{-1}B + D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & S \end{pmatrix}, \quad (PM)Q = \begin{pmatrix} A & -AA^{-1}B + B \\ 0 & S \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}.$$

P is invertible: $\begin{pmatrix} I_k & 0 \\ CA^{-1} & I_\ell \end{pmatrix}$ is a two-sided inverse, since both products have off-diagonal block $CA^{-1} - CA^{-1} = 0$ and identity diagonal blocks. Likewise $\begin{pmatrix} I_k & A^{-1}B \\ 0 & I_\ell \end{pmatrix}$ is a two-sided inverse of Q (Definition 2.5.12).

(b) Left multiplication by an invertible matrix does not change the rank (Proposition 3.5.2(c)), and neither does right multiplication (Exercise 3.5.41(c)). So by (a) and Exercise 3.5.36(a),

$$\operatorname{rank} M = \operatorname{rank}\begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} = \operatorname{rank} A + \operatorname{rank} S = k + \operatorname{rank} S,$$

where $\operatorname{rank} A = k$ because A is invertible (Theorem 3.5.14). M is square of size $k + \ell$, so it is invertible if and only if $\operatorname{rank} M = k + \ell$ (Corollary 3.5.25), if and only if $\operatorname{rank} S = \ell$, if and only if S is invertible.

(c) Here $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = C = I_2$, and $D = \begin{pmatrix} t & 1 \\ 1 & t \end{pmatrix}$. Row reducing $(A \mid I_2)$ gives $A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$ (check: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} = \begin{pmatrix} -2+3 & 1-1 \\ -6+6 & 3-2 \end{pmatrix} = I_2$). So

$$S = D - A^{-1} = \begin{pmatrix} t+2 & 0 \\ -1/2 & t+1/2 \end{pmatrix}.$$

If $t \neq -2$ and $t \neq -1/2$, the first column of S is nonzero, and the second column $(0, t+1/2)$ is not a multiple of it (a multiple $c(t+2, -1/2)$ with first entry 0 needs $c = 0$), so $\operatorname{rank} S = 2$ (Proposition 3.5.2(d)). If $t = -2$, then $S = \begin{pmatrix} 0 & 0 \\ -1/2 & -3/2 \end{pmatrix}$, and if $t = -1/2$, then $S = \begin{pmatrix} 3/2 & 0 \\ -1/2 & 0 \end{pmatrix}$; in both cases the two columns are proportional and $\operatorname{rank} S = 1$. By (b), $\operatorname{rank} M_t = 4$ for $t \notin \{-2, -1/2\}$, and $\operatorname{rank} M_t = 3$ for those two values.

Solution to Problem 3.15

Uses. Corollary 3.4.20 (Square systems)
Techniques. Uniqueness gives existence (Technique 3.4.21), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. The system is square. By Corollary 3.4.20 it suffices to show that with all $s_i = 0$ the only solution is $\mathbf{v} = \mathbf{0}$.
Hint 2. For a homogeneous solution, multiply equation i by v_i and add over $i = 1, \dots, n$. You may add the term for $i = 0$ for free, since $v_0 = 0$. Pair the term (i, j) with the term (j, i).
Hint 3. You should get $\sum_{i,j} c_{ij}(v_i - v_j)^2 = 0$. A sum of nonnegative terms vanishes only if every term does. Then use connectedness.
Sketch. (a) n equations, n unknowns. For $\mathbf{s} = \mathbf{0}$: $0 = \sum_{i=0}^{n} v_i \sum_j c_{ij}(v_i - v_j)$, and symmetrizing in i, j gives $\frac{1}{2}\sum_{i,j} c_{ij}(v_i - v_j)^2 = 0$. So $v_i = v_j$ whenever $c_{ij} > 0$, hence all potentials equal $v_0 = 0$ by connectedness, and Corollary 3.4.20 gives existence and uniqueness. (b) The double sum $\sum_{i,j} c_{ij}(v_i - v_j)$ over all nodes is 0 by antisymmetry. (c) $4v_1 - v_2 - v_3 = 0$, $-v_1 + 2v_2 - v_3 = 1$, $-v_1 - v_2 + 3v_3 = 0$ give $\mathbf{v} = (4, 11, 5)/13$; the diagonal carries $1/13$ A from 3 to 1; the effective resistance is $11/13\ \Omega$.

Throughout, sums over i or j run over all nodes $0, \dots, n$ unless stated otherwise, and $v_0 = 0$.

(a) The equations are linear in $v_1, \dots, v_n$ (the terms with $j = 0$ contribute $c_{i0}v_i$), and there are n of them, so the system is square. By Corollary 3.4.20 ((iv) $\Rightarrow$ (vi)) it suffices to show that the homogeneous system, with every $s_i = 0$, has only the trivial solution.

Let $\mathbf{v}$ solve the homogeneous system. Multiply equation i by v_i and add over $i = 1, \dots, n$. We may also add the term with $i = 0$, which is $v_0 \sum_j c_{0j}(v_0 - v_j) = 0$ because $v_0 = 0$, whatever the inner sum is. So

$$E := \sum_i \sum_j c_{ij} v_i (v_i - v_j) = 0.$$

Renaming i and j and using $c_{ji} = c_{ij}$, the same sum is $E = \sum_i \sum_j c_{ij} v_j (v_j - v_i)$. Adding the two expressions,

$$0 = 2E = \sum_i \sum_j c_{ij}\bigl(v_i(v_i - v_j) - v_j(v_i - v_j)\bigr) = \sum_i \sum_j c_{ij}(v_i - v_j)^2.$$

Every term is a nonnegative real number, so every term is 0: whenever i and j are joined, $c_{ij} > 0$ forces $v_i = v_j$. Since the network is connected, every node i is joined to node 0 by a path of resistors, along which the potential does not change; so $v_i = v_0 = 0$. Thus the homogeneous system has only the trivial solution, and for every $\mathbf{s}$ there is exactly one solution.

(b) The current flowing out of node 0 through the resistors is $\sum_j c_{0j}(v_0 - v_j)$. Now

$$\sum_i \sum_j c_{ij}(v_i - v_j) = 0,$$

because the terms (i,j) and (j,i) cancel: $c_{ij}(v_i - v_j) + c_{ji}(v_j - v_i) = 0$. Splitting off $i = 0$ and using the equations at the other nodes,

$$\sum_j c_{0j}(v_0 - v_j) = -\sum_{i=1}^{n} s_i.$$

So a current $s_1 + \cdots + s_n$ flows into the ground node, as conservation of charge requires; the equation at node 0 is a consequence of the others, just as one balance equation was redundant in Exercise 3.1.38.

(c) The conductances are $c_{01} = 2$ and $c_{12} = c_{23} = c_{30} = c_{13} = 1$, and $\mathbf{s} = (0,1,0)$. The equations at nodes 1, 2, 3 are

$$\begin{aligned} 2v_1 + (v_1 - v_2) + (v_1 - v_3) = 0, &\quad \text{that is,} \quad 4v_1 - v_2 - v_3 = 0,\\ (v_2 - v_1) + (v_2 - v_3) = 1, &\quad \text{that is,} \quad -v_1 + 2v_2 - v_3 = 1,\\ (v_3 - v_2) + v_3 + (v_3 - v_1) = 0, &\quad \text{that is,} \quad -v_1 - v_2 + 3v_3 = 0. \end{aligned}$$

The first gives $v_2 = 4v_1 - v_3$; then the third gives $-5v_1 + 4v_3 = 0$, so $v_3 = 5v_1/4$ and $v_2 = 11v_1/4$. The second becomes $(-4 + 22 - 5)v_1/4 = 1$, so $v_1 = 4/13$. Hence

$$v_1 = \frac{4}{13}, \quad v_2 = \frac{11}{13}, \quad v_3 = \frac{5}{13}.$$

Check: $-4 + 22 - 5 = 13$ in the second equation, and $-4 - 11 + 15 = 0$ in the third. The current through the diagonal is $c_{31}(v_3 - v_1) = 1/13$ A, from node 3 to node 1. The effective resistance between nodes 2 and 0 is $v_2/1 = 11/13\ \Omega$.

The argument in (a) follows the same strategy as the uniqueness of the plate temperatures in Exercise 3.1.47, with an energy in place of a maximum. In physical terms, $\frac{1}{2}\sum_{i,j} c_{ij}(v_i - v_j)^2$ is the power dissipated in the resistors, and a network with no sources dissipates no power only if no current flows.

Solution to Problem 3.16

Uses. Corollary 3.4.20 (Square systems), Exercise 3.5.37, Proposition 2.4.23 (Transpose of a product)
Techniques. Uniqueness gives existence (Technique 3.4.21), Particular plus homogeneous (Technique 3.1.18)
Hint 1. P_n is square, so it is invertible if and only if $P_n\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = \mathbf{0}$. Treat x_1 as a parameter: the first $n-1$ equations then determine $x_2, \ldots, x_n$ one at a time.
Hint 2. The forced values are x_1 times the sequence $u_1 = 1$, $u_2 = -1$, $u_{i+1} = -u_i - u_{i-1}$, which is $1, -1, 0, 1, -1, 0, \ldots$. The last equation then says $-x_1 u_{n+1} = 0$, where u_{n+1} continues the sequence.
Hint 3. For (b), P_n is symmetric, so $\mathbf{z}^{\mathsf{T}} P_n = \mathbf{0}$ means $P_n\mathbf{z} = \mathbf{0}$. Over $\mathbb{F}_2$ the only nonzero such $\mathbf{z}$ is $(1,1,0,1,1,0,\ldots,1,1)$.
Sketch. (a) For $n \geq 2$, the first $n-1$ equations of $P_n\mathbf{x} = \mathbf{0}$ hold exactly for $\mathbf{x} = x_1\mathbf{u}$ with $\mathbf{u} = (u_1, \ldots, u_n)$, $u_i = 1, -1, 0$ according as $i \equiv 1, 2, 0 \bmod 3$. The last equation is $x_1(u_{n-1} + u_n) = -x_1 u_{n+1} = 0$. If $n \not\equiv 2$, $u_{n+1} = \pm 1$ forces $x_1 = 0$, so P_n is invertible (Corollary 3.4.20); if $n \equiv 2$, $\mathbf{u}$ is a nonzero kernel vector. (b) $P_n\mathbf{x} = \mathbf{y}$ is unsolvable iff some $\mathbf{z}$ has $P_n\mathbf{z} = \mathbf{0}$

and $\mathbf{z}^{\mathsf{T}}\mathbf{y} \neq 0$; the only candidate is $\mathbf{z} = \mathbf{u} = (1, 1, 0, \ldots, 1, 1)$, and $\mathbf{u}^{\mathsf{T}}\mathbf{y}$ counts the lit lamps in positions not divisible by 3, modulo 2.

(a) For $n = 1$, $P_1 = (1)$ is invertible, and $1 \not\equiv 2 \bmod 3$. Let $n \geq 2$. Since P_n is square, it is invertible if and only if $P_n\mathbf{x} = \mathbf{0}$ has only the trivial solution (Corollary 3.4.20). The equations of $P_n\mathbf{x} = \mathbf{0}$ are

$$x_1 + x_2 = 0, \quad x_{i-1} + x_i + x_{i+1} = 0 \quad (2 \leq i \leq n-1), \quad x_{n-1} + x_n = 0.$$

Define $u_1 = 1$, $u_2 = -1$, and $u_{i+1} = -u_i - u_{i-1}$ for $i \geq 2$. Then $u_3 = 1 - 1 = 0$, $u_4 = 0 + 1 = 1$, $u_5 = -1 - 0 = -1$, and by induction the sequence repeats with period 3:

$$u_i = 1, -1, 0 \quad \text{according as} \quad i \equiv 1, 2, 0 \bmod 3.$$

(Indeed, if u_{i-1}, u_i, u_{i+1} is $1, -1, 0$ or $-1, 0, 1$ or $0, 1, -1$, then $u_{i+2} = -u_{i+1} - u_i$ continues the pattern.)

The first $n-1$ equations. We claim that they hold exactly for the vectors $\mathbf{x} = t\mathbf{u}$, $t \in F$, where $\mathbf{u} = (u_1, \ldots, u_n)$. If they hold, put $t = x_1$. The first equation gives $x_2 = -t = tu_2$, and equation i gives $x_{i+1} = -x_i - x_{i-1}$, so by induction $x_i = tu_i$ for all $i \leq n$. Conversely, $t\mathbf{u}$ satisfies them by the definition of the sequence.

The last equation. For $\mathbf{x} = t\mathbf{u}$ it reads $t(u_{n-1} + u_n) = 0$, that is, $-tu_{n+1} = 0$. If $n \not\equiv 2 \bmod 3$, then $n + 1 \not\equiv 0$, so $u_{n+1} = \pm 1$ and $t = 0$: only the trivial solution, and P_n is invertible. If $n \equiv 2 \bmod 3$, then $u_{n+1} = 0$, so $\mathbf{u} \neq \mathbf{0}$ solves $P_n\mathbf{x} = \mathbf{0}$, and P_n is not invertible. Nothing in the argument depends on the field.

(b) As in Exercise 3.1.40, a set of presses $\mathbf{x} \in \mathbb{F}_2^n$ switches off the pattern $\mathbf{y}$ exactly when $P_n\mathbf{x} = \mathbf{y}$: lamp i is toggled by presses at $i-1, i, i+1$. Let $n \equiv 2 \bmod 3$. By (a), $\ker T_{P_n} = \{t\mathbf{u} \mid t \in \mathbb{F}_2\} = \{\mathbf{0}, \mathbf{u}\}$, where, since $-1 = 1$ in $\mathbb{F}_2$,

$$\mathbf{u} = (1, 1, 0, 1, 1, 0, \ldots, 1, 1),$$

with $u_i = 0$ exactly when 3 divides i.

By Exercise 3.5.37, $P_n\mathbf{x} = \mathbf{y}$ has no solution if and only if there is $\mathbf{z}$ with $\mathbf{z}^{\mathsf{T}}P_n = \mathbf{0}$ and $\mathbf{z}^{\mathsf{T}}\mathbf{y} \neq 0$. The matrix P_n is symmetric, so transposing (Proposition 2.4.23), $\mathbf{z}^{\mathsf{T}}P_n = \mathbf{0}$ says $P_n\mathbf{z} = \mathbf{0}$, that is, $\mathbf{z} \in \{\mathbf{0}, \mathbf{u}\}$. And $\mathbf{0}^{\mathsf{T}}\mathbf{y} = 0$. So the pattern cannot be switched off if and only if $\mathbf{u}^{\mathsf{T}}\mathbf{y} \neq 0$. Finally, $\mathbf{u}^{\mathsf{T}}\mathbf{y} = \sum_{3\nmid i} y_i$ in $\mathbb{F}_2$, which is 0 exactly when the number of lit lamps in positions not divisible by 3 is even.

Solution to Problem 3.17

Uses. Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 2.4.34 (Block multiplication), Theorem 3.5.14 (Third Fundamental Theorem, completed), Corollary 1.5.17 (Sublists), Corollary 3.5.16 (Lists of vectors in F^m), Theorem 3.3.12 (Matrices in RREF), Proposition 3.5.8 (Right inverses from the pivots), Proposition 3.5.9 (Left inverses), Corollary 3.3.18 (RREF and row equivalence), Corollary 3.5.21 (Invertibility and row equivalence), Exercise 1.5.43, Theorem 3.5.13 (Second Fundamental Theorem, completed)

Techniques. Reduce to RREF (Technique 3.3.22), Uniqueness via two candidates (Technique 1.1.8)

Hint 1. For existence in (a), write $GA = \begin{pmatrix} S \\ 0 \end{pmatrix}$ and split G^{-1} into its first r columns and the rest. For uniqueness, S has a right inverse.

Hint 2. For the converse, a matrix C with r columns and $\operatorname{rank} C = r$ can be reduced to $\begin{pmatrix} I_r \\ 0 \end{pmatrix}$ by row operations (as in the first proof of Proposition 3.5.9). Apply the same operations to CS.

Hint 3. For (b), (a) sets up a bijection $(C, S) \mapsto CS$. For (c), count $r \times n$ matrices of rank r twice: with (b) for $m = r$, and as lists of r independent rows.

Sketch. (a) $GA = \begin{pmatrix} S \\ 0 \end{pmatrix}$, so $A = G^{-1}\begin{pmatrix} S \\ 0 \end{pmatrix} = CS$ with C the first r columns of G^{-1}; they are independent, so $\operatorname{rank} C = r$. If $CS = C'S$, multiply by a right inverse of S (rank r = number of rows). Conversely, $KC = \begin{pmatrix} I_r \\ 0 \end{pmatrix}$ for an invertible K, so $K(CS) = \begin{pmatrix} S \\ 0 \end{pmatrix}$, which is in RREF; hence $\operatorname{rref}(CS) = \begin{pmatrix} S \\ 0 \end{pmatrix}$. (b) $(C, S) \mapsto CS$ is a bijection onto the rank r matrices, and the C's are the independent lists of r vectors in F^m, counted by Exercise 1.5.43. (c) With $m = r$: $N(r, n) \cdot |\text{rank-}\, r \text{ matrices in } F^{r\times r}| = |\text{independent lists of } r \text{ rows in } F^n|$. For $r = 2$, $n = 3$ the quotient is $q^2 + q + 1$.

(a) *Existence.* Let G be invertible with $GA = \operatorname{rref}(A)$ (Theorem 3.3.10). Since $\operatorname{rank} A = r$, the nonzero rows of $\operatorname{rref}(A)$ are its first r rows, so $GA = \begin{pmatrix} S \\ 0 \end{pmatrix}$. Split $G^{-1} = \begin{pmatrix} C & C' \end{pmatrix}$ with $C \in F^{m\times r}$ its first r columns. By block multiplication (Theorem 2.4.34),

$$A = G^{-1}\begin{pmatrix} S \\ 0 \end{pmatrix} = \begin{pmatrix} C & C' \end{pmatrix}\begin{pmatrix} S \\ 0 \end{pmatrix} = CS.$$

The columns of G^{-1} form a basis of F^m (Theorem 3.5.14, as G^{-1} is invertible), so the columns of C are independent (Corollary 1.5.17), and $\operatorname{rank} C = r$ by Corollary 3.5.16(a).

Uniqueness of C. S satisfies (R1)–(R4), so it is in RREF and its pivots are the leading entries of its r nonzero rows (Theorem 3.3.12); so $\operatorname{rank} S = r$, the number of its rows. By Proposition 3.5.8, $SZ = I_r$ for some Z. If $CS = C'S$, then $C = CSZ = C'SZ = C'$.

Converse. Let $C \in F^{m\times r}$ have rank r and S be in RREF with no zero row. Since $\operatorname{rank} C$ is the number of columns of C, the first proof of Proposition 3.5.9 shows that $KC = \begin{pmatrix} I_r \\ 0 \end{pmatrix}$ for an invertible K (one with $KC = \operatorname{rref}(C)$). Then

$$K(CS) = \begin{pmatrix} I_r \\ 0 \end{pmatrix} S = \begin{pmatrix} S \\ 0 \end{pmatrix}.$$

This matrix satisfies (R1)–(R4): S does, and the zero rows come last. By Corollary 3.5.21(b), $CS \sim \begin{pmatrix} S \\ 0 \end{pmatrix}$, so $\operatorname{rref}(CS) = \begin{pmatrix} S \\ 0 \end{pmatrix}$ by Corollary 3.3.18(c). It has r pivots, so $\operatorname{rank}(CS) = r$.

(b) Let $\mathcal{C}$ be the set of $C \in F^{m\times r}$ of rank r, $\mathcal{S}$ the set of $r \times n$ matrices in RREF with no zero row, and $\mathcal{A}$ the set of $A \in F^{m\times n}$ of rank r. By the converse in (a), $(C, S) \mapsto CS$ maps $\mathcal{C} \times \mathcal{S}$ into $\mathcal{A}$. It is onto, by the existence part. It is injective: if $CS = C'S' = A$, then by the converse $\begin{pmatrix} S \\ 0 \end{pmatrix} = \operatorname{rref}(A) = \begin{pmatrix} S' \\ 0 \end{pmatrix}$, so $S = S'$, and then $C = C'$ by uniqueness. Hence $|\mathcal{A}| = |\mathcal{C}| \cdot |\mathcal{S}| = |\mathcal{C}| N(r, n)$. A matrix in $F^{m\times r}$ has rank r if and only if its r columns are independent (Corollary 3.5.16(a)), so by Exercise 1.5.43(b), $|\mathcal{C}| = (q^m - 1)(q^m - q)\cdots(q^m - q^{r-1})$.

(c) Take $m = r$ in (b): the number of $r \times n$ matrices of rank r is $N(r, n)(q^r - 1)\cdots(q^r - q^{r-1})$. We count them a second way. An $r \times n$ matrix A has rank r if and only if its rows are linearly independent:

- if $\operatorname{rank} A = r$, then A has a right inverse X (Theorem 3.5.13), and a relation $\mathbf{z}A = \mathbf{0}$ among the rows gives $\mathbf{z} = \mathbf{z}AX = \mathbf{0}$;
- if $\operatorname{rank} A < r$, the last row of $\operatorname{rref}(A) = GA$ is zero, so the last row $\mathbf{g}$ of G is nonzero (as G is invertible) and satisfies $\mathbf{g}A = \mathbf{0}$, a nontrivial relation among the rows.

So by Exercise 1.5.43(b), applied to rows in F^n, there are $(q^n - 1)(q^n - q)\cdots(q^n - q^{r-1})$ such matrices. Equating the two counts and dividing by the nonzero integer $(q^r - 1)\cdots(q^r - q^{r-1})$ gives the formula.

For $r = 2$, $n = 3$:

$$N(2, 3) = \frac{(q^3 - 1)(q^3 - q)}{(q^2 - 1)(q^2 - q)} = \frac{(q^3 - 1)\cdot q(q^2 - 1)}{(q^2 - 1)\cdot q(q - 1)} = \frac{q^3 - 1}{q - 1} = q^2 + q + 1,$$

which is the number of 2×3 matrices in RREF with two pivots found in Exercise 3.3.48 over $\mathbb{F}_p$ (p^2, p, and 1 for the pivot columns $(1, 2)$, $(1, 3)$, $(2, 3)$), with $q = p$. The same count works over any finite field.

With $m = n = r$, part (b) counts the invertible $n \times n$ matrices, since $N(n, n) = 1$ (only I_n): there are $(q^n - 1)(q^n - q)\cdots(q^n - q^{n-1})$ of them. The numbers $N(r, n)$ are the *Gaussian binomial coefficients*; by Exercise 3.3.50 they also count the subspaces of F^n spanned by r independent vectors.

Solution to Problem 3.18

Uses. Corollary 3.4.20 (Square systems), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$), Exercise 3.3.54, Exercise 3.3.52
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)

Hint 1. Suppose A is not invertible over $\mathbb{Q}$. Then $A\mathbf{x} = \mathbf{0}$ for some nonzero rational $\mathbf{x}$. Clear denominators, and divide by the largest power of 2 dividing all entries.
Hint 2. Now $\mathbf{x}$ has integer entries, not all even. Reduce the equations $A\mathbf{x} = \mathbf{0}$ modulo 2.
Hint 3. For (c), $\overline{A}$ has zeros on the diagonal and ones elsewhere: see Exercise 3.3.52. For (d), try $2I_n$.
Sketch. (a) If $A\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \in \mathbb{Q}^n$ nonzero, scale $\mathbf{x}$ to an integer vector with an odd entry. Reducing modulo 2 gives $\overline{A}\,\overline{\mathbf{x}} = \mathbf{0}$ with $\overline{\mathbf{x}} \neq \mathbf{0}$, so $\overline{A}$ is not invertible. Invertibility over $\mathbb{Q}$ gives invertibility over $\mathbb{R}$ since the reduced form does not depend on the field. (b) $\overline{A} = I_n$. (c) $\overline{A} = J - I_n$, invertible over $\mathbb{F}_2$ iff $n-1$ is odd; $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ has third row the sum of the first two. (d) $2I_n$.

(a) We prove the contrapositive. Suppose A is not invertible over $\mathbb{Q}$. By Corollary 3.4.20 there is a nonzero $\mathbf{x} \in \mathbb{Q}^n$ with $A\mathbf{x} = \mathbf{0}$. Multiplying by a common denominator of its entries, we may assume $\mathbf{x} \in \mathbb{Z}^n$, still nonzero. Let 2^e be the largest power of 2 dividing every entry of $\mathbf{x}$ (it exists, since some entry is a nonzero integer), and replace $\mathbf{x}$ by $2^{-e}\mathbf{x}$. Now $\mathbf{x}$ has integer entries, $A\mathbf{x} = \mathbf{0}$, and some entry of $\mathbf{x}$ is odd.

Each equation $\sum_j a_{ij}x_j = 0$ is an identity between integers, so it remains true after reducing modulo 2, because reduction modulo 2 respects sums and products (Lemma 1.1.21). That is, $\overline{A}\,\overline{\mathbf{x}} = \mathbf{0}$ in $\mathbb{F}_2^n$, where $\overline{\mathbf{x}}$ is the vector of remainders of $\mathbf{x}$. Since some entry of $\mathbf{x}$ is odd, $\overline{\mathbf{x}} \neq \mathbf{0}$. So $\overline{A}$ is not invertible over $\mathbb{F}_2$ (Corollary 3.4.20).

So if $\overline{A}$ is invertible, then A is invertible over $\mathbb{Q}$, that is, $\operatorname{rref}(A) = I_n$ computed over $\mathbb{Q}$. By Exercise 3.3.54, the reduced form is the same over $\mathbb{R}$, so A is invertible over $\mathbb{R}$ too.

(b) For such A, $\overline{A} = I_n$, which is invertible.

(c) Now $\overline{A}$ has 0 on the diagonal and 1 elsewhere. By Exercise 3.3.52, this matrix is invertible over $\mathbb{F}_2$ if and only if $(n-1) * 1 \neq 0$ in $\mathbb{F}_2$, that is, if and only if n is even. Then A is invertible by (a). For $n = 3$, the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

has even diagonal entries and odd off-diagonal entries, but its third column is the sum of the first two, so $A(1, 1, -1) = \mathbf{0}$ and A is not invertible.

(d) $A = 2I_n$ is invertible over $\mathbb{Q}$, with inverse $\frac{1}{2}I_n$, but $\overline{A} = 0$.

The point of (a) is that a single computation over a small field certifies invertibility over an infinite one. Choosing $\mathbf{x}$ with an odd entry is an extremal choice, as in Technique 1.5.9: without it, reducing modulo 2 could turn $\mathbf{x}$ into $\mathbf{0}$ and prove nothing. Section 5.3 revisits this with determinants (Technique 5.3.21).

Solution to Problem 3.19

Uses. Corollary 3.4.20 (Square systems), Lemma 3.4.23 (Too many roots), Definition 1.2.17 (Evaluation), Proposition 1.1.17 (No zero divisors)
Techniques. Uniqueness gives existence (Technique 3.4.21)
Hint 1. It is enough to show that $C\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = \mathbf{0}$. Row i of $C\mathbf{x} = \mathbf{0}$ says that the function $t \mapsto \sum_j x_j/(t + b_j)$ vanishes at $t = a_i$.
Hint 2. Clear denominators: multiply by $\prod_k (t + b_k)$. The result $P(t) = \sum_j x_j \prod_{k \neq j}(t + b_k)$ is the value at t of a polynomial of degree at most $n - 1$.
Hint 3. P vanishes at the n distinct points $a_1, \ldots, a_n$. Then evaluate P at $t = -b_j$.
Sketch. If $C\mathbf{x} = \mathbf{0}$, let $P(t) = \sum_j x_j \prod_{k \neq j}(t + b_k)$, the value of a polynomial of degree at most $n - 1$. Then $P(a_i) = \prod_k (a_i + b_k) \cdot (C\mathbf{x})_i = 0$ for all i, so this polynomial is zero (Lemma 3.4.23) and P vanishes everywhere. But $P(-b_j) = x_j \prod_{k \neq j}(b_k - b_j)$, so $x_j = 0$. Square system, trivial kernel: invertible. (b) $a_i = i$, $b_j = j - 1$.

(a) C is square, so by Corollary 3.4.20 it suffices to show that $C\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. Let $C\mathbf{x} = \mathbf{0}$.

A polynomial. For a list of scalars $\beta_1, \ldots, \beta_s$, the function $t \mapsto \prod_{k=1}^{s}(t + \beta_k)$ is the polynomial function of a polynomial of degree at most s: this is clear for $s = 0$ (the constant 1), and if $t \mapsto f(t) = \sum_{i \leq d} e_i t^i$, then

$$(t+\beta)f(t) = \sum_{i\le d} e_i t^{i+1} + \sum_{i\le d} \beta e_i t^i,$$

which is again a polynomial expression in t, of degree at most $d+1$. So for each j there is $q_j \in F[x]_{\le n-1}$ with $q_j(t) = \prod_{k\neq j}(t+b_k)$ for all $t \in F$, and $p = \sum_j x_j q_j \in F[x]_{\le n-1}$ satisfies

$$p(t) = \sum_{j=1}^{n} x_j \prod_{k\neq j}(t+b_k) \quad \text{for all } t \in F,$$

since evaluation is linear (Definition 1.2.17).

It vanishes at the a_i. Fix i. Every factor $a_i + b_k$ is nonzero, and $\prod_{k\neq j}(a_i+b_k) = \prod_k (a_i+b_k)/(a_i+b_j)$. So

$$p(a_i) = \prod_{k=1}^{n}(a_i+b_k) \cdot \sum_{j=1}^{n} \frac{x_j}{a_i+b_j} = \prod_{k=1}^{n}(a_i+b_k) \cdot (C\mathbf{x})_i = 0.$$

Thus p has degree at most $n-1$ and vanishes at the n distinct points $a_1, \ldots, a_n$, so $p = \mathbf{0}$ by Lemma 3.4.23 (with $n-1$ in place of n). Hence $p(t) = 0$ for every $t \in F$.

Evaluating at $-b_j$. For $l \neq j$, the product $\prod_{k\neq l}(-b_j + b_k)$ contains the factor $k = j$, which is 0. So only the term $l = j$ survives:

$$0 = p(-b_j) = x_j \prod_{k\neq j}(b_k - b_j).$$

The b_k are distinct, so the product is nonzero (Proposition 1.1.17), and $x_j = 0$. As j was arbitrary, $\mathbf{x} = \mathbf{0}$. So C is invertible.

(b) Take $F = \mathbb{Q}$, $a_i = i$, and $b_j = j - 1$ for $i, j = 1, \ldots, n$. The a_i are distinct, the b_j are distinct, and $a_i + b_j = i + j - 1 \ge 1$, so it is nonzero. The Cauchy matrix is then H_n, which is therefore invertible.

The proof never computes C^{-1}; it shows that the only combination of the functions $t \mapsto 1/(t+b_j)$ vanishing at n points is the trivial one, and lets Corollary 3.4.20 produce the inverse. This is Technique 3.4.21 in its pure form: interpolation by the functions $1/(t+b_j)$ in place of the monomials t^j. (A formula for C^{-1} exists, but is much harder to find this way.)

Section 4.1

Solution to Exercise 4.1.9

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), Definition 4.1.5 (Dimension)
Hint 1. Split $(a, b, a+b, a-b)$ into a part with a and a part with b.
Hint 2. Check that the two vectors you get are independent by looking at their first two coordinates.
Sketch. $(a, b, a+b, a-b) = a(1,0,1,1) + b(0,1,1,-1)$, so these two vectors span U. A relation $a(1,0,1,1) + b(0,1,1,-1) = \mathbf{0}$ has first two coordinates a and b, so it is trivial. Hence $\dim U = 2$.

For all $a, b \in F$,

$$(a, b, a+b, a-b) = a(1,0,1,1) + b(0,1,1,-1),$$

so $(1,0,1,1), (0,1,1,-1)$ lies in U and spans it. If $a(1,0,1,1) + b(0,1,1,-1) = \mathbf{0}$, then the first two coordinates give $a = 0$ and $b = 0$, so the list is independent (Theorem 1.5.5). It is a basis of U by Proposition 1.6.2, and $\dim U = 2$. This holds over every field, $\mathbb{F}_2$ included, where $a - b = a + b$.

Solution to Exercise 4.1.15

Uses. Theorem 4.1.10 (Independent lists are short), Example 4.1.7 (Dimensions of the standard spaces), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. The list has $4 = \dim F[x]_{\le 3}$ entries, so it suffices to check independence.

Hint 2. The degrees are $1, 2, 3, 3$, so Proposition 1.5.24 does not apply directly. Write out a relation and compare coefficients, starting from the constant and x coefficients.
Sketch. Length $4 = \dim F[x]_{\leq 3}$, so independence suffices. In $a(1+x) + b\big(1+2x+x^2\big) + c\big(x^2+x^3\big) + dx^3 = 0$, the constant term gives $a + b = 0$, the x term $a + 2b = 0$, so $b = 0 = a$; then x^2 gives $c = 0$ and x^3 gives $d = 0$.

Since $\dim F[x]_{\leq 3} = 4$ (Example 4.1.7), by Theorem 4.1.10(b) it suffices to show that the four polynomials are independent. Suppose

$$a(1+x) + b\big(1+2x+x^2\big) + c\big(x^2+x^3\big) + dx^3 = 0.$$

Comparing coefficients of 1, x, x^2, and x^3 gives

$$a + b = 0, \quad a + 2b = 0, \quad b + c = 0, \quad c + d = 0.$$

Subtracting the first equation from the second gives $b = 0$, and then $a = 0$, $c = 0$, and $d = 0$. So the only relation is trivial (Theorem 1.5.5), and the list is a basis. Every step was valid over every field.

Solution to Exercise 4.1.20

Uses. Theorem 4.1.16 (Extending to a basis), Proposition 1.5.14 (Appending a vector), Proposition 1.5.15 (Short lists, zeros, and repeats)
Hint 1. Sift the list $(1, 2, 1), \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$: keep each entry that is not in the span of the entries kept before it.
Hint 2. $\mathbf{e}_1$ is never a multiple of $(1, 2, 1)$. For the third vector, describe $\operatorname{span}((1, 2, 1), \mathbf{e}_1)$ and test $\mathbf{e}_2$.
Sketch. In both cases $\mathbf{e}_1$ is kept. $\operatorname{span}((1, 2, 1), \mathbf{e}_1) = \{(a + b, 2a, a)\}$, whose vectors have second coordinate 2 times the third; $\mathbf{e}_2$ fails this, so it is kept, and $(1, 2, 1), \mathbf{e}_1, \mathbf{e}_2$ is a basis in both cases, by length.

We follow the proof of Theorem 4.1.16 with $\mathcal{S} = \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. The vector $\mathbf{e}_1$ is not a multiple of $(1, 2, 1)$, since a multiple $t(1, 2, 1)$ with third coordinate 0 has $t = 0$. So $(1, 2, 1), \mathbf{e}_1$ is independent (Proposition 1.5.15(d)). Its span consists of the vectors $a(1, 2, 1) + b\mathbf{e}_1 = (a + b, 2a, a)$, whose second coordinate is 2 times the third. The vector $\mathbf{e}_2 = (0, 1, 0)$ fails this test, since $1 \neq 2 \cdot 0$ in every field. So $(1, 2, 1), \mathbf{e}_1, \mathbf{e}_2$ is independent (Proposition 1.5.14), and as it has length $3 = \dim F^3$, it is a basis. The same vectors are appended in (a) and (b): the computation used only that $1 \neq 0$.

Solution to Exercise 4.1.24

Uses. Proposition 1.5.24 (Distinct degrees), Theorem 4.1.10 (Independent lists are short), Proposition 4.1.21 (Dimensions of subspaces)
Hint 1. Find three independent polynomials in U; powers of $x - 2$ are natural.
Hint 2. Find one polynomial of degree at most 3 that is not in U, and use Proposition 4.1.21(b).
Sketch. $x - 2$, $(x - 2)^2$, $(x - 2)^3$ lie in U and have distinct degrees, so $\dim U \geq 3$. The constant 1 is not in U, so $\dim U < 4$. Hence $\dim U = 3$, and the three polynomials form a basis.

The polynomials $x - 2$, $(x - 2)^2$, $(x - 2)^3$ vanish at 2, so they lie in U, and they are independent by Proposition 1.5.24. So $\dim U \geq 3$ by Theorem 4.1.10(a), applied in U. The constant polynomial 1 does not vanish at 2, so $U \neq \mathbb{R}[x]_{\leq 3}$, and $\dim U < 4$ by Proposition 4.1.21(b). Hence $\dim U = 3$, and the three polynomials, an independent list of length $\dim U$ in U, form a basis of U by Theorem 4.1.10(b).

Solution to Exercise 4.1.29

Uses. Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps)
Hint 1. If $n \leq 4$, what does Corollary 4.1.27 say about A^n, and hence about A^4?
Hint 2. For $n \geq 5$, generalize the shift: send $\mathbf{e}_5 \mapsto \mathbf{e}_4 \mapsto \cdots \mapsto \mathbf{e}_1 \mapsto \mathbf{0}$ and kill the remaining standard basis vectors.
Sketch. If $n \leq 4$, then $A^5 = 0$ gives $A^n = 0$, so $A^4 = A^{4-n}A^n = 0$. For $n \geq 5$, take $A = U_{12} + U_{23} + U_{34} + U_{45}$, with $A\mathbf{e}_{j+1} = \mathbf{e}_j$ for $j \leq 4$ and $A\mathbf{e}_j = \mathbf{0}$ otherwise; then $A^4\mathbf{e}_5 = \mathbf{e}_1$ and $A^5 = 0$.

Suppose $n \leq 4$ and $A^5 = 0$. By Corollary 4.1.27, $A^n = 0$, so $A^4 = A^{4-n}A^n = 0$. So no such A exists for $n \leq 4$.

For $n \geq 5$, let $A \in F^{n\times n}$ have 1 in positions $(1, 2), (2, 3), (3, 4), (4, 5)$ and 0 elsewhere. Then $A\mathbf{e}_1 = \mathbf{0}$, $A\mathbf{e}_j = \mathbf{e}_{j-1}$ for $2 \leq j \leq 5$, and $A\mathbf{e}_j = \mathbf{0}$ for $j \geq 6$. Hence $A^4\mathbf{e}_5 = \mathbf{e}_1 \neq \mathbf{0}$, so $A^4 \neq 0$. And $A^5\mathbf{e}_j = \mathbf{0}$ for every j: for $j \leq 5$ the vector $\mathbf{e}_j$ reaches $\mathbf{0}$ after $j \leq 5$ steps, and for $j \geq 6$ after one. Since A^5 kills a basis, $A^5 = 0$. So the answer is: exactly for $n \geq 5$.

Solution to Exercise 4.1.33

Uses. Proposition 4.1.31 (Every square matrix satisfies a polynomial)
Techniques. Too many vectors force a dependence (Technique 4.1.30)
Hint 1. Compute A^2 and compare it with A and I_2.
Hint 2. From $A^2 = 2A - I_2$, show by induction that $A^k = kA - (k-1)I_2$.
Sketch. $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = 2A - I_2$, so $A^2 - 2A + I_2 = 0$. Induction gives $A^k = kA - (k-1)I_2 = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, so $A^{10} = \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix}$.

We have $A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = 2A - I_2$, so $A^2 - 2A + I_2 = 0$ is a nontrivial relation. We show $A^k = kA - (k-1)I_2$ for $k \geq 1$ by induction. For $k = 1$ it reads $A = A$. If it holds for k, then

$$A^{k+1} = A(kA - (k-1)I_2) = kA^2 - (k-1)A = k(2A - I_2) - (k-1)A = (k+1)A - kI_2.$$

Hence $A^{10} = 10A - 9I_2 = \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix}$.

Solution to Exercise 4.1.39

Uses. Proposition 4.1.34 (Size of a vector space over a finite field), Corollary 4.1.35 (Counting independent lists and bases)
Hint 1. By Proposition 4.1.34, a finite vector space over $\mathbb{F}_5$ has 5^n elements for some n.
Hint 2. For (c), use $N_3(2,2) = (3^2 - 1)(3^2 - 3)$.
Sketch. (a) No: 100 is not a power of 5. (b) $5^3 = 125$. (c) $(9-1)(9-3) = 48$.

(a) No. A vector space over $\mathbb{F}_5$ with finitely many elements is finite-dimensional, so it has $5^{\dim V}$ elements by Proposition 4.1.34, and $100 = 2^2 \cdot 5^2$ is not a power of 5.

(b) A 3-dimensional subspace is a 3-dimensional vector space over $\mathbb{F}_5$, so it has $5^3 = 125$ elements.

(c) By Corollary 4.1.35, $\mathbb{F}_3^2$ has $N_3(2,2) = (9-1)(9-3) = 48$ bases.

Solution to Exercise 4.1.40

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.5.24 (Distinct degrees), Corollary 1.5.17 (Sublists), Proposition 1.6.7 (Standard bases)
Hint 1. In each case solve the defining conditions for some of the coordinates and write a general element as a combination of fixed vectors with the remaining coordinates as coefficients.
Hint 2. In (a), $x_4 = -x_1 - x_2 - x_3$. In (b), $p(0)$ is the constant term and $p(1)$ is the sum of the coefficients. In (c), use matrix units.
Sketch. (a) $\mathbf{x} = x_1(\mathbf{e}_1 - \mathbf{e}_4) + x_2(\mathbf{e}_2 - \mathbf{e}_4) + x_3(\mathbf{e}_3 - \mathbf{e}_4)$; dimension 3. (b) $p = a_2(x^2 - x) + a_3(x^3 - x)$; dimension 2. (c) The six U_{ij} with $i \leq j$; dimension 6.

(a) For $\mathbf{x} \in U_1$ we have $x_4 = -x_1 - x_2 - x_3$, so

$$\mathbf{x} = x_1(\mathbf{e}_1 - \mathbf{e}_4) + x_2(\mathbf{e}_2 - \mathbf{e}_4) + x_3(\mathbf{e}_3 - \mathbf{e}_4).$$

The three vectors $\mathbf{e}_i - \mathbf{e}_4$ lie in U_1, so they span U_1. A combination $\sum_{i=1}^{3} c_i(\mathbf{e}_i - \mathbf{e}_4)$ has first three coordinates c_1, c_2, c_3, so it is $\mathbf{0}$ only if all $c_i = 0$. So they form a basis (Proposition 1.6.2), and $\dim U_1 = 3$.

(b) Let $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3$. Then $p(0) = a_0$ and $p(1) = a_0 + a_1 + a_2 + a_3$, so $p \in U_2$ if and only if $a_0 = 0$ and $a_1 = -a_2 - a_3$, that is,

$$p = a_2(x^2 - x) + a_3(x^3 - x).$$

So $x^2 - x, x^3 - x$ spans U_2, and it is independent because its entries are nonzero with distinct degrees (Proposition 1.5.24). Hence $\dim U_2 = 2$.

(c) An upper triangular A is $\sum_{i \leq j} a_{ij} U_{ij}$, so the six matrix units $U_{11}, U_{12}, U_{13}, U_{22}, U_{23}, U_{33}$ span U_3. They are a sublist of the standard basis of $F^{3\times 3}$ (Proposition 1.6.7), hence independent (Corollary 1.5.17). So $\dim U_3 = 6$.

Solution to Exercise 4.1.41

Uses. Corollary 4.1.11 (Spanning lists are long), Proposition 4.1.21 (Dimensions of subspaces), Proposition 1.5.15 (Short lists, zeros, and repeats), Definition 1.4.1 (Subspace)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)

Hint 1. Length alone can force a property only in one direction: a list longer than $\dim V$ cannot be independent, and a list shorter than $\dim V$ cannot span. It never forces a list to span or to be independent.
Hint 2. For (d), think of two different lines in a plane. For (e), apply Proposition 4.1.21(b) inside W. For (f), what are the elements of $\mathbb{R}^2$?
Sketch. (a) False: five copies of $\mathbf{0}$. (b) False: three copies of $\mathbf{0}$. (c) True by Corollary 4.1.11(b). (d) False: two different 1-dimensional subspaces. (e) True: U is a subspace of the vector space W of full dimension. (f) False: $\mathbb{R}^2$ is not a subset of $\mathbb{R}^3$, though $\{(x, y, 0)\}$ is a 2-dimensional subspace.

(a) False. The list $\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}$ spans $\{\mathbf{0}\} \neq V$.

(b) False. The list $\mathbf{0}, \mathbf{0}, \mathbf{0}$ is dependent (Proposition 1.5.15(c)).

(c) True. A spanning list of length $4 = \dim V$ is a basis by Corollary 4.1.11(b), so it is independent.

(d) False. Let $\mathbf{v}_1, ..., \mathbf{v}_4$ be a basis of V. Then $\operatorname{span}(\mathbf{v}_1)$ and $\operatorname{span}(\mathbf{v}_2)$ are both 1-dimensional, and they differ, since $\mathbf{v}_2 \notin \operatorname{span}(\mathbf{v}_1)$ by independence.

(e) True. W is a vector space with the operations of V, and U is a subspace of W: it is a subset closed under the same operations (Definition 1.4.1). Since $\dim U = \dim W$, Proposition 4.1.21(b), applied in W, gives $U = W$.

(f) False. The elements of $\mathbb{R}^2$ are pairs and those of $\mathbb{R}^3$ are triples, so $\mathbb{R}^2$ is not a subset of $\mathbb{R}^3$. The subset $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$ is a 2-dimensional subspace of $\mathbb{R}^3$, with basis $\mathbf{e}_1, \mathbf{e}_2$, and it is a copy of $\mathbb{R}^2$ in a sense made precise in Section 4.5.

Solution to Exercise 4.1.42

Uses. Theorem 4.1.16 (Extending to a basis), Proposition 1.5.14 (Appending a vector), Theorem 4.1.10 (Independent lists are short)
Hint 1. The span of $\mathbf{v}_1, \mathbf{v}_2$ consists of the vectors $(b, 0, a + 2b, 2a + b)$. Is $\mathbf{e}_1$ of this form?
Hint 2. For $\mathbf{e}_1$ you need $b = 1$, $a = -2$, and $2a + b = 0$, that is, $-3 = 0$. Which field allows this? Stop sifting as soon as you have 4 vectors.
Sketch. $\mathbf{e}_1 \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ exactly when $3 = 0$; indeed $\mathbf{v}_1 + \mathbf{v}_2 = (1, 0, 3, 3)$, which is $\mathbf{e}_1$ in $\mathbb{F}_3^4$. Over $\mathbb{Q}$ sifting keeps $\mathbf{e}_1$ and $\mathbf{e}_2$ (all of $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)$ has second coordinate 0). Over $\mathbb{F}_3$ it deletes $\mathbf{e}_1$ and keeps $\mathbf{e}_2$ and $\mathbf{e}_3$.

The vectors $\mathbf{v}_1, \mathbf{v}_2$ are independent over every field, since $a\mathbf{v}_1 + b\mathbf{v}_2 = (b, 0, a + 2b, 2a + b)$ is $\mathbf{0}$ only if $b = 0$ and then $a = 0$. Their span consists of the vectors $(b, 0, a + 2b, 2a + b)$.

Is $\mathbf{e}_1$ in the span? We need $b = 1$, $a + 2 = 0$, and $2a + 1 = 0$. The second gives $a = -2$, and then the third reads $-3 = 0$. Over $\mathbb{Q}$ this is false, so $\mathbf{e}_1$ is kept. Over $\mathbb{F}_3$ it is true, and indeed $\mathbf{v}_1 + \mathbf{v}_2 = (1, 0, 3, 3) = \mathbf{e}_1$ there, so $\mathbf{e}_1$ is deleted.

Over $\mathbb{Q}$. The vectors of $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)$ have second coordinate 0, so $\mathbf{e}_2$ is not among them and is kept. Now $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2$ is independent (Proposition 1.5.14) of length $4 = \dim \mathbb{Q}^4$, so it is a basis (Theorem 4.1.10(b)), and $\mathbf{e}_3, \mathbf{e}_4$ are deleted.

Over $\mathbb{F}_3$. Again $\mathbf{e}_2$ is kept, because every vector of $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ has second coordinate 0. The span of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_2$ consists of the vectors $(b, d, a + 2b, 2a + b)$. For $\mathbf{e}_3 = (0, 0, 1, 0)$ we would need $b = 0$, $a = 1$, and $2a = 0$, which is false in $\mathbb{F}_3$. So $\mathbf{e}_3$ is kept, and $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_2, \mathbf{e}_3$ is a basis of $\mathbb{F}_3^4$.

The answers differ because whether a vector lies in a span is a question about solving linear equations, and here the equations are solvable exactly when $3 = 0$ in F.

Solution to Exercise 4.1.43

Uses. Example 1.4.11 (Matrices), Theorem 4.1.10 (Independent lists are short), Proposition 4.1.21 (Dimensions of subspaces), Proposition 1.6.7 (Standard bases)
Hint 1. Generalize Example 4.1.22. Which matrix units lie in U, and which combinations of diagonal ones do?
Hint 2. Use U_{ij} for $i \neq j$ and $U_{ii} - U_{nn}$ for $i < n$ to get a lower bound, and $U_{11} \notin U$ for an upper bound.
Sketch. The $n^2 - n$ matrices U_{ij} $(i \neq j)$ and the $n - 1$ matrices $U_{ii} - U_{nn}$ $(i < n)$ lie in U and are independent, since a combination of them has these coefficients as its off-diagonal and first $n - 1$ diagonal entries. So $\dim U \geq n^2 - 1$. Since $\operatorname{tr} U_{11} = 1$, $U \neq F^{n \times n}$, so $\dim U < n^2$. Hence $\dim U = n^2 - 1$ and the list is a basis.

U is a subspace by Example 1.4.11. The $n^2 - n$ matrices U_{ij} with $i \neq j$ have trace 0, and so do the $n-1$ matrices $U_{ii} - U_{nn}$ with $1 \leq i < n$. Consider a combination

$$A = \sum_{i \neq j} a_{ij} U_{ij} + \sum_{i=1}^{n-1} c_i (U_{ii} - U_{nn}).$$

Its (i,j) entry is a_{ij} for $i \neq j$, its (i,i) entry is c_i for $i < n$, and its (n,n) entry is $-(c_1 + \cdots + c_{n-1})$. If $A = 0$, all a_{ij} and c_i are 0. So these $n^2 - 1$ matrices form an independent list in U, and $\dim U \geq n^2 - 1$ by Theorem 4.1.10(a), applied in U. On the other hand, $\operatorname{tr} U_{11} = 1 \neq 0$, so $U \neq F^{n \times n}$ and $\dim U < n^2$ by Proposition 4.1.21(b). Hence $\dim U = n^2 - 1$, and the independent list above, of length $\dim U$, is a basis of U (Theorem 4.1.10(b)). For $n = 1$ the list is empty and $U = \{0\}$.

Solution to Exercise 4.1.44

Uses. Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Proposition 2.4.14 (Algebra of matrices), Definition 2.5.12 (Invertible)
Hint 1. First replace m by n using Corollary 4.1.27.
Hint 2. Multiply out $(I_n - A)\big(I_n + A + \cdots + A^{n-1}\big)$; the sum telescopes.
Sketch. By Corollary 4.1.27, $A^n = 0$. Then $(I_n - A)\big(I_n + A + \cdots + A^{n-1}\big) = I_n - A^n = I_n$, and the product in the other order telescopes in the same way.

By Corollary 4.1.27, $A^n = 0$. Let $B = I_n + A + \cdots + A^{n-1}$. By distributivity (Proposition 2.4.14),

$$(I_n - A)B = \big(I_n + A + \cdots + A^{n-1}\big) - \big(A + A^2 + \cdots + A^n\big) = I_n - A^n = I_n,$$

and in the same way $B(I_n - A) = I_n - A^n = I_n$, since A commutes with its powers. So B is a two-sided inverse of $I_n - A$ (Definition 2.5.12). Exercise 2.5.39 gave the inverse $I_n + A + \cdots + A^{m-1}$; the point here is that the terms from A^n on are all 0, so at most n terms are needed.

Solution to Exercise 4.1.45

Uses. Definition 1.2.9 (Function spaces), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.6.2 (Bases are independent spanning lists), Corollary 4.1.13 (Detecting infinite dimension)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. To show that $\delta_{x_1}, \ldots, \delta_{x_k}$ is independent for distinct points, evaluate a relation at x_j. Which terms survive?
Hint 2. For spanning in (a), compare f with $\sum_i f(x_i)\delta_{x_i}$ at each point. For (b), use Corollary 4.1.13.
Sketch. For distinct $x_1, \ldots, x_k$, evaluating $\sum a_i \delta_{x_i} = 0$ at x_j kills every term but the j-th and gives $a_j = 0$. In (a), $f = \sum_i f(x_i)\delta_{x_i}$ because both sides agree at every x_j. In (b), X has k distinct points for every k, so F^X has independent lists of every length.

Let $x_1, \ldots, x_k \in X$ be distinct, and suppose $\sum_{i=1}^k a_i \delta_{x_i} = 0$, the zero function. Evaluating at x_j, and using that $\delta_{x_i}(x_j)$ is 1 for $i = j$ and 0 for $i \neq j$, gives $a_j = 0$. So $\delta_{x_1}, \ldots, \delta_{x_k}$ is independent (Theorem 1.5.5). Evaluation at x_j is a linear map $F^X \to F$ that kills every δ_{x_i} except δ_{x_j}, which is the move of Technique 4.1.25.

(a) For $f \in F^X$, the function $g = \sum_{i=1}^n f(x_i)\delta_{x_i}$ satisfies $g(x_j) = f(x_j)$ for every j, so $g = f$. So $\delta_{x_1}, \ldots, \delta_{x_n}$ spans F^X, and with the independence above it is a basis (Proposition 1.6.2). Hence $\dim F^X = n$. For $X = \{1, \ldots, n\}$, these functions are the standard basis vectors of F^n (Remark 1.2.11).

(b) If X is infinite, then for every k it contains k distinct points, and the corresponding δ's form an independent list of length k. By Corollary 4.1.13, F^X is infinite-dimensional.

Solution to Exercise 4.1.46

Uses. Proposition 1.6.2 (Bases are independent spanning lists), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.1.17 (No zero divisors)
Hint 1. A symmetric matrix is determined by its entries on and above the diagonal. Find one matrix in S for each such position.
Hint 2. For K, look at the diagonal: $a_{ii} = -a_{ii}$ means $(1+1)a_{ii} = 0$. When $1 + 1 = 0$, compare the conditions defining S and K.

Sketch. (a) Basis U_{ii} $(1 \le i \le n)$ and $U_{ij} + U_{ji}$ $(i < j)$: $n + n(n-1)/2$ matrices. (b) If $1 + 1 \neq 0$, the diagonal of a skew-symmetric matrix is 0, and $U_{ij} - U_{ji}$ $(i < j)$ is a basis. If $1 + 1 = 0$, then $-a = a$ for all a, so $K = S$.

(a) If $A \in S$, then

$$A = \sum_{i=1}^{n} a_{ii} U_{ii} + \sum_{i<j} a_{ij}\big(U_{ij} + U_{ji}\big),$$

since both sides have (i, j) entry a_{ij} for $i \le j$, and (j, i) entry $a_{ij} = a_{ji}$ for $i < j$. The matrices U_{ii} and $U_{ij} + U_{ji}$ $(i < j)$ are symmetric, so they span S. A combination $\sum_i c_{ii} U_{ii} + \sum_{i<j} c_{ij}\big(U_{ij} + U_{ji}\big)$ has (i, j) entry c_{ij} for $i \le j$, so it is 0 only if all coefficients are 0. So these matrices form a basis of S, and

$$\dim S = n + n\frac{n-1}{2} = n\frac{n+1}{2}.$$

(b) Suppose $1 + 1 \neq 0$. For $A \in K$, $a_{ii} = -a_{ii}$ gives $(1 + 1)a_{ii} = 0$, so $a_{ii} = 0$ (Proposition 1.1.17). Then $A = \sum_{i<j} a_{ij}\big(U_{ij} - U_{ji}\big)$, as both sides have (i, j) entry a_{ij} and (j, i) entry $-a_{ij} = a_{ji}$ for $i < j$, and zero diagonal. The matrices $U_{ij} - U_{ji}$ lie in K and are independent, since a combination of them has (i, j) entry equal to the coefficient of $U_{ij} - U_{ji}$ for $i < j$. So $\dim K = n(n-1)/2$.

Suppose $1 + 1 = 0$. Then $-1 = 1$, so $-a_{ji} = a_{ji}$ and the conditions defining K and S coincide. Hence $K = S$ and $\dim K = n(n+1)/2$ by (a).

Solution to Exercise 4.1.47

Uses. Definition B.1.1 (Complex conjugate), Theorem 1.4.3 (Subspace criterion), Proposition 1.6.2 (Bases are independent spanning lists), Example 4.1.7 (Dimensions of the standard spaces)
Hint 1. Conjugation satisfies $\overline{z + w} = \overline{z} + \overline{w}$ and $\overline{tz} = t\overline{z}$ for real t, but $\overline{iz} = -i\overline{z}$. What does $a_{ii} = \overline{a_{ii}}$ say about diagonal entries?
Hint 2. For (b), a Hermitian matrix is determined by n real diagonal entries and the complex entries above the diagonal, each of which carries two real parameters. Find a matrix in H for each real parameter.
Sketch. (a) The conditions are preserved by sums and real multiples because conjugation is additive and fixes reals; $I_n \in H$ but $iI_n \notin H$. (b) Basis over $\mathbb{R}$: U_{kk}, and for $k < l$ both $U_{kl} + U_{lk}$ and $i(U_{kl} - U_{lk})$: $n + 2 \cdot n(n-1)/2 = n^2$ matrices. And $\dim_{\mathbb{R}} \mathbb{C}^{n\times n} = 2 \dim_{\mathbb{C}} \mathbb{C}^{n\times n} = 2n^2$.

(a) The zero matrix is Hermitian. If $A, B \in H$ and $t \in \mathbb{R}$, then for all i, j,

$$(A + tB)_{ji} = a_{ji} + tb_{ji} = \overline{a_{ij}} + t\overline{b_{ij}} = \overline{a_{ij} + tb_{ij}},$$

since conjugation is additive and $\bar{t} = t$ for real t. So H is a subspace over $\mathbb{R}$ (Theorem 1.4.3). It is not a subspace over $\mathbb{C}$: $I_n \in H$, but iI_n has diagonal entry i, while a Hermitian matrix has $a_{11} = \overline{a_{11}}$, so a_{11} is real.

(b) For $k < l$, let $X_{kl} = U_{kl} + U_{lk}$ and $Y_{kl} = i(U_{kl} - U_{lk})$. These are Hermitian: the (k, l) and (l, k) entries are $1, 1$ and $i, -i$, which are conjugate pairs, and the other entries are 0. So are the U_{kk}. Let $A \in H$. Its diagonal entries are real, and for $k < l$ write $a_{kl} = x_{kl} + iy_{kl}$ with $x_{kl}, y_{kl} \in \mathbb{R}$, so that $a_{lk} = x_{kl} - iy_{kl}$. Then

$$A = \sum_{k=1}^{n} a_{kk} U_{kk} + \sum_{k<l} (x_{kl} X_{kl} + y_{kl} Y_{kl}),$$

since both sides have the same (k, k), (k, l), and (l, k) entries. So these matrices span H over $\mathbb{R}$. If a real combination $\sum_k r_k U_{kk} + \sum_{k<l} (s_{kl} X_{kl} + t_{kl} Y_{kl})$ is 0, its (k, k) entries give $r_k = 0$, and its (k, l) entry $s_{kl} + it_{kl} = 0$ gives $s_{kl} = t_{kl} = 0$, because s_{kl}, t_{kl} are real. So they form a basis of H over $\mathbb{R}$, and

$$\dim_{\mathbb{R}} H = n + 2 \cdot \frac{n(n-1)}{2} = n^2.$$

Since $\dim_{\mathbb{C}} \mathbb{C}^{n\times n} = n^2$, Example 4.1.7(c) gives $\dim_{\mathbb{R}} \mathbb{C}^{n\times n} = 2n^2$. So the Hermitian matrices are a real subspace of exactly half the real dimension.

Solution to Exercise 4.1.48

Uses. Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.16 (Extending to a basis)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. If a sum lies in a subspace, must its terms?
Hint 2. Try $V = \mathbb{R}^2$ with the standard basis and $U = \operatorname{span}((1,1))$.
Sketch. The step "each term $a_i\mathbf{v}_i$ lies in U" is wrong: a sum can lie in a subspace without its terms doing so. The claim is false: in $\mathbb{R}^2$, no standard basis vector lies in $U = \operatorname{span}((1,1))$, so $\mathcal{B}' = ()$, which does not span U. What is true is the reverse: a basis of U extends to a basis of V (Theorem 4.1.16).

The error is the sentence "each term $a_i\mathbf{v}_i$ lies in U". A subspace is closed under sums, but a sum lying in U says nothing about its terms. For instance $(1,0) + (0,1) \in \operatorname{span}((1,1))$, while neither term is.

The claim is false. Take $V = \mathbb{R}^2$, $\mathcal{B} = \mathbf{e}_1, \mathbf{e}_2$, and $U = \operatorname{span}((1,1))$. Neither $\mathbf{e}_1$ nor $\mathbf{e}_2$ is a multiple of $(1,1)$, so $\mathcal{B}' = ()$, whose span $\{\mathbf{0}\}$ is not U. A basis of V need not have any vectors in a given subspace, which is why the proof of Proposition 4.1.21 had to build a basis of U from scratch. The correct relation goes the other way: a basis of U can be extended to a basis of V (Theorem 4.1.16), so *some* basis of V contains a basis of U.

Solution to Exercise 4.1.49

Uses. Theorem 1.5.5 (Independence is decided at zero), Theorem 4.1.10 (Independent lists are short), Lemma 1.3.9 (Arithmetic of linear combinations)
Hint 1. Rewrite $\sum_i a_i(\mathbf{v}_i - \mathbf{w}) = \mathbf{0}$ as $\sum_i a_i\mathbf{v}_i = \left(\sum_i a_i\right)\mathbf{w}$. Split into the cases $\sum_i a_i = 0$ and $\sum_i a_i \neq 0$.
Hint 2. For (b), the differences $(\mathbf{v}_i - \mathbf{w}) - (\mathbf{v}_m - \mathbf{w})$ lie in the span. Are they independent?
Sketch. (a) A relation $\sum a_i(\mathbf{v}_i - \mathbf{w}) = \mathbf{0}$ says $\sum a_i\mathbf{v}_i = s\mathbf{w}$ with $s = \sum a_i$. If $s = 0$ the relation is trivial by independence of the $\mathbf{v}_i$; so a nontrivial one has $s \neq 0$ and gives $\mathbf{w} = \sum(a_i/s)\mathbf{v}_i$. Conversely, coefficients t_i with sum 1 give the nontrivial relation $\sum t_i(\mathbf{v}_i - \mathbf{w}) = \mathbf{w} - \mathbf{w} = \mathbf{0}$. (b) The $m-1$ vectors $\mathbf{v}_i - \mathbf{v}_m$ $(i < m)$ lie in the span and are independent.

(a) For scalars $a_1, \ldots, a_m$ with $s = a_1 + \cdots + a_m$, Lemma 1.3.9 gives

$$\sum_{i=1}^{m} a_i(\mathbf{v}_i - \mathbf{w}) = \sum_{i=1}^{m} a_i\mathbf{v}_i - s\mathbf{w}.$$

($\Rightarrow$) Suppose the a_i form a nontrivial relation of the list, so $\sum_i a_i\mathbf{v}_i = s\mathbf{w}$. If $s = 0$, then $\sum_i a_i\mathbf{v}_i = \mathbf{0}$, and independence forces every $a_i = 0$, which is false. So $s \neq 0$, and $\mathbf{w} = \sum_i(a_i/s)\mathbf{v}_i$, where the coefficients $t_i = a_i/s$ satisfy $\sum_i t_i = s/s = 1$.

($\Leftarrow$) If $\mathbf{w} = \sum_i t_i\mathbf{v}_i$ with $\sum_i t_i = 1$, then $\sum_i t_i(\mathbf{v}_i - \mathbf{w}) = \mathbf{w} - 1\mathbf{w} = \mathbf{0}$, and not all t_i are 0 since they sum to 1. So the list is dependent.

(b) Let $W = \operatorname{span}(\mathbf{v}_1 - \mathbf{w}, \ldots, \mathbf{v}_m - \mathbf{w})$, which is finite-dimensional. If $m = 1$ there is nothing to prove. For $m \geq 2$ and $i < m$, the vector $\mathbf{v}_i - \mathbf{v}_m = (\mathbf{v}_i - \mathbf{w}) - (\mathbf{v}_m - \mathbf{w})$ lies in W. These $m-1$ vectors are independent: if $\sum_{i<m} c_i(\mathbf{v}_i - \mathbf{v}_m) = \mathbf{0}$, then $\sum_{i<m} c_i\mathbf{v}_i - \left(\sum_{i<m} c_i\right)\mathbf{v}_m = \mathbf{0}$, and independence of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ gives $c_i = 0$ for all $i < m$. By Theorem 4.1.10(a), applied in W, $\dim W \geq m - 1$.

Geometrically, (a) says the translated list is dependent exactly when $\mathbf{w}$ lies on the affine hull of the $\mathbf{v}_i$, the "flat" through them (Section 4.7).

Solution to Exercise 4.1.50

Uses. Proposition 4.1.21 (Dimensions of subspaces), Corollary 1.5.13 (Independence, one entry at a time), Definition 1.4.1 (Subspace)
Hint 1. Each U_{i-1} is a subspace of U_i and differs from it. What does Proposition 4.1.21(b), applied in U_i, say about their dimensions?
Hint 2. The dimensions form a strictly increasing sequence of integers between 0 and n. For existence, use the spans of the first i vectors of a basis.
Sketch. U_{i-1} is a proper subspace of U_i, so $\dim U_{i-1} < \dim U_i$. Hence $\dim U_i \geq \dim U_0 + i \geq i$, and $k \leq \dim U_k \leq n$. For a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$, the spans $U_i = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)$ increase strictly. If $k = n$, then $n \geq \dim U_n \geq \dim U_0 + n$ forces $\dim U_0 = 0$ and equality throughout.

(a) For $1 \leq i \leq k$, U_{i-1} is a subset of U_i closed under the operations, so it is a subspace of U_i (Definition 1.4.1), and $U_{i-1} \neq U_i$. By Proposition 4.1.21, applied in the finite-dimensional space U_i, $\dim U_{i-1} < \dim U_i$. So the

integers $\dim U_0 < \dim U_1 < \cdots < \dim U_k$ increase strictly, which gives $\dim U_i \geq \dim U_0 + i \geq i$. In particular $k \leq \dim U_k \leq n$, again by Proposition 4.1.21.

(b) Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of V and $U_i = \operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_i)$ for $0 \leq i \leq n$. Then $U_{i-1} \subseteq U_i$, and $U_{i-1} \neq U_i$ because $\mathbf{b}_i \notin U_{i-1}$ (Corollary 1.5.13). So this is a chain with $k = n$.

Now let $U_0 \subsetneq \cdots \subsetneq U_n$ be any chain with $k = n$. By (a), $n \geq \dim U_n \geq \dim U_0 + n$, so $\dim U_0 = 0$, and then $i \leq \dim U_i \leq \dim U_n - (n - i) \leq i$ for every i, using the strict increase once more. So $\dim U_i = i$. In particular $U_0 = \{\mathbf{0}\}$, and $U_n = V$ by Proposition 4.1.21(b).

Solution to Exercise 4.1.51

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Example 2.1.9 (Maps of polynomials), Proposition 1.1.17 (No zero divisors), Proposition 1.5.15 (Short lists, zeros, and repeats)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Show that $D^j\left(x^k\right) = k(k-1)\cdots(k-j+1)x^{k-j}$ for $j \leq k$, and $D^j\left(x^k\right) = 0$ for $j > k$. What are $D^n p$ and $D^{n+1} p$?
Hint 2. $D^n p$ is the constant $n! a_n$, where $a_n \neq 0$ is the leading coefficient. Apply Proposition 4.1.26 with $T = D$ and $k = n$, then count.
Sketch. By induction on j, $D^j x^k = k(k-1)\cdots(k-j+1)x^{k-j}$ for $j \leq k$, so $D^n p = n! a_n \neq 0$ and $D^{n+1} p = 0$. By Proposition 4.1.26 the $n+1$ polynomials $p, \ldots, D^n p$ are independent. They lie in $F[x]_{\leq n}$, which has dimension $n+1$, so they form a basis. Over $\mathbb{F}_2$, $D\left(x^2\right) = 2x = 0$, and the list $x^2, 0, 0$ contains $\mathbf{0}$.

(a) By Example 2.1.9, $D\left(x^k\right) = kx^{k-1}$ for $k \geq 1$ and $D(1) = 0$. By induction on j, $D^j\left(x^k\right) = k(k-1)\cdots(k-j+1)x^{k-j}$ for $0 \leq j \leq k$; in particular $D^k\left(x^k\right) = k!$ is a constant, and hence $D^j\left(x^k\right) = 0$ for $j > k$.

Write $p = a_0 + a_1 x + \cdots + a_n x^n$ with $a_n \neq 0$. By linearity, $D^n p = \sum_k a_k D^n\left(x^k\right) = n! a_n$, since the terms with $k < n$ vanish. This is nonzero, since $n! \neq 0$ and $a_n \neq 0$ (Proposition 1.1.17). And $D^{n+1} p = D(n! a_n) = 0$. By Proposition 4.1.26, applied to $T = D \in \mathrm{L}(F[x])$, $\mathbf{v} = p$, and $k = n$, the list $p, Dp, \ldots, D^n p$ is independent. Each $D^j p$ has degree at most n, so the list lies in $F[x]_{\leq n}$. It has $n + 1 = \dim F[x]_{\leq n}$ entries, so it is a basis by Theorem 4.1.10(b).

(b) Over $\mathbb{F}_2$, $D\left(x^2\right) = 2x = 0$, so the list is $x^2, 0, 0$. It contains $\mathbf{0}$, so it is dependent (Proposition 1.5.15(c)). Here $2! = 2 = 0$, so the hypothesis fails, and the chain from x^2 reaches $\mathbf{0}$ after one step instead of two.

Solution to Exercise 4.1.52

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Definition 2.3.1 (Matrix of a linear map), (1.6.1)
Hint 1. Since $T^{n-1} \neq 0$, some $\mathbf{v}$ has $T^{n-1}\mathbf{v} \neq \mathbf{0}$. What is $T^n \mathbf{v}$?
Hint 2. For (b), write $\mathbf{b}_i = T^{n-i}\mathbf{v}$ and compute $T\mathbf{b}_i$ in terms of the $\mathbf{b}$'s. The i-th column of the matrix is $[T\mathbf{b}_i]_{\mathcal{B}}$.
Sketch. (a) Choose $\mathbf{v}$ with $T^{n-1}\mathbf{v} \neq \mathbf{0}$; since $T^n \mathbf{v} = \mathbf{0}$, the chain $\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is independent (Proposition 4.1.26) of length n, hence a basis. (b) With $\mathbf{b}_i = T^{n-i}\mathbf{v}$: $T\mathbf{b}_1 = \mathbf{0}$ and $T\mathbf{b}_i = \mathbf{b}_{i-1}$ for $i \geq 2$. So the matrix has 1's just above the diagonal and 0's elsewhere.

(a) Since $T^{n-1} \neq 0$, there is $\mathbf{v} \in V$ with $T^{n-1}\mathbf{v} \neq \mathbf{0}$, and $T^n \mathbf{v} = \mathbf{0}$. By Proposition 4.1.26 with $k = n - 1$, the list $\mathbf{v}, T\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is independent. It has $n = \dim V$ entries, so it is a basis (Theorem 4.1.10(b)).

(b) The list $\mathcal{B}$ is a reordering of the basis in (a), so it is a basis. Write $\mathbf{b}_i = T^{n-i}\mathbf{v}$ for $1 \leq i \leq n$. Then $T\mathbf{b}_1 = T^n \mathbf{v} = \mathbf{0}$, and $T\mathbf{b}_i = T^{n-i+1}\mathbf{v} = \mathbf{b}_{i-1}$ for $2 \leq i \leq n$. So $[T\mathbf{b}_1]_{\mathcal{B}} = \mathbf{0}$ and, by (1.6.1), $[T\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_{i-1}$ for $i \geq 2$. These are the columns of $[T]_{\mathcal{B}\mathcal{B}}$ (Definition 2.3.1), so

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

the matrix with 1 in positions $(i, i+1)$ and 0 elsewhere. It is the matrix of the shift N from the start of Section 4.1.7, so every such T looks like the shift in a suitable basis. This is the simplest case of the Jordan form (Section 10.3).

Solution to Exercise 4.1.53

Uses. Corollary 4.1.35 (Counting independent lists and bases), Theorem 3.5.14 (Third Fundamental Theorem, completed), Proposition 4.1.34 (Size of a vector space over a finite field)
Hint 1. A matrix is invertible exactly when its columns form a basis (Theorem 3.5.14), so there are $N_2(n,n)$ invertible matrices out of 2^{n^2}.
Hint 2. Write the fraction as $\prod_{j=1}^{n}(1-2^{-j})$. Keep the factor $1/2$ and bound the rest with $\prod(1-a_j) \geq 1-\sum a_j$ for $0 \leq a_j \leq 1$.
Sketch. (a) $168/512 = 21/64$. (b) The fraction is $\prod_{i=0}^{n-1}(2^n-2^i)/2^n = \prod_{j=1}^{n}(1-2^{-j})$. By induction, $\prod(1-a_j) \geq 1-\sum a_j$ when all $a_j \in [0,1]$, so $\prod_{j=2}^{n}(1-2^{-j}) \geq 1-(1/2-2^{-n}) > 1/2$, and the fraction exceeds $1/4$.

By Theorem 3.5.14, $A \in F^{n\times n}$ is invertible if and only if its columns form a basis of F^n, and A is determined by its list of columns. So by Corollary 4.1.35 there are $N_2(n,n) = \prod_{i=0}^{n-1}(2^n-2^i)$ invertible matrices in $\mathbb{F}_2^{n\times n}$, out of 2^{n^2} in all (Proposition 4.1.34).

(a) For $n=3$: $(8-1)(8-2)(8-4) = 168$, out of $2^9 = 512$, so the fraction is $168/512 = 21/64 \approx 0.33$.

(b) Dividing each of the n factors by 2^n,

$$N_2\frac{n,n}{2^{n^2}} = \prod_{i=0}^{n-1}(1-2^{i-n}) = \prod_{j=1}^{n}(1-2^{-j}).$$

First, if $a_2,\dots,a_n \in [0,1]$, then $\prod_{j=2}^{n}(1-a_j) \geq 1-\sum_{j=2}^{n} a_j$. This is clear for an empty product, and if P is the product of the first factors and S their sum, with $P \geq 1-S$, then, since $1-a \geq 0$,

$$P(1-a) \geq (1-S)(1-a) = 1-(S+a)+Sa \geq 1-(S+a).$$

With $a_j = 2^{-j}$ we get $\sum_{j=2}^{n} 2^{-j} = 1/2 - 2^{-n} < 1/2$, so

$$\prod_{j=1}^{n}(1-2^{-j}) = \frac{1}{2}\prod_{j=2}^{n}(1-2^{-j}) \geq \frac{1}{2}\left(\frac{1}{2}+2^{-n}\right) > \frac{1}{4}.$$

(As $n \to \infty$ the fraction decreases to about 0.289.)

Solution to Exercise 4.1.54

Uses. Proposition 4.1.37 (Counting subspaces), Corollary 4.1.35 (Counting independent lists and bases)
Techniques. Double counting (Technique 4.1.36)
Hint 1. Use Proposition 4.1.37 with $q=2$, $n=4$.
Hint 2. For (b), factor q^i out of each factor q^n-q^i in $N_q(n,k)$, and write the count as a ratio of products of factors q^j-1.
Sketch. (a) $1, 15, 35, 15, 1$, so 67 subspaces in all. (b) $N_q(n,k) = q^{k(k-1)/2}\prod_{j=n-k+1}^{n}(q^j-1)$, so the count is $[n]!/([k]![n-k]!)$ with $[m]! = \prod_{j=1}^{m}(q^j-1)$, which is symmetric in k and $n-k$.

(a) With $q=2$ and $n=4$, Proposition 4.1.37 gives

$$k=1: \frac{15}{1} = 15, \quad k=2: \frac{15\cdot 14}{3\cdot 2} = 35, \quad k=3: \frac{15\cdot 14\cdot 12}{7\cdot 6\cdot 4} = 15,$$

and of course one subspace each of dimensions 0 and 4. So $\mathbb{F}_2^4$ has $1+15+35+15+1 = 67$ subspaces.

(b) Factoring q^i out of q^n-q^i,

$$N_q(n,k) = \prod_{i=0}^{k-1} q^i(q^{n-i}-1) = q^{k(k-1)/2}\prod_{j=n-k+1}^{n}(q^j-1).$$

For $n=k$ this reads $N_q(k,k) = q^{k(k-1)/2}\prod_{j=1}^{k}(q^j-1)$. Writing $[m]! = \prod_{j=1}^{m}(q^j-1)$, with $[0]!=1$, the number of k-dimensional subspaces is therefore

$$N_q\frac{n,k}{N_q}(k,k) = \frac{\prod_{j=n-k+1}^{n}(q^j-1)}{[k]!} = \frac{[n]!}{[k]!\,[n-k]!},$$

since $\prod_{j=n-k+1}^{n}(q^j-1) = [n]!/[n-k]!$. The last expression does not change when k is replaced by $n-k$. (Duality explains the symmetry: sending U to the functionals that vanish on U matches the k-dimensional subspaces of V with the $(n-k)$-dimensional subspaces of the dual space, which has the same dimension as V (Section 13.1).)

Solution to Exercise 4.1.55

Uses. Exercise 4.1.52, Corollary 2.2.6 (Agreeing on a spanning list), Proposition 2.1.24 (Algebra of composition)
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. By Exercise 4.1.52 there is $\mathbf{v}$ with $\mathbf{v}, N\mathbf{v}, \dots, N^{n-1}\mathbf{v}$ a basis. Expand $S\mathbf{v}$ in this basis.
Hint 2. If $S\mathbf{v} = \sum c_i N^i\mathbf{v}$, let $P = \sum c_i N^i$. Compare S and P on each basis vector $N^j\mathbf{v}$, using that both commute with N^j.
Sketch. Polynomials in N commute with N. Conversely, take $\mathbf{v}$ with $\mathbf{v}, \dots, N^{n-1}\mathbf{v}$ a basis, write $S\mathbf{v} = \sum_i c_i N^i\mathbf{v}$, and let $P = \sum_i c_i N^i$. Since S and P commute with N^j, $SN^j\mathbf{v} = N^jS\mathbf{v} = N^jP\mathbf{v} = PN^j\mathbf{v}$. So S and P agree on a basis, and $S = P$.

($\Leftarrow$) N commutes with each power N^i, since $NN^i = N^{i+1} = N^iN$. By the distributive laws of Proposition 2.1.24, N commutes with every combination $\sum_i c_i N^i$.

($\Rightarrow$) Suppose $SN = NS$. By induction on j, $SN^j = N^jS$ for all $j \geq 0$: if $SN^j = N^jS$, then $SN^{j+1} = (SN)N^j = NSN^j = NN^jS = N^{j+1}S$. By Exercise 4.1.52 there is $\mathbf{v} \in V$ such that $\mathbf{v}, N\mathbf{v}, \dots, N^{n-1}\mathbf{v}$ is a basis of V. Expand $S\mathbf{v}$ in it:

$$S\mathbf{v} = c_0\mathbf{v} + c_1N\mathbf{v} + \cdots + c_{n-1}N^{n-1}\mathbf{v},$$

and let $P = c_0\mathrm{id}_V + c_1N + \cdots + c_{n-1}N^{n-1}$, so that $S\mathbf{v} = P\mathbf{v}$. For each j, using that S and P both commute with N^j,

$$S\big(N^j\mathbf{v}\big) = N^j(S\mathbf{v}) = N^j(P\mathbf{v}) = P\big(N^j\mathbf{v}\big).$$

So S and P agree on every vector of a basis of V, and therefore $S = P$ (Corollary 2.2.6).

Without the hypothesis $N^{n-1} \neq 0$ the statement fails: for $n \geq 2$ and $N = 0$, every S commutes with N, but the combinations of powers of 0 are only the multiples of id_V.

Solution to Exercise 4.1.56

Uses. Proposition 4.1.31 (Every square matrix satisfies a polynomial), Proposition 2.1.24 (Algebra of composition), Definition 2.6.1 (Injective), Definition 2.5.12 (Invertible)
Techniques. Too many vectors force a dependence (Technique 4.1.30), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. By Proposition 4.1.31 there is a nontrivial relation $\sum_{j=0}^{N} c_jT^j = 0$. Let r be the smallest index with $c_r \neq 0$ and factor out T^r.
Hint 2. T^r is injective, so $T^rR = 0$ forces $R = 0$. Now $R = c_r\mathrm{id}_V + TQ$ for a combination Q of powers of T. Solve for id_V.
Sketch. Take $\sum_{j=r}^{N} c_jT^j = 0$ with $c_r \neq 0$. Then $T^rR = 0$ with $R = \sum_{j\geq r} c_jT^{j-r}$, and T^r is injective, so $R = 0$. Writing $R = c_r\mathrm{id}_V + TQ$ gives $T\big(-c_r^{-1}Q\big) = \mathrm{id}_V$, and Q commutes with T, so $-c_r^{-1}Q$ is a two-sided inverse. For M: $1 \notin \operatorname{im} M$; the step that fails is the existence of a relation among $\mathrm{id}, M, M^2, \dots$, since these are independent ($M^j(1) = x^j$).

(a) If $V = \{\mathbf{0}\}$, then $T = \mathrm{id}_V$ and $T^{-1} = \mathrm{id}_V$. Otherwise $n = \dim V \geq 1$, and by Proposition 4.1.31 there are $c_0, \dots, c_N$ ($N = n^2$), not all 0, with $\sum_{j=0}^{N} c_jT^j = 0$. Let r be the smallest index with $c_r \neq 0$, and let

$$R = c_r\mathrm{id}_V + c_{r+1}T + \cdots + c_NT^{N-r},$$

so that $\sum_{j=r}^{N} c_jT^j = T^rR$ by Proposition 2.1.24. Thus $T^rR = 0$. The map T^r is injective, as a composite of injective maps. For every $\mathbf{u} \in V$, $T^r(R\mathbf{u}) = \mathbf{0} = T^r(\mathbf{0})$, so $R\mathbf{u} = \mathbf{0}$. Hence $R = 0$. Let $Q = c_{r+1}\mathrm{id}_V + c_{r+2}T + \cdots + c_NT^{N-r-1}$ (an empty sum is 0), so that $R = c_r\mathrm{id}_V + TQ$, and let $S = -c_r^{-1}Q$. Then $R = 0$ gives

$$TS = -c_r^{-1}TQ = -c_r^{-1}(R - c_r\mathrm{id}_V) = \mathrm{id}_V.$$

Since S is a combination of powers of T, it commutes with T, so $ST = TS = \mathrm{id}_V$ as well. So T is invertible (Definition 2.5.12), and $T^{-1} = S$ is a combination of $\mathrm{id}_V, T, \dots, T^{N-r-1}$.

(b) If $xp = 0$, then every coefficient of p, which is a coefficient of xp shifted by one place, is 0; so M is injective. Every xp has constant coefficient 0, so $1 \notin \operatorname{im} M$, and M is not surjective. The argument of (a) fails at its first step: there is no nontrivial relation $\sum_j c_j M^j = 0$, because applying it to 1 gives $\sum_j c_j x^j = 0$, which forces all $c_j = 0$. The space $\mathrm{L}(F[x])$ is infinite-dimensional, and the powers of M never become too many.

In particular, an injective operator on a finite-dimensional space is surjective. Section 4.2 proves this again with dimension counting (Corollary 4.2.18).

Solution to Exercise 4.1.57

> *Uses.* Theorem 4.1.10 (Independent lists are short), Theorem 1.5.5 (Independence is decided at zero), Example 4.1.7 (Dimensions of the standard spaces)
> *Techniques.* Too many vectors force a dependence (Technique 4.1.30)
> *Hint 1.* Encode A_i as the vector $\mathbf{x}_i \in \mathbb{R}^n$ whose t-th entry is 1 if $t \in A_i$ and 0 otherwise. There are $n+1$ of these.
> *Hint 2.* Take a nontrivial relation and move the terms with negative coefficients to the other side. When is the t-th entry of a combination of the $\mathbf{x}_i$ with positive coefficients positive?
> *Sketch.* The $n+1$ indicator vectors in $\mathbb{R}^n$ are dependent. Split a nontrivial relation as $\sum_{i\in I} c_i\mathbf{x}_i = \sum_{j\in J} d_j\mathbf{x}_j$ with all $c_i, d_j > 0$, where I and J are the indices with positive and negative coefficients. Both are nonempty, since a positive combination of nonzero vectors with entries ≥ 0 is nonzero. The t-th entry of the left side is positive exactly when $t \in \bigcup_I A_i$, and likewise on the right.

For each i, let $\mathbf{x}_i \in \mathbb{R}^n$ have t-th entry 1 if $t \in A_i$ and 0 otherwise; $\mathbf{x}_i \neq \mathbf{0}$ since $A_i \neq \emptyset$. These are $n+1$ vectors in $\mathbb{R}^n$, and $\dim \mathbb{R}^n = n$, so they are dependent (Theorem 4.1.10(a)): there are real $a_1, \ldots, a_{n+1}$, not all 0, with $\sum_i a_i\mathbf{x}_i = \mathbf{0}$. Let $I = \{i \mid a_i > 0\}$ and $J = \{j \mid a_j < 0\}$, which are disjoint. Moving the negative terms across,

$$\mathbf{y} = \sum_{i\in I} a_i\mathbf{x}_i = \sum_{j\in J} (-a_j)\mathbf{x}_j,$$

where all coefficients now are positive.

Consider the t-th entry of $\mathbf{y}$. On the left it is $\sum_{i\in I, t\in A_i} a_i$, a sum of positive numbers, which is positive if $t \in \bigcup_{i\in I} A_i$ and 0 otherwise. On the right, in the same way, it is positive exactly when $t \in \bigcup_{j\in J} A_j$. So the two unions are both equal to the set of t with $y_t > 0$, and hence equal to each other.

It remains to see that I and J are nonempty. Not all a_i are 0, so $I \cup J \neq \emptyset$; say $I \neq \emptyset$. Pick $i_0 \in I$ and $t \in A_{i_0}$. Then $y_t \geq a_{i_0} > 0$, so $\mathbf{y} \neq \mathbf{0}$, and the right side, an empty sum if $J = \emptyset$, would be $\mathbf{0}$. So $J \neq \emptyset$. The case $J \neq \emptyset$ is symmetric.

Solution to Exercise 4.1.58

> *Uses.* Theorem 1.6.10 (Every spanning list contains a basis), Lemma 2.4.24 (Matrix units), Definition 3.2.2 (Elementary matrices), Proposition 3.2.10 (Inverses of elementary matrices), Proposition 2.5.17 (Algebra of inverses), Theorem 4.1.10 (Independent lists are short)
> *Hint 1.* It is enough to show that the invertible matrices span $F^{n\times n}$, since a spanning list contains a basis. Write each matrix unit as a difference of two invertible matrices.
> *Hint 2.* For $i \neq j$, $U_{ij} = \big(I + U_{ij}\big) - I$. For U_{ii}, pick $j \neq i$ and let $P = E_{ij}$ be the swap. Show $P + U_{ii} = P\big(I + U_{ji}\big)$.
> *Sketch.* For $n = 1$ take (1). For $n \geq 2$: if $i \neq j$ then $\big(I + U_{ij}\big)\big(I - U_{ij}\big) = I$ because $U_{ij}^2 = 0$, so $U_{ij} = \big(I + U_{ij}\big) - I$ is a difference of invertibles. For U_{ii}, take $j \neq i$ and the swap $P = E_{ij}$; then $PU_{ii} = U_{ji}$, so $P + U_{ii} = P\big(I + U_{ji}\big)$ is invertible and $U_{ii} = (P + U_{ii}) - P$. So finitely many invertible matrices span, and a basis can be extracted from them. Over $\mathbb{F}_2$: I_2, $I_2 + U_{12}$, $I_2 + U_{21}$, $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

For $n = 1$, the matrix (1) is an invertible basis of $F^{1\times 1}$. Let $n \geq 2$, and write $I = I_n$. By Lemma 2.4.24, $U_{ij}U_{kl} = U_{il}$ if $j = k$ and 0 otherwise.

Off-diagonal units. For $i \neq j$, $U_{ij}^2 = 0$, so $\big(I + U_{ij}\big)\big(I - U_{ij}\big) = I - U_{ij}^2 = I$ and likewise $\big(I - U_{ij}\big)\big(I + U_{ij}\big) = I$. So $I + U_{ij}$ is invertible, and $U_{ij} = \big(I + U_{ij}\big) - I$.

Diagonal units. Fix i and choose $j \neq i$. Let $P = E_{ij}$ be the elementary matrix that swaps rows i and j (Definition 3.2.2); it is invertible with $P^{-1} = P$ (Proposition 3.2.10). Its i-th column is $\mathbf{e}_j$, so $PU_{ii} = P\mathbf{e}_i\mathbf{e}_i^{\mathsf{T}} = \mathbf{e}_j\mathbf{e}_i^{\mathsf{T}} = U_{ji}$. Hence

$$P\big(I + U_{ji}\big) = P + PU_{ji} = P + P(PU_{ii}) = P + U_{ii}.$$

By the previous paragraph $I + U_{ji}$ is invertible, so $P + U_{ii}$ is a product of invertible matrices and is invertible (Proposition 2.5.17). And $U_{ii} = (P + U_{ii}) - P$.

So every matrix unit is a combination of finitely many invertible matrices, and these invertible matrices span $F^{n\times n}$, since the matrix units do. By Theorem 1.6.10, some sublist of them is a basis.

For $\mathbb{F}_2^{2\times 2}$, take

$$I_2 = \begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \begin{pmatrix}1&1\\0&1\end{pmatrix},\quad \begin{pmatrix}1&0\\1&1\end{pmatrix},\quad \begin{pmatrix}1&1\\1&0\end{pmatrix}.$$

Each is invertible: the first three are I_2 and $I_2 + U_{12}$, $I_2 + U_{21}$, and the last is $E_{12} + U_{11}$. A combination $aI_2 + b(I_2 + U_{12}) + c(I_2 + U_{21}) + d(E_{12} + U_{11})$ has $(1,1)$ entry $a + b + c + d$ and $(2,2)$ entry $a + b + c$, so if it is 0 then $d = 0$; its $(1,2)$ and $(2,1)$ entries are then b and c, so $b = c = 0$, and finally $a = 0$. Four independent vectors in the 4-dimensional space $\mathbb{F}_2^{2\times 2}$ form a basis (Theorem 4.1.10(b)).

Solution to Exercise 4.1.59

Uses. Exercise 1.5.33, Proposition 1.6.19 (Coordinates respect linear combinations), Definition 1.6.14 (Coordinates)
Hint 1. By Exercise 1.5.33, replacing $\mathbf{b}_i$ by $\mathbf{w}$ gives a basis exactly when the i-th $\mathcal{B}$-coordinate of $\mathbf{w}$ is nonzero. Express both conditions (a) and (b) by coordinates.
Hint 2. Let β_j be the j-th $\mathcal{C}$-coordinate of $\mathbf{b}_i$ and γ_j the i-th $\mathcal{B}$-coordinate of $\mathbf{c}_j$. Take the i-th $\mathcal{B}$-coordinate of both sides of $\mathbf{b}_i = \sum_j \beta_j\mathbf{c}_j$.
Sketch. By Exercise 1.5.33, (a) holds iff $\gamma_j \neq 0$ and (b) holds iff $\beta_j \neq 0$, where $\mathbf{b}_i = \sum_j \beta_j\mathbf{c}_j$ and γ_j is the i-th $\mathcal{B}$-coordinate of $\mathbf{c}_j$. Taking i-th $\mathcal{B}$-coordinates of $\mathbf{b}_i = \sum_j \beta_j\mathbf{c}_j$ gives $1 = \sum_j \beta_j\gamma_j$, so some $\beta_j\gamma_j \neq 0$.

For $\mathbf{w} \in V$, Exercise 1.5.33 says that replacing $\mathbf{b}_i$ by $\mathbf{w}$ in the independent list $\mathcal{B}$ gives an independent list if and only if the coefficient of $\mathbf{b}_i$ in $\mathbf{w}$, that is, the i-th coordinate of $[\mathbf{w}]_{\mathcal{B}}$, is nonzero, and that the span is then still V. So the new list is a basis exactly in that case. Let γ_j be the i-th coordinate of $\big[\mathbf{c}_j\big]_{\mathcal{B}}$. Then (a) holds for j if and only if $\gamma_j \neq 0$.

In the same way, with the roles of the bases exchanged, write $[\mathbf{b}_i]_{\mathcal{C}} = (\beta_1, \ldots, \beta_n)$, that is, $\mathbf{b}_i = \sum_j \beta_j\mathbf{c}_j$. Then (b) holds for j if and only if $\beta_j \neq 0$.

Take the i-th $\mathcal{B}$-coordinate of both sides of $\mathbf{b}_i = \sum_j \beta_j\mathbf{c}_j$. On the left it is 1, since $[\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_i$. On the right, by Proposition 1.6.19(c), it is $\sum_j \beta_j\gamma_j$. So $\sum_{j=1}^{n} \beta_j\gamma_j = 1 \neq 0$, and some term $\beta_j\gamma_j$ is nonzero. For that j, both $\beta_j \neq 0$ and $\gamma_j \neq 0$, so (a) and (b) both hold.

Solution to Exercise 4.1.60

Uses. Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Theorem 4.1.10 (Independent lists are short), Theorem 1.5.5 (Independence is decided at zero), Proposition 2.1.17 (Compositions of linear maps are linear), Proposition 1.2.35 (Zero products), Definition 1.1.28 (Characteristic)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Too many vectors force a dependence (Technique 4.1.30)
Hint 1. Let $\varphi : F^{n\times n} \to F^{n\times n}$, $\varphi(X) = XB - BX$. Show that φ is linear and that $\varphi\big(A^k\big) = kA^k$ for all $k \geq 1$.
Hint 2. Suppose $A^1, \ldots, A^N$ are all nonzero, with $N = n^2 + 1$. Show they are independent: to isolate the coefficient of A^m in a relation, apply the composite of the maps $\varphi - j\,\mathrm{id}$ for $j \neq m$.
Hint 3. So some A^k with $k \leq n^2 + 1$ is 0. Finish with Corollary 4.1.27.
Sketch. $\varphi(X) = XB - BX$ is linear, and $\varphi\big(A^{k+1}\big) = A\varphi\big(A^k\big) + \varphi(A)A^k$ gives $\varphi\big(A^k\big) = kA^k$. If $A, \ldots, A^N$ ($N = n^2 + 1$) were all nonzero, applying $\prod_{j\neq m}(\varphi - j\,\mathrm{id})$ to a relation $\sum c_kA^k = 0$ would leave $c_m \prod_{j\neq m}(m - j)A^m = 0$, and $\prod_{j\neq m}(m - j)$ is a nonzero integer, nonzero in F. So $c_m = 0$ for all m, and $N > n^2$ independent matrices is impossible. Hence A is nilpotent, and $A^n = 0$.

Let $\varphi : F^{n\times n} \to F^{n\times n}$ be $\varphi(X) = XB - BX$. It is linear, since $(X + cY)B - B(X + cY) = (XB - BX) + c(YB - BY)$. We show $\varphi\big(A^k\big) = kA^k$ for $k \geq 1$ by induction. For $k = 1$ this is the hypothesis. If it holds for k, then

$$\varphi\big(A^{k+1}\big) = AA^kB - BAA^k = A\big(A^kB - BA^k\big) + (AB - BA)A^k = A\big(kA^k\big) + AA^k = (k + 1)A^{k+1}.$$

For an integer j, let $L_j = \varphi - j\,\mathrm{id}$, a linear map on $F^{n\times n}$; here j stands for $j \cdot 1 \in F$. Then $L_j\big(A^k\big) = (k - j)A^k$, and therefore $L_j\big(cA^k\big) = c(k - j)A^k$ for every scalar c.

Let $N = n^2 + 1$, and suppose, to the contrary, that $A^k \neq 0$ for $k = 1, \ldots, N$. We show that $A, A^2, \ldots, A^N$ is independent. Let $\sum_{k=1}^{N} c_kA^k = 0$, and fix $m \in \{1, \ldots, N\}$. Apply the composite $L = L_1 \cdots L_{m-1}L_{m+1} \cdots L_N$, which is linear (Proposition 2.1.17). Applying the factors one at a time,

$$L\big(A^k\big) = \left(\prod_{j\neq m}(k - j)\right)A^k,$$

which is 0 for $k \neq m$, because the product contains the factor $k - k = 0$. So applying L to the relation leaves

$$c_m\left(\prod_{j\neq m}(m - j)\right)A^m = 0.$$

The product is a product of nonzero integers, hence a nonzero integer, and in a field of characteristic 0 a nonzero integer is a nonzero element. Since $A^m \neq 0$, Proposition 1.2.35 gives $c_m = 0$. As m was arbitrary, the relation is trivial, and $A, \ldots, A^N$ is an independent list of $n^2 + 1$ vectors in $F^{n\times n}$, contradicting Theorem 4.1.10(a).

So $A^k = 0$ for some k with $1 \leq k \leq n^2 + 1$, and Corollary 4.1.27 gives $A^n = 0$.

The hypothesis on F is needed. Over $\mathbb{F}_3$, the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

satisfy $AB - BA = A$ (check it), while $A^3 = I_3$. In the proof, the integers $1, 2, 3, \ldots$ are no longer distinct in $\mathbb{F}_3$, so φ cannot tell the powers of A apart.

Section 4.2

Solution to Exercise 4.2.4

> *Uses.* Proposition 4.2.3 (Properties of restriction), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Example 2.1.9 (Maps of polynomials)
> *Hint 1.* $\ker D$ is the constants. Intersect it with each subspace.
> *Hint 2.* The image of a restriction to $\mathrm{span}(\mathcal{L})$ is spanned by the values on $\mathcal{L}$ (Proposition 2.1.16(a)).
> *Sketch.* $\ker D = \mathrm{span}(1)$. (a) $U \cap \ker D = \{\mathbf{0}\}$, so the kernel is $\{\mathbf{0}\}$ and the image is $\mathrm{span}\big(2x, 3x^2\big) = \mathrm{span}\big(x, x^2\big)$. (b) $U' \cap \ker D = \mathrm{span}(1)$, and the image is $\mathrm{span}(0, 1) = \mathrm{span}(1)$. (c) $U'' = \mathrm{span}\big(x, x^2, x^3\big)$.

Over $\mathbb{R}$, $D\big(a_0 + a_1x + a_2x^2 + a_3x^3\big) = a_1 + 2a_2x + 3a_3x^2$ is $\mathbf{0}$ exactly when $a_1 = a_2 = a_3 = 0$, so $\ker D = \mathrm{span}(1)$.

(a) A polynomial $ax^2 + bx^3$ is constant only if $a = b = 0$, so $U \cap \ker D = \{\mathbf{0}\}$, and $\ker(D|_U) = \{\mathbf{0}\}$ by Proposition 4.2.3(c). By Proposition 2.1.16(a), $\mathrm{im}(D|_U) = \mathrm{span}\big(D\big(x^2\big), D\big(x^3\big)\big) = \mathrm{span}\big(2x, 3x^2\big) = \mathrm{span}\big(x, x^2\big)$.

(b) Here $U' \cap \ker D = \mathrm{span}(1)$, which is $\ker(D|_{U'})$, and $\mathrm{im}(D|_{U'}) = \mathrm{span}(D(1), D(x)) = \mathrm{span}(\mathbf{0}, 1) = \mathrm{span}(1)$.

(c) Let $U'' = \mathrm{span}\big(x, x^2, x^3\big)$. Its only constant is $\mathbf{0}$, so $D|_{U''}$ is injective, and $\mathrm{im}(D|_{U''}) = \mathrm{span}\big(1, 2x, 3x^2\big) = \mathbb{R}[x]_{\leq 2} = \mathrm{im}\, D$. (So does $\mathrm{span}\big(x + 1, x^2, x^3\big)$; such a subspace is not unique.)

Solution to Exercise 4.2.10

Uses. Theorem 4.2.7 (The image of T_A), Remark 3.4.11 (A basis of the kernel), Definition 3.4.8 (Basic solutions)
Hint 1. Row reduce. Subtract twice row 1 from row 2 first.
Hint 2. For (b), use the basis from (a) and compare entries, starting with the third.
Sketch. $A_r = \begin{pmatrix} 1 & -1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, with pivot columns 1, 3. So $(1,2,0), (2,5,1)$ is a basis of $\operatorname{im} T_A$, and the basic solutions $(1,1,0,0), (2,0,-1,1)$ are a basis of $\ker T_A$. (b) No: $a(1,2,0) + b(2,5,1)$ has third entry b, so $b = 0$, and then the second entry is $2a = 2 \neq 0$.

(a) Adding -2 times row 1 to row 2 gives rows $(1,-1,2,0)$, $(0,0,1,1)$, $(0,0,1,1)$. Adding -1 times row 2 to row 3 and -2 times row 2 to row 1 gives

$$A_r = \begin{pmatrix} 1 & -1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The pivot columns are 1 and 3, so by Theorem 4.2.7 the columns $(1,2,0)$ and $(2,5,1)$ of A form a basis of $\operatorname{im} T_A$. The free columns are 2 and 4, and the basic solutions are $\mathbf{s}_1 = (1,1,0,0)$ and $\mathbf{s}_2 = (2,0,-1,1)$, a basis of $\ker T_A$ (Remark 3.4.11). Check: $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$ and $2\mathbf{a}_1 - \mathbf{a}_3 + \mathbf{a}_4 = (2-2+0, 4-5+1, 0-1+1) = \mathbf{0}$.

(b) No. If $a(1,2,0) + b(2,5,1) = (1,0,0)$, the third entry gives $b = 0$, the first then gives $a = 1$, and the second entry is $2 \neq 0$. So $(1,0,0) \notin \operatorname{im} T_A$, although it is a pivot column of A_r.

Solution to Exercise 4.2.13

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Proposition 1.6.2 (Bases are independent spanning lists)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. For (a), set each entry of $T(\mathbf{x})$ to 0.
Hint 2. For (b), show one containment directly, and compare dimensions using rank–nullity.
Sketch. (a) $T(\mathbf{x}) = \mathbf{0}$ iff $x_1 = x_2 = x_3$, so $\ker T = \operatorname{span}((1,1,1))$. (b) The entries of $T(\mathbf{x})$ sum to 0, so $\operatorname{im} T \subseteq P = \{y_1 + y_2 + y_3 = 0\}$. By rank–nullity $\dim \operatorname{im} T = 3 - 1 = 2$, and P has the basis $(1,-1,0), (0,1,-1)$, so $\dim P = 2$ and $\operatorname{im} T = P$.

(a) $T(\mathbf{x}) = \mathbf{0}$ means $x_1 = x_2$, $x_2 = x_3$, and $x_3 = x_1$, that is, $x_1 = x_2 = x_3$. So $\ker T = \operatorname{span}((1,1,1))$, which has dimension 1.

(b) Let $P = \{\mathbf{y} \in F^3 \mid y_1 + y_2 + y_3 = 0\}$. The entries of $T(\mathbf{x})$ add up to $(x_1 - x_2) + (x_2 - x_3) + (x_3 - x_1) = 0$, so $\operatorname{im} T \subseteq P$. By Theorem 4.2.11 and (a), $\dim \operatorname{im} T = 3 - 1 = 2$. On the other hand, every $\mathbf{y} \in P$ satisfies

$$\mathbf{y} = y_1(1,-1,0) + (y_1 + y_2)(0,1,-1),$$

since the third entry of the right side is $-y_1 - y_2 = y_3$. The two vectors lie in P and are independent (look at the first entries, then the second), so they form a basis of P and $\dim P = 2$. By Proposition 4.1.21(b), applied in P, $\operatorname{im} T = P$. Every step is valid over every field.

Solution to Exercise 4.2.17

Uses. Proposition 4.2.15 (Rank through any matrix), Definition 2.3.1 (Matrix of a linear map), Theorem 4.2.11 (Rank–nullity)
Hint 1. The columns of the matrix are $T(1)$, $T(x)$, $T(x^2)$.
Hint 2. Over $\mathbb{F}_2$, the entry 2 is 0 and so is 4.
Sketch. The matrix is $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$. Over $\mathbb{Q}$ it reduces to I_3, so $\operatorname{rank} T = 3$ and $\operatorname{nullity} T = 0$. Over $\mathbb{F}_2$, $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ has rank 2, so $\operatorname{nullity} T = 1$, and $\ker T_A = \operatorname{span}((0,1,1))$ gives $p = x + x^2$.

The columns of $A = [T]$ are $T(1) = (1,1,1)$, $T(x) = (0,1,2)$, and $T(x^2) = (0,1,4)$:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}.$$

Over $\mathbb{Q}$, subtracting row 1 from rows 2 and 3 gives rows $(0,1,1)$ and $(0,2,4)$, and subtracting twice the new row 2 from row 3 leaves $(0,0,2)$. There are three pivots, so $\operatorname{rank} A = 3$. By Proposition 4.2.15, $\operatorname{rank} T = 3$, and $\operatorname{nullity} T = 3 - 3 = 0$ by Theorem 4.2.11.

Over $\mathbb{F}_2$, $2 = 4 = 0$, so $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Rows 1 and 3 agree, and subtracting row 1 from rows 2 and 3 gives $A_r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. So $\operatorname{rank} T = 2$ and $\operatorname{nullity} T = 1$. The basic solution is $(0,1,1)$, which is $\left[x + x^2\right]$; indeed $x + x^2$ takes the values 0, $1 + 1 = 0$, and 0 at 0, 1, and $2 = 0$. Over $\mathbb{F}_2$, a nonzero polynomial can vanish at every point of the field.

Solution to Exercise 4.2.22

Uses. Corollary 4.2.21 (Dimension obstructions), Example 4.1.7 (Dimensions of the standard spaces)
Hint 1. Compute the dimensions involved first: $\dim F^{2\times 3} = 6$, $\dim F[x]_{\le 4} = 5$, $\dim F[x]_{\le 3} = 4 = \dim F^{2\times 2}$.
Hint 2. In (c), how small can $\dim \ker T$ be?
Sketch. (a) None: $6 > 5$. (b) $a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mapsto \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix}$. (c) None: $\dim \ker T \ge 5 - 2 = 3$. (d) $T(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4)$.

(a) None. By Example 4.1.7, $\dim F^{2\times 3} = 6 > 5 = \dim F[x]_{\le 4}$, so no linear map between them is injective (Corollary 4.2.21(b)).

(b) $T\left(a_0 + a_1 x + a_2 x^2 + a_3 x^3\right) = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix}$ is linear (coefficients are added and scaled entrywise) and surjective, indeed bijective.

(c) None. By Corollary 4.2.21(a), every linear map $\mathbb{R}^5 \to \mathbb{R}^2$ has $\dim \ker T \ge 5 - 2 = 3$, while $\dim \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2) = 2$.

(d) $T(x_1, x_2, x_3, x_4) = (x_2, x_3, x_4)$ is linear, and $T(\mathbf{x}) = \mathbf{0}$ if and only if $x_2 = x_3 = x_4 = 0$, so $\ker T = \operatorname{span}(\mathbf{e}_1)$. (Rank–nullity allows it: $1 + 3 = 4$.)

Solution to Exercise 4.2.35

Uses. Theorem 4.2.7 (The image of T_A), Remark 3.4.11 (A basis of the kernel), Theorem 4.2.11 (Rank–nullity)
Hint 1. Clear the first column with $-i$ times row 1.
Hint 2. The third row becomes a multiple of the second.
Sketch. $A_r = \begin{pmatrix} 1 & i & 0 & 1 \\ 0 & 0 & 1 & i \\ 0 & 0 & 0 & 0 \end{pmatrix}$, pivot columns $1, 3$. So $(1, i, 0), (0, 1, 1 + i)$ is a basis of $\operatorname{im} T_A$, and the basic solutions $(-i, 1, 0, 0), (-1, 0, -i, 1)$ form a basis of $\ker T_A$; $2 + 2 = 4$.

Adding $-i$ times row 1 to row 2 gives $\left(i - i, -1 - i^2, 1 - 0, 2i - i\right) = (0, 0, 1, i)$. Adding $-(1 + i)$ times this row to row 3 gives $(0, 0, 0, -1 + i - (1 + i)i) = (0, 0, 0, -1 + i - i + 1) = \mathbf{0}$. So

$$A_r = \begin{pmatrix} 1 & i & 0 & 1 \\ 0 & 0 & 1 & i \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

with pivot columns $1, 3$ and free columns $2, 4$. By Theorem 4.2.7, columns 1 and 3 of A, $(1, i, 0)$ and $(0, 1, 1 + i)$, form a basis of $\operatorname{im} T_A$. The basic solutions are $(-i, 1, 0, 0)$ and $(-1, 0, -i, 1)$, a basis of $\ker T_A$ (Remark 3.4.11). Check: $-i\mathbf{a}_1 + \mathbf{a}_2 = (-i + i, 1 - 1, 0) = \mathbf{0}$, and $-\mathbf{a}_1 - i\mathbf{a}_3 + \mathbf{a}_4 = \mathbf{0}$: the first entries give $-1 + 0 + 1 = 0$, the second $-i - i + 2i = 0$, and the third $0 - i(1 + i) + (-1 + i) = 0$. Rank–nullity: $2 + 2 = 4$.

Solution to Exercise 4.2.36

Uses. Theorem 4.2.11 (Rank–nullity), Remark 3.4.11 (A basis of the kernel), Corollary 3.3.19 (Pivot columns, intrinsically)
Hint 1. Over $\mathbb{Q}$, swap rows 1 and 2 to get a pivot in position $(1, 1)$.
Hint 2. Over $\mathbb{F}_3$, reduce the entries first: $-1 = 2$ and $-2 = 1$. Compare the first two columns.

Sketch. Over $\mathbb{Q}$, $M_r = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$: rank 3, kernel spanned by $(-2, 0, -1, 1)$. Over $\mathbb{F}_3$, $M_r = \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$: rank 2, kernel basis $(1, 1, 0, 0), (1, 0, 2, 1)$. An integer relation among the columns stays a relation modulo 3; but $\mathbf{m}_2 - 2\mathbf{m}_1 = (0, 0, 6)$ is zero only modulo 3.

Over $\mathbb{Q}$. Swap rows 1 and 2 and scale the new row 1 by -1, giving $(1, 2, -2, 0)$. Adding 2 times it to row 3 gives $(0, 6, -1, -1)$. Swap this into row 2 and scale by $1/6$; clearing column 2 and then column 3 (with the pivot row $(0, 0, 1, 1)$) gives

$$M_r = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

So $\operatorname{rank} M = 3$, $\dim \ker T_M = 4 - 3 = 1$ by Theorem 4.2.11, and the basic solution $(-2, 0, -1, 1)$ is a basis of $\ker T_M$.

Over $\mathbb{F}_3$. Reducing the entries, $M = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 2 & 1 & 2 & 0 \\ 1 & 2 & 0 & 2 \end{pmatrix}$. Swap rows 1 and 3; adding -2 times the new row 1 to row 2 gives $(0, -3, 2, -4) = (0, 0, 2, 2)$. Scaling by $2^{-1} = 2$ gives $(0, 0, 1, 1)$, and subtracting it from row 3 leaves $\mathbf{0}$:

$$M_r = \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

So $\operatorname{rank} M = 2$ and $\dim \ker T_M = 2$, with basic solutions $(-2, 1, 0, 0) = (1, 1, 0, 0)$ and $(-2, 0, -1, 1) = (1, 0, 2, 1)$.

If $\mathbf{x}$ has integer entries and $M\mathbf{x} = \mathbf{0}$ over $\mathbb{Q}$, then every entry of $M\mathbf{x}$ is the integer 0, so it is 0 modulo 3 as well. That is why $(-2, 0, -1, 1)$, read in $\mathbb{F}_3$, is the second basic solution. The converse fails because a combination of the columns can be a nonzero integer vector divisible by 3: $\mathbf{m}_2 - 2\mathbf{m}_1 = (0, 0, 6)$, so $(-2, 1, 0, 0)$ is in the kernel over $\mathbb{F}_3$ but not over $\mathbb{Q}$. (By Corollary 3.3.19, column 2 is free over $\mathbb{F}_3$ precisely because $\mathbf{m}_2 = 2\mathbf{m}_1$ there.)

Solution to Exercise 4.2.37

Uses. Example 4.1.7 (Dimensions of the standard spaces), Theorem 4.2.11 (Rank–nullity), Theorem 4.2.7 (The image of T_A)
Hint 1. The sets $\ker T_A$ and $\operatorname{im} T_A$ do not depend on which scalars we use. They are complex subspaces.
Hint 2. Example 4.1.7(c): a complex space of complex dimension d has real dimension $2d$.
Sketch. $\operatorname{im} T_A$ is a complex subspace of complex dimension $r = \operatorname{rank} A$, so its real dimension is $2r$. Over $\mathbb{R}$, $\mathbb{C}^n$ has dimension $2n$, so rank–nullity gives $\dim_{\mathbb{R}} \ker T_A = 2n - 2r = 2(n - r) = 2 \dim_{\mathbb{C}} \ker T_A$.

The set $\operatorname{im} T_A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n\}$ is the same whether we regard T_A over $\mathbb{C}$ or over $\mathbb{R}$, and it is a subspace over $\mathbb{C}$ of dimension $r = \operatorname{rank} A$ (Theorem 4.2.7). By Example 4.1.7(c), its dimension over $\mathbb{R}$ is $2r$. So the rank of T_A over $\mathbb{R}$ is $2r$. The domain $\mathbb{C}^n$ has real dimension $2n$, so by Theorem 4.2.11 over $\mathbb{R}$, $\dim_{\mathbb{R}} \ker T_A = 2n - 2r$. Over $\mathbb{C}$, $\dim_{\mathbb{C}} \ker T_A = n - r$. Hence $\dim_{\mathbb{R}} \ker T_A = 2 \dim_{\mathbb{C}} \ker T_A$, which also follows from Example 4.1.7(c) applied to the complex subspace $\ker T_A$.

Solution to Exercise 4.2.38

Uses. Corollary 4.2.21 (Dimension obstructions), Theorem 4.2.11 (Rank–nullity), Example 4.1.7 (Dimensions of the standard spaces)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Items (a), (b), (c), and (e) are about dimensions only. Use rank–nullity.
Hint 2. For (d), look for two maps $\mathbb{R}^2 \to \mathbb{R}^2$ that kill the same line but send the other directions to different lines.
Sketch. (a) True: $\dim \ker T \geq 7 - 4$. (b) True: $\dim W = \operatorname{rank} T \leq 5$. (c) False: $6 > 5$. (d) False: $(x, y) \mapsto (x, 0)$ and $(x, y) \mapsto (0, x)$. (e) True: $\dim V = 2 \dim \ker T$.

(a) True, by Corollary 4.2.21(a): $\dim \ker T \geq 7 - 4 = 3$.

(b) True. $W = \operatorname{im} T$, so $\dim W = \operatorname{rank} T = 5 - \dim \ker T \leq 5$ by Theorem 4.2.11.

(c) False. $\dim F^{2\times 3} = 6 > 5 = \dim F[x]_{\leq 4}$ (Example 4.1.7), so no linear map between them is injective (Corollary 4.2.21(b)).

(d) False. Let $S(x, y) = (x, 0)$ and $T(x, y) = (0, x)$ on $\mathbb{R}^2$. Both have kernel $\operatorname{span}(\mathbf{e}_2)$, but $\operatorname{im} S = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} T = \operatorname{span}(\mathbf{e}_2)$. (They do have images of the same dimension, by rank–nullity.)

(e) True. By Theorem 4.2.11, $\dim V = \dim \ker T + \dim \operatorname{im} T = 2 \dim \ker T$.

Solution to Exercise 4.2.39

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 1.5.24 (Distinct degrees), Theorem 4.1.10 (Independent lists are short), Example 2.1.9 (Maps of polynomials)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Find polynomials with $T(p) = (1, 0)$ and $T(p) = (0, 1)$.
Hint 2. Rank–nullity gives $\dim \ker T$. Two independent vectors in a 2-dimensional space are a basis.
Sketch. $T(1) = (1, 0)$ and $T(x - 1) = (0, 1)$, so T is surjective and $\dim \ker T = 4 - 2 = 2$. Expanding, $(x - 1)^2$ and $(x - 1)^3$ vanish at 1 together with their derivatives, and they have distinct degrees, so they are an independent list of length 2 in $\ker T$.

T is linear, as evaluation and the formal derivative are (Example 2.1.9). Since $T(1) = (1, 0)$ and $T(x - 1) = (0, 1)$, the image contains a basis of F^2, so T is surjective, and $\dim \ker T = 4 - 2 = 2$ by Theorem 4.2.11.

Now $(x - 1)^2 = x^2 - 2x + 1$ has derivative $2x - 2$, and $(x - 1)^3 = x^3 - 3x^2 + 3x - 1$ has derivative $3x^2 - 6x + 3$. Both polynomials and both derivatives vanish at 1, so both polynomials lie in $\ker T$. They are nonzero with degrees 2 and 3, so they are independent (Proposition 1.5.24). An independent list of length $2 = \dim \ker T$ in $\ker T$ is a basis of it (Theorem 4.1.10(b)). Every step is valid over every field.

Solution to Exercise 4.2.40

Uses. Example 2.1.9 (Maps of polynomials), Proposition 1.1.17 (No zero divisors), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Proposition 1.6.7 (Standard bases)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $D\left(\sum_j a_j x^j\right) = \sum_{j\geq 1} j a_j x^{j-1}$. When is $j a_j = 0$ in $\mathbb{F}_p$?
Hint 2. Which monomials can be values of D? Count them and compare with rank D.
Sketch. $j a_j = 0$ iff $p \mid j$ or $a_j = 0$, so $\ker D = \operatorname{span}\left(x^j \mid p \mid j\right)$, of dimension $\lfloor n/p \rfloor + 1$, and $\operatorname{rank} D = n - \lfloor n/p \rfloor$. Every value of D is a combination of the x^{j-1} with $1 \leq j \leq n$ and $p \nmid j$; there are $n - \lfloor n/p \rfloor$ of these, so they span $\operatorname{im} D$. For $p = 3$, $n = 7$: $\ker D = \operatorname{span}\left(1, x^3, x^6\right)$, $\operatorname{im} D = \operatorname{span}\left(1, x, x^3, x^4, x^6\right)$.

For $f = \sum_{j=0}^{n} a_j x^j$, $D(f) = \sum_{j=1}^{n} j a_j x^{j-1}$, where $j a_j$ means $(j \cdot 1) a_j$ in $\mathbb{F}_p$. So $D(f) = \mathbf{0}$ if and only if $j a_j = 0$ for $1 \leq j \leq n$. If $p \mid j$, then $j \cdot 1 = 0$ and this holds for every a_j; if $p \nmid j$, then $j \cdot 1 \neq 0$ and it forces $a_j = 0$ (Proposition 1.1.17). Hence

$$\ker D = \operatorname{span}\left(x^j \mid 0 \leq j \leq n,\ p \mid j\right) = \operatorname{span}\left(1, x^p, x^{2p}, \ldots, x^{\lfloor n/p \rfloor p}\right),$$

and these monomials are independent, being part of the standard basis (Proposition 1.6.7). So $\operatorname{nullity} D = \lfloor n/p \rfloor + 1$, and by Theorem 4.2.11, $\operatorname{rank} D = (n + 1) - (\lfloor n/p \rfloor + 1) = n - \lfloor n/p \rfloor$.

Let $J = \{j \mid 1 \leq j \leq n,\ p \nmid j\}$, which has $n - \lfloor n/p \rfloor$ elements, and $S = \operatorname{span}\left(x^{j-1} \mid j \in J\right)$, of dimension $|J|$. Every value $D(p) = \sum_{j\in J} j a_j x^{j-1}$ lies in S, so $\operatorname{im} D \subseteq S$, and both have dimension $n - \lfloor n/p \rfloor$. By Proposition 4.1.21(b), $\operatorname{im} D = S$: the image is spanned by the monomials x^i with $0 \leq i \leq n - 1$ and $i + 1$ not divisible by p.

For $p = 3$ and $n = 7$: $\ker D = \operatorname{span}\left(1, x^3, x^6\right)$, $\operatorname{nullity} D = 3$, $\operatorname{rank} D = 5$, and $J = \{1, 2, 4, 5, 7\}$, so $\operatorname{im} D = \operatorname{span}\left(1, x, x^3, x^4, x^6\right)$. The monomials x^2 and x^5 are not derivatives, because the only candidates x^3 and x^6 have derivative 0.

Solution to Exercise 4.2.41

Uses. Definition 2.4.21 (Transpose), Exercise 4.1.46, Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)

Techniques. Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. T is linear because transposition is. The dimensions of the symmetric and skew-symmetric matrices are in Exercise 4.1.46.
Hint 2. Show $\operatorname{im} T \subseteq$ (the claimed space) directly, then compare dimensions.
Hint 3. For (c), what is the (i,i) entry of $X - X^{\mathsf{T}}$?
Sketch. (a) $T(X) = 0$ iff $X = X^{\mathsf{T}}$; symmetric matrices have dimension $n(n+1)/2$, so $\operatorname{rank} T = n^2 - n(n+1)/2 = n(n-1)/2$. (b) $\left(X - X^{\mathsf{T}}\right)^{\mathsf{T}} = -\left(X - X^{\mathsf{T}}\right)$, so $\operatorname{im} T$ lies in the skew-symmetric matrices, which have dimension $n(n-1)/2$ when $1+1 \neq 0$. (c) Values of T are symmetric (as $-1 = 1$) with zero diagonal; that space has basis $U_{ij} + U_{ji}$, $i < j$, so dimension $n(n-1)/2$.

Since $(aX+Y)^{\mathsf{T}} = aX^{\mathsf{T}} + Y^{\mathsf{T}}$ entry by entry (Definition 2.4.21), T is linear.

(a) $T(X) = 0$ if and only if $X = X^{\mathsf{T}}$, so $\ker T$ is the space S of symmetric matrices, and $\dim S = n(n+1)/2$ by Exercise 4.1.46(a). By Theorem 4.2.11, $\operatorname{rank} T = n^2 - n(n+1)/2 = n(n-1)/2$.

(b) Since $\left(X^{\mathsf{T}}\right)^{\mathsf{T}} = X$, we have $\left(X - X^{\mathsf{T}}\right)^{\mathsf{T}} = X^{\mathsf{T}} - X = -\left(X - X^{\mathsf{T}}\right)$, so every value of T is skew-symmetric: $\operatorname{im} T \subseteq K$. If $1+1 \neq 0$, then $\dim K = n(n-1)/2 = \operatorname{rank} T$ (Exercise 4.1.46(b)), so $\operatorname{im} T = K$ by Proposition 4.1.21(b).

(c) Suppose $1+1 = 0$, so $-1 = 1$. By the computation in (b), every value of T satisfies $Y^{\mathsf{T}} = -Y = Y$, so it is symmetric, and its (i,i) entry is $x_{ii} - x_{ii} = 0$. So $\operatorname{im} T \subseteq Z$, the symmetric matrices with zero diagonal. A matrix in Z equals $\sum_{i<j} y_{ij}\left(U_{ij} + U_{ji}\right)$, and these $n(n-1)/2$ matrices are independent (the (i,j) entry of a combination, for $i < j$, is the coefficient of $U_{ij} + U_{ji}$). So $\dim Z = n(n-1)/2 = \operatorname{rank} T$, and $\operatorname{im} T = Z$. Since U_{11} is symmetric but not in Z, the image is a proper subspace of the symmetric matrices.

Solution to Exercise 4.2.42

Uses. Theorem 4.2.11 (Rank–nullity), Example 4.2.25 (Kernels of linear functionals), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.1.10 (Independent lists are short), Lemma 2.4.24 (Matrix units)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Every value $(\mathbf{r}, \mathbf{c})$ of Φ satisfies $\sum_i r_i = \sum_j c_j$. Why? These vectors form the kernel of a nonzero linear functional on F^{2n}.
Hint 2. To see that every such $(\mathbf{r}, \mathbf{c})$ is a value, try $X = \sum_i r_i U_{i1} + \sum_j c_j U_{1j} - sU_{11}$ with $s = \sum_i r_i$.
Hint 3. For (b), there are $(n-1)^2$ matrices; check that they lie in Z_n and are independent.
Sketch. (a) $Z_n = \ker \Phi$. Values of Φ lie in $H = \left\{(\mathbf{r}, \mathbf{c}) \mid \sum r_i = \sum c_j\right\}$, of dimension $2n-1$, and the matrix in the hint shows every element of H is a value. So $\operatorname{rank} \Phi = 2n-1$ and $\dim Z_n = n^2 - 2n + 1$. (b) Each listed matrix has row and column sums 0, and the (i,j) entry of a combination, for $i, j < n$, is the coefficient of the (i,j)th matrix; an independent list of length $\dim Z_n$ is a basis.

(a) Φ is linear, since row and column sums are, and $Z_n = \ker \Phi$. Let $H = \left\{(\mathbf{r}, \mathbf{c}) \in F^{2n} \;\middle|\; r_1 + \cdots + r_n = c_1 + \cdots + c_n\right\}$. Every value of Φ lies in H, since both sums equal the sum of all entries of X. H is the kernel of the linear functional $(\mathbf{r}, \mathbf{c}) \mapsto \sum_i r_i - \sum_j c_j$, which is not zero (it sends $\mathbf{e}_1$ to 1), so $\dim H = 2n-1$ (Example 4.2.25).

Conversely, let $(\mathbf{r}, \mathbf{c}) \in H$ with $s = \sum_i r_i = \sum_j c_j$, and let

$$X = \sum_{i=1}^{n} r_i U_{i1} + \sum_{j=1}^{n} c_j U_{1j} - sU_{11}.$$

For $i \geq 2$, the only entry of row i of X that can be nonzero is r_i; row 1 has sum $r_1 + \sum_j c_j - s = r_1$. For $j \geq 2$, the only entry of column j that can be nonzero is c_j; column 1 has sum $\sum_i r_i + c_1 - s = c_1$. So $\Phi(X) = (\mathbf{r}, \mathbf{c})$, and $\operatorname{im} \Phi = H$. By Theorem 4.2.11,

$$\dim Z_n = n^2 - (2n-1) = (n-1)^2.$$

(b) Let $V_{ij} = U_{ij} - U_{in} - U_{nj} + U_{nn}$ for $i, j \leq n-1$. Row i and row n of V_{ij} contain one 1 and one -1, and so do columns j and n; all other rows and columns are zero. So $V_{ij} \in Z_n$. In a combination $\sum_{i,j<n} c_{ij} V_{ij}$, the only term with a nonzero (i,j) entry for $i, j < n$ is $c_{ij} V_{ij}$, so the combination is 0 only if all $c_{ij} = 0$. These $(n-1)^2$ independent matrices form a basis of Z_n by (a) and Theorem 4.1.10(b).

Solution to Exercise 4.2.43

Uses. Definition 2.4.1 (Matrix product), Theorem 4.2.11 (Rank–nullity), Proposition 2.4.14 (Algebra of matrices), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.6.2 (Bases are independent spanning lists)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. Column j of AX is A times column j of X. So when is $AX = 0$?
Hint 2. If $\mathbf{k}_1, \dots, \mathbf{k}_d$ is a basis of $\ker T_A$, let K_{ij} be the $n \times p$ matrix whose column j is $\mathbf{k}_i$ and whose other columns are $\mathbf{0}$. Show these dp matrices form a basis of $\ker L_A$.
Sketch. $AX = 0$ iff every column of X lies in $\ker T_A$. With a basis $\mathbf{k}_1, \dots, \mathbf{k}_d$ of $\ker T_A$, $d = n - \operatorname{rank} A$, the matrices K_{ij} (column j equal to $\mathbf{k}_i$, others zero) form a basis of $\ker L_A$: expand each column, and compare columns for independence. So $\dim \ker L_A = p(n - \operatorname{rank} A)$, and rank–nullity gives $\operatorname{rank} L_A = np - p(n - \operatorname{rank} A) = p \operatorname{rank} A$.

L_A is linear by Proposition 2.4.14. By Definition 2.4.1, column j of AX is $A\mathbf{x}_j$, where $\mathbf{x}_j$ is column j of X. So $AX = 0$ if and only if $\mathbf{x}_j \in \ker T_A$ for every j.

Let $d = \dim \ker T_A = n - \operatorname{rank} A$ (Theorem 4.2.11) and let $\mathbf{k}_1, \dots, \mathbf{k}_d$ be a basis of $\ker T_A$. For $1 \le i \le d$ and $1 \le j \le p$, let $K_{ij} \in F^{n \times p}$ have column j equal to $\mathbf{k}_i$ and all other columns zero. Each K_{ij} lies in $\ker L_A$. If $X \in \ker L_A$, write $\mathbf{x}_j = \sum_i c_{ij}\mathbf{k}_i$; then X and $\sum_{i,j} c_{ij}K_{ij}$ have the same column j for every j, so they are equal, and the K_{ij} span $\ker L_A$. If $\sum_{i,j} c_{ij}K_{ij} = 0$, then column j gives $\sum_i c_{ij}\mathbf{k}_i = \mathbf{0}$, so $c_{ij} = 0$ for all i (Theorem 1.5.5). So the K_{ij} form a basis of $\ker L_A$, and $\dim \ker L_A = dp$. By Theorem 4.2.11,

$$\operatorname{rank} L_A = np - p(n - \operatorname{rank} A) = p \cdot \operatorname{rank} A.$$

Solution to Exercise 4.2.44

Uses. Proposition 4.2.3 (Properties of restriction), Theorem 4.2.11 (Rank–nullity), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. (a) is rank–nullity for $T{\restriction_U}$.
Hint 2. (b) Take a basis $\mathbf{w}_1, \dots, \mathbf{w}_r$ of $\operatorname{im} T$, pick $\mathbf{v}_i$ with $T(\mathbf{v}_i) = \mathbf{w}_i$, and let $U = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_r)$.
Sketch. (a) $\ker(T{\restriction_U}) = U \cap \ker T$; apply rank–nullity to $T{\restriction_U}$. (b) With U as in the hint, $\operatorname{im}(T{\restriction_U}) = \operatorname{span}(\mathbf{w}_i) = \operatorname{im} T$, and $\sum c_i\mathbf{v}_i \in \ker T$ gives $\sum c_i\mathbf{w}_i = \mathbf{0}$, so all $c_i = 0$. For any such U, (a) gives $\dim U = \operatorname{rank} T$. (c) $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$ and $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$.

(a) By Proposition 4.2.3(c), $T{\restriction_U} \in \mathrm{L}(U, W)$ has kernel $U \cap \ker T$. By Theorem 4.2.11 applied to it, $\dim U = \dim(U \cap \ker T) + \dim \operatorname{im}(T{\restriction_U})$.

(b) By Theorem 4.2.11, $\operatorname{im} T$ is finite-dimensional; let $\mathbf{w}_1, \dots, \mathbf{w}_r$ be a basis, $r = \operatorname{rank} T$, and choose $\mathbf{v}_i \in V$ with $T(\mathbf{v}_i) = \mathbf{w}_i$. Let $U = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_r)$. By Proposition 2.1.16(a), $\operatorname{im}(T{\restriction_U}) = \operatorname{span}(\mathbf{w}_1, \dots, \mathbf{w}_r) = \operatorname{im} T$. If $\mathbf{u} = \sum_i c_i\mathbf{v}_i \in \ker T$, then $\mathbf{0} = T(\mathbf{u}) = \sum_i c_i\mathbf{w}_i$, so all $c_i = 0$ (Theorem 1.5.5) and $\mathbf{u} = \mathbf{0}$. So $U \cap \ker T = \{\mathbf{0}\}$.

If U is any subspace with these two properties, it is finite-dimensional because V is, and (a) gives $\dim U = \dim \operatorname{im} T + 0 = \operatorname{rank} T$.

(c) $T(x, y, z) = \mathbf{0}$ means $x = -y$ and $z = -y$, so $\ker T = \operatorname{span}((1, -1, 1))$, and $\operatorname{rank} T = 2$, so $\operatorname{im} T = \mathbb{R}^2$. For $U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$: $T(\mathbf{e}_1) = (1, 0)$ and $T(\mathbf{e}_3) = (0, 1)$ span $\mathbb{R}^2$, and a vector $(a, 0, c)$ is a multiple $t(1, -1, 1)$ only if $t = 0$. For $U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$: $T(\mathbf{e}_1) = (1, 0)$ and $T(\mathbf{e}_2) = (1, 1)$ span $\mathbb{R}^2$, and $(a, b, 0) = t(1, -1, 1)$ forces $t = 0$. So U is not determined by T; only its dimension is.

Solution to Exercise 4.2.45

Uses. Theorem 4.2.11 (Rank–nullity), Theorem 4.1.10 (Independent lists are short), Corollary 1.5.17 (Sublists), Corollary 4.1.11 (Spanning lists are long)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Theorem 4.1.10(b) applies to independent lists *in* $\ker T$. Is each $\mathbf{v}_i$ with $i \le k$ in $\ker T$?
Hint 2. Try $T(x, y) = x$ on $\mathbb{R}^2$ with the basis $(1, 1), (1, 0)$.
Hint 3. To repair, subtract from each $\mathbf{v}_i$ ($i \le k$) a combination of $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ with the same image under T.
Sketch. The proof never checks that $\mathbf{v}_1, \dots, \mathbf{v}_k$ lie in $\ker T$. Counterexample: $T(x, y) = x$, $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (1, 0)$; $T(\mathbf{v}_2) = 1$ is a basis of $\operatorname{im} T = \mathbb{R}$, but $\mathbf{v}_1 \notin \ker T$. Repair: write $T(\mathbf{v}_i) = \sum_{j>k} c_{ij}T(\mathbf{v}_j)$ and set $\mathbf{v}_{i'} = \mathbf{v}_i - \sum_{j>k} c_{ij}\mathbf{v}_j$; then $\mathbf{v}_{1'}, \dots, \mathbf{v}_{k'}$ is a basis of $\ker T$.

The error. Theorem 4.1.10(b), applied in the vector space $\ker T$, needs an independent list of vectors *of* $\ker T$. The proof never shows that $\mathbf{v}_1, \dots, \mathbf{v}_k$ lie in $\ker T$, and they need not.

Counterexample. Let $T : \mathbb{R}^2 \to \mathbb{R}$, $T(x, y) = x$, with the basis $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (1, 0)$ of $\mathbb{R}^2$ and $k = 1$. Then $T(\mathbf{v}_2) = 1$ is a basis of $\operatorname{im} T = \mathbb{R}$, but $T(\mathbf{v}_1) = 1 \neq 0$, so $\mathbf{v}_1$ is not in $\ker T = \operatorname{span}(\mathbf{e}_2)$.

Repair. The claim holds if we also assume $\mathbf{v}_1, \dots, \mathbf{v}_k \in \ker T$; then the proof above is correct. Without that assumption we can modify the $\mathbf{v}_i$. For $i \leq k$, $T(\mathbf{v}_i) \in \operatorname{im} T$, so $T(\mathbf{v}_i) = \sum_{j>k} c_{ij} T(\mathbf{v}_j)$ for some scalars c_{ij}. Let $\mathbf{v}_{i'} = \mathbf{v}_i - \sum_{j>k} c_{ij} \mathbf{v}_j$. Then $T(\mathbf{v}_{i'}) = \mathbf{0}$, and the list $\mathbf{v}_{1'}, \dots, \mathbf{v}_{k'}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ spans V (each $\mathbf{v}_i$ is recovered by adding back a combination of the later vectors) and has length n, so it is a basis (Corollary 4.1.11(b)). Its sublist $\mathbf{v}_{1'}, \dots, \mathbf{v}_{k'}$ is independent (Corollary 1.5.17), lies in $\ker T$, and has length $k = \dim \ker T$, so it is a basis of $\ker T$. In the counterexample, $\mathbf{v}_{1'} = (1, 1) - (1, 0) = (0, 1)$.

Solution to Exercise 4.2.46

> *Uses.* Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations)
> *Techniques.* Dimension counting (Technique 4.2.24)
> *Hint 1.* $T^2 = 0$ says that T kills every value of T. Compare $\operatorname{im} T$ and $\ker T$.
> *Hint 2.* For the construction, send $\mathbf{e}_{k+i} \mapsto \mathbf{e}_i$ for $1 \leq i \leq k$ and every other standard basis vector to $\mathbf{0}$.
> *Sketch.* $T^2 = 0$ iff $\operatorname{im} T \subseteq \ker T$, which forces $k \leq n - k$. Conversely, if $2k \leq n$, the map with $T(\mathbf{e}_{k+i}) = \mathbf{e}_i$ $(i \leq k)$ and $T(\mathbf{e}_j) = \mathbf{0}$ otherwise has image $\operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$, which it kills.

Only if. If $T^2 = 0$, then $T(T(\mathbf{v})) = \mathbf{0}$ for all $\mathbf{v}$, so $\operatorname{im} T \subseteq \ker T$. By Proposition 4.1.21(a) and Theorem 4.2.11, $k = \dim \operatorname{im} T \leq \dim \ker T = n - k$, so $2k \leq n$.

If. Suppose $2k \leq n$. By Theorem 2.2.1 there is $T \in \mathrm{L}(F^n)$ with $T(\mathbf{e}_{k+i}) = \mathbf{e}_i$ for $1 \leq i \leq k$ (these indices are at most $2k \leq n$) and $T(\mathbf{e}_j) = \mathbf{0}$ for every other j. By Proposition 2.1.16(a), $\operatorname{im} T = \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$, so $\operatorname{rank} T = k$. For $i \leq k$, the index i is not of the form $k + i'$ with $i' \geq 1$, so $T(\mathbf{e}_i) = \mathbf{0}$. Hence $T^2(\mathbf{e}_j) = T(T(\mathbf{e}_j)) = \mathbf{0}$ for every j, since $T(\mathbf{e}_j)$ is $\mathbf{0}$ or some $\mathbf{e}_i$ with $i \leq k$. A linear map that kills a basis is 0, so $T^2 = 0$.

Solution to Exercise 4.2.47

> *Uses.* Corollary 4.2.18 (Injective if and only if surjective), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Example 2.1.9 (Maps of polynomials)
> *Techniques.* Uniqueness gives existence (Technique 3.4.21), Dimension counting (Technique 4.2.24)
> *Hint 1.* Both spaces have dimension 4, so it suffices to show $\ker T = \{\mathbf{0}\}$.
> *Hint 2.* Write $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ and compute the four entries of $T(p)$.
> *Sketch.* $T(p) = (a_0, a_1, a_0 + a_1 + a_2 + a_3, a_1 + 2a_2 + 3a_3)$. If this is $\mathbf{0}$, then $a_0 = a_1 = 0$, $a_2 + a_3 = 0$, and $2a_2 + 3a_3 = 0$; subtracting twice the third equation from the fourth gives $a_3 = 0$, then $a_2 = 0$. So T is injective, hence bijective. (b) Solving $a_2 + a_3 = 1$, $2a_2 + 3a_3 = 0$ gives $p = 3x^2 - 2x^3$.

(a) T is linear since evaluation and the formal derivative are (Example 2.1.9). For $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, $p' = a_1 + 2a_2 x + 3a_3 x^2$, so

$$T(p) = (a_0, a_1, a_0 + a_1 + a_2 + a_3, a_1 + 2a_2 + 3a_3).$$

If $T(p) = \mathbf{0}$, then $a_0 = a_1 = 0$, $a_2 + a_3 = 0$, and $2a_2 + 3a_3 = 0$. Subtracting twice the third equation from the fourth gives $a_3 = 0$, and then $a_2 = 0$. So $\ker T = \{\mathbf{0}\}$, and T is injective (Theorem 2.6.11). Both spaces have dimension 4, so T is bijective by Corollary 4.2.18. No step divided by anything, so this holds over every field.

(b) We need $a_0 = a_1 = 0$, $a_2 + a_3 = 1$, and $2a_2 + 3a_3 = 0$. Subtracting twice the first of the last two from the second gives $a_3 = -2$, so $a_2 = 3$, and $p = 3x^2 - 2x^3$. Check: $p(1) = 1$, $p' = 6x - 6x^2$ vanishes at 0 and 1.

Solution to Exercise 4.2.48

> *Uses.* Example 4.2.28 (Sums of powers)
> *Hint 1.* By Example 4.2.28 such a p exists. Try $p = ax^4 + bx^3 + cx^2 + dx$, and compute $\Delta(x^4) = 4x^3 + 6x^2 + 4x + 1$.
> *Hint 2.* Telescoping gives $\sum_{t=0}^{N-1} t^3 = p(N) - p(0)$. Then replace N by $N + 1$.
> *Sketch.* $\Delta p = 4ax^3 + (6a + 3b)x^2 + (4a + 3b + 2c)x + (a + b + c + d)$, so $a = 1/4$, $b = -1/2$, $c = 1/4$, $d = 0$, and $p = (x(x-1)/2)^2$. Then $\sum_{t=0}^{N-1} t^3 = (N(N-1)/2)^2$; replace N by $N + 1$.

Using $\Delta(x^4) = 4x^3 + 6x^2 + 4x + 1$, $\Delta(x^3) = 3x^2 + 3x + 1$, $\Delta(x^2) = 2x + 1$, and $\Delta(x) = 1$, for $p = ax^4 + bx^3 + cx^2 + dx$ (a constant term may be dropped, since constants lie in $\ker \Delta$) we get

$$\Delta p = 4ax^3 + (6a + 3b)x^2 + (4a + 3b + 2c)x + (a + b + c + d).$$

Setting this equal to x^3: $a = 1/4$; $6a + 3b = 0$ gives $b = -1/2$; $4a + 3b + 2c = 0$ gives $c = 1/4$; and $a + b + c + d = 0$ gives $d = 0$. So

$$p = \frac{1}{4}x^4 - \frac{1}{2}x^3 + \frac{1}{4}x^2 = \frac{1}{4}x^2(x-1)^2 = \left(\frac{x(x-1)}{2}\right)^2.$$

By the telescoping identity of Example 4.2.28, $0^3 + 1^3 + \cdots + (N-1)^3 = p(N) - p(0) = (N(N-1)/2)^2$ for every $N \geq 1$. Applying this with $N + 1$ in place of N gives $1^3 + \cdots + N^3 = (N(N+1)/2)^2$.

Solution to Exercise 4.2.49

Uses. Proposition 4.2.26 (Subspaces that must meet), Proposition 1.4.16 (Intersections of subspaces), Example 4.2.25 (Kernels of linear functionals)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Induct on k, using Proposition 4.2.26 for the intersection of $U_1 \cap \cdots \cap U_k$ with U_{k+1}.
Hint 2. For (c), in F^n, use the coordinate hyperplanes $x_i = 0$.
Sketch. (a) For $k = 1$ it is trivial; if $X = U_1 \cap \cdots \cap U_k$ satisfies the bound, then $\dim(X \cap U_{k+1}) \geq \dim X + \dim U_{k+1} - n$, which is the bound for $k + 1$. (b) The bound is $(n-1)^2 - (n-2)n = 1$. (c) $U_i = \{\mathbf{x} \in F^n \mid x_i = 0\}$ has dimension $n - 1$, and their intersection is $\{\mathbf{0}\}$.

(a) By induction on k. For $k = 1$ the claim is $\dim U_1 \geq \dim U_1$. Suppose it holds for k, and let $X = U_1 \cap \cdots \cap U_k$, a subspace (Proposition 1.4.16). By Proposition 4.2.26 and the inductive hypothesis,

$$\dim(X \cap U_{k+1}) \geq \dim X + \dim U_{k+1} - n \geq \sum_{i=1}^{k+1} \dim U_i - (k-1)n - n,$$

which is the claim for $k + 1$.

(b) With $k = n - 1$ subspaces of dimension $n - 1$, the bound is $(n-1)(n-1) - (n-2)n = n^2 - 2n + 1 - n^2 + 2n = 1$. So the intersection has dimension at least 1 and contains a nonzero vector.

(c) Choose a basis and identify V with F^n by coordinates. The subspace $U_i = \{\mathbf{x} \in F^n \mid x_i = 0\}$ is the kernel of the nonzero functional $\mathbf{x} \mapsto x_i$, so it has dimension $n - 1$ (Example 4.2.25). A vector in all n of them has every coordinate 0, so $U_1 \cap \cdots \cap U_n = \{\mathbf{0}\}$.

Solution to Exercise 4.2.50

Uses. Exercise 2.6.39, Proposition 2.6.15 (Kernels and images of composites), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24), Proving TFAE (Technique 2.7.1)
Hint 1. Exercise 2.6.39(c) already gives (i) ⇔ (ii) and (iv) ⇔ (v), with no dimension needed.
Hint 2. $\ker T \subseteq \ker T^2$ and $\operatorname{im} T^2 \subseteq \operatorname{im} T$ always. Containment plus equal dimension gives equality.
Hint 3. For the example, try differentiation on $\mathbb{R}[x]$.
Sketch. (i) ⇔ (ii) and (iv) ⇔ (v) are Exercise 2.6.39(c). Since $\ker T \subseteq \ker T^2$, (ii) holds iff the dimensions agree, iff $\operatorname{rank} T^2 = \operatorname{rank} T$ by rank–nullity for T and T^2. Since $\operatorname{im} T^2 \subseteq \operatorname{im} T$, (iv) holds iff the ranks agree. On $\mathbb{R}[x]$, D is surjective, so $\operatorname{im} D^2 = \operatorname{im} D$, but $\ker D \cap \operatorname{im} D$ is the constants.

By Exercise 2.6.39(c), (i) ⇔ (ii) and (iv) ⇔ (v). It remains to connect (ii), (iii), and (iv).

(ii) ⇔ (iii). By Proposition 2.6.15(a), $\ker T \subseteq \ker T^2$. By Proposition 4.1.21(b), applied in $\ker T^2$, they are equal if and only if $\dim \ker T = \dim \ker T^2$. By Theorem 4.2.11 for T and for T^2, $\dim \ker T = \dim V - \operatorname{rank} T$ and $\dim \ker T^2 = \dim V - \operatorname{rank} T^2$. So (ii) holds if and only if $\operatorname{rank} T = \operatorname{rank} T^2$.

(iii) ⇔ (iv). By Proposition 2.6.15(b), $\operatorname{im} T^2 \subseteq \operatorname{im} T$, so by Proposition 4.1.21(b) they are equal if and only if $\operatorname{rank} T^2 = \operatorname{rank} T$.

Infinite dimensions. Let D be differentiation on $\mathbb{R}[x]$. Every polynomial is a derivative, so $\operatorname{im} D = \mathbb{R}[x]$ and $\operatorname{im} D^2 = D(\operatorname{im} D) = \operatorname{im} D$, which is (iv). But $\ker D$ is the constants, which lie in $\operatorname{im} D$, so $\ker D \cap \operatorname{im} D \neq \{\mathbf{0}\}$ and (i) fails. The dimension count in (ii) ⇔ (iii) ⇔ (iv) has nothing to count here.

Solution to Exercise 4.2.51

Uses. Proposition 2.6.15 (Kernels and images of composites), Proposition 4.1.21 (Dimensions of subspaces), Exercise 2.6.39, Theorem 4.2.11 (Rank–nullity), Theorem 2.4.2 (The matrix of a composite)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Dimension counting (Technique 4.2.24)
Hint 1. $\ker T^0 \subseteq \ker T^1 \subseteq \cdots$. If every inclusion up to $\ker T^{n+1}$ were strict, how large would $\dim \ker T^{n+1}$ be?
Hint 2. For (b), if $T^{m+2}\mathbf{v} = \mathbf{0}$, then $T\mathbf{v} \in \ker T^{m+1}$. Induct on m.
Hint 3. For (c), apply Exercise 2.6.39(c) to $S = T^n$.
Sketch. (a) A strict inclusion of subspaces raises the dimension by at least 1, so if all inclusions $\ker T^k \subseteq \ker T^{k+1}$ for $k = 0, \ldots, n$ were strict, $\dim \ker T^{n+1} \geq n+1$. (b) If $\ker T^{m+1} = \ker T^m$ and $T^{m+2}\mathbf{v} = \mathbf{0}$, then $T\mathbf{v} \in \ker T^{m+1} = \ker T^m$, so $\mathbf{v} \in \ker T^{m+1}$; induct. Since $k \leq n \leq 2n$, $\ker T^n = \ker T^k = \ker T^{2n}$. (c) With $S = T^n$, $\ker S^2 = \ker S$, so $\ker S \cap \operatorname{im} S = \{\mathbf{0}\}$; rank–nullity gives the dimensions.

For all $j \geq 0$, $\ker T^j \subseteq \ker T^{j+1}$, since $T^{j+1} = T \circ T^j$ (Proposition 2.6.15(a)).

(a) Suppose, to the contrary, that $\ker T^k \neq \ker T^{k+1}$ for every $k = 0, 1, \ldots, n$. A proper subspace has strictly smaller dimension (Proposition 4.1.21(b)), so $\dim \ker T^{k+1} \geq \dim \ker T^k + 1$ for each such k, and since $\dim \ker T^0 = \dim \ker \mathrm{id}_V = 0$, induction gives $\dim \ker T^{n+1} \geq n+1$. This contradicts $\dim \ker T^{n+1} \leq n$ (Proposition 4.1.21(a)).

(b) We show by induction on $m \geq k$ that $\ker T^{m+1} = \ker T^m$; the case $m = k$ is the hypothesis. Suppose $\ker T^{m+1} = \ker T^m$, and let $\mathbf{v} \in \ker T^{m+2}$. Then $T^{m+1}(T\mathbf{v}) = \mathbf{0}$, so $T\mathbf{v} \in \ker T^{m+1} = \ker T^m$, that is, $T^{m+1}\mathbf{v} = \mathbf{0}$. So $\ker T^{m+2} \subseteq \ker T^{m+1}$, and the other containment always holds. Chaining the equalities, $\ker T^m = \ker T^k$ for all $m \geq k$. With $k \leq n$ from (a), $\ker T^n = \ker T^k = \ker T^{2n}$.

(c) Let $S = T^n$. By (b), $\ker S^2 = \ker T^{2n} = \ker T^n = \ker S$, so $\ker S \cap \operatorname{im} S = \{\mathbf{0}\}$ by Exercise 2.6.39(c). By Theorem 4.2.11, $\dim \ker S + \dim \operatorname{im} S = n$. For $A \in F^{n\times n}$, apply this to $T = T_A$, using $(T_A)^n = T_{A^n}$ (Theorem 2.4.2).

With the direct sums of Section 4.4, (c) says $V = \ker T^n \oplus \operatorname{im} T^n$. This is Fitting's decomposition, the starting point of Section 10.1.

Solution to Exercise 4.2.52

Uses. Proposition 4.2.3 (Properties of restriction), Theorem 4.2.11 (Rank–nullity), Proposition 2.6.15 (Kernels and images of composites), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Restrict to a subspace (Technique 4.2.5), Dimension counting (Technique 4.2.24)
Hint 1. Let $S = T^{k-1}|_{\ker T^k}$. Its kernel is $\ker T^{k-1}$. Show that its image is $\operatorname{im} T^{k-1} \cap \ker T$.
Hint 2. For (b), $\operatorname{im} T^k \subseteq \operatorname{im} T^{k-1}$.
Hint 3. For (c), $n = \dim \ker T^m = d_1 + \cdots + d_m$.
Sketch. (a) $S = T^{k-1}|_{\ker T^k}$ has kernel $\ker T^k \cap \ker T^{k-1} = \ker T^{k-1}$ and image $\operatorname{im} T^{k-1} \cap \ker T$ (a vector $T^{k-1}\mathbf{v}$ lies in $\ker T$ iff $\mathbf{v} \in \ker T^k$). Rank–nullity for S is the claim. (b) The subspaces $\operatorname{im} T^{k-1} \cap \ker T$ shrink as k grows. (c) $n = d_1 + \cdots + d_m \leq m d_1 = m \dim \ker T$.

(a) Let $S = T^{k-1}|_{\ker T^k} : \ker T^k \to V$, which is linear (Proposition 4.2.3(c)). Its kernel is $\ker T^k \cap \ker T^{k-1} = \ker T^{k-1}$, since $\ker T^{k-1} \subseteq \ker T^k$. Its image is $\operatorname{im} T^{k-1} \cap \ker T$: if $\mathbf{v} \in \ker T^k$, then $T^{k-1}\mathbf{v} \in \operatorname{im} T^{k-1}$ and $T\big(T^{k-1}\mathbf{v}\big) = T^k\mathbf{v} = \mathbf{0}$; conversely, if $\mathbf{w} = T^{k-1}\mathbf{v}$ lies in $\ker T$, then $T^k\mathbf{v} = T\mathbf{w} = \mathbf{0}$, so $\mathbf{v} \in \ker T^k$ and $\mathbf{w} = S(\mathbf{v})$. By Theorem 4.2.11 for S,

$$\dim \ker T^k = \dim \ker T^{k-1} + \dim\big(\operatorname{im} T^{k-1} \cap \ker T\big).$$

(b) Since $T^k = T^{k-1} \circ T$, $\operatorname{im} T^k \subseteq \operatorname{im} T^{k-1}$ (Proposition 2.6.15(b)). So $\operatorname{im} T^k \cap \ker T \subseteq \operatorname{im} T^{k-1} \cap \ker T$, and $d_{k+1} \leq d_k$ by (a) and Proposition 4.1.21(a).

(c) If $T^m = 0$, then $\ker T^m = V$, and $\ker T^0 = \{\mathbf{0}\}$, so the sum telescopes:

$$n = \dim \ker T^m - \dim \ker T^0 = d_1 + d_2 + \cdots + d_m \leq m d_1 = m \dim \ker T,$$

using (b). Hence $\dim \ker T \geq n/m$, and $\operatorname{rank} T = n - \dim \ker T \leq n - n/m = (m-1)n/m$.

For $m = 2$ this is the bound $\operatorname{rank} T \leq n/2$ of Exercise 4.2.46. For the shift N on F^4 of Section 4.1.7, $d_1 = d_2 = d_3 = d_4 = 1$; how the d_k can decrease is what Section 10.3 classifies.

Solution to Exercise 4.2.53

Uses. Proposition 4.2.3 (Properties of restriction), Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.1.21 (Dimensions of subspaces), Theorem 3.5.12 (First Fundamental Theorem, completed), Theorem 3.5.14 (Third Fundamental Theorem, completed), Remark 3.5.7 (Inverses of T_A and inverses of A), Proposition 2.4.14 (Algebra of matrices)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. If $(I_m - AB)\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = AB\mathbf{x}$. Compute $(I_n - BA)(B\mathbf{x})$.
Hint 2. If also $B\mathbf{x} = \mathbf{0}$, what is $\mathbf{x}$?
Hint 3. For (b), swap the roles of A and B.
Sketch. If $\mathbf{x} = AB\mathbf{x}$, then $(I_n - BA)B\mathbf{x} = B\mathbf{x} - B(AB\mathbf{x}) = \mathbf{0}$, and $B\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = A(B\mathbf{x}) = \mathbf{0}$. So $T_B\restriction_{\ker(I_m - AB)}$ is injective into $\ker(I_n - BA)$, giving $\leq$; by symmetry, equality. A square matrix is invertible iff its kernel is $\{\mathbf{0}\}$.

(a) Let $\mathbf{x} \in \ker T_{I_m - AB}$, so $\mathbf{x} = AB\mathbf{x}$. Then

$$(I_n - BA)(B\mathbf{x}) = B\mathbf{x} - B(AB\mathbf{x}) = B\mathbf{x} - B\mathbf{x} = \mathbf{0},$$

so $B\mathbf{x} \in \ker T_{I_n - BA}$. If $B\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = A(B\mathbf{x}) = \mathbf{0}$. So the restriction of T_B to $\ker T_{I_m - AB}$ is linear with kernel $\{\mathbf{0}\}$ (Proposition 4.2.3(c)), hence injective, with image inside $\ker T_{I_n - BA}$.

(b) By (a) and Proposition 4.2.6(b), $\dim\ker(I_m - AB) = \dim\operatorname{im}\left(T_B\restriction_{\ker(I_m - AB)}\right) \leq \dim\ker(I_n - BA)$, using Proposition 4.1.21(a). Applying (a) with the roles of A and B exchanged (and m, n exchanged) gives the reverse inequality.

(c) For a square matrix $C \in F^{k\times k}$, $\ker T_C = \{\mathbf{0}\}$ if and only if $\operatorname{rank} C = k$ (Theorem 3.5.12), if and only if T_C is invertible (Theorem 3.5.14), if and only if C is invertible (Remark 3.5.7). By (b), $\ker(I_m - AB) = \{\mathbf{0}\}$ if and only if $\ker(I_n - BA) = \{\mathbf{0}\}$.

Replacing A by $\lambda^{-1}A$ for a scalar $\lambda \neq 0$ shows $\dim\ker(\lambda I_m - AB) = \dim\ker(\lambda I_n - BA)$: in the language of Section 7.2, AB and BA have the same nonzero eigenvalues, with eigenspaces of the same dimensions, even when $m \neq n$.

Solution to Exercise 4.2.54

Uses. Corollary 4.2.21 (Dimension obstructions), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$), Definition 2.5.12 (Invertible)
Hint 1. Find an injective linear map from S into F^n. What does a nonzero invertible matrix do to $\mathbf{e}_1$?
Hint 2. For (b), multiply $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ by $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.
Hint 3. For (c), you need a matrix M such that M and $I_2 + M$ are both invertible. Recall Exercise 1.1.56.
Sketch. (a) $X \mapsto X\mathbf{e}_1$ is linear $S \to F^n$, and injective: if $X\mathbf{e}_1 = \mathbf{0}$ with $X \neq 0$, then X is invertible and $\mathbf{e}_1 = X^{-1}\mathbf{0} = \mathbf{0}$. So $\dim S \leq n$. (b) $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = (a^2 + b^2)I_2$. (c) $\operatorname{span}(I_2, M)$ with $M = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$: $M(M + I_2) = I_2$.

(a) The map $\varphi : S \to F^n$, $\varphi(X) = X\mathbf{e}_1$, is linear. If $\varphi(X) = \mathbf{0}$ and $X \neq 0$, then X is invertible, so $\mathbf{e}_1 = X^{-1}(X\mathbf{e}_1) = \mathbf{0}$, which is false. So $\ker\varphi = \{0\}$, φ is injective (Theorem 2.6.11), and $\dim S \leq n$ by Corollary 4.2.21(b).

(b) $S = \operatorname{span}(I_2, J)$ with $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and I_2, J are independent, so $\dim S = 2$. For $(a, b) \neq (0, 0)$,

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & 0 \\ 0 & a^2 + b^2 \end{pmatrix},$$

and the product in the other order is the same, so $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ has inverse $\left(a^2 + b^2\right)^{-1}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, since $a^2 + b^2 > 0$. (This is $\mathbb{C}$ in disguise.)

(c) Let $M = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$ and $S = \operatorname{span}(I_2, M) = \{0, I_2, M, I_2 + M\}$. Over $\mathbb{F}_2$, $M^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = I_2 + M$, so $M(M + I_2) = M^2 + M = I_2 + M + M = I_2$, and likewise $(M + I_2)M = I_2$. So M and $I_2 + M$ are invertible, as is I_2, and I_2, M are independent, so $\dim S = 2$. (The four matrices multiply like the field $\mathbb{F}_4$ of Exercise 1.1.56, with M as α.)

Solution to Exercise 4.2.55

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.1.19 (Extending linear maps and separating vectors), Theorem 1.5.5 (Independence is decided at zero), Proposition 4.1.34 (Size of a vector space over a finite field)
Hint 1. For (a), with a basis $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of C, consider $\varphi : \mathbb{F}_2^n \to \mathbb{F}_2^k$, $\mathbf{y} \mapsto (\mathbf{b}_1 \cdot \mathbf{y}, \ldots, \mathbf{b}_k \cdot \mathbf{y})$. Its kernel is $C^\perp$. Why is it surjective?
Hint 2. If φ were not surjective, a nonzero linear functional on $\mathbb{F}_2^k$ would vanish on $\operatorname{im}\varphi$ (Proposition 4.1.19(b)). Turn that into a linear relation among the $\mathbf{b}_i$.
Hint 3. For (b), let C be the span of the indicator vectors of the clubs. Show $C \subseteq C^\perp$ and count.
Sketch. (a) $\ker\varphi = C^\perp$ by bilinearity. If $\operatorname{im}\varphi \neq \mathbb{F}_2^k$, some nonzero $f(\mathbf{z}) = \sum a_i z_i$ vanishes on $\operatorname{im}\varphi$, so $(\sum a_i \mathbf{b}_i) \cdot \mathbf{y} = 0$ for all $\mathbf{y}$; taking $\mathbf{y} = \mathbf{e}_j$ gives $\sum a_i \mathbf{b}_i = \mathbf{0}$, contradicting independence. So $\dim C^\perp = n - k$. (b) Indicator vectors satisfy $\mathbf{v}_A \cdot \mathbf{v}_B = |A \cap B| = 0$ in $\mathbb{F}_2$, so their span C satisfies $C \subseteq C^\perp$, whence $\dim C \leq n - \dim C$ and $|C| \leq 2^{\lfloor n/2 \rfloor}$. Unions of the pairs $\{1,2\}, \{3,4\}, \ldots$ attain the bound.

The product is linear in each argument: $(a\mathbf{x} + \mathbf{x}') \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y}) + \mathbf{x}' \cdot \mathbf{y}$, and symmetrically, directly from the formula.

(a) Let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ be a basis of C and define $\varphi : \mathbb{F}_2^n \to \mathbb{F}_2^k$ by $\varphi(\mathbf{y}) = (\mathbf{b}_1 \cdot \mathbf{y}, \ldots, \mathbf{b}_k \cdot \mathbf{y})$, which is linear. If $\mathbf{y} \in C^\perp$, then $\varphi(\mathbf{y}) = \mathbf{0}$. Conversely, if $\varphi(\mathbf{y}) = \mathbf{0}$ and $\mathbf{c} = \sum_i a_i \mathbf{b}_i \in C$, then $\mathbf{c} \cdot \mathbf{y} = \sum_i a_i (\mathbf{b}_i \cdot \mathbf{y}) = 0$. So $C^\perp = \ker\varphi$, a subspace.

We show that φ is surjective. If not, $\operatorname{im}\varphi$ is a subspace of $\mathbb{F}_2^k$ missing some vector $\mathbf{z}_0$, and by Proposition 4.1.19(b) there is a linear $f : \mathbb{F}_2^k \to \mathbb{F}_2$ with $f(\mathbf{z}_0) = 1$ and $f = 0$ on $\operatorname{im}\varphi$. Writing $a_i = f(\mathbf{e}_i)$, we have $f(\mathbf{z}) = \sum_i a_i z_i$, and not all a_i are 0 since $f \neq 0$. For every $\mathbf{y}$,

$$0 = f(\varphi(\mathbf{y})) = \sum_{i=1}^{k} a_i (\mathbf{b}_i \cdot \mathbf{y}) = \left(\sum_{i=1}^{k} a_i \mathbf{b}_i \right) \cdot \mathbf{y}.$$

Taking $\mathbf{y} = \mathbf{e}_j$ shows that the jth entry of $\sum_i a_i \mathbf{b}_i$ is 0, for every j. So $\sum_i a_i \mathbf{b}_i = \mathbf{0}$ with not all $a_i = 0$, contradicting the independence of the $\mathbf{b}_i$ (Theorem 1.5.5). Hence $\operatorname{rank}\varphi = k$, and by Theorem 4.2.11, $\dim C^\perp = n - k$.

(b) Number the residents $1, \ldots, n$, and for a club $A \subseteq \{1, \ldots, n\}$ let $\mathbf{v}_A \in \mathbb{F}_2^n$ have 1 in the positions of the members and 0 elsewhere. Then $\mathbf{v}_A \cdot \mathbf{v}_B$ counts the common members modulo 2, so it is 0 for any two clubs A, B, including $A = B$ (each club has an even number of members). Let C be the span of all the $\mathbf{v}_A$. By linearity in each argument, $\mathbf{c} \cdot \mathbf{c}' = 0$ for all $\mathbf{c}, \mathbf{c}' \in C$, so $C \subseteq C^\perp$. With $k = \dim C$, (a) gives $k \leq n - k$, so $k \leq n/2$, and as k is an integer, $k \leq \lfloor n/2 \rfloor$. Different clubs have different vectors $\mathbf{v}_A$, all in C, so the number of clubs is at most $|C| = 2^k \leq 2^{\lfloor n/2 \rfloor}$ (Proposition 4.1.34).

For the construction, let $m = \lfloor n/2 \rfloor$ and take as clubs all unions of some of the pairs $\{1,2\}, \{3,4\}, \ldots, \{2m-1, 2m\}$, including the empty union. There are 2^m of them, each a union of pairs, so of even size, and the intersection of two of them is the union of their common pairs, also of even size.

Solution to Exercise 4.2.56

Uses. Theorem 4.2.11 (Rank–nullity), Corollary 2.6.24 (The columns decide), Theorem 4.2.7 (The image of T_A), Proposition 4.2.33 (Distinct columns give distance three), Theorem 4.2.34 (Syndrome decoding)
Hint 1. (a) The columns include $\mathbf{e}_1, \ldots, \mathbf{e}_r$.
Hint 2. (b) Proposition 4.2.33 gives at least 3. For a codeword of weight 3, use three columns adding up to $\mathbf{0}$, such as $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$.
Hint 3. (c) If $H_r\mathbf{x} \neq \mathbf{0}$, the syndrome is one of the columns.
Sketch. (a) T_{H_r} is surjective, so $\operatorname{rank} H_r = r$ and $\dim C_r = n - r$. (b) Distinct nonzero columns give weight ≥ 3; the positions of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$ give a codeword of weight 3. (c) If $H_r\mathbf{x} = \mathbf{h}_j$, then $\mathbf{x} + \mathbf{e}_j \in C_r$; if $H_r\mathbf{x} = \mathbf{0}$, $\mathbf{x} \in C_r$. Two codewords within distance 1 of $\mathbf{x}$ are within distance 2 of each other, hence equal.

(a) The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_r$ are among the columns, so the columns span $\mathbb{F}_2^r$, T_{H_r} is surjective (Corollary 2.6.24), and $\operatorname{rank} H_r = \dim \operatorname{im} T_{H_r} = r$ (Theorem 4.2.7). By Theorem 4.2.11, $\dim C_r = n - r$.

(b) The columns are nonzero and pairwise distinct, so every nonzero codeword has weight at least 3 (Proposition 4.2.33). Since $r \geq 2$, the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_1 + \mathbf{e}_2$ are three different columns, in positions i, j, l say, and their sum is $\mathbf{0}$. So the vector with 1s in positions i, j, l is a codeword of weight 3.

(c) *Existence.* If $H_r\mathbf{x} = \mathbf{0}$, then $\mathbf{x}$ is a codeword. Otherwise $H_r\mathbf{x}$ is a nonzero vector of $\mathbb{F}_2^r$, hence equal to some column $\mathbf{h}_j$, and $H_r(\mathbf{x} + \mathbf{e}_j) = \mathbf{h}_j + \mathbf{h}_j = \mathbf{0}$. So $\mathbf{x} + \mathbf{e}_j$ is a codeword at distance 1 from $\mathbf{x}$.

Uniqueness. If $\mathbf{c}, \mathbf{c}' \in C_r$ are both within distance 1 of $\mathbf{x}$, then $d(\mathbf{c}, \mathbf{c}') \leq 2$ by the triangle inequality, so $\mathbf{c} - \mathbf{c}'$ is a codeword of weight at most 2, and it is $\mathbf{0}$ by (b).

So C_r is perfect: 2^{n-r} codewords, each with $n + 1 = 2^r$ words within distance 1, account for all 2^n words.

Solution to Exercise 4.2.57

Uses. Example 4.2.32 (The Hamming code), Theorem 4.2.34 (Syndrome decoding)
Hint 1. Put the message in positions $3, 5, 6, 7$ and compute $x_1 = x_3 + x_5 + x_7$, $x_2 = x_3 + x_6 + x_7$, $x_4 = x_5 + x_6 + x_7$.
Hint 2. The syndrome is the sum of the columns $\mathbf{h}_j$ at the 1s of the received word; read it as a binary number.
Sketch. (a) 1100110. (b) 1000110 has syndrome $(0, 1, 0) = 2$, so it decodes to 1100110 and the message 0110; 1101011 has syndrome $(1, 1, 0) = 6$, so it decodes to 1101001 and the message 0001. (c) Syndrome $\mathbf{h}_1 + \mathbf{h}_2 = (0, 1, 1) = 3$, so the decoder flips bit 3 and returns 1110000, the wrong codeword.

(a) With $(x_3, x_5, x_6, x_7) = (0, 1, 1, 0)$: $x_1 = 0 + 1 + 0 = 1$, $x_2 = 0 + 1 + 0 = 1$, $x_4 = 1 + 1 + 0 = 0$. The codeword is 1100110.

(b) The word 1000110 has 1s in positions $1, 5, 6$, so its syndrome is $\mathbf{h}_1 + \mathbf{h}_5 + \mathbf{h}_6 = (0, 0, 1) + (1, 0, 1) + (1, 1, 0) = (0, 1, 0)$, binary for 2. By Theorem 4.2.34, flipping bit 2 gives the codeword 1100110, and the message is 0110. The word 1101011 has 1s in positions $1, 2, 4, 6, 7$, with syndrome $(0, 0, 1) + (0, 1, 0) + (1, 0, 0) + (1, 1, 0) + (1, 1, 1) = (1, 1, 0)$, binary for 6. Flipping bit 6 gives 1101001, and the message (positions $3, 5, 6, 7$) is 0001.

(c) The received word 1100000 has syndrome $\mathbf{h}_1 + \mathbf{h}_2 = (0, 1, 1)$, binary for 3, so the decoder flips bit 3 and returns 1110000. That is a codeword (Section 4.2.7), but not the one sent. The code corrects one error; two errors can be mistaken for one error at a third position, since the minimum distance is only 3.

Section 4.3

Solution to Exercise 4.3.2

Uses. Proposition 4.3.1 (Invertible factors preserve rank), Proposition 2.6.18 (The column space is spanned by the columns)
Hint 1. G adds -2 times row 1 to row 2, and H adds -2 times column 1 to column 2.
Hint 2. All three matrices have rank 1, so each kernel and each image is a line. Find a spanning vector for each.
Sketch. $GA = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ and $AH = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$. The kernels are $\operatorname{span}((-2, 1))$, $\operatorname{span}((-2, 1))$, $\operatorname{span}(\mathbf{e}_2)$, and the images are $\operatorname{span}((1, 2))$, $\operatorname{span}(\mathbf{e}_1)$, $\operatorname{span}((1, 2))$, for A, GA, AH respectively. So GA keeps the kernel of A and moves the image, and AH keeps the image and moves the kernel.

Multiplying out,

$$GA = \begin{pmatrix} 1 & 2 \\ 2-4 & 4-8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$

$$AH = \begin{pmatrix} 1 & -2+2 \\ 2 & -4+4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}.$$

The columns of A are $(1, 2)$ and $2(1, 2)$, so $\operatorname{im} A = \operatorname{span}((1, 2))$ (Proposition 2.6.18), and $A\mathbf{x} = \mathbf{0}$ means $x_1 + 2x_2 = 0$, so $\ker A = \operatorname{span}((-2, 1))$. For GA, the columns are $\mathbf{e}_1$ and $2\mathbf{e}_1$, so $\operatorname{im}(GA) = \operatorname{span}(\mathbf{e}_1)$, and $GA\mathbf{x} = \mathbf{0}$ again means $x_1 + 2x_2 = 0$, so $\ker(GA) = \operatorname{span}((-2, 1))$. For AH, the columns are $(1, 2)$ and $\mathbf{0}$, so $\operatorname{im}(AH) = \operatorname{span}((1, 2))$, and $AH\mathbf{x} = x_1(1, 2)$ vanishes exactly when $x_1 = 0$, so $\ker(AH) = \operatorname{span}(\mathbf{e}_2)$.

So $\ker(GA) = \ker A$ but $\operatorname{im}(GA) \neq \operatorname{im} A$, and $\operatorname{im}(AH) = \operatorname{im} A$ but $\ker(AH) \neq \ker A$, as Proposition 4.3.1 and Intuition (Outputs and inputs) predict. All three matrices have rank 1. Note that $\ker(AH) = H^{-1}(\ker A)$: indeed $H^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $H^{-1}(-2, 1) = (0, 1)$.

Solution to Exercise 4.3.7

Uses. Theorem 4.3.3 (Rank of a product), Proposition 3.5.2 (First properties of rank), Corollary 3.5.25 (The invertible matrix theorem), Theorem 2.4.34 (Block multiplication)
Hint 1. $\operatorname{rank}(BA) \le \operatorname{rank} A$, and A has only m rows.
Hint 2. For (b), try block matrices built from I_m and zeros.
Hint 3. For (c), sandwich $I_m = (AB)(AB) = A(BA)B$.
Sketch. (a) $\operatorname{rank}(BA) \le \operatorname{rank} A \le m < n$, so BA is not invertible. (b) $A = \begin{pmatrix} I_m & 0 \end{pmatrix}$, $B = \begin{pmatrix} I_m \\ 0 \end{pmatrix}$. (c) $m = \operatorname{rank}(I_m) = \operatorname{rank}(A(BA)B) \le \operatorname{rank}(BA) \le m$.

(a) By Theorem 4.3.3 and Proposition 3.5.2(a), $\operatorname{rank}(BA) \le \operatorname{rank} A \le m < n$. An invertible $n \times n$ matrix has rank n (Corollary 3.5.25), so BA is not invertible.

(b) Let $A = \begin{pmatrix} I_m & 0 \end{pmatrix} \in F^{m \times n}$ and $B = \begin{pmatrix} I_m \\ 0 \end{pmatrix} \in F^{n \times m}$, where the zero blocks have $n - m$ columns and $n - m$ rows. By Theorem 2.4.34, $AB = I_m I_m + 0 \cdot 0 = I_m$. (Here $BA = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$, of rank m.)

(c) If $AB = I_m$, then $I_m = (AB)(AB) = A(BA)B$. By Theorem 4.3.3, applied twice,

$$m = \operatorname{rank}(I_m) = \operatorname{rank}(A(BA)B) \le \operatorname{rank}(BA) \le m,$$

so $\operatorname{rank}(BA) = m$.

Solution to Exercise 4.3.11

Uses. Theorem 4.3.8 (Row rank equals column rank), Proposition 4.3.1 (Invertible factors preserve rank), Corollary 4.1.11 (Spanning lists are long)
Hint 1. Subtract twice row 1 from row 2, then scale by $2^{-1} = 3$.
Hint 2. For (b), the row space has dimension rank A. Are rows 1 and 2 multiples of each other?
Sketch. $\operatorname{rref}(A) = \begin{pmatrix} 1 & 0 & 4 & 4 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, so $\operatorname{rank} A = 2$ and $(1, 0, 4, 4), (0, 1, 3, 2)$ is a basis of the row space. Rows 1 and 2 are not proportional, so they are independent, and two independent vectors in a 2-dimensional space form a basis. Row 3 is row 1 plus twice row 2.

(a) Adding -2 times row 1 to row 2 gives $(0, 1 - 4, 1, -6) = (0, 2, 1, 4)$ in $\mathbb{F}_5$, and scaling by $2^{-1} = 3$ gives $(0, 6, 3, 12) = (0, 1, 3, 2)$. Adding -4 times this row to row 3 gives $(0, 0, 2 - 12, 3 - 8) = (0, 0, 0, 0)$, and adding -2 times it to row 1 gives $(1, 0, -6, -1) = (1, 0, 4, 4)$. So

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & 0 & 4 & 4 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$\operatorname{rank} A = 2$, and by the second proof of Theorem 4.3.8 the nonzero rows $(1, 0, 4, 4)$ and $(0, 1, 3, 2)$ form a basis of the row space of A.

(b) The row space has dimension 2. Rows 1 and 2 of A are not proportional (a multiple $c(1, 2, 0, 3)$ with third entry 1 does not exist), so they are linearly independent, and two independent vectors in a 2-dimensional space form a basis. Row 3 must be a combination of them, say row 3 $= a$ (row 1) $+ b$ (row 2). The first entries give $0 = a + 2b$ and the second entries $4 = 2a + b$; so $a = -2b = 3b$ and $4 = 7b = 2b$ in $\mathbb{F}_5$, giving $b = 2$ and $a = 1$. Indeed $(1, 2, 0, 3) + 2(2, 1, 1, 0) = (5, 4, 2, 3) = (0, 4, 2, 3)$ in $\mathbb{F}_5$.

Solution to Exercise 4.3.19

Uses. Proposition 4.3.16 (Rank-one matrices), Example 4.3.18 (Powers of a rank-one matrix)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. Every row is a multiple of $(3, 1, 1)$.
Hint 2. $A^k = (\operatorname{tr} A)^{k-1} A$. Over $\mathbb{F}_2$, reduce the trace modulo 2.
Sketch. (a) $A = (1, 2, -1)(3, 1, 1)^{\mathsf{T}}$, so $\operatorname{rank} A = 1$, $\operatorname{tr} A = 3 + 2 - 1 = 4$, and $A^{10} = 4^9 A = 262144A$. (b) Over $\mathbb{F}_2$, $\operatorname{tr} A = 0$, so $A^2 = 0$.

(a) The rows are 1, 2, and -1 times $(3, 1, 1)$, so $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (3, 1, 1)$, both nonzero, and $\operatorname{rank} A = 1$ by Proposition 4.3.16. Its trace is $3 + 2 - 1 = 4$, so by Example 4.3.18, $A^{10} = 4^9 A = 262144A$.

(b) Over $\mathbb{F}_2$, $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = (1,0,1)(1,1,1)^{\mathsf{T}}$, of rank one, with $\operatorname{tr} A = 1+0+1 = 0$. So $A^2 = (\operatorname{tr} A)A = 0$. Directly: $(1,1,1)\cdot(1,0,1) = 2 = 0$ in $\mathbb{F}_2$.

Solution to Exercise 4.3.25

Uses. Theorem 4.3.21 (Column–row factorization), Corollary 4.3.24 (Rank as a number of rank-one pieces)
Hint 1. Clear the first column with row 1. Row 3 is the sum of rows 1 and 2.
Hint 2. The pivot columns are 1 and 3.
Sketch. $\operatorname{rref}(A) = \begin{pmatrix} 1 & i & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, with pivot columns $1, 3$. So $C = \begin{pmatrix} 1 & 0 \\ i & 1 \\ 1+i & 1 \end{pmatrix}$, $R = \begin{pmatrix} 1 & i & 0 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, and $A = (1, i, 1+i)(1\ \ i\ \ 0\ \ 2) + (0,1,1)(0\ \ 0\ \ 1\ \ 0)$.

Adding $-i$ times row 1 to row 2 gives $\left(i - i, -1 - i^2, 1, 2i - 2i\right) = (0,0,1,0)$. Row 3 is the sum of rows 1 and 2 of A, so adding -1 times each of them to it gives $\mathbf{0}$. Hence

$$\operatorname{rref}(A) = \begin{pmatrix} 1 & i & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

with pivot columns 1 and 3, and $\operatorname{rank} A = 2$. By Theorem 4.3.21,

$$A = CR, \quad C = \begin{pmatrix} 1 & 0 \\ i & 1 \\ 1+i & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & i & 0 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Check: column 2 of CR is $i(1, i, 1+i) = (i, -1, -1+i)$ and column 4 is $2(1, i, 1+i) = (2, 2i, 2+2i)$, the columns of A. By Corollary 4.3.24 and its proof,

$$A = \begin{pmatrix} 1 \\ i \\ 1+i \end{pmatrix}(1\ \ i\ \ 0\ \ 2) + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}(0\ \ 0\ \ 1\ \ 0).$$

Solution to Exercise 4.3.26

Uses. Proposition 4.3.5 (Rank of a product, exactly), Theorem 4.2.11 (Rank–nullity)
Hint 1. Row 3 of A is the sum of rows 1 and 2, so $\operatorname{rank} A = 2$ and $\ker A$ is a line.
Hint 2. For (b), how small can $\operatorname{rank}(AB)$ be when $\operatorname{rank} B = 2$ and $\dim \ker A = 1$?
Sketch. $\ker A = \operatorname{span}((1,1,-1))$. $\operatorname{im} B_1$ contains $(1,1,-1)$, so $\operatorname{im} B_1 \cap \ker A = \ker A$ and $\operatorname{rank}(AB_1) = 2 - 1 = 1$; $\operatorname{im} B_2 = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ misses it, so $\operatorname{rank}(AB_2) = 2$. (b) No: $\operatorname{rank}(AB) = 2 - \dim(\operatorname{im} B \cap \ker A) \geq 2 - 1 = 1$.

(a) $A\mathbf{x} = \mathbf{0}$ means $x_1 + x_3 = 0$, $x_2 + x_3 = 0$, and $x_1 + x_2 + 2x_3 = 0$, the third being the sum of the first two. So $\ker A = \operatorname{span}((1,1,-1))$ and $\operatorname{rank} A = 2$.

The first column of B_1 is $(1,1,-1)$, so $\ker A \subseteq \operatorname{im} B_1$ and $\operatorname{im} B_1 \cap \ker A = \ker A$, of dimension 1. The columns of B_1 are independent, so $\operatorname{rank} B_1 = 2$, and (4.3.1) gives $\operatorname{rank}(AB_1) = 2 - 1 = 1$. Indeed

$$AB_1 = \begin{pmatrix} 1-1 & 0 \\ 1-1 & 1 \\ 1+1-2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

which has rank 1.

$\operatorname{im} B_2 = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ consists of the vectors with third entry 0, and the only such multiple of $(1,1,-1)$ is $\mathbf{0}$. So $\operatorname{im} B_2 \cap \ker A = \{\mathbf{0}\}$ and $\operatorname{rank}(AB_2) = 2$. Indeed $AB_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, the first two columns of A, which are independent.

(b) No. If $\operatorname{rank} B = 2$, then by (4.3.1), $\operatorname{rank}(AB) = 2 - \dim(\operatorname{im} B \cap \ker A) \geq 2 - \dim \ker A = 1$, so $AB \neq 0$.

Solution to Exercise 4.3.27

Uses. Proposition 4.3.13 (Dimensions of the four subspaces), Definition 4.3.12 (The four fundamental subspaces)
Hint 1. Over either field, row 3 is the sum of rows 1 and 2, and the first two rows are already in RREF.
Hint 2. Over $\mathbb{F}_2$, compare the basic solutions with the rows. Over $\mathbb{Q}$, apply A to a general vector $a(1,1,0,0)+b(0,0,1,1)$ of the row space.
Sketch. Over both fields $\operatorname{rref}(A)=\begin{pmatrix}1&1&0&0\\0&0&1&1\\0&0&0&0\end{pmatrix}$, $r=2$, pivot columns $1,3$. Over $\mathbb{F}_2$: $\operatorname{im}A=\operatorname{span}((1,0,1),(0,1,1))$, $\ker A=\operatorname{span}((1,1,0,0),(0,0,1,1))=\operatorname{im}A^{\mathsf{T}}$, $\ker A^{\mathsf{T}}=\operatorname{span}((1,1,1))$. Over $\mathbb{Q}$: $\ker A=\operatorname{span}((-1,1,0,0),(0,0,-1,1))$, $\ker A^{\mathsf{T}}=\operatorname{span}((1,1,-1))$, and $A(a,a,b,b)=(2a,2b,2a+2b)$ vanishes only for $a=b=0$.

Over any field, row 3 of A is the sum of rows 1 and 2, so subtracting rows 1 and 2 from row 3 gives

$$\operatorname{rref}(A)=\begin{pmatrix}1&1&0&0\\0&0&1&1\\0&0&0&0\end{pmatrix},$$

with $r=2$, pivot columns $1,3$, and free columns $2,4$. Here $m=3$ and $n=4$, so Proposition 4.3.13 predicts dimensions 2, 2, 2, 1.

(a) Over $\mathbb{F}_2$:

- $\operatorname{im}A$ has the basis $(1,0,1),(0,1,1)$, columns 1 and 3 of A.
- The basic solutions are $(-1,1,0,0)=(1,1,0,0)$ and $(0,0,-1,1)=(0,0,1,1)$, a basis of $\ker A$.
- $\operatorname{im}A^{\mathsf{T}}$ has the basis $(1,1,0,0),(0,0,1,1)$, the nonzero rows of $\operatorname{rref}(A)$.
- Row 1 plus row 2 plus row 3 is $(2,2,2,2)=\mathbf{0}$, so $(1,1,1)\in\ker A^{\mathsf{T}}$, which has dimension $3-2=1$; so $\ker A^{\mathsf{T}}=\operatorname{span}((1,1,1))$.

The bases of $\ker A$ and $\operatorname{im}A^{\mathsf{T}}$ are the same list, so $\operatorname{im}A^{\mathsf{T}}=\ker A$. Each row is "orthogonal" to itself: $(1,1,0,0)\cdot(1,1,0,0)=2=0$ in $\mathbb{F}_2$.

(b) Over $\mathbb{Q}$, $\operatorname{im}A=\operatorname{span}((1,0,1),(0,1,1))$ and $\operatorname{im}A^{\mathsf{T}}=\operatorname{span}((1,1,0,0),(0,0,1,1))$ as before. The basic solutions are now $(-1,1,0,0)$ and $(0,0,-1,1)$, and row 1 plus row 2 minus row 3 is $\mathbf{0}$, so $\ker A^{\mathsf{T}}=\operatorname{span}((1,1,-1))$. A vector of the row space is $\mathbf{w}=(a,a,b,b)$, and $A\mathbf{w}=(2a,2b,2a+2b)$. This is $\mathbf{0}$ only if $a=b=0$, since $2\neq 0$ in $\mathbb{Q}$. So $\operatorname{im}A^{\mathsf{T}}\cap\ker A=\{\mathbf{0}\}$. The two computations differ only in whether $1+1=0$.

Solution to Exercise 4.3.28

Uses. Proposition 4.3.16 (Rank-one matrices), Proposition 4.3.13 (Dimensions of the four subspaces), Theorem 4.2.11 (Rank–nullity), Exercise 3.3.50
Hint 1. In (a), A has rank 1, so $A=\mathbf{u}\mathbf{v}^{\mathsf{T}}$. What must $\mathbf{u}$ be, and which condition on $\mathbf{v}$ makes $\ker A$ correct?
Hint 2. In (b) and (e), count dimensions. In (c), every row has dot product 0 with every vector of the null space (Exercise 3.3.50(a)).
Sketch. (a) $A=(1,0,1)(1,-1,1)^{\mathsf{T}}$. (b) None: $\dim\ker A\geq 4-2=2$. (c) None: $(1,1,0)\cdot(1,0,0)=1\neq 0$. (d) $A=\begin{pmatrix}0&1\\0&0\end{pmatrix}$. (e) None: $r=3-r$ is impossible.

(a) Such an A has rank 1, so $A=\mathbf{u}\mathbf{v}^{\mathsf{T}}$, and we may take $\mathbf{u}=(1,0,1)$ (Proposition 4.3.16). Then $A\mathbf{x}=\left(\mathbf{v}^{\mathsf{T}}\mathbf{x}\right)\mathbf{u}$, so $\ker A=\left\{\mathbf{x}\;\middle|\;\mathbf{v}^{\mathsf{T}}\mathbf{x}=0\right\}$. We need $\mathbf{v}^{\mathsf{T}}(1,1,0)=0$ and $\mathbf{v}^{\mathsf{T}}(0,1,1)=0$, that is, $v_1+v_2=0=v_2+v_3$; take $\mathbf{v}=(1,-1,1)$. Then

$$A=\begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}1&-1&1\end{pmatrix}=\begin{pmatrix}1&-1&1\\0&0&0\\1&-1&1\end{pmatrix}.$$

Its kernel contains $(1,1,0)$ and $(0,1,1)$, which are independent, and has dimension $3-1=2$ (Theorem 4.2.11), so it is their span; and $\operatorname{im}A=\operatorname{span}((1,0,1))$.

(b) None. By Theorem 4.2.11, $\dim\ker A=4-\operatorname{rank}A\geq 4-2=2$, while $\dim\operatorname{span}(\mathbf{e}_1)=1$.

(c) None, over every field. By Exercise 3.3.50(a), every vector of the row space has dot product 0 with every vector of the null space, but $(1,1,0)\cdot(1,0,0)=1\neq 0$.

(d) $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$: its image is spanned by its nonzero column $\mathbf{e}_1$, and $A(x_1, x_2) = (x_2, 0)$ vanishes exactly when $x_2 = 0$, so $\ker A = \operatorname{span}(\mathbf{e}_1)$ too.

(e) None. If $\operatorname{im} A = \ker A$, then $\operatorname{rank} A = \dim \ker A = 3 - \operatorname{rank} A$ by Theorem 4.2.11, so $2 \operatorname{rank} A = 3$, which no integer satisfies. In general, a square matrix with $\operatorname{im} A = \ker A$ must have even size.

Solution to Exercise 4.3.29

Uses. Exercise 3.3.50, Proposition 2.4.22 (Transpose is a linear involution), Definition 4.3.12 (The four fundamental subspaces)
Techniques. Transpose to trade rows for columns (Technique 3.2.27), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), think of invertible matrices.
Hint 2. For (b), Exercise 3.3.50(a) describes $\ker A$ in terms of the row space alone. Then transpose for (c).
Hint 3. For (d), find two 2×2 matrices of rank 1 with the same kernel.
Sketch. (a) False: I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ are both invertible. (b) True: $\ker A$ is the set of $\mathbf{x}$ with $\mathbf{w} \cdot \mathbf{x} = 0$ for all $\mathbf{w}$ in the row space. (c) True: apply (b) to A^{T} and B^{T}, whose row spaces are $\operatorname{im} A$ and $\operatorname{im} B$. (d) False: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

(a) False. I_2 and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ are both invertible, so both have column space and row space F^2 and null space and left null space $\{\mathbf{0}\}$; but $I_2 \neq B$.

(b) True. By Exercise 3.3.50(a), $\ker A$ is the set of $\mathbf{x} \in F^n$ with $\mathbf{w} \cdot \mathbf{x} = 0$ for every $\mathbf{w}$ in the row space of A. This description involves only the row space, so equal row spaces give equal null spaces.

(c) True. The row space of A^{T} is $\operatorname{im}\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = \operatorname{im} A$ (Proposition 2.4.22), and likewise for B. So if $\operatorname{im} A = \operatorname{im} B$, then A^{T} and B^{T} have the same row space, and (b) gives $\ker A^{\mathsf{T}} = \ker B^{\mathsf{T}}$.

(d) False. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Both send (x_1, x_2) to a vector that is zero exactly when $x_1 = 0$, so $\ker A = \ker B = \operatorname{span}(\mathbf{e}_2)$, but $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} B = \operatorname{span}(\mathbf{e}_2)$. Parts (b)–(d) show that the row space determines the null space and the column space determines the left null space, but that the null space does not determine the column space.

Solution to Exercise 4.3.30

Uses. Proposition 4.3.5 (Rank of a product, exactly), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. By (4.3.1), $\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A)$. Bound the subtracted term.
Hint 2. For (c), induct on k, applying (a) to $A_1 \cdots A_{k-1}$ and A_k.
Hint 3. For the example in (d), take the shift N with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$.
Sketch. (a) $\dim(\operatorname{im} B \cap \ker A) \leq \dim \ker A = n - \operatorname{rank} A$. (b) Put $\operatorname{rank}(AB) = 0$. (c) Induction: $\operatorname{rank}(PA_k) \geq \operatorname{rank} P + \operatorname{rank} A_k - n$ with $P = A_1 \cdots A_{k-1}$. (d) With all $A_i = A$, $k(n-1) - (k-1)n = n - k$; the shift has rank $N^k = n - k$.

(a) By (4.3.1), $\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A)$. Since $\operatorname{im} B \cap \ker A \subseteq \ker A$, its dimension is at most $\dim \ker A = n - \operatorname{rank} A$ (Proposition 4.1.21(a), Theorem 4.2.11(a)). Hence

$$\begin{aligned} \operatorname{rank}(AB) &\geq \operatorname{rank} B - (n - \operatorname{rank} A) \\ &= \operatorname{rank} A + \operatorname{rank} B - n. \end{aligned}$$

(b) If $AB = 0$, then $0 = \operatorname{rank}(AB) \geq \operatorname{rank} A + \operatorname{rank} B - n$.

(c) By induction on k. For $k = 1$ there is nothing to prove. For $k \geq 2$, let $P = A_1 \cdots A_{k-1}$. By (a), $\operatorname{rank}(PA_k) \geq \operatorname{rank} P + \operatorname{rank} A_k - n$, and by the induction hypothesis,

$$\operatorname{rank} P \geq \operatorname{rank} A_1 + \cdots + \operatorname{rank} A_{k-1} - (k-2)n.$$

Substituting the second inequality into the first gives the claim for k.

(d) With $A_1 = \cdots = A_k = A$, (c) gives $\operatorname{rank}\left(A^k\right) \geq k(n-1) - (k-1)n = n - k$. For the shift $N \in F^{n\times n}$ with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \geq 2$, the power N^k sends $\mathbf{e}_j$ to $\mathbf{e}_{j-k}$ for $j > k$ and to $\mathbf{0}$ for $j \leq k$, so $\operatorname{im} N^k = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{n-k})$ and $\operatorname{rank} N^k = n - k$ for $k \leq n$. In particular $\operatorname{rank} N = n - 1$.

Sylvester's inequality is the lower bound that the exact formula hides: the product loses at most $\dim\ker A$ dimensions, because the lost directions lie in $\ker A$.

Solution to Exercise 4.3.31

Uses. Proposition 4.3.5 (Rank of a product, exactly), Proposition 2.6.15 (Kernels and images of composites), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Rewrite the inequality as $\operatorname{rank}(BC) - \operatorname{rank}(A(BC)) \leq \operatorname{rank} B - \operatorname{rank}(AB)$: what A destroys of $\operatorname{im}(BC)$ is at most what it destroys of $\operatorname{im} B$.
Hint 2. Express both sides with (4.3.1). How are $\operatorname{im}(BC)$ and $\operatorname{im} B$ related?
Hint 3. For (b), take $B = I_n$. For (c), take the three matrices A, A^k, A.
Sketch. By (4.3.1), $\operatorname{rank} B - \operatorname{rank}(AB) = \dim(\operatorname{im} B \cap \ker A)$ and $\operatorname{rank}(BC) - \operatorname{rank}(ABC) = \dim(\operatorname{im}(BC) \cap \ker A)$. Since $\operatorname{im}(BC) \subseteq \operatorname{im} B$, the second is at most the first. (b) $B = I_n$. (c) With A, A^k, A: $2\operatorname{rank}\left(A^{k+1}\right) \leq \operatorname{rank}\left(A^k\right) + \operatorname{rank}\left(A^{k+2}\right)$, which rearranges to $d_{k+1} \leq d_k$.

(a) By (4.3.1), applied to A and B and to A and BC,

$$\operatorname{rank} B - \operatorname{rank}(AB) = \dim(\operatorname{im} B \cap \ker A),$$
$$\operatorname{rank}(BC) - \operatorname{rank}(ABC) = \dim(\operatorname{im}(BC) \cap \ker A).$$

By Proposition 2.6.15(b), $\operatorname{im}(BC) \subseteq \operatorname{im} B$, so $\operatorname{im}(BC) \cap \ker A \subseteq \operatorname{im} B \cap \ker A$, and Proposition 4.1.21(a) gives

$$\operatorname{rank}(BC) - \operatorname{rank}(ABC) \leq \operatorname{rank} B - \operatorname{rank}(AB).$$

Rearranging gives (a). In words: of the directions in $\operatorname{im}(BC)$, which is part of $\operatorname{im} B$, the matrix A can destroy no more than it destroys of all of $\operatorname{im} B$.

(b) Let $A \in F^{m\times n}$ and $C \in F^{n\times q}$, and take $B = I_n$ in (a): $\operatorname{rank} A + \operatorname{rank} C \leq \operatorname{rank} I_n + \operatorname{rank}(AC) = n + \operatorname{rank}(AC)$.

(c) Apply (a) to the three $n \times n$ matrices A, A^k, A:

$$\operatorname{rank}\left(A^{k+1}\right) + \operatorname{rank}\left(A^{k+1}\right) \leq \operatorname{rank}\left(A^k\right) + \operatorname{rank}\left(A^{k+2}\right).$$

Rearranged, $\operatorname{rank}\left(A^{k+1}\right) - \operatorname{rank}\left(A^{k+2}\right) \leq \operatorname{rank}\left(A^k\right) - \operatorname{rank}\left(A^{k+1}\right)$, that is, $d_{k+1} \leq d_k$. (For $k = 0$, $A^0 = I_n$.) By rank–nullity, $d_k = \dim\ker A^{k+1} - \dim\ker A^k$, so this is Exercise 4.2.52(b) once more.

Solution to Exercise 4.3.32

Uses. Theorem 4.3.3 (Rank of a product), Exercise 4.3.30, Lemma 2.4.24 (Matrix units), Proposition 2.6.18 (The column space is spanned by the columns), Corollary 1.5.17 (Sublists), Theorem 4.2.7 (The image of T_A)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. The bounds are Theorem 4.3.3 and Sylvester's inequality. For the converse, use matrices made of 0s and a few 1s.
Hint 2. Let A have 1 in positions $(1,1), \ldots, (a,a)$ and 0 elsewhere. Let B have, in its columns $1, \ldots, b$, distinct standard basis vectors $\mathbf{e}_{t_1}, \ldots, \mathbf{e}_{t_b}$. What is AB?
Hint 3. AB keeps exactly the columns $\mathbf{e}_{t_k}$ with $t_k \leq a$. How many of the t_k can you place in $\{1, \ldots, a\}$?
Sketch. Necessity: Theorem 4.3.3 and Exercise 4.3.30(a). Sufficiency: $A = \sum_{i\leq a} \mathbf{e}_i\mathbf{e}_i^{\mathsf{T}}$ and $B = \sum_{k\leq b} \mathbf{e}_{t_k}\mathbf{e}_k^{\mathsf{T}}$ with distinct t_k; then $AB = \sum_{k:t_k\leq a} \mathbf{e}_{t_k}\mathbf{e}_k^{\mathsf{T}}$ has rank $|\{k \mid t_k \leq a\}|$. Take $t_k = k$ for $k \leq r$ and $t_k = a + k - r$ for $r < k \leq b$; this needs $r \leq a$ and $a + b - r \leq n$. For $(5; 3, 4)$, $r \in \{2, 3\}$.

Necessity. By Theorem 4.3.3, $\operatorname{rank}(AB) \leq \min(a, b)$, and by Exercise 4.3.30(a), $\operatorname{rank}(AB) \geq a + b - n$; ranks are nonnegative.

Sufficiency. Let r satisfy the bounds. A matrix whose nonzero columns are distinct standard basis vectors has rank equal to the number of those columns, since its image is spanned by them (Proposition 2.6.18),

they are independent as part of the standard basis (Corollary 1.5.17), and rank is the dimension of the image (Theorem 4.2.7). Let

$$A = \sum_{i=1}^{a} \mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \in F^{m\times n},$$

whose columns are $\mathbf{e}_1, \dots, \mathbf{e}_a \in F^m$ followed by zeros, so $\operatorname{rank} A = a$. Choose distinct indices $t_1, \dots, t_b \in \{1, \dots, n\}$, specified below, and let

$$B = \sum_{k=1}^{b} \mathbf{e}_{t_k} \mathbf{e}_k^{\mathsf{T}} \in F^{n\times p},$$

whose columns are $\mathbf{e}_{t_1}, \dots, \mathbf{e}_{t_b} \in F^n$ followed by zeros, so $\operatorname{rank} B = b$. By Lemma 2.4.24(c), $\mathbf{e}_i \mathbf{e}_i^{\mathsf{T}} \mathbf{e}_{t_k} \mathbf{e}_k^{\mathsf{T}}$ is $\mathbf{e}_i \mathbf{e}_k^{\mathsf{T}}$ if $t_k = i$ and 0 otherwise, so

$$AB = \sum_{k:t_k \le a} \mathbf{e}_{t_k} \mathbf{e}_k^{\mathsf{T}},$$

whose nonzero columns are the distinct vectors $\mathbf{e}_{t_k}$ with $t_k \le a$. So $\operatorname{rank}(AB)$ is the number of k with $t_k \le a$.

Now take $t_k = k$ for $1 \le k \le r$ and $t_k = a + (k - r)$ for $r < k \le b$. These are distinct, since $r \le a$. They lie in $\{1, \dots, n\}$, since the largest is $a + b - r \le n$ when $r \ge a + b - n$. Exactly $t_1, \dots, t_r$ are at most a, as $r \le b$ and $t_k > a$ for $k > r$. So $\operatorname{rank}(AB) = r$.

The example. With $n = 5$, $a = 3$, $b = 4$: $\max(0, 3 + 4 - 5) = 2$ and $\min(3, 4) = 3$, so $r \in \{2, 3\}$. Take $A = \begin{pmatrix} I_3 & 0 \\ 0 & 0 \end{pmatrix} \in F^{5\times 5}$. For $r = 3$, take $t = (1, 2, 3, 4)$, that is, $B = \begin{pmatrix} I_4 & 0 \\ 0 & 0 \end{pmatrix}$; then $AB = A$ has rank 3. For $r = 2$, take $t = (1, 2, 4, 5)$, that is, B has columns $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_5, \mathbf{0}$; then AB has columns $\mathbf{e}_1, \mathbf{e}_2, \mathbf{0}, \mathbf{0}, \mathbf{0}$, of rank 2.

Solution to Exercise 4.3.33

Uses. Proposition 4.3.5 (Rank of a product, exactly), Theorem 4.3.3 (Rank of a product), Proposition 2.4.23 (Transpose of a product), Proposition 2.4.22 (Transpose is a linear involution)
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. (a) is (4.3.1) for the product $A^{\mathsf{T}}A$.
Hint 2. For (b), if $\mathbf{y} = A\mathbf{x}$ and $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$, compute $\mathbf{y}^{\mathsf{T}}\mathbf{y}$ in two ways.
Hint 3. For (c), $\operatorname{rank} A = \operatorname{rank}\left(A^{\mathsf{T}}A\right) \le \operatorname{rank} A^{\mathsf{T}}$. Now apply the same to A^{T}.
Sketch. (a) is (4.3.1). (b) If $\mathbf{y} = A\mathbf{x} \in \ker A^{\mathsf{T}}$, then $\sum_i y_i^2 = \mathbf{y}^{\mathsf{T}}\mathbf{y} = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}\mathbf{y} = 0$, so $\mathbf{y} = \mathbf{0}$. (c) $\operatorname{rank} A = \operatorname{rank}\left(A^{\mathsf{T}}A\right) \le \operatorname{rank} A^{\mathsf{T}}$ by Theorem 4.3.3, and the same for A^{T} gives the reverse; then $\operatorname{rank}\left(AA^{\mathsf{T}}\right) = \operatorname{rank}\left(\left(A^{\mathsf{T}}\right)^{\mathsf{T}}A^{\mathsf{T}}\right) = \operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$.

(a) This is (4.3.1) with A^{T} in the role of A and A in the role of B.

(b) Let $\mathbf{y} \in \operatorname{im} A \cap \ker A^{\mathsf{T}}$, say $\mathbf{y} = A\mathbf{x}$ with $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$. By Proposition 2.4.23, $\mathbf{y}^{\mathsf{T}} = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}$, so

$$y_1^2 + \cdots + y_m^2 = \mathbf{y}^{\mathsf{T}}\mathbf{y} = \mathbf{x}^{\mathsf{T}}\left(A^{\mathsf{T}}\mathbf{y}\right) = \mathbf{x}^{\mathsf{T}}\mathbf{0} = 0.$$

A sum of squares of real numbers is 0 only if every term is 0 (Technique 6.1.9), so $\mathbf{y} = \mathbf{0}$. By (a), $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$.

(c) By (b) and Theorem 4.3.3, $\operatorname{rank} A = \operatorname{rank}\left(A^{\mathsf{T}}A\right) \le \operatorname{rank} A^{\mathsf{T}}$. Applying this to the real matrix A^{T}, whose transpose is A (Proposition 2.4.22), gives $\operatorname{rank} A^{\mathsf{T}} \le \operatorname{rank} A$. So $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$. Finally, (b) for A^{T} gives $\operatorname{rank}\left(AA^{\mathsf{T}}\right) = \operatorname{rank}\left(\left(A^{\mathsf{T}}\right)^{\mathsf{T}}A^{\mathsf{T}}\right) = \operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$.

Over $\mathbb{C}$ the same argument works with the conjugate transpose A^* in place of A^{T}, since $\overline{y_1}y_1 + \cdots + \overline{y_m}y_m = |y_1|^2 + \cdots + |y_m|^2$; this is Proposition 6.5.9. Exercise 4.3.34 shows what goes wrong with A^{T}.

Solution to Exercise 4.3.34

Uses. Exercise 4.3.33, Exercise 3.3.54, Corollary 3.5.25 (The invertible matrix theorem)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)

Hint 1. Look for a nonzero vector $\mathbf{y}$ with $y_1^2 + \cdots + y_m^2 = 0$, and use it as a one-column matrix.
Hint 2. For (c), a rational matrix is a real matrix, and its rank does not depend on which of the two fields we compute in.
Sketch. (a) False: $A = (1,1)^{\mathsf{T}}$ gives $A^{\mathsf{T}}A = 1 + 1 = 0$. (b) False: $A = (1,i)^{\mathsf{T}}$ gives $1 + i^2 = 0$. (c) True, by Exercise 4.3.33 and Exercise 3.3.54. (d) False: $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ gives $A^{\mathsf{T}}A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The failing step is "a sum of squares is 0 only if each term is 0".

(a) False. Let $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 1}$, of rank 1. Then $A^{\mathsf{T}}A = (1 \cdot 1 + 1 \cdot 1) = (0)$, of rank 0.

(b) False. Let $A = \begin{pmatrix} 1 \\ i \end{pmatrix} \in \mathbb{C}^{2\times 1}$, of rank 1. Then $A^{\mathsf{T}}A = \left(1 + i^2\right) = (0)$.

(c) True. A matrix with rational entries is also a real matrix, and by Exercise 3.3.54(a) its reduced row echelon form, hence its rank, is the same over $\mathbb{Q}$ and over $\mathbb{R}$. The same holds for $A^{\mathsf{T}}A$, which also has rational entries. So $\operatorname{rank}\left(A^{\mathsf{T}}A\right) = \operatorname{rank} A$ over $\mathbb{Q}$ follows from Exercise 4.3.33(b) over $\mathbb{R}$.

(d) False. Over $\mathbb{F}_2$, let $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, whose two columns are independent. Then

$$A^{\mathsf{T}}A = \begin{pmatrix} 1+1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

which has rank 1 and is not invertible (Corollary 3.5.25).

In each false case, the vector $\mathbf{y} = A\mathbf{e}_1$, that is, $(1,1)$, $(1,i)$, or $(1,1,0)$, is nonzero, satisfies $\mathbf{y}^{\mathsf{T}}\mathbf{y} = 0$, and lies in $\operatorname{im} A \cap \ker A^{\mathsf{T}}$, which is exactly what Exercise 4.3.33(a) says must happen when the rank drops. The proof of Exercise 4.3.33(b) used that a sum of squares vanishes only when every term does, which is a property of $\mathbb{R}$ (and of its subfields), not of fields in general.

Solution to Exercise 4.3.35

Uses. Non-example 4.2.8 (The columns of A_r are the wrong vectors), Proposition 4.3.1 (Invertible factors preserve rank), Theorem 4.2.7 (The image of T_A)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Which operations preserve the column space: row operations or column operations?
Hint 2. Try $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.
Hint 3. The conclusion $\operatorname{rank} A = r$ is still right. Why do $\operatorname{im} A$ and $\operatorname{im} A_r$ have the same dimension even though they are different?
Sketch. The false step is $\operatorname{im} A = \operatorname{im} A_r$: row operations preserve the row space and the relations among the columns, not the column space. For $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $\operatorname{im} A = \operatorname{span}((1,2))$ but $\operatorname{im} A_r = \operatorname{span}(\mathbf{e}_1)$. The claim is true. Repair: $A = G^{-1}A_r$, so $\operatorname{rank} A = \operatorname{rank} A_r = r$ by Proposition 4.3.1(a) (or directly by Theorem 4.2.7); the rest of the argument is the second proof of Theorem 4.3.8.

The error. The step "row operations do not change the column space" is false. Row operations preserve the row space and the linear relations among the columns, but they move the columns themselves. For $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ we have $A_r = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, so $\operatorname{im} A = \operatorname{span}((1,2))$ while $\operatorname{im} A_r = \operatorname{span}(\mathbf{e}_1)$ (Non-example 4.2.8). In particular the intermediate conclusion $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_r)$ is false for this A.

The conclusion. The claim is true; it is Theorem 4.3.8. Only the first half of the argument needs repair. With G invertible and $GA = A_r$, Proposition 4.3.1(a) gives $\operatorname{rank} A = \operatorname{rank}(GA) = \operatorname{rank} A_r = r$, although $\operatorname{im} A \neq \operatorname{im} A_r$ in general. (This is also the definition of rank A as the number of pivots, together with Theorem 4.2.7.) The second half is correct as stated: A and A_r have the same row space, and the nonzero rows of A_r are independent. It is the second proof of Theorem 4.3.8.

The two halves are not symmetric: row operations keep the row space as a set, but keep the column space only up to an invertible map, so only its dimension survives.

Solution to Exercise 4.3.36

Uses. Proposition 4.3.5 (Rank of a product, exactly), Theorem 4.2.11 (Rank–nullity), Proposition 2.6.15 (Kernels and images of composites), Example 4.3.4 (A product of rank zero)

Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), see Example 4.3.4.
Hint 2. For (b), what are $\ker A$ and $\operatorname{im} B$ when both ranks are n?
Hint 3. For (c), the columns may avoid each other while the rows do not. Try 2×2 matrices with a single nonzero entry each, in the same column.
Hint 4. For (d), use (4.3.1) with A^k and A: the loss is $\dim\left(\operatorname{im} A^k \cap \ker A\right)$.
Sketch. (a) False: $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ give $MN = N$, $NM = 0$. (b) True: $\ker A = \{\mathbf{0}\}$, so $\operatorname{rank}(AB) = \operatorname{rank} B = n$. (c) False: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ has rank 1. (d) True: $\operatorname{im} A \cap \ker A = \{\mathbf{0}\}$, and $\operatorname{im} A^k \subseteq \operatorname{im} A$, so $\operatorname{rank}\left(A^{k+1}\right) = \operatorname{rank}\left(A^k\right) - \dim\left(\operatorname{im} A^k \cap \ker A\right) = \operatorname{rank}\left(A^k\right)$.

(a) False. With $M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we have $MN = N$, of rank 1, and $NM = 0$ (Example 4.3.4).

(b) True. A has n columns and rank n, so $\dim \ker A = 0$ (Theorem 4.2.11). By (4.3.1), $\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A) = n - 0 = n$.

(c) False. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Then $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} B = \operatorname{span}(\mathbf{e}_2)$ meet only in $\mathbf{0}$, but $A + B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ has rank $1 < 2$. Here the row spaces coincide. (If both $\operatorname{im} A \cap \operatorname{im} B = \{\mathbf{0}\}$ and $\operatorname{im} A^{\mathsf{T}} \cap \operatorname{im} B^{\mathsf{T}} = \{\mathbf{0}\}$, then the equality does hold, but that needs more work.)

(d) True. By (4.3.1) with $B = A$, the hypothesis says $\operatorname{im} A \cap \ker A = \{\mathbf{0}\}$. For $k \geq 1$, (4.3.1) applied to A and A^k gives

$$\operatorname{rank}\left(A^{k+1}\right) = \operatorname{rank}\left(A^k\right) - \dim\left(\operatorname{im} A^k \cap \ker A\right).$$

Since $A^k = A \circ A^{k-1}$ as maps, $\operatorname{im} A^k \subseteq \operatorname{im} A$ (Proposition 2.6.15(b)), so $\operatorname{im} A^k \cap \ker A \subseteq \operatorname{im} A \cap \ker A = \{\mathbf{0}\}$, and $\operatorname{rank}\left(A^{k+1}\right) = \operatorname{rank}\left(A^k\right)$. By induction, $\operatorname{rank}\left(A^k\right) = \operatorname{rank} A$ for all $k \geq 1$.

Solution to Exercise 4.3.37

Uses. Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 4.3.20 (Rank is subadditive), Proposition 4.3.16 (Rank-one matrices), Corollary 3.5.25 (The invertible matrix theorem)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17), Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. Multiply out, using $\mathbf{u}\mathbf{v}^{\mathsf{T}}\mathbf{u}\mathbf{v}^{\mathsf{T}} = s\mathbf{u}\mathbf{v}^{\mathsf{T}}$.
Hint 2. For (b), apply $I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}$ to $\mathbf{u}$. For the rank, compare I_n with $\left(I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) - \mathbf{u}\mathbf{v}^{\mathsf{T}}$.
Hint 3. $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$, where $\mathbf{1}$ is the vector of ones.
Sketch. (a) $\left(I + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)\left(I - (1+s)^{-1}\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = I + \mathbf{u}\mathbf{v}^{\mathsf{T}} - (1+s)^{-1}(1+s)\mathbf{u}\mathbf{v}^{\mathsf{T}} = I$. (b) $\left(I + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right)\mathbf{u} = (1+s)\mathbf{u} = \mathbf{0}$ with $\mathbf{u} \neq \mathbf{0}$; and $n = \operatorname{rank} I \leq \operatorname{rank}\left(I + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) + 1$. (c) $cJ = (c\mathbf{1})\mathbf{1}^{\mathsf{T}}$, $s = cn$: invertible iff $1 + cn \neq 0$, inverse $I - c(1+cn)^{-1}J$. $I_4 + J$: invertible over $\mathbb{Q}$ (inverse $I - J/5$), not over $\mathbb{F}_5$.

(a) Let $N = \mathbf{u}\mathbf{v}^{\mathsf{T}}$. Then $N^2 = \mathbf{u}\left(\mathbf{v}^{\mathsf{T}}\mathbf{u}\right)\mathbf{v}^{\mathsf{T}} = sN$ (Technique 4.3.17), so

$$\begin{aligned}(I_n + N)\left(I_n - (1+s)^{-1}N\right) &= I_n + N - (1+s)^{-1}(N + sN) \\ &= I_n + N - N = I_n.\end{aligned}$$

A one-sided inverse of a square matrix is two-sided (Theorem 3.5.23), so $I_n + N$ is invertible with the stated inverse.

(b) Suppose $1 + s = 0$. Then $\mathbf{u} \neq \mathbf{0}$, since otherwise $s = 0$. Now $(I_n + N)\mathbf{u} = \mathbf{u} + \mathbf{u}\left(\mathbf{v}^{\mathsf{T}}\mathbf{u}\right) = (1+s)\mathbf{u} = \mathbf{0}$, so $I_n + N$ has a nonzero vector in its kernel and is not invertible (Corollary 3.5.25). Its rank is at most $n - 1$. Also $\mathbf{v} \neq \mathbf{0}$, since $\mathbf{v}^{\mathsf{T}}\mathbf{u} = -1$, so N has rank one (Proposition 4.3.16), and by Proposition 4.3.20,

$$\begin{aligned}n = \operatorname{rank} I_n &= \operatorname{rank}((I_n + N) + (-N)) \\ &\leq \operatorname{rank}(I_n + N) + 1.\end{aligned}$$

So $\operatorname{rank}(I_n + N) = n - 1$.

(c) $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$, so $cJ = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u} = c\mathbf{1}$ and $\mathbf{v} = \mathbf{1}$, and $s = \mathbf{1}^{\mathsf{T}}(c\mathbf{1}) = cn$, where n means $n \cdot 1 \in F$. By (a) and (b), $I_n + cJ$ is invertible if and only if $1 + cn \neq 0$, and then

$$(I_n + cJ)^{-1} = I_n - \frac{c}{1+cn}J.$$

For $I_4 + J$, $c = 1$ and $1 + cn = 5$. Over $\mathbb{Q}$, $5 \neq 0$, so $(I_4 + J)^{-1} = I_4 - J/5$. Over $\mathbb{F}_5$, $5 = 0$, so $I_4 + J$ is not invertible, and indeed $(I_4 + J)\mathbf{1} = \mathbf{1} + 4\mathbf{1} = 5\mathbf{1} = \mathbf{0}$.

Solution to Exercise 4.3.38

Uses. Proposition 4.3.16 (Rank-one matrices), Example 4.3.18 (Powers of a rank-one matrix), Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17), Cyclicity of trace (Technique 2.4.29)
Hint 1. A rank-one matrix X satisfies $X^2 = (\operatorname{tr} X)X$.
Hint 2. What is the trace of a commutator?
Sketch. $X = AB - BA$ has rank one, so $X^2 = (\operatorname{tr} X)X$ by Example 4.3.18, and $\operatorname{tr} X = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. Example: $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $AB - BA = B$.

Let $X = AB - BA$. Since $\operatorname{rank} X = 1$, Example 4.3.18 gives $X^2 = (\operatorname{tr} X)X$. By linearity and cyclicity of the trace (Proposition 2.4.27, Proposition 2.4.28), $\operatorname{tr} X = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. Hence $X^2 = 0$.

For an example, let $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Then $AB = B$ and $BA = 0$, so $AB - BA = B$, which has rank 1, and indeed $B^2 = 0$.

The hypothesis cannot be weakened to $\operatorname{rank}(AB - BA) \leq 2$: with $A = \mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}$ and $B = \mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}$ in $F^{2\times 2}$, $AB - BA = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, which has rank 2 and square I_2 over every field.

Solution to Exercise 4.3.39

Uses. Proposition 4.3.20 (Rank is subadditive), Proposition 4.3.16 (Rank-one matrices), Lemma 2.4.24 (Matrix units), Proposition 4.3.1 (Invertible factors preserve rank)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. Write $A = (A + B) + (-B)$ and apply subadditivity.
Hint 2. Changing entry (i, j) by c adds $c\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$.
Hint 3. $M_t = (1 - t)I_5 + tJ$, where J is the all-ones matrix, of rank one.
Sketch. (a) $\operatorname{rank} A \leq \operatorname{rank}(A + B) + \operatorname{rank}(-B)$ and symmetrically. (b) The change adds $c\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$, of rank at most 1; apply (a). (c) $M_t = (1 - t)I_5 + tJ$; for $t \neq 1$, $(1 - t)I_5$ has rank 5 and tJ rank at most 1, so $\operatorname{rank} M_t \geq 4$.

(a) $-B$ has the same image as B, so $\operatorname{rank}(-B) = \operatorname{rank} B$. By Proposition 4.3.20, $\operatorname{rank} A = \operatorname{rank}((A + B) + (-B)) \leq \operatorname{rank}(A + B) + \operatorname{rank} B$, so $\operatorname{rank}(A + B) \geq \operatorname{rank} A - \operatorname{rank} B$. Exchanging A and B gives $\operatorname{rank}(A + B) \geq \operatorname{rank} B - \operatorname{rank} A$.

(b) Changing entry (i, j) of A from a_{ij} to $a_{ij} + c$ replaces A by $A' = A + c\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Lemma 2.4.24). The matrix $c\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ has rank 0 (if $c = 0$) or 1 (Proposition 4.3.16). By (a) and Proposition 4.3.20, $\operatorname{rank} A - 1 \leq \operatorname{rank} A' \leq \operatorname{rank} A + 1$.

(c) Let J be the all-ones matrix, $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$, of rank 1. Then $M_t = (1 - t)I_5 + tJ$. For $t \neq 1$, $(1 - t)I_5$ is invertible, of rank 5, and $\operatorname{rank}(tJ) \leq 1$, so by (a), $\operatorname{rank} M_t \geq 5 - 1 = 4$. (For $t = 1$, $M_1 = J$ has rank 1. By Exercise 4.3.37(c), applied to $I_5 + (t/(1 - t))J$, the rank is exactly 4 only for $t = -1/4$.)

Solution to Exercise 4.3.40

Uses. Proposition 4.3.20 (Rank is subadditive), Proposition 4.3.16 (Rank-one matrices), Proposition 4.3.1 (Invertible factors preserve rank), Theorem 4.3.8 (Row rank equals column rank), Theorem 3.5.13 (Second Fundamental Theorem, completed), Proposition 2.6.15 (Kernels and images of composites), Theorem 3.4.24 (Polynomial interpolation), Corollary 3.5.16 (Lists of vectors in F^m), Theorem 2.4.34 (Block multiplication)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. $(i - j)^2 = i^2 - 2ij + j^2$. Let $\mathbf{1}$, $\mathbf{w}$, $\mathbf{q}$ be the vectors with entries 1, i, i^2.
Hint 2. Write $A = XMX^{\mathsf{T}}$ with $X = (\mathbf{1}\ \mathbf{w}\ \mathbf{q}) \in F^{n\times 3}$ and a 3×3 matrix M. When is M invertible, and when are the columns of X independent?

Hint 3. Over $\mathbb{F}_2$, $-2 = 0$ and $i^2 = i$.
Sketch. (a) $A = \mathbf{q}\mathbf{1}^\mathsf{T} - 2\mathbf{w}\mathbf{w}^\mathsf{T} + \mathbf{1}\mathbf{q}^\mathsf{T}$. (b) $A = XMX^\mathsf{T}$ with $M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, invertible when $2 \neq 0$. Rows $1, 2, 3$ of X form the Vandermonde matrix for the distinct nodes $1, 2, 3$, so rank $X = 3$, and T_{X^T} is onto F^3. Hence $\operatorname{im} A = \operatorname{im}(XM) = \operatorname{im} X$, of dimension 3. (c) Over $\mathbb{F}_2$, $a_{ij} = i + j$, so $A = \mathbf{w}\mathbf{1}^\mathsf{T} + \mathbf{1}\mathbf{w}^\mathsf{T}$ has rank 2.

Let $\mathbf{1}, \mathbf{w}, \mathbf{q} \in F^n$ have entries 1, i, and i^2 in position i.

(a) Since $(i-j)^2 = i^2 \cdot 1 - 2ij + 1 \cdot j^2$, entry by entry

$$A = \mathbf{q}\mathbf{1}^\mathsf{T} - 2\mathbf{w}\mathbf{w}^\mathsf{T} + \mathbf{1}\mathbf{q}^\mathsf{T}.$$

Each term has rank at most one (Proposition 4.3.16), so rank $A \leq 3$ by Proposition 4.3.20.

(b) Let $X = (\mathbf{1} \;\; \mathbf{w} \;\; \mathbf{q}) \in F^{n\times 3}$ and $M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. By block multiplication (Theorem 2.4.34), $XMX^\mathsf{T} = \mathbf{q}\mathbf{1}^\mathsf{T} - 2\mathbf{w}\mathbf{w}^\mathsf{T} + \mathbf{1}\mathbf{q}^\mathsf{T} = A$. Suppose $2 \neq 0$ in F.

M is invertible. $M^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, which is invertible since $4 \neq 0$; so M has the inverse $M\left(M^2\right)^{-1}$.

rank $X = 3$. The nodes $1, 2, 3$ of F are distinct, since their differences are 1 and 2, both nonzero. The first three rows of X form the matrix $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$, which is invertible by Theorem 3.4.24, so its columns are independent. A relation among the columns of X restricts to a relation among the columns of this 3×3 matrix, so the columns of X are independent, and rank $X = 3$ (Corollary 3.5.16(a)).

Conclusion. By Theorem 4.3.8, rank $X^\mathsf{T} = 3$, so $T_{X^\mathsf{T}} : F^n \to F^3$ is surjective. Hence $\operatorname{im} A = \operatorname{im}\left(XMX^\mathsf{T}\right) = \operatorname{im}(XM)$ (Proposition 2.6.15(b)), and $\operatorname{rank}(XM) = \operatorname{rank} X = 3$ (Proposition 4.3.1). So rank $A = 3$.

(c) Over $\mathbb{F}_2$, $-2 = 0$, and i^2 has the same parity as i, so $\mathbf{q} = \mathbf{w}$ and $a_{ij} = i + j$. So $A = \mathbf{w}\mathbf{1}^\mathsf{T} + \mathbf{1}\mathbf{w}^\mathsf{T}$ has rank at most 2. Rows 1 and 2 are $(0, 1, 0, 1, \ldots)$ and $(1, 0, 1, 0, \ldots)$, which are independent, so rank $A \geq 2$ by Theorem 4.3.8. Hence rank $A = 2$ over $\mathbb{F}_2$.

Solution to Exercise 4.3.41

Uses. Proposition 4.3.16 (Rank-one matrices)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^\mathsf{T}$ (Technique 4.3.17)
Hint 1. For "only if", write $A = \mathbf{u}\mathbf{v}^\mathsf{T}$, so $a_{ij} = u_i v_j$.
Hint 2. For "if", assume $A \neq 0$ and fix an entry $a_{kl} \neq 0$. Solve the condition for a_{ij}.
Sketch. If $A = \mathbf{u}\mathbf{v}^\mathsf{T}$, both sides are $u_i u_k v_j v_l$. Conversely, if $a_{kl} \neq 0$, the condition gives $a_{ij} = a_{il}\left(a_{kj}/a_{kl}\right)$, so $A = \mathbf{u}\mathbf{v}^\mathsf{T}$ with $\mathbf{u}$ column l and $v_j = a_{kj}/a_{kl}$. (b) $4 - t^2 = 0$ and $2t^2 - 8 = 0$, and the remaining condition holds for all t: $t = \pm 2$.

(a) *Only if.* If $A = 0$, both sides are 0. If rank $A = 1$, then $A = \mathbf{u}\mathbf{v}^\mathsf{T}$ (Proposition 4.3.16), so $a_{ij} = u_i v_j$, and both sides equal $u_i u_k v_j v_l$.

If. Suppose the condition holds. If $A = 0$, then rank $A = 0$. Otherwise fix k, l with $a_{kl} \neq 0$. For all i, j, the condition gives

$$a_{ij} = a_{il} \cdot \frac{a_{kj}}{a_{kl}}.$$

Let $\mathbf{u} = (a_{1l}, \ldots, a_{ml})$, column l of A, and $\mathbf{v} = (a_{k1}, \ldots, a_{kn})/a_{kl}$. Then $a_{ij} = u_i v_j$ for all i, j, so $A = \mathbf{u}\mathbf{v}^\mathsf{T}$, and rank $A \leq 1$ by Proposition 4.3.16 (in fact $= 1$, since $u_k = a_{kl} \neq 0$ and $v_l = 1$).

(b) The matrix is nonzero, so it has rank 1 exactly when all the conditions in (a) hold. With rows $1, 2$ and columns $(1, 2)$, $(1, 3)$, $(2, 3)$ they read $1 \cdot 4 = t \cdot t$, $1 \cdot 2t = 2 \cdot t$, and $t \cdot 2t = 2 \cdot 4$ (the other choices of indices repeat these or are trivial). The second always holds, and the first and third both say $t^2 = 4$. So $t = 2$ or $t = -2$, giving $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \end{pmatrix}$.

Solution to Exercise 4.3.42

Uses. Proposition 4.3.20 (Rank is subadditive), Exercise 4.3.30, Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. One inequality always holds: $I_n = A + (I_n - A)$.
Hint 2. If $A^2 = A$, then $A(I_n - A) = 0$. Which inequality bounds $\operatorname{rank} A + \operatorname{rank}(I_n - A)$ from above?
Hint 3. Conversely, show $\ker A \subseteq \operatorname{im}(I_n - A)$, compare dimensions, and conclude $A(I_n - A) = 0$.
Sketch. Always $n = \operatorname{rank} I_n \le \operatorname{rank} A + \operatorname{rank}(I_n - A)$. If $A^2 = A$, then $A(I_n - A) = 0$, and Sylvester gives $\operatorname{rank} A + \operatorname{rank}(I_n - A) \le n$. Conversely, $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = (I_n - A)\mathbf{x}$, so $\ker A \subseteq \operatorname{im}(I_n - A)$; the hypothesis says $\dim \ker A = n - \operatorname{rank} A = \operatorname{rank}(I_n - A)$, so $\ker A = \operatorname{im}(I_n - A)$, that is, $A(I_n - A) = 0$.

By Proposition 4.3.20, $n = \operatorname{rank} I_n = \operatorname{rank}(A + (I_n - A)) \le \operatorname{rank} A + \operatorname{rank}(I_n - A)$ for every A.

If. Suppose $A^2 = A$. Then $A(I_n - A) = A - A^2 = 0$, so by Exercise 4.3.30(b), $\operatorname{rank} A + \operatorname{rank}(I_n - A) \le n$. With the inequality above, equality holds.

Only if. Suppose $\operatorname{rank} A + \operatorname{rank}(I_n - A) = n$. If $A\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = \mathbf{x} - A\mathbf{x} = (I_n - A)\mathbf{x}$, so $\ker A \subseteq \operatorname{im}(I_n - A)$. By Theorem 4.2.11 and the hypothesis, $\dim \ker A = n - \operatorname{rank} A = \operatorname{rank}(I_n - A) = \dim \operatorname{im}(I_n - A)$. By Proposition 4.1.21(b), $\ker A = \operatorname{im}(I_n - A)$. So $A(I_n - A)\mathbf{x} = \mathbf{0}$ for every $\mathbf{x}$, that is, $A - A^2 = 0$.

So a matrix is idempotent exactly when the ranks of A and $I_n - A$ are as small as subadditivity allows. Geometrically, F^n then splits into $\operatorname{im} A$, on which A acts as the identity, and $\ker A = \operatorname{im}(I_n - A)$; this is the decomposition of Section 4.4.

Solution to Exercise 4.3.43

Uses. Theorem 4.3.21 (Column–row factorization), Proposition 3.5.9 (Left inverses), Proposition 3.5.8 (Right inverses from the pivots), Proposition 2.4.28 (Cyclicity of the trace), Proposition 4.3.23 (Factorizations through F^k)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. C has independent columns and R has independent rows. What one-sided inverses do they have?
Hint 2. Multiply $CRCR = CR$ on the left by a left inverse of C and on the right by a right inverse of R.
Hint 3. $\operatorname{tr}(CR) = \operatorname{tr}(RC)$.
Sketch. $\operatorname{rank} C = r$ = number of columns, so $LC = I_r$ for some L; $\operatorname{rank} R = r$ = number of rows, so $RZ = I_r$ for some Z. Multiplying $C(RC)R = CR$ by L on the left and Z on the right gives $RC = I_r$. Then $\operatorname{tr} A = \operatorname{tr}(CR) = \operatorname{tr}(RC) = \operatorname{tr} I_r = r \cdot 1$. Over $\mathbb{F}_2$, I_2 and 0 both have trace 0.

(a) By Theorem 4.3.21, $\operatorname{rank} C = r$, the number of columns of C, and $\operatorname{rank} R = r$, the number of rows of R. By Proposition 3.5.9 and Proposition 3.5.8, there are $L \in F^{r \times n}$ with $LC = I_r$ and $Z \in F^{n \times r}$ with $RZ = I_r$. Since $A^2 = A$, $C(RC)R = CR$. Multiplying on the left by L and on the right by Z,

$$\begin{aligned} RC &= (LC)(RC)(RZ) = L(CRCR)Z \\ &= L(CR)Z = (LC)(RZ) = I_r. \end{aligned}$$

(The same argument works for any factorization $A = XY$ with inner size r, by Proposition 4.3.23.)

(b) By Proposition 2.4.28, $\operatorname{tr} A = \operatorname{tr}(CR) = \operatorname{tr}(RC) = \operatorname{tr} I_r = r \cdot 1$.

(c) In $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, $r \cdot 1$ is the integer r, so $\operatorname{tr} A = \operatorname{rank} A$ by (b); the idempotent $A = 0$ has trace and rank 0. Over $\mathbb{F}_2$, the idempotents I_2 and 0 have trace $1 + 1 = 0$ and 0, but ranks 2 and 0. In general, (b) determines $\operatorname{rank} A$ only modulo the characteristic of F. For example, the rank of the idempotent $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ is $1 = \operatorname{tr} A$, which (b) predicts over any field.

Solution to Exercise 4.3.44

Uses. Proposition 4.3.23 (Factorizations through F^k), Proposition 2.4.8 (Four views of the product), Proposition 3.5.9 (Left inverses), Proposition 3.5.8 (Right inverses from the pivots), Corollary 3.5.21 (Invertibility and row equivalence), Corollary 3.3.18 (RREF and row equivalence)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. By Proposition 4.3.23(b), the columns of C and of C' are both bases of $\operatorname{im} A$. Express one basis in terms of the other.
Hint 2. C has a left inverse, so $CX = CY$ implies $X = Y$.

Hint 3. For (b), $R' = P^{-1}R$ says that R' and R are row equivalent.
Sketch. The columns of C form a basis of $\operatorname{im} A$ and the columns of C' lie in it, so $C' = CP$ for a unique P; likewise $C = C'Q$, and $CPQ = C$ gives $PQ = I_r$ by cancelling C, so P is invertible. Then $CR = CPR'$ gives $R = PR'$. (b) If R' is in RREF, then $R' = P^{-1}R$ is row equivalent to the RREF matrix R, so $R' = R$, and then $P = I_r$ and $C' = C$.

By Proposition 4.3.23(b), the columns of C form a basis of $\operatorname{im} A$, and so do the columns of C'. Since $\operatorname{rank} C = r$ is the number of columns of C, C has a left inverse L (Proposition 3.5.9), so $CX = CY$ implies $X = LCX = LCY = Y$. That is, C can be cancelled on the left.

(a) Each column of C' lies in $\operatorname{im} A$, so it is a combination of the columns of C; collecting the coefficients into the columns of $P \in F^{r\times r}$ gives $C' = CP$ (Proposition 2.4.8(b)). Symmetrically, $C = C'Q$ for some Q. Then $C(PQ) = C'Q = C = CI_r$, so $PQ = I_r$ by cancellation. Likewise $C'(QP) = C'$ and C' can be cancelled, so $QP = I_r$. So P is invertible. It is unique: if also $C' = CP'$, then $CP = CP'$ gives $P = P'$. Finally, $CR = A = C'R' = CPR'$, so $R = PR'$ by cancellation, and $R' = P^{-1}R$.

(b) Let R be the first r rows of $\operatorname{rref}(A)$ and suppose R' is in RREF. By (a), $R' = P^{-1}R$ with P^{-1} invertible, so $R' \sim R$ (Corollary 3.5.21); both are in RREF, so $R' = R$ by Corollary 3.3.18(c). Then $R = PR$, and since R has a right inverse Z (Proposition 3.5.8), $I_r = RZ = PRZ = P$. So $C' = C$. (Problem 3.17 proved the uniqueness of C for this R by the same cancellation.)

Solution to Exercise 4.3.45

Uses. Proposition 4.3.23 (Factorizations through F^k), Proposition 3.5.9 (Left inverses), Proposition 3.5.8 (Right inverses from the pivots), Theorem 4.3.3 (Rank of a product), Proposition 4.3.1 (Invertible factors preserve rank), Corollary 3.5.25 (The invertible matrix theorem)
Hint 1. $A^2 = C(RC)R$. Use one-sided inverses of C and R to show $\operatorname{rank}(RC) = \operatorname{rank}(A^2)$.
Hint 2. For (b), compute AX and XA first; both equal $C(RC)^{-1}R$.
Hint 3. For (c), $A = AXA = A^2X$.
Sketch. (a) With $LC = I_r$ and $RZ = I_r$, $RC = L(CRCR)Z = LA^2Z$, so $\operatorname{rank}(RC) \geq \operatorname{rank}(A^2) = r$ and RC is invertible. (b) $AX = CRC(RC)^{-2}R = C(RC)^{-1}R$ and $XA = C(RC)^{-2}RCR = C(RC)^{-1}R$; then $AXA = C(RC)^{-1}RCR = CR$ and $XAX = C(RC)^{-1}RC(RC)^{-2}R = X$. (c) $\operatorname{rank} A = \operatorname{rank}(A^2X) \leq \operatorname{rank}(A^2) \leq \operatorname{rank} A$.

By Proposition 4.3.23(b), C has r independent columns and R has r independent rows, so $\operatorname{rank} C = \operatorname{rank} R = r$, and there are L and Z with $LC = I_r$ and $RZ = I_r$ (Proposition 3.5.9, Proposition 3.5.8).

(a) $A^2 = CRCR$, so $LA^2Z = (LC)(RC)(RZ) = RC$. By Theorem 4.3.3, $\operatorname{rank}(RC) = \operatorname{rank}(LA^2Z) \leq \operatorname{rank}(A^2)$, and also $\operatorname{rank}(A^2) = \operatorname{rank}(C(RC)R) \leq \operatorname{rank}(RC)$. So $\operatorname{rank}(RC) = \operatorname{rank}(A^2) = r$, and the $r \times r$ matrix RC is invertible (Corollary 3.5.25).

(b) Let $K = RC$, invertible by (a), and $X = CK^{-2}R$. Then

$$AX = CRCK^{-2}R = CKK^{-2}R = CK^{-1}R,$$
$$XA = CK^{-2}RCR = CK^{-2}KR = CK^{-1}R,$$

so $AX = XA$. Next, $AXA = CK^{-1}RCR = CK^{-1}KR = CR = A$, and

$$XAX = CK^{-1}RCK^{-2}R = CK^{-1}KK^{-2}R = CK^{-2}R = X.$$

(c) If $AXA = A$ and $AX = XA$, then $A = A(XA) = A(AX) = A^2X$. By Theorem 4.3.3, $\operatorname{rank} A = \operatorname{rank}(A^2X) \leq \operatorname{rank}(A^2) \leq \operatorname{rank} A$. So $\operatorname{rank}(A^2) = \operatorname{rank} A$.

The matrix X is called the *group inverse* of A. If A is invertible, then $r = n$, C and R are invertible, and $X = C(RC)^{-2}R = CC^{-1}R^{-1}C^{-1}R^{-1}R = A^{-1}$. For the projection $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$, $RC = (1)$ and $X = A$. For $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\operatorname{rank}(N^2) = 0 < 1$, and no such X exists.

Section 4.4

Solution to Exercise 4.4.11

Uses. Theorem 4.4.7 (Criteria for direct sums), Proposition 1.1.17 (No zero divisors)

Techniques. Compare components (Technique 4.4.8)
Hint 1. For (a), compare coordinates: each of the three vectors has a coordinate equal to 1 where another has 0.
Hint 2. For (b), use (IV) of Theorem 4.4.7: it remains to decide when $U_3 \cap (U_1 + U_2) = \{\mathbf{0}\}$. Solve $c(0,1,1) = a(1,1,0) + b(1,0,1)$.
Sketch. (a) If $a(1,1,0) = b(1,0,1)$, the second coordinates give $a = 0$ and the third give $b = 0$; the other pairs are similar. (b) $c(0,1,1) = (a+b, a, b)$ forces $a = b = c$ and $2c = 0$. If $1+1 \neq 0$, then $c = 0$, so (IV) holds. Over $\mathbb{F}_2$, $(1,1,0) + (1,0,1) + (0,1,1) = \mathbf{0}$ violates (II).

(a) If $a(1,1,0) = b(1,0,1)$, then comparing second coordinates gives $a = 0$, and comparing third coordinates gives $b = 0$, so the common vector is $\mathbf{0}$. In the same way, comparing coordinates shows $U_1 \cap U_3 = \{\mathbf{0}\}$ (from $a(1,1,0) = c(0,1,1)$, the first coordinates give $a = 0$ and the third give $c = 0$) and $U_2 \cap U_3 = \{\mathbf{0}\}$ (from $b(1,0,1) = c(0,1,1)$, the first coordinates give $b = 0$ and the second give $c = 0$).

(b) We check (IV) of Theorem 4.4.7. By (a), $U_2 \cap U_1 = \{\mathbf{0}\}$. Suppose $c(0,1,1) \in U_1 + U_2$, say

$$(0, c, c) = a(1,1,0) + b(1,0,1) = (a+b, a, b).$$

Then $a = c$, $b = c$, and $a + b = 0$, so $(1+1)c = 0$. If $1+1 \neq 0$, then $c = 0$ (Proposition 1.1.17), so $U_3 \cap (U_1 + U_2) = \{\mathbf{0}\}$ and the sum is direct. Over $\mathbb{F}_2$, taking $c = 1$ gives

$$(1,1,0) + (1,0,1) + (0,1,1) = (0,0,0),$$

a decomposition of $\mathbf{0}$ with nonzero components, so (II) fails and the sum is not direct.

So the pairwise intersections are trivial over every field, but directness depends on the field. Over $\mathbb{F}_2$, the three lines lie in the plane $x_1 + x_2 + x_3 = 0$, like the three lines of Non-example 4.4.6(b).

Solution to Exercise 4.4.23

Uses. Example 4.2.25 (Kernels of linear functionals), Proposition 4.4.16 (Complements exist), Proposition 4.4.21 (Two out of three), Proposition 1.2.35 (Zero products)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. U is the kernel of a nonzero linear functional. What is $\dim U$, and what must the dimension of a complement be?
Hint 2. For a nonzero $\mathbf{v}$, when is $U \cap \operatorname{span}(\mathbf{v}) = \{\mathbf{0}\}$? Then apply Proposition 4.4.21.
Sketch. $\dim U = 3$ (Example 4.2.25), so every complement is a line $\operatorname{span}(\mathbf{v})$, $\mathbf{v} \neq \mathbf{0}$ (Proposition 4.4.16(b)). If $\mathbf{v} \in U$, then $\operatorname{span}(\mathbf{v}) \subseteq U$; if $\mathbf{v} \notin U$, then $a\mathbf{v} \in U$ forces $a = 0$. So $\operatorname{span}(\mathbf{v})$ is a complement iff $\mathbf{v} \notin U$, by (i) and (iii) of Proposition 4.4.21. (b) $(1,1,1,1) \in U$ iff $1+1+1+1 = 0$: no over $\mathbb{Q}$ and $\mathbb{F}_3$, yes over $\mathbb{F}_2$.

(a) U is the kernel of the nonzero linear functional $\mathbf{x} \mapsto x_1 + x_2 + x_3 + x_4$, so $\dim U = 3$ (Example 4.2.25). By Proposition 4.4.16(b), every complement of U has dimension $4 - 3 = 1$, so it is $\operatorname{span}(\mathbf{v})$ for some $\mathbf{v} \neq \mathbf{0}$. Conversely, let $\mathbf{v} \neq \mathbf{0}$. If $\mathbf{v} \in U$, then $\operatorname{span}(\mathbf{v}) \subseteq U$ meets U in a nonzero vector and is not a complement. If $\mathbf{v} \notin U$ and $a\mathbf{v} \in U$ with $a \neq 0$, then $\mathbf{v} = a^{-1}(a\mathbf{v}) \in U$, a contradiction; so $U \cap \operatorname{span}(\mathbf{v}) = \{\mathbf{0}\}$. Since also $\dim U + \dim \operatorname{span}(\mathbf{v}) = 3 + 1 = 4$, Proposition 4.4.21 gives $F^4 = U \oplus \operatorname{span}(\mathbf{v})$. So the complements of U are exactly the lines $\operatorname{span}(\mathbf{v})$ with $\mathbf{v} \notin U$ (and then automatically $\mathbf{v} \neq \mathbf{0}$).

(b) By (a), $\operatorname{span}((1,1,1,1))$ is a complement exactly when $(1,1,1,1) \notin U$, that is, when $1+1+1+1 \neq 0$ in F. In $\mathbb{Q}$, $1+1+1+1 = 4 \neq 0$, so yes. In $\mathbb{F}_2$, $1+1+1+1 = 0$, so no: there $(1,1,1,1) \in U$. In $\mathbb{F}_3$, $1+1+1+1 = 1 \neq 0$, so yes.

Solution to Exercise 4.4.30

Uses. Definition 4.4.24 (Projection along a complement), Theorem 4.4.7 (Criteria for direct sums), Example 4.2.25 (Kernels of linear functionals)
Techniques. Split the space with an idempotent (Technique 4.4.28)
Hint 1. Subtract from $\mathbf{x}$ the multiple of $(1,1,1)$ that makes the coordinates sum to 0.
Hint 2. Over $\mathbb{F}_3$, is $(1,1,1)$ in U?
Sketch. With $m = (x_1 + x_2 + x_3)/3$, $\mathbf{x} = (\mathbf{x} - m(1,1,1)) + m(1,1,1)$ is a decomposition, and $t(1,1,1) \in U$ forces $3t = 0$. So $\operatorname{Pr}_{W,U}$ has all entries $1/3$ and $\operatorname{Pr}_{U,W} = I_3 - \operatorname{Pr}_{W,U}$. Over $\mathbb{F}_3$, $1+1+1 = 0$ puts W inside U; the division by 3 is impossible.

(a) For $\mathbf{x} \in \mathbb{Q}^3$, let $m = (x_1 + x_2 + x_3)/3$. The coordinates of $\mathbf{x} - m(1,1,1)$ sum to $(x_1 + x_2 + x_3) - 3m = 0$, so

$$\mathbf{x} = (\mathbf{x} - m(1,1,1)) + m(1,1,1)$$

is a decomposition of $\mathbf{x}$ along U and W. If $t(1,1,1) \in U$, then $3t = 0$ and $t = 0$, so $U \cap W = \{\mathbf{0}\}$. By Theorem 4.4.7, $\mathbb{Q}^3 = U \oplus W$, with $\mathrm{Pr}_{W,U}(\mathbf{x}) = m(1,1,1)$ and $\mathrm{Pr}_{U,W}(\mathbf{x}) = \mathbf{x} - m(1,1,1)$. Their matrices are

$$\mathrm{Pr}_{W,U} : \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathrm{Pr}_{U,W} : \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

(b) In $\mathbb{F}_3$, $1 + 1 + 1 = 0$, so $(1,1,1) \in U$ and $W \subseteq U$. Then $U + W = U$, which has dimension 2 (Example 4.2.25), so $U + W \neq \mathbb{F}_3^3$. The step that fails is the division by 3 in the definition of m, since $3 = 0$ in $\mathbb{F}_3$. Correspondingly, the uniqueness argument fails too: $3t = 0$ holds for every $t \in \mathbb{F}_3$.

Solution to Exercise 4.4.33

Uses. Theorem 4.4.19 (Dimension of a sum), Theorem 4.2.11 (Rank–nullity), Theorem 4.2.7 (The image of T_A), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Write a general vector of U as $(a, a+b, b+c, c)$ and substitute it into the two equations that define W.
Hint 2. Over $\mathbb{Q}$ the first equation becomes $2(a+b+c) = 0$. What does it become over $\mathbb{F}_2$?
Sketch. Over every field, the spanning vectors of U are independent, so $\dim U = 3$, and W is cut out by two independent equations, so $\dim W = 2$. A vector $(a, a+b, b+c, c)$ of U lies in W iff $2(a+b+c) = 0$ and $a = b + c$. Over $\mathbb{Q}$ this forces $a = 0$ and $c = -b$: $U \cap W = \mathrm{span}((0,1,0,-1))$, and $\dim(U + W) = 4$. Over $\mathbb{F}_2$ the first condition is empty, $\dim(U \cap W) = 2$, so $W \subseteq U$ and $\dim(U + W) = 3$.

(a) If $a(1,1,0,0) + b(0,1,1,0) + c(0,0,1,1) = (a, a+b, b+c, c) = \mathbf{0}$, then $a = c = 0$ from the first and last coordinates, and then $b = 0$. So the spanning list of U is independent and $\dim U = 3$; this works over every field. The subspace W is the kernel of the linear map $F^4 \to F^2$ with matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

Its first two columns $(1,1)$ and $(1,0)$ are independent, so it has rank 2 (Theorem 4.2.7), and $\dim W = 4 - 2 = 2$ by Theorem 4.2.11.

A general vector of U is $(a, a+b, b+c, c)$. It lies in W if and only if

$$2a + 2b + 2c = 0 \quad \text{and} \quad a = b + c.$$

Over $\mathbb{Q}$, the first condition says $a + b + c = 0$, and together with $a = b + c$ it gives $2a = 0$, so $a = 0$ and $c = -b$. Hence $U \cap W = \{(0, b, 0, -b) \mid b \in \mathbb{Q}\} = \mathrm{span}((0,1,0,-1))$, of dimension 1. By Theorem 4.4.19, $\dim(U + W) = 3 + 2 - 1 = 4$, so $U + W = \mathbb{Q}^4$.

(b) Over $\mathbb{F}_2$ the same arguments give $\dim U = 3$ and $\dim W = 2$ (the matrix is now $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$, with the same first two columns). But $2 = 0$ in $\mathbb{F}_2$, so the first condition holds for every vector of U, and $U \cap W$ consists of the vectors with $a = b + c$:

$$(b + c, c, b + c, c) = b(1,0,1,0) + c(1,1,1,1), \quad b, c \in \mathbb{F}_2.$$

So $\dim(U \cap W) = 2 = \dim W$, which means $W \subseteq U$, and $\dim(U + W) = 3 + 2 - 2 = 3$, so $U + W = U$. This is visible directly: over $\mathbb{F}_2$, each spanning vector of U has an even number of entries 1, so U is the 3-dimensional subspace $x_1 + x_2 + x_3 + x_4 = 0$, and W lies in it by definition. (Over $\mathbb{Q}$, U is the subspace $x_1 - x_2 + x_3 - x_4 = 0$.)

Solution to Exercise 4.4.34

Uses. Non-example 4.4.6 (Sums that are not direct), Proposition 4.4.9 (Grouping summands), Example 4.4.17 (Complements are not unique), Proposition 4.4.16 (Complements exist), Proposition 4.4.21 (Two out of three)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Dimension counting (Technique 4.2.24)
Hint 1. Three of the statements are false, and lines through $\mathbf{0}$ in $\mathbb{R}^2$ give counterexamples to all three.
Hint 2. For the true ones, use Proposition 4.4.9, Proposition 4.4.16(b), and Proposition 4.4.21.

Sketch. (a) False: three lines in $\mathbb{R}^2$. (b) True: a selection of independent subspaces is independent. (c) False: the y-axis and the line $y = x$ are both complements of the x-axis. (d) True: both dimensions are $\dim V - \dim U$. (e) False: U, W the axes of $\mathbb{R}^2$ and X the line $y = x$. (f) True: two out of three.

(a) False. In $\mathbb{R}^2$, the lines $\operatorname{span}(\mathbf{e}_1)$, $\operatorname{span}(\mathbf{e}_2)$, $\operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$ meet pairwise in $\{\mathbf{0}\}$, but their sum is not direct (Non-example 4.4.6(b)).

(b) True, by the last part of Proposition 4.4.9: a selection of independent subspaces is independent.

(c) False. In $\mathbb{R}^2$, with $U = \operatorname{span}(\mathbf{e}_1)$, both $W = \operatorname{span}(\mathbf{e}_2)$ and $X = \operatorname{span}((1,1))$ are complements of U (Example 4.4.17), and $W \neq X$.

(d) True. By Proposition 4.4.16(b), $\dim W = \dim V - \dim U = \dim X$.

(e) False. In $\mathbb{R}^2$, let $U = \operatorname{span}(\mathbf{e}_1)$, $W = \operatorname{span}(\mathbf{e}_2)$, and $X = \operatorname{span}((1,1))$. Then $\mathbb{R}^2 = U \oplus W$, but $X \cap U = X \cap W = \{\mathbf{0}\}$, so $(X \cap U) \oplus (X \cap W) = \{\mathbf{0}\} \neq X$. The components of a vector of X lie in U and W, but not in X.

(f) True. Conditions (ii) and (iii) of Proposition 4.4.21 hold, so $V = U \oplus W$.

Solution to Exercise 4.4.35

Uses. Definition 4.2.1 (Restriction), Theorem 4.4.7 (Criteria for direct sums), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Test injectivity at $\mathbf{0}$ (Technique 2.6.12)
Hint 1. For (a), $T(\mathbf{u}_1 + \cdots + \mathbf{u}_k) = T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k)$.
Hint 2. For (b), use criterion (II): pull a decomposition of $\mathbf{0}$ back through T.
Sketch. (a) Both sides consist of the vectors $T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k)$ with $\mathbf{u}_i \in U_i$. (b) If $T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k) = \mathbf{0}$, then $T\left(\sum_i \mathbf{u}_i\right) = \mathbf{0}$, so $\sum_i \mathbf{u}_i = \mathbf{0}$, so every $\mathbf{u}_i = \mathbf{0}$ and every $T(\mathbf{u}_i) = \mathbf{0}$. (c) $T(x, y) = x + y$ maps both axes of F^2 onto F.

(a) An element of the left side is $T(\mathbf{u}_1 + \cdots + \mathbf{u}_k)$ with $\mathbf{u}_i \in U_i$, and by linearity this equals $T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k)$, a typical element of the right side. Reading the same equation backwards gives the other containment.

(b) We check (II) of Theorem 4.4.7. Let $T(\mathbf{u}_1) + \cdots + T(\mathbf{u}_k) = \mathbf{0}$ with $\mathbf{u}_i \in U_i$. Then $T(\mathbf{u}_1 + \cdots + \mathbf{u}_k) = \mathbf{0}$, so $\mathbf{u}_1 + \cdots + \mathbf{u}_k = \mathbf{0}$ since T is injective (Theorem 2.6.11). As $U_1, \ldots, U_k$ are independent, every $\mathbf{u}_i = \mathbf{0}$, so every $T(\mathbf{u}_i) = \mathbf{0}$.

(c) Let $T : F^2 \to F$, $T(x, y) = x + y$, and $U_1 = \operatorname{span}(\mathbf{e}_1)$, $U_2 = \operatorname{span}(\mathbf{e}_2)$, which are independent. Both images are all of F, and $F + F$ is not direct, since $F \cap F = F \neq \{\mathbf{0}\}$.

Solution to Exercise 4.4.36

Uses. Example 4.4.5 (Direct sums), Definition 4.4.24 (Projection along a complement), Theorem 4.1.10 (Independent lists are short)
Techniques. Compare components (Technique 4.4.8)
Hint 1. Solve $(x, y) = a(1, i) + b(1, -i)$ for a and b.
Hint 2. Compute $R(1, i)$ and $R(1, -i)$.
Sketch. The vectors $(1, i)$ and $(1, -i)$ are not proportional, so they form a basis and $\mathbb{C}^2 = U \oplus W$. Solving gives $a = (x - iy)/2$, so $\operatorname{Pr}_{U,W}(x, y) = ((x - iy)/2, (ix + y)/2)$ and $P = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$. (b) $R(1, i) = (-i, 1) = -i(1, i)$ and $R(1, -i) = (i, 1) = i(1, -i)$, so $\lambda = -i$ and $\mu = i$.

(a) If $(1, -i) = c(1, i)$, then $c = 1$ and $-i = i$, which is false; so $(1, i), (1, -i)$ is independent, hence a basis of $\mathbb{C}^2$ (Theorem 4.1.10(b)), and $\mathbb{C}^2 = U \oplus W$ by Example 4.4.5(a). To decompose (x, y), solve $a(1, i) + b(1, -i) = (x, y)$, that is, $a + b = x$ and $i(a - b) = y$. Then $a - b = -iy$, so

$$a = \frac{x - iy}{2}, \quad b = \frac{x + iy}{2}.$$

Hence $\operatorname{Pr}_{U,W}(x, y) = a(1, i) = ((x - iy)/2, (ix + y)/2)$, and

$$P = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad P^2 = \frac{1}{4}\begin{pmatrix} 1 + (-i)i & -i - i \\ i + i & i(-i) + 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix} = P.$$

(b) $R(1, i) = (-i, 1) = -i(1, i)$, since $(-i)i = 1$; and $R(1, -i) = (i, 1) = i(1, -i)$, since $i(-i) = 1$. By linearity, $R(\mathbf{u}) = -i\mathbf{u}$ for $\mathbf{u} \in U$ and $R(\mathbf{w}) = i\mathbf{w}$ for $\mathbf{w} \in W$, so $\lambda = -i$ and $\mu = i$.

Over $\mathbb{R}$, the quarter turn maps no line through $\mathbf{0}$ into itself. Over $\mathbb{C}$, it splits $\mathbb{C}^2$ into two lines on which it acts as multiplication by a scalar. This is diagonalization (Section 7.4) in its simplest form.

Solution to Exercise 4.4.37

> *Uses.* Non-example 4.4.6 (Sums that are not direct), Theorem 4.4.7 (Criteria for direct sums), Theorem 4.4.13 (Bases and dimension of a sum)
> *Techniques.* Induction on the number of summands (Technique 4.4.10)
> *Hint 1.* Test the claim on three lines in $\mathbb{R}^2$ before looking for the error.
> *Hint 2.* Where does a combination of entries of $\mathcal{B}_1$ and $\mathcal{B}_2$ lie?
> *Sketch.* The claim is false: three lines in $\mathbb{R}^2$ that meet pairwise in $\mathbf{0}$ have dimensions adding up to $3 > 2$. The error is "it lies in U_1 or in U_2": a combination of entries of $\mathcal{B}_1$ and $\mathcal{B}_2$ lies in $U_1 + U_2$, which is usually much larger than $U_1 \cup U_2$. The argument is correct under criterion (IV), $U_3 \cap (U_1 + U_2) = \{\mathbf{0}\}$.

The claim is false. The lines $\operatorname{span}(\mathbf{e}_1)$, $\operatorname{span}(\mathbf{e}_2)$, and $\operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$ in $\mathbb{R}^2$ meet pairwise in $\{\mathbf{0}\}$, their dimensions add up to 3, and their sum is $\mathbb{R}^2$, of dimension 2.

The error is the sentence "it lies in U_1 or in U_2". A combination of entries of $\mathcal{B}_1$ and $\mathcal{B}_2$ is a vector of U_1 plus a vector of U_2, so it lies in $U_1 + U_2$, not in $U_1 \cup U_2$. To conclude that both sides vanish, the argument needs $U_3 \cap (U_1 + U_2) = \{\mathbf{0}\}$. In the example, $\mathbf{e}_1 + \mathbf{e}_2$ lies in U_3 and also in $U_1 + U_2$, and the argument breaks exactly there.

With the hypotheses $U_2 \cap U_1 = \{\mathbf{0}\}$ and $U_3 \cap (U_1 + U_2) = \{\mathbf{0}\}$ instead, the argument is valid. (For an entry of $\mathcal{B}_2$, the right side is a combination of entries of $\mathcal{B}_1$ and lies in U_1, so the pairwise hypothesis is enough there.) These are condition (IV) of Theorem 4.4.7, and the corrected argument is a proof of (IV) $\Rightarrow$ (I) combined with Theorem 4.4.13(c).

Solution to Exercise 4.4.38

> *Uses.* Theorem 4.4.19 (Dimension of a sum)
> *Techniques.* Extend a basis of a subspace (Technique 4.1.18)
> *Hint 1.* Do a basis of U and a basis of W have to share any vectors at all?
> *Hint 2.* Try $U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ with basis $\mathbf{e}_1, \mathbf{e}_2$, and $W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_3)$ with basis $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_3$, in $\mathbb{R}^3$.
> *Sketch.* The claim that the shared vectors form a basis of $U \cap W$ is unjustified and usually false. With $\mathcal{B} = \mathbf{e}_1, \mathbf{e}_2$ and $\mathcal{C} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_3$, no vector is shared and no entry of either basis lies in $U \cap W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$. The repair is to choose the bases: extend one basis of $U \cap W$ to both, as in the first proof of Theorem 4.4.19.

The error is the claim that the vectors occurring in both $\mathcal{B}$ and $\mathcal{C}$ form a basis of $U \cap W$. Arbitrary bases need not share any vector, and the entries of $\mathcal{B}$ that happen to lie in $U \cap W$ need not span it.

For example, in $\mathbb{R}^3$ let $U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ with $\mathcal{B} = \mathbf{e}_1, \mathbf{e}_2$, and $W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_3)$ with $\mathcal{C} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_3$. The bases share no vector, so the procedure deletes nothing and returns four vectors in $\mathbb{R}^3$, which cannot be a basis. Here $U \cap W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$, since a vector $a(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) + b\mathbf{e}_3 = (a, a, a + b)$ of W lies in U exactly when $b = -a$; and none of $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{e}_3$ lies in this line.

The argument becomes correct if the bases are chosen with care: start from a basis of $U \cap W$ and extend it to a basis of U and to a basis of W. Then the shared vectors are exactly the chosen basis of $U \cap W$, and what remains to be shown is that the merged list is independent. That is the first proof of Theorem 4.4.19.

Solution to Exercise 4.4.39

> *Uses.* Remark 4.4.22 (No formula for three subspaces), Proposition 1.4.16 (Intersections of subspaces), Proposition 4.4.2 (Sums are the smallest containing subspaces), Theorem 4.4.19 (Dimension of a sum)
> *Hint 1.* For (a), see Remark 4.4.22.
> *Hint 2.* For (c), apply the dimension formula three times: to $U + W$ and X, to U and W, and to $U \cap X$ and $W \cap X$.
> *Sketch.* (a) Three distinct lines: $\dim(U + W + X) = 2$ but $e = 3$. (b) $U \cap X$ and $W \cap X$ lie in the subspace $(U + W) \cap X$. (c) $\dim(U + W + X) = \dim(U + W) + \dim X - \dim((U \cap X) + (W \cap X))$; expand $\dim(U + W)$ and $\dim((U \cap X) + (W \cap X))$ by the formula, using $(U \cap X) \cap (W \cap X) = U \cap W \cap X$.

(a) Let $U = \operatorname{span}(\mathbf{e}_1)$, $W = \operatorname{span}(\mathbf{e}_2)$, and $X = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$. All pairwise intersections and the triple intersection are $\{\mathbf{0}\}$, so $e(U, W, X) = 3$, but $\dim(U + W + X) = \dim \mathbb{R}^2 = 2$ (Remark 4.4.22). Here $(U + W) \cap X = X$, while $(U \cap X) + (W \cap X) = \{\mathbf{0}\}$.

(b) $U \cap X \subseteq U \subseteq U + W$ and $U \cap X \subseteq X$, so $U \cap X \subseteq (U + W) \cap X$; likewise $W \cap X \subseteq (U + W) \cap X$. Since $(U + W) \cap X$ is a subspace (Proposition 1.4.16), it contains $(U \cap X) + (W \cap X)$ by Proposition 4.4.2(b).

(c) All the subspaces involved are finite-dimensional, being contained in $U + W + X$ (Theorem 4.4.13(a)). By Proposition 4.4.2(a), $U + W + X = (U + W) + X$, so Theorem 4.4.19 and the hypothesis give

$$\dim(U + W + X) = \dim(U + W) + \dim X - \dim((U \cap X) + (W \cap X)).$$

By Theorem 4.4.19 again, $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, and since $(U \cap X) \cap (W \cap X) = U \cap W \cap X$,

$$\dim((U \cap X) + (W \cap X)) = \dim(U \cap X) + \dim(W \cap X) - \dim(U \cap W \cap X).$$

Substituting both gives $\dim(U + W + X) = e(U, W, X)$.

For instance, the hypothesis holds when $X \subseteq U$, since then both sides equal X.

Solution to Exercise 4.4.40

Uses. Theorem 4.4.19 (Dimension of a sum), Proposition 4.1.21 (Dimensions of subspaces), Example 4.2.25 (Kernels of linear functionals)
Techniques. Induction on the number of summands (Technique 4.4.10), Dimension counting (Technique 4.2.24)
Hint 1. Induct on k. In the inductive step, apply the dimension formula to $X = U_1 \cap \cdots \cap U_{k-1}$ and U_k.
Hint 2. $\dim(X + U_k) \le n$.
Sketch. (a) For $k = 1$ it is an equality. With $X = U_1 \cap \cdots \cap U_{k-1}$, $\dim(X \cap U_k) = \dim X + \dim U_k - \dim(X + U_k) \ge \dim X + \dim U_k - n$, and the inductive hypothesis bounds $\dim X$. (b) The intersection is $\operatorname{span}(\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$, of dimension $n - k = k(n-1) - (k-1)n$. (c) For $n - 1$ subspaces of dimension $n - 1$, the bound is $(n-1)^2 - (n-2)n = 1$.

(a) We induct on k. For $k = 1$ both sides equal $\dim U_1$. Let $k \ge 2$, and put $X = U_1 \cap \cdots \cap U_{k-1}$; by the inductive hypothesis, $\dim X \ge \dim U_1 + \cdots + \dim U_{k-1} - (k-2)n$. By Theorem 4.4.19 and $\dim(X + U_k) \le n$ (Proposition 4.1.21(a)),

$$\begin{aligned} \dim(X \cap U_k) &= \dim X + \dim U_k - \dim(X + U_k) \\ &\ge \dim U_1 + \cdots + \dim U_{k-1} - (k-2)n + \dim U_k - n \\ &= \dim U_1 + \cdots + \dim U_k - (k-1)n. \end{aligned}$$

Since $X \cap U_k = U_1 \cap \cdots \cap U_k$, this is the claim.

(b) Each U_i is the kernel of the nonzero functional $\mathbf{x} \mapsto x_i$, so $\dim U_i = n - 1$ (Example 4.2.25), and the bound in (a) is $k(n-1) - (k-1)n = n - k$. The intersection consists of the vectors with $x_1 = \cdots = x_k = 0$, which is $\operatorname{span}(\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$, of dimension $n - k$.

(c) For $k = n - 1$ subspaces of dimension $n - 1$, (a) gives

$$\dim(U_1 \cap \cdots \cap U_{n-1}) \ge (n-1)^2 - (n-2)n = 1.$$

Solution to Exercise 4.4.41

Uses. Proposition 2.6.18 (The column space is spanned by the columns), Proposition 4.4.2 (Sums are the smallest containing subspaces), Theorem 4.4.19 (Dimension of a sum), Theorem 4.2.7 (The image of T_A), Proposition 4.1.21 (Dimensions of subspaces)
Hint 1. The column space of $\begin{pmatrix} A & B \end{pmatrix}$ is spanned by the columns of A together with those of B.
Hint 2. $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$.
Sketch. (a) $\operatorname{im}\begin{pmatrix} A & B \end{pmatrix} = \operatorname{im} A + \operatorname{im} B$ by Proposition 2.6.18 and Proposition 4.4.2(c), so (a) is the dimension formula. (b) $\operatorname{im}(A + B) \subseteq \operatorname{im} A + \operatorname{im} B$. (c) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ have images $\operatorname{span}(\mathbf{e}_1)$ and $\operatorname{span}(\mathbf{e}_2)$, but $A + B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ has rank 1.

(a) By Proposition 2.6.18, $\operatorname{im} A$, $\operatorname{im} B$, and $\operatorname{im}(A \;\; B)$ are the spans of the columns of A, of B, and of both together. So $\operatorname{im}(A \;\; B) = \operatorname{im} A + \operatorname{im} B$ by Proposition 4.4.2(c), and since rank is the dimension of the image (Theorem 4.2.7), Theorem 4.4.19 gives

$$\operatorname{rank}(A \;\; B) = \dim(\operatorname{im} A + \operatorname{im} B) = \operatorname{rank} A + \operatorname{rank} B - \dim(\operatorname{im} A \cap \operatorname{im} B).$$

(b) For every $\mathbf{x} \in F^n$, $(A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} \in \operatorname{im} A + \operatorname{im} B$. So $\operatorname{im}(A+B) \subseteq \operatorname{im} A + \operatorname{im} B$, and by Proposition 4.1.21(a) and (a),

$$\operatorname{rank}(A+B) \le \dim(\operatorname{im} A + \operatorname{im} B) = \operatorname{rank} A + \operatorname{rank} B - \dim(\operatorname{im} A \cap \operatorname{im} B).$$

(c) Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Then $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} B = \operatorname{span}(\mathbf{e}_2)$ meet only in $\mathbf{0}$, and $\operatorname{rank} A + \operatorname{rank} B = 2$. But $A + B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ has rank 1. The loss comes from the other side: A and B both kill $\mathbf{e}_2$, so $A + B$ does too.

Solution to Exercise 4.4.42

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.4.16 (Complements exist), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Theorem 4.4.27 (Idempotents are projections), Proposition 4.4.26 (Properties of projections)
Techniques. Extend a basis of a subspace (Technique 4.1.18), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. One direction of (a) is rank–nullity. For the other, a complement X of U has the same dimension as W.
Hint 2. Define T to kill a basis of U and to send a basis of X to a basis of W. For (b), use Theorem 4.4.27.
Sketch. (a) ($\Rightarrow$) is rank–nullity. ($\le$) Let X be a complement of U, so $\dim X = \dim V - \dim U = \dim W$. Glue bases $\mathbf{u}_i$ of U and $\mathbf{x}_j$ of X into a basis of V, and let $T(\mathbf{u}_i) = \mathbf{0}$ and $T(\mathbf{x}_j) = \mathbf{w}_j$, a basis of W. Then $\operatorname{im} T = W$ and $\ker T = U$. (b) An idempotent has $V = \operatorname{im} T \oplus \ker T = W \oplus U$; conversely, $\mathrm{Pr}_{W,U}$ works. (c) $T(x,y) = (y,0)$. An idempotent fixes its image and kills its kernel, so these meet only in $\mathbf{0}$.

(a) If $\ker T = U$ and $\operatorname{im} T = W$, then $\dim U + \dim W = \dim V$ by Theorem 4.2.11. Conversely, suppose $\dim U + \dim W = \dim V$. Let X be a complement of U (Proposition 4.4.16), so $\dim X = \dim V - \dim U = \dim W$; call this number r. Let $\mathbf{u}_1, \dots, \mathbf{u}_k$, $\mathbf{x}_1, \dots, \mathbf{x}_r$, and $\mathbf{w}_1, \dots, \mathbf{w}_r$ be bases of U, X, and W. Since $V = U \oplus X$, the list $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{x}_1, \dots, \mathbf{x}_r$ is a basis of V (Theorem 4.4.13(b)), so there is $T \in \mathrm{L}(V)$ with $T(\mathbf{u}_i) = \mathbf{0}$ and $T(\mathbf{x}_j) = \mathbf{w}_j$ for all i, j (Theorem 2.2.1). Its image is $\operatorname{span}(\mathbf{w}_1, \dots, \mathbf{w}_r) = W$ (Proposition 2.1.16). For $\mathbf{v} = \sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{x}_j$, we get $T(\mathbf{v}) = \sum_j b_j \mathbf{w}_j$, which is $\mathbf{0}$ if and only if all $b_j = 0$, that is, if and only if $\mathbf{v} \in U$. So $\ker T = U$.

(b) If $T^2 = T$, then $V = \operatorname{im} T \oplus \ker T = W \oplus U$ by Theorem 4.4.27. Conversely, if $V = U \oplus W$, then $T = \mathrm{Pr}_{W,U}$ satisfies $T^2 = T$, $\operatorname{im} T = W$, and $\ker T = U$ by Proposition 4.4.26.

(c) Here $\dim U + \dim W = 2$, and $T(x,y) = (y,0)$ has $\ker T = \operatorname{span}(\mathbf{e}_1) = \operatorname{im} T$. Suppose some T with $T^2 = T$ had $\ker T = \operatorname{im} T = \operatorname{span}(\mathbf{e}_1)$. Then $\mathbf{e}_1 = T(\mathbf{x})$ for some $\mathbf{x}$, so $T(\mathbf{e}_1) = T^2(\mathbf{x}) = T(\mathbf{x}) = \mathbf{e}_1$; but $\mathbf{e}_1 \in \ker T$, so $T(\mathbf{e}_1) = \mathbf{0}$, a contradiction.

Solution to Exercise 4.4.43

Uses. Proposition 2.6.8 (Kernels and images are subspaces), Theorem 4.4.7 (Criteria for direct sums), Proposition 1.2.35 (Zero products), Theorem 4.4.27 (Idempotents are projections), Definition 2.4.21 (Transpose), Lemma 2.4.24 (Matrix units), Theorem 4.4.13 (Bases and dimension of a sum)
Techniques. Split the space with an idempotent (Technique 4.4.28)
Hint 1. Write $2 = 1 + 1$, which is invertible in F. Split $\mathbf{v} = 2^{-1}(\mathbf{v} + T(\mathbf{v})) + 2^{-1}(\mathbf{v} - T(\mathbf{v}))$ and check where each piece lies.
Hint 2. Equivalently, $P = 2^{-1}(\mathrm{id}_V + T)$ is idempotent. For (b), apply (a) to $A \mapsto A^{\mathrm{T}}$.
Sketch. (a) $T(\mathbf{v} + T\mathbf{v}) = \mathbf{v} + T\mathbf{v}$ and $T(\mathbf{v} - T\mathbf{v}) = -(\mathbf{v} - T\mathbf{v})$, so the split shows $V = V_+ + V_-$; if $T\mathbf{v} = \mathbf{v} = -\mathbf{v}$, then $2\mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$. (b) Transposition is linear with square the identity, and V_+, V_- are the symmetric and skew-symmetric matrices. The symmetric ones have basis U_{ii} and $U_{ij} + U_{ji}$ ($i < j$), so dimension $n(n+1)/2$, and the skew-symmetric ones have dimension $n^2 - n(n+1)/2 = n(n-1)/2$. (c) $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: $V_+ = V_- = \operatorname{span}(\mathbf{e}_1)$.

(a) $V_+ = \ker(T - \mathrm{id}_V)$ and $V_- = \ker(T + \mathrm{id}_V)$ are subspaces (Proposition 2.6.8). Write $2 = 1 + 1$, which is nonzero by hypothesis. For $\mathbf{v} \in V$, put $\mathbf{v}_+ = 2^{-1}(\mathbf{v} + T(\mathbf{v}))$ and $\mathbf{v}_- = 2^{-1}(\mathbf{v} - T(\mathbf{v}))$. Then $\mathbf{v}_+ + \mathbf{v}_- = 2^{-1}(2\mathbf{v}) = \mathbf{v}$, and since $T^2 = \mathrm{id}_V$,

$$T(\mathbf{v}_+) = 2^{-1}(T(\mathbf{v}) + \mathbf{v}) = \mathbf{v}_+, \quad T(\mathbf{v}_-) = 2^{-1}(T(\mathbf{v}) - \mathbf{v}) = -\mathbf{v}_-.$$

So $\mathbf{v}_+ \in V_+$ and $\mathbf{v}_- \in V_-$, and $V = V_+ + V_-$. If $\mathbf{v} \in V_+ \cap V_-$, then $\mathbf{v} = T(\mathbf{v}) = -\mathbf{v}$, so $2\mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$ (Proposition 1.2.35). By Theorem 4.4.7, $V = V_+ \oplus V_-$.

In the language of Theorem 4.4.27, $P = 2^{-1}(\mathrm{id}_V + T)$ satisfies $P^2 = 4^{-1}\big(\mathrm{id}_V + 2T + T^2\big) = 2^{-1}(\mathrm{id}_V + T) = P$, with $\operatorname{im} P = V_+$ and $\ker P = V_-$.

(b) The map $A \mapsto A^{\mathsf{T}}$ on $F^{n\times n}$ is linear, and $\big(A^{\mathsf{T}}\big)^{\mathsf{T}} = A$ (Definition 2.4.21). Its V_+ consists of the symmetric matrices and its V_- of the skew-symmetric ones, so (a) gives the direct sum, with $A = 2^{-1}\big(A + A^{\mathsf{T}}\big) + 2^{-1}\big(A - A^{\mathsf{T}}\big)$. Write $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$. A matrix $A = \sum_{i,j} a_{ij}U_{ij}$ is symmetric if and only if $a_{ij} = a_{ji}$ for all i, j, that is, if and only if

$$A = \sum_{i=1}^{n} a_{ii}U_{ii} + \sum_{i<j} a_{ij}\big(U_{ij} + U_{ji}\big).$$

So the $n + n(n-1)/2$ matrices U_{ii} and $U_{ij} + U_{ji}$ $(i < j)$ span the symmetric matrices, and they are independent, since a combination of them has entry a_{ij} in position (i, j) for $i \le j$, and the matrix units are independent (Lemma 2.4.24). Hence the symmetric matrices have dimension $n(n+1)/2$, and by Theorem 4.4.13(c) the skew-symmetric matrices have dimension $n^2 - n(n+1)/2 = n(n-1)/2$.

(c) Over $\mathbb{F}_2$, let T have matrix $\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$. Then T^2 has matrix $\left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right) = I_2$. $T(x, y) = (x + y, y)$ equals (x, y) if and only if $y = 0$, so $V_+ = \operatorname{span}(\mathbf{e}_1)$, and since $-1 = 1$ in $\mathbb{F}_2$, $V_- = V_+$. So $V_+ + V_- = \operatorname{span}(\mathbf{e}_1) \neq \mathbb{F}_2^2$. (Likewise, over $\mathbb{F}_2$ the skew-symmetric matrices are exactly the symmetric ones.)

Solution to Exercise 4.4.44

Uses. Theorem 4.4.7 (Criteria for direct sums), Theorem 4.4.13 (Bases and dimension of a sum), Definition 4.2.14 (Rank and nullity of a linear map), Proposition 4.4.2 (Sums are the smallest containing subspaces)
Techniques. Compare components (Technique 4.4.8)
Hint 1. If $T(\mathbf{u} + \mathbf{w}) = \mathbf{0}$, then $T(\mathbf{u}) + T(\mathbf{w}) = \mathbf{0}$ with $T(\mathbf{u}) \in U$ and $T(\mathbf{w}) \in W$. Compare components.
Hint 2. Subspaces of independent subspaces are independent: $(U \cap \ker T) \cap (W \cap \ker T) \subseteq U \cap W$.
Sketch. (a) If $\mathbf{u} + \mathbf{w} \in \ker T$, then $T\mathbf{u} + T\mathbf{w} = \mathbf{0}$ is a decomposition of $\mathbf{0}$ along $U \oplus W$, so $T\mathbf{u} = T\mathbf{w} = \mathbf{0}$. The image: $T(\mathbf{u} + \mathbf{w}) = T\mathbf{u} + T\mathbf{w}$. Both sums are direct because their summands lie in U and W. (b) Dimensions add in a direct sum. (c) $T(x, y) = (x + y, 0)$ has kernel $\operatorname{span}((1, -1))$, which meets both axes trivially.

(a) Let $\mathbf{v} = \mathbf{u} + \mathbf{w} \in \ker T$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Then $\mathbf{0} = T(\mathbf{u}) + T(\mathbf{w})$ with $T(\mathbf{u}) \in U$ and $T(\mathbf{w}) \in W$, a decomposition of $\mathbf{0}$ along $U \oplus W$, so $T(\mathbf{u}) = T(\mathbf{w}) = \mathbf{0}$ by (II) of Theorem 4.4.7. Hence $\mathbf{u} \in U \cap \ker T$ and $\mathbf{w} \in W \cap \ker T$, so $\ker T \subseteq (U \cap \ker T) + (W \cap \ker T)$. The reverse containment holds because $\ker T$ is a subspace containing both summands (Proposition 4.4.2(b)). The sum is direct, since $(U \cap \ker T) \cap (W \cap \ker T) \subseteq U \cap W = \{\mathbf{0}\}$.

For the image, every $T(\mathbf{u} + \mathbf{w}) = T(\mathbf{u}) + T(\mathbf{w})$ lies in $\operatorname{im}(T{\restriction_U}) + \operatorname{im}(T{\restriction_W})$, and conversely $T(\mathbf{u}) + T(\mathbf{w}) = T(\mathbf{u} + \mathbf{w}) \in \operatorname{im} T$. By hypothesis $\operatorname{im}(T{\restriction_U}) \subseteq U$ and $\operatorname{im}(T{\restriction_W}) \subseteq W$, so their intersection lies in $U \cap W = \{\mathbf{0}\}$, and the sum is direct.

(b) By (a) and Theorem 4.4.13(c), $\dim \operatorname{im} T = \dim \operatorname{im}(T{\restriction_U}) + \dim \operatorname{im}(T{\restriction_W})$, which is the claim (Definition 4.2.14).

(c) Let $T(x, y) = (x + y, 0)$. Then $\ker T = \operatorname{span}((1, -1))$, and a vector $(t, -t)$ lies on an axis only if $t = 0$. So $(U \cap \ker T) + (W \cap \ker T) = \{\mathbf{0}\} \neq \ker T$. Here $T(\mathbf{e}_2) = \mathbf{e}_1 \notin W$, so the hypothesis fails.

Solution to Exercise 4.4.45

Uses. Theorem 4.2.11 (Rank–nullity), Theorem 4.3.8 (Row rank equals column rank), Theorem 4.2.7 (The image of T_A), Proposition 4.4.21 (Two out of three), Proposition 2.4.23 (Transpose of a product)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Over every field the dimensions add up to n, so by Proposition 4.4.21 only the intersection matters.
Hint 2. If $\mathbf{x} = A^{\mathsf{T}}\mathbf{y}$ and $A\mathbf{x} = \mathbf{0}$, compute $\mathbf{x}^{\mathsf{T}}\mathbf{x}$.

Sketch. $\dim\ker A + \dim\operatorname{im} A^{\mathsf{T}} = (n - \operatorname{rank} A) + \operatorname{rank} A = n$. Over $\mathbb{R}$, if $\mathbf{x} = A^{\mathsf{T}}\mathbf{y} \in \ker A$, then $\mathbf{x}^{\mathsf{T}}\mathbf{x} = \mathbf{y}^{\mathsf{T}}A\mathbf{x} = 0$, and $x_1^2 + \cdots + x_n^2 = 0$ forces $\mathbf{x} = \mathbf{0}$. Over $\mathbb{C}$, $A = \begin{pmatrix}1 & i\end{pmatrix}$ has $\ker A = \operatorname{span}((-i, 1)) = \operatorname{im} A^{\mathsf{T}}$. Over $\mathbb{F}_2$, $A = \begin{pmatrix}1 & 1\end{pmatrix}$ has $\ker A = \operatorname{span}((1,1)) = \operatorname{im} A^{\mathsf{T}}$.

Over every field, $\dim\ker A = n - \operatorname{rank} A$ by Theorem 4.2.11, and $\dim\operatorname{im} A^{\mathsf{T}} = \operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$ by Theorem 4.2.7 and Theorem 4.3.8. So condition (iii) of Proposition 4.4.21 always holds, and $F^n = \ker A \oplus \operatorname{im} A^{\mathsf{T}}$ if and only if $\ker A \cap \operatorname{im} A^{\mathsf{T}} = \{\mathbf{0}\}$.

Over $\mathbb{R}$. Let $\mathbf{x} \in \ker A \cap \operatorname{im} A^{\mathsf{T}}$, say $\mathbf{x} = A^{\mathsf{T}}\mathbf{y}$. By Proposition 2.4.23, $\mathbf{x}^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}A$, so

$$x_1^2 + \cdots + x_n^2 = \mathbf{x}^{\mathsf{T}}\mathbf{x} = \mathbf{y}^{\mathsf{T}}(A\mathbf{x}) = \mathbf{y}^{\mathsf{T}}\mathbf{0} = 0.$$

A sum of squares of real numbers vanishes only if each term does, so $\mathbf{x} = \mathbf{0}$.

Over $\mathbb{C}$. Let $A = \begin{pmatrix}1 & i\end{pmatrix} \in \mathbb{C}^{1\times 2}$. Then $A\mathbf{x} = x_1 + ix_2 = 0$ if and only if $x_1 = -ix_2$, so $\ker A = \operatorname{span}((-i, 1))$. Also $\operatorname{im} A^{\mathsf{T}} = \operatorname{span}((1, i))$, and $(-i, 1) = -i(1, i)$. So $\ker A = \operatorname{im} A^{\mathsf{T}}$ is a line, and the sum is not direct. The real argument breaks because $(1,i)^{\mathsf{T}}(1,i) = 1 + i^2 = 0$. Over $\mathbb{C}$ the correct statement uses the conjugate transpose (Section 6.6).

Over $\mathbb{F}_2$. Let $A = \begin{pmatrix}1 & 1\end{pmatrix} \in \mathbb{F}_2^{1\times 2}$. Then $\ker A = \{\mathbf{x} \mid x_1 + x_2 = 0\} = \operatorname{span}((1,1)) = \operatorname{im} A^{\mathsf{T}}$, and again the sum is not direct; here $(1,1)^{\mathsf{T}}(1,1) = 1 + 1 = 0$.

Solution to Exercise 4.4.46

Uses. Proposition 4.4.16 (Complements exist), Proposition 4.4.21 (Two out of three), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. If $U = \{\mathbf{0}\}$ or $U = V$, what must a complement be?
Hint 2. Otherwise, take a complement W with basis $\mathbf{w}_1, \ldots, \mathbf{w}_r$ and a nonzero $\mathbf{u} \in U$. Tilt $\mathbf{w}_1$ towards $\mathbf{u}$: try $\operatorname{span}(\mathbf{w}_1 + \mathbf{u}, \mathbf{w}_2, \ldots, \mathbf{w}_r)$.
Sketch. A complement W of $\{\mathbf{0}\}$ satisfies $W = \{\mathbf{0}\} + W = V$, and a complement of V satisfies $W = V \cap W = \{\mathbf{0}\}$. If $\{\mathbf{0}\} \neq U \neq V$, take a complement W with basis $\mathbf{w}_1, \ldots, \mathbf{w}_r$ ($r \geq 1$) and $\mathbf{u} \in U$ nonzero. If $a(\mathbf{w}_1 + \mathbf{u}) + \sum_{j\geq 2} a_j\mathbf{w}_j \in U$, then $a\mathbf{w}_1 + \sum_{j\geq 2} a_j\mathbf{w}_j \in U \cap W$, so all coefficients vanish. So $W' = \operatorname{span}(\mathbf{w}_1 + \mathbf{u}, \mathbf{w}_2, \ldots, \mathbf{w}_r)$ has dimension r and meets U trivially, hence is a complement, and $\mathbf{w}_1 + \mathbf{u} \notin W$.

If $U = \{\mathbf{0}\}$, then $V = \{\mathbf{0}\} \oplus V$, and any complement W satisfies $W = \{\mathbf{0}\} + W = V$. If $U = V$, then $V = V \oplus \{\mathbf{0}\}$, and any complement W satisfies $W = V \cap W = \{\mathbf{0}\}$. In both cases the complement is unique.

Now suppose $\{\mathbf{0}\} \neq U \neq V$. By Proposition 4.4.16, U has a complement W, of dimension $r = \dim V - \dim U \geq 1$; let $\mathbf{w}_1, \ldots, \mathbf{w}_r$ be a basis of W, and let $\mathbf{u} \in U$ be nonzero. Let $W' = \operatorname{span}(\mathbf{w}_1 + \mathbf{u}, \mathbf{w}_2, \ldots, \mathbf{w}_r)$. Suppose

$$\mathbf{y} = a(\mathbf{w}_1 + \mathbf{u}) + a_2\mathbf{w}_2 + \cdots + a_r\mathbf{w}_r \in U.$$

Then $a\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_r\mathbf{w}_r = \mathbf{y} - a\mathbf{u}$ lies in W and in U, hence is $\mathbf{0}$, and since $\mathbf{w}_1, \ldots, \mathbf{w}_r$ is independent, $a = a_2 = \cdots = a_r = 0$ (Theorem 1.5.5), so $\mathbf{y} = \mathbf{0}$. This shows $U \cap W' = \{\mathbf{0}\}$, and, taking $\mathbf{y} = \mathbf{0}$, that the spanning list of W' is independent, so $\dim W' = r$. By Proposition 4.4.21, $V = U \oplus W'$. Finally $W' \neq W$: if $\mathbf{w}_1 + \mathbf{u} \in W$, then $\mathbf{u} = (\mathbf{w}_1 + \mathbf{u}) - \mathbf{w}_1 \in U \cap W = \{\mathbf{0}\}$, a contradiction. So U has at least two complements.

Solution to Exercise 4.4.47

Uses. Example 2.1.9 (Maps of polynomials), Lemma 3.4.23 (Too many roots), Theorem 3.4.24 (Polynomial interpolation), Theorem 4.4.7 (Criteria for direct sums), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 4.2.11 (Rank–nullity)
Hint 1. A polynomial of degree at most k that vanishes at $k + 1$ distinct points is $\mathbf{0}$.
Hint 2. Given p, interpolate its values at $c_0, \ldots, c_k$ by a polynomial q of degree at most k. Then $p - q \in Z$.
Hint 3. For the second count, Z is the kernel of $p \mapsto (p(c_0), \ldots, p(c_k))$.
Sketch. $Z \cap F[x]_{\leq k} = \{\mathbf{0}\}$ by Lemma 3.4.23. For p, let $q \in F[x]_{\leq k}$ interpolate the values $p(c_i)$ (Theorem 3.4.24); then $p = (p - q) + q$ with $p - q \in Z$. So $\dim Z = (n + 1) - (k + 1) = n - k$ by Theorem 4.4.13(c). Alternatively, evaluation $F[x]_{\leq n} \to F^{k+1}$ is onto by interpolation and has kernel Z; rank–nullity gives the same answer.

The map $\mathrm{ev} : F[x]_{\leq n} \to F^{k+1}$, $p \mapsto (p(c_0), \ldots, p(c_k))$, is linear (Example 2.1.9), and $Z = \ker \mathrm{ev}$ is a subspace.

If $p \in Z \cap F[x]_{\le k}$, then p has degree at most k and vanishes at the $k+1$ distinct points $c_0, \dots, c_k$, so $p = \mathbf{0}$ by Lemma 3.4.23. Next, let $p \in F[x]_{\le n}$. By Theorem 3.4.24 there is $q \in F[x]_{\le k}$ with $q(c_i) = p(c_i)$ for $i = 0, \dots, k$. Since $k \le n$, $p - q \in F[x]_{\le n}$, and it vanishes at every c_i, so $p - q \in Z$ and $p = (p-q) + q \in Z + F[x]_{\le k}$. By Theorem 4.4.7, $F[x]_{\le n} = Z \oplus F[x]_{\le k}$.

First count. By Theorem 4.4.13(c), $n + 1 = \dim Z + (k+1)$, so $\dim Z = n - k$.

Second count. The map ev is surjective: for $(d_0, \dots, d_k) \in F^{k+1}$, the interpolating polynomial of Theorem 3.4.24 lies in $F[x]_{\le k} \subseteq F[x]_{\le n}$ and has these values. By Theorem 4.2.11, $\dim Z = \dim \ker \mathrm{ev} = (n+1) - (k+1) = n - k$.

The decomposition $p = (p-q) + q$ says that the projection onto $F[x]_{\le k}$ along Z sends p to its interpolating polynomial at $c_0, \dots, c_k$. Note that the hypothesis requires F to have at least $k+1$ elements.

Solution to Exercise 4.4.48

Uses. Theorem 4.4.13 (Bases and dimension of a sum), Theorem 4.2.11 (Rank–nullity), Theorem 4.2.7 (The image of T_A), Proposition 4.1.21 (Dimensions of subspaces), Proposition 4.4.21 (Two out of three)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Since $A + B$ is invertible, every $\mathbf{x}$ equals $(A+B)\mathbf{y} = A\mathbf{y} + B\mathbf{y}$ for some $\mathbf{y}$. What does that say about $\operatorname{im} A + \operatorname{im} B$?
Hint 2. $AB = 0$ means $\operatorname{im} B \subseteq \ker A$. Combine this with rank–nullity for A.
Sketch. (a) Every $\mathbf{x} = (A+B)\mathbf{y}$ lies in $\operatorname{im} A + \operatorname{im} B$, so $\operatorname{im} A + \operatorname{im} B = F^n$ and $\operatorname{rank} A + \operatorname{rank} B \ge n$. Also $\operatorname{im} B \subseteq \ker A$, so $\operatorname{rank} B \le n - \operatorname{rank} A$. Hence $\operatorname{rank} A + \operatorname{rank} B = n$, and two out of three gives the direct sum. (b) $\operatorname{im} B \subseteq \ker A$, with equal dimensions. (c) Symmetrically $\operatorname{im} A = \ker B$. Without $BA = 0$: $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

(a) For $\mathbf{x} \in F^n$, invertibility of $A + B$ gives $\mathbf{y}$ with $\mathbf{x} = (A+B)\mathbf{y} = A\mathbf{y} + B\mathbf{y}$, so $\operatorname{im} A + \operatorname{im} B = F^n$. By Theorem 4.4.13(a) and Theorem 4.2.7, $n \le \operatorname{rank} A + \operatorname{rank} B$. On the other hand, $A(B\mathbf{y}) = \mathbf{0}$ for every $\mathbf{y}$, so $\operatorname{im} B \subseteq \ker A$, and by Proposition 4.1.21(a) and Theorem 4.2.11, $\operatorname{rank} B \le \dim \ker A = n - \operatorname{rank} A$. Hence $\operatorname{rank} A + \operatorname{rank} B = n$. Conditions (ii) and (iii) of Proposition 4.4.21 hold for $\operatorname{im} A$ and $\operatorname{im} B$, so $F^n = \operatorname{im} A \oplus \operatorname{im} B$.

(b) $\operatorname{im} B \subseteq \ker A$, and $\dim \operatorname{im} B = \operatorname{rank} B = n - \operatorname{rank} A = \dim \ker A$, so $\operatorname{im} B = \ker A$ by Proposition 4.1.21(b).

(c) If $BA = 0$, the hypotheses are symmetric in A and B (note $B + A = A + B$), so (b) with the roles exchanged gives $\operatorname{im} A = \ker B$.

Without $BA = 0$, let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$. Then $AB = 0$ and $A + B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ is invertible, with inverse $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, but $BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ne 0$, and $\operatorname{im} A = \operatorname{span}(\mathbf{e}_1)$ while $\ker B = \operatorname{span}((1,-1))$.

The kernels are complementary in any case: $\ker A \cap \ker B \subseteq \ker(A+B) = \{\mathbf{0}\}$, and by (a), $\dim \ker A + \dim \ker B = (n - \operatorname{rank} A) + (n - \operatorname{rank} B) = n$, so $F^n = \ker A \oplus \ker B$ by Proposition 4.4.21, whether or not $BA = 0$. In the example, $\ker A = \operatorname{span}(\mathbf{e}_2)$ and $\ker B = \operatorname{span}((1,-1))$.

Solution to Exercise 4.4.49

Uses. Proposition 4.2.6 (Injective maps preserve dimension), Proposition 4.4.21 (Two out of three), Proposition 4.4.16 (Complements exist), Definition 4.4.24 (Projection along a complement), Proposition 4.4.26 (Properties of projections), Proposition 4.1.21 (Dimensions of subspaces), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 4.1.34 (Size of a vector space over a finite field)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. (a) Γ_S is the image of the linear map $\mathbf{w} \mapsto \mathbf{w} + S(\mathbf{w})$. Show that this map is injective and that $\Gamma_S \cap U = \{\mathbf{0}\}$, and count dimensions.
Hint 2. (b) Given a complement W, decompose each $\mathbf{w} \in W_0$ along $V = U \oplus W$. Its W-component should be $\mathbf{w} + S(\mathbf{w})$.
Hint 3. (c) A linear map $W_0 \to U$ is determined by the images of a basis of W_0, which can be any vectors of U.
Sketch. (a) If $\mathbf{w} + S\mathbf{w} = \mathbf{0}$ or $\mathbf{w} + S\mathbf{w} \in U$, then $\mathbf{w} \in U \cap W_0 = \{\mathbf{0}\}$; so $\dim \Gamma_S = \dim W_0 = n - k$ and two out of three applies. (b) For a complement W, let $S = -\operatorname{Pr}_{U,W}|_{W_0}$, with values in U. Then $\mathbf{w} + S(\mathbf{w}) = \operatorname{Pr}_{W,U}(\mathbf{w}) \in W$, so $\Gamma_S \subseteq W$, with equal dimensions. If $\Gamma_S = \Gamma_{S'}$, compare components in $W_0 \oplus U$. (c) $|\mathrm{L}(W_0, U)| = \left(q^k\right)^{n-k}$. (d) The four planes $x_1 = 0$, $x_2 = 0$, $x_1 + x_3 = 0$, $x_2 + x_3 = 0$.

Let $n = \dim V$ and $k = \dim U$, so $\dim W_0 = n - k$ (Proposition 4.4.16(b)).

(a) The map $g : W_0 \to V$, $g(\mathbf{w}) = \mathbf{w} + S(\mathbf{w})$, is linear with image Γ_S, so Γ_S is a subspace. If $g(\mathbf{w}) \in U$, then $\mathbf{w} = g(\mathbf{w}) - S(\mathbf{w}) \in U$, as $S(\mathbf{w}) \in U$; so $\mathbf{w} \in U \cap W_0 = \{\mathbf{0}\}$. Taking $g(\mathbf{w}) = \mathbf{0}$ shows that g is injective, so $\dim \Gamma_S = n - k$ (Proposition 4.2.6(a)); and in general it shows $U \cap \Gamma_S = \{\mathbf{0}\}$. By Proposition 4.4.21, $V = U \oplus \Gamma_S$.

(b) *Existence.* Let W be a complement of U. Define $S : W_0 \to U$ by $S(\mathbf{w}) = -\mathrm{Pr}_{U,W}(\mathbf{w})$; it is linear with values in U (Proposition 4.4.26). For $\mathbf{w} \in W_0$,

$$\mathbf{w} + S(\mathbf{w}) = \mathbf{w} - \mathrm{Pr}_{U,W}(\mathbf{w}) = \mathrm{Pr}_{W,U}(\mathbf{w}) \in W,$$

by Proposition 4.4.26(d). So $\Gamma_S \subseteq W$. Both have dimension $n - k$, by (a) and Proposition 4.4.16(b), so $\Gamma_S = W$ (Proposition 4.1.21(b)).

Uniqueness. Suppose $\Gamma_S = \Gamma_{S'}$, and let $\mathbf{w} \in W_0$. Then $\mathbf{w} + S(\mathbf{w}) \in \Gamma_{S'}$, so $\mathbf{w} + S(\mathbf{w}) = \mathbf{w}' + S'(\mathbf{w}')$ for some $\mathbf{w}' \in W_0$. Both sides are decompositions along $V = W_0 \oplus U$, so comparing components gives $\mathbf{w} = \mathbf{w}'$ and $S(\mathbf{w}) = S'(\mathbf{w}') = S'(\mathbf{w})$. Hence $S = S'$.

(c) By (a) and (b), $S \mapsto \Gamma_S$ is a bijection from $\mathrm{L}(W_0, U)$ onto the set of complements of U. Fix a basis $\mathbf{w}_1, \ldots, \mathbf{w}_{n-k}$ of W_0. By Theorem 2.2.1, $S \mapsto (S(\mathbf{w}_1), \ldots, S(\mathbf{w}_{n-k}))$ is a bijection from $\mathrm{L}(W_0, U)$ onto the set of $(n-k)$-tuples of vectors of U. Since $|U| = q^k$ (Proposition 4.1.34), there are $\left(q^k\right)^{n-k} = q^{k(n-k)}$ complements.

(d) Here $n = 3$ and $k = 1$, so there are $2^2 = 4$ complements. Take $W_0 = \mathrm{span}(\mathbf{e}_2, \mathbf{e}_3) = \{\mathbf{x} \mid x_1 = 0\}$. Since $(1,1,0) \notin W_0$, $U \cap W_0 = \{\mathbf{0}\}$, and $1 + 2 = 3$, so W_0 is a complement of U (Proposition 4.4.21). A map $S : W_0 \to U$ is given by $S(\mathbf{e}_2) = a(1,1,0)$ and $S(\mathbf{e}_3) = b(1,1,0)$ with $a, b \in \mathbb{F}_2$, and $\Gamma_S = \mathrm{span}(\mathbf{e}_2 + a(1,1,0), \mathbf{e}_3 + b(1,1,0))$. The four choices give

$$\begin{aligned}
(a,b) = (0,0) &: \mathrm{span}((0,1,0),(0,0,1)) = \{x_1 = 0\},\\
(a,b) = (1,0) &: \mathrm{span}((1,0,0),(0,0,1)) = \{x_2 = 0\},\\
(a,b) = (0,1) &: \mathrm{span}((0,1,0),(1,1,1)) = \{x_1 + x_3 = 0\},\\
(a,b) = (1,1) &: \mathrm{span}((1,0,0),(1,1,1)) = \{x_2 + x_3 = 0\}.
\end{aligned}$$

These are exactly the planes of $\mathbb{F}_2^3$ that do not contain $(1,1,0)$.

Solution to Exercise 4.4.50

Uses. Exercise 1.4.21, Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.4.21 (Two out of three), Proposition 4.4.16 (Complements exist)
Techniques. Dimension counting (Technique 4.2.24), Extend a basis of a subspace (Technique 4.1.18)
Hint 1. Induct on $c = \dim V - \dim U$. If $c \geq 1$, find a vector $\mathbf{v}$ that lies in neither U nor W.
Hint 2. Two proper subspaces never cover V (Exercise 1.4.21). Enlarge U and W by $\mathbf{v}$, apply the inductive hypothesis, and add $\mathbf{v}$ back to the common complement.
Hint 3. For (b), a complement of a line in $\mathbb{F}_2^2$ is another line. How many lines are there?
Sketch. (a) Induct on c. If $c = 0$, take $X = \{\mathbf{0}\}$. Otherwise U, W are proper, so some $\mathbf{v}$ lies in neither. $U' = U + \mathrm{span}(\mathbf{v})$ and $W' = W + \mathrm{span}(\mathbf{v})$ have dimension $\dim U + 1$; take a common complement X' and set $X = X' + \mathrm{span}(\mathbf{v})$. Then $\dim X = c$, and if $\mathbf{x}' + a\mathbf{v} \in U$ with $\mathbf{x}' \in X'$, then $\mathbf{x}' \in U' \cap X' = \{\mathbf{0}\}$ and $a\mathbf{v} \in U$ forces $a = 0$. Two out of three. (b) There are only three lines, and a common complement would be a fourth.

(a) Let $n = \dim V$ and $d = \dim U = \dim W$. We prove the claim by induction on $c = n - d$, for all pairs of subspaces of V of equal dimension. If $c = 0$, then $U = W = V$ (Proposition 4.1.21(b)) and $X = \{\mathbf{0}\}$ works.

Let $c \geq 1$. Then U and W are proper subspaces, so $U \cup W \neq V$ by Exercise 1.4.21; choose $\mathbf{v} \in V$ with $\mathbf{v} \notin U$ and $\mathbf{v} \notin W$. If $a\mathbf{v} \in U$ with $a \neq 0$, then $\mathbf{v} = a^{-1}(a\mathbf{v}) \in U$; so $U \cap \mathrm{span}(\mathbf{v}) = \{\mathbf{0}\}$, and $U' = U + \mathrm{span}(\mathbf{v})$ has dimension $d + 1$ by Theorem 4.4.13(c). Likewise $W' = W + \mathrm{span}(\mathbf{v})$ has dimension $d + 1$. By the inductive hypothesis there is X' with $V = U' \oplus X' = W' \oplus X'$, and $\dim X' = n - d - 1$. Let $X = X' + \mathrm{span}(\mathbf{v})$. Since $\mathbf{v} \in U'$, we have $X' \cap \mathrm{span}(\mathbf{v}) \subseteq X' \cap U' = \{\mathbf{0}\}$, so $\dim X = n - d$.

We show $U \cap X = \{\mathbf{0}\}$. Let $\mathbf{y} = \mathbf{x}' + a\mathbf{v} \in U$ with $\mathbf{x}' \in X'$ and $a \in F$. Then $\mathbf{x}' = \mathbf{y} - a\mathbf{v}$ lies in U', since $\mathbf{y} \in U \subseteq U'$ and $\mathbf{v} \in U'$. So $\mathbf{x}' \in U' \cap X' = \{\mathbf{0}\}$, and $a\mathbf{v} = \mathbf{y} \in U$, which forces $a = 0$ since $\mathbf{v} \notin U$. Hence $\mathbf{y} = \mathbf{0}$.

As $\dim U + \dim X = n$, Proposition 4.4.21 gives $V = U \oplus X$. The same argument with W and W' in place of U and U' gives $V = W \oplus X$.

(b) Each line through $\mathbf{0}$ in $\mathbb{F}_2^2$ contains exactly one nonzero vector, so there are exactly three lines: $\operatorname{span}(\mathbf{e}_1)$, $\operatorname{span}(\mathbf{e}_2)$, $\operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$. A complement of a line has dimension 1 (Proposition 4.4.16(b)) and is a line different from it. A common complement of all three would be a fourth line, which does not exist.

So (a) does not extend to three subspaces over every field. Over an infinite field it does extend to any finite number of subspaces of equal dimension, by the same induction together with Exercise 1.4.43.

Solution to Exercise 4.4.51

Uses. Exercise 4.3.43, Proposition 2.4.27 (The trace is linear), Theorem 4.2.7 (The image of T_A), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.4.29 (Projections of a direct sum)
Techniques. Compare components (Technique 4.4.8), Split the space with an idempotent (Technique 4.4.28)
Hint 1. Every $\mathbf{x}$ equals $\sum_i P_i\mathbf{x}$, so the images add up to F^n. Directness is then a dimension count. How is $\operatorname{rank} P_i$ related to $\operatorname{tr} P_i$ over these fields?
Hint 2. For (b), decompose $P_j\mathbf{x}$ in two ways: as itself, a vector of $\operatorname{im} P_j$, and as $\sum_i P_i(P_j\mathbf{x})$.
Sketch. (a) $\mathbf{x} = \sum_i P_i\mathbf{x}$, so $\sum_i \operatorname{im} P_i = F^n$. By Exercise 4.3.43, $\operatorname{rank} P_i = \operatorname{tr} P_i$, so $\sum_i \dim \operatorname{im} P_i = \operatorname{tr}\left(\sum_i P_i\right) = \operatorname{tr} I_n = n$, and Theorem 4.4.13(c) makes the sum direct. (b) $P_j\mathbf{x} = \sum_i P_iP_j\mathbf{x}$ has components $P_iP_j\mathbf{x} \in \operatorname{im} P_i$, while $P_j\mathbf{x} \in \operatorname{im} P_j$ also has the decomposition with $P_j\mathbf{x}$ in slot j and $\mathbf{0}$ elsewhere; compare components. (c) Over $\mathbb{F}_2$, take $P_1 = P_2 = P_3 = I_1$.

(a) For $\mathbf{x} \in F^n$, $\mathbf{x} = I_n\mathbf{x} = P_1\mathbf{x} + \cdots + P_k\mathbf{x}$, so $\operatorname{im} P_1 + \cdots + \operatorname{im} P_k = F^n$. Over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, an idempotent matrix has rank equal to its trace (Exercise 4.3.43(c)). So, using Theorem 4.2.7 and Proposition 2.4.27,

$$\dim \operatorname{im} P_1 + \cdots + \dim \operatorname{im} P_k = \operatorname{tr} P_1 + \cdots + \operatorname{tr} P_k = \operatorname{tr}(P_1 + \cdots + P_k) = \operatorname{tr} I_n = n.$$

This is the dimension of the sum, so the sum is direct by Theorem 4.4.13(c).

(b) Fix j and $\mathbf{x} \in F^n$, and let $\mathbf{y} = P_j\mathbf{x} \in \operatorname{im} P_j$. Since $\sum_i P_i = I_n$,

$$\mathbf{y} = P_1\mathbf{y} + \cdots + P_k\mathbf{y},$$

a decomposition of $\mathbf{y}$ along $\operatorname{im} P_1, \ldots, \operatorname{im} P_k$. Another decomposition of $\mathbf{y}$ has component $\mathbf{y}$ in $\operatorname{im} P_j$ and $\mathbf{0}$ in every other $\operatorname{im} P_i$. By (a) the two agree, so $P_i\mathbf{y} = P_iP_j\mathbf{x} = \mathbf{0}$ for $i \neq j$. As $\mathbf{x}$ was arbitrary, $P_iP_j = 0$.

By (b) and Proposition 4.4.29(b), the P_i are the projections of the decomposition in (a).

(c) Over $\mathbb{F}_2$, let $n = 1$, $k = 3$, and $P_1 = P_2 = P_3 = I_1 = (1)$. Each is idempotent, and $P_1 + P_2 + P_3 = (1{+}1{+}1) = I_1$, but $P_1P_2 = I_1 \neq 0$. What fails is (a): the ranks add up to $3 > 1$, while the trace only records them modulo 2.

Solution to Exercise 4.4.52

Uses. Proposition 4.1.34 (Size of a vector space over a finite field), Theorem 4.4.7 (Criteria for direct sums), Proposition 4.4.21 (Two out of three), Proposition 4.2.6 (Injective maps preserve dimension)
Techniques. Double counting (Technique 4.1.36)
Hint 1. Two members of a spread share only $\mathbf{0}$. Count nonzero vectors.
Hint 2. The subspace Γ_M of vectors $(\mathbf{x}, M\mathbf{x})$ meets $\Gamma_{M'}$ only in $\mathbf{0}$ when $M - M'$ is invertible. Add the subspace of vectors $(\mathbf{0}, \mathbf{y})$.
Hint 3. Over $\mathbb{F}_2$, $C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ satisfies $C^2 = C + I_2$ and $C^3 = I_2$. Try $M \in \{0, I_2, C, C^2\}$.
Sketch. (a) Each member has $q^2 - 1$ nonzero vectors, no nonzero vector lies in two members, and F^4 has $q^4 - 1$ nonzero vectors; so $m(q^2 - 1) \leq q^4 - 1$, i.e. $m \leq q^2 + 1$. (b) Take Γ_M for $M \in \{0, I_2, C, C^2\}$ and $\Gamma_\infty = \{(\mathbf{0}, \mathbf{y})\}$. Every difference of two of the four matrices is I_2, C, or C^2, all invertible, so $\Gamma_M \cap \Gamma_{M'} = \{\mathbf{0}\}$; and $\Gamma_M \cap \Gamma_\infty = \{\mathbf{0}\}$. Two-dimensional subspaces of $\mathbb{F}_2^4$ meeting trivially are complements.

Write vectors of F^4 as $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x}, \mathbf{y} \in F^2$.

(a) Let $U_1, \ldots, U_m$ be a spread. For $i \neq j$, $U_i \cap U_j = \{\mathbf{0}\}$, so the sets $U_i \setminus \{\mathbf{0}\}$ are pairwise disjoint subsets of $F^4 \setminus \{\mathbf{0}\}$. Each U_i has q^2 elements (Proposition 4.1.34), so

$$m(q^2 - 1) \leq q^4 - 1 = (q^2 - 1)(q^2 + 1),$$

and $m \leq q^2 + 1$.

(b) For $M \in \mathbb{F}_2^{2\times 2}$ let $\Gamma_M = \{(\mathbf{x}, M\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^2\}$, the image of the injective linear map $\mathbf{x} \mapsto (\mathbf{x}, M\mathbf{x})$, so $\dim \Gamma_M = 2$ (Proposition 4.2.6(a)). Let $\Gamma_\infty = \{(\mathbf{0}, \mathbf{y}) \mid \mathbf{y} \in \mathbb{F}_2^2\}$, also of dimension 2. Let

$$C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad C^2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = C + I_2, \quad C^3 = CC^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

So C and C^2 are invertible, with $C^{-1} = C^2$. We claim that $\Gamma_0, \Gamma_{I_2}, \Gamma_C, \Gamma_{C^2}, \Gamma_\infty$ is a spread.

Since $-1 = 1$ in $\mathbb{F}_2$, the differences of two distinct matrices among $0, I_2, C, C^2$ are

$$I_2 - 0 = I_2, \quad C - 0 = C, \quad C^2 - 0 = C^2, \quad C - I_2 = C^2, \quad C^2 - I_2 = C, \quad C^2 - C = I_2,$$

all invertible. If $(\mathbf{x}, M\mathbf{x}) = (\mathbf{x}', M'\mathbf{x}')$ with $M \neq M'$ among them, then $\mathbf{x} = \mathbf{x}'$ and $(M - M')\mathbf{x} = \mathbf{0}$, so $\mathbf{x} = (M - M')^{-1}\mathbf{0} = \mathbf{0}$. Hence $\Gamma_M \cap \Gamma_{M'} = \{\mathbf{0}\}$. Also, $(\mathbf{x}, M\mathbf{x}) = (\mathbf{0}, \mathbf{y})$ forces $\mathbf{x} = \mathbf{0}$ and then $\mathbf{y} = \mathbf{0}$, so $\Gamma_M \cap \Gamma_\infty = \{\mathbf{0}\}$. Any two of the five subspaces have dimensions adding up to 4 and meet only in $\mathbf{0}$, so they are complements of each other by Proposition 4.4.21. This spread has $2^2 + 1$ members, so the bound in (a) is attained.

The four matrices $0, I_2, C, C^2$, with matrix addition and multiplication, form the field with four elements (Section 14.3). The same construction, with $\mathbb{F}_{q^2}$ realized as q^2 matrices in $F^{2\times 2}$, gives a spread of size $q^2 + 1$ for every q.

Section 4.5

Solution to Exercise 4.5.9

Uses. Theorem 4.5.6 (Classification by dimension), Theorem 4.5.8 (Dimension of $\mathrm{L}(V, W)$), Example 4.1.7 (Dimensions of the standard spaces), Definition 2.4.26 (Trace)
Hint 1. By Theorem 4.5.6, it is enough to compute dimensions.
Hint 2. For (d), (e), and (g), write down a basis: which matrix units or monomials are allowed?
Sketch. The dimensions are $6, 6, 6, 6, 6, 3, 3$. So (a)–(e) form one class and (f), (g) another.

By Theorem 4.5.6, two finite-dimensional spaces are isomorphic exactly when their dimensions agree. We compute the dimensions.

(a) $2 \cdot 3 = 6$ (Example 4.1.7). **(b)** 6, with basis $1, x, \ldots, x^5$. **(c)** $3 \cdot 2 = 6$ by Theorem 4.5.8. **(d)** An upper triangular matrix is $\sum_{i \le j} a_{ij} U_{ij}$ with unique coefficients, so the six matrix units U_{ij} with $i \le j$ form a basis: dimension 6. **(e)** $p(0) = 0$ means that the constant coefficient vanishes, so $x, x^2, \ldots, x^6$ is a basis: dimension 6. **(f)** $1 \cdot 3 = 3$ by Theorem 4.5.8. **(g)** $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has trace 0 when $d = -a$, and then it equals $a(U_{11} - U_{22}) + bU_{12} + cU_{21}$, with unique coefficients. So $U_{11} - U_{22}, U_{12}, U_{21}$ is a basis: dimension 3.

So the classes are $\{(a), (b), (c), (d), (e)\}$ and $\{(f), (g)\}$.

Solution to Exercise 4.5.13

Uses. Proposition 4.5.12 (Calculus of change of coordinates), Definition 4.5.10 (Change-of-coordinates matrix), Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime)
Hint 1. $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = M^{-1}N$, where N and M have the vectors of $\mathcal{B}$ and $\mathring{\mathcal{B}}$ as columns. For a 2×2 matrix with $ad - bc \neq 0$, the inverse is $(ad - bc)^{-1}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ (Proposition 2.5.28).
Hint 2. In $\mathbb{F}_5$, $3^{-1} = 2$. Alternatively, find the columns $[\mathbf{b}_j]_{\mathring{\mathcal{B}}}$ by solving $x(1,1) + y(1,4) = \mathbf{b}_j$ directly.
Sketch. $M = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$ has $ad - bc = 3$, so $M^{-1} = 2\begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 2 \end{pmatrix}$, and $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = M^{-1}N = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix}$. Then $\mathbf{v} = (4, 6) = (4, 1)$ and $[\mathbf{v}]_{\mathring{\mathcal{B}}} = (5, 4) = (0, 4)$; indeed $4(1, 4) = (4, 16) = (4, 1)$.

(a) Let $N = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ and $M = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$. Then $1 \cdot 4 - 3 \cdot 2 = -2 = 3$ and $1 \cdot 4 - 1 \cdot 1 = 3$ are nonzero in $\mathbb{F}_5$, so both matrices are invertible (Proposition 2.5.28), their columns are independent (Corollary 3.5.25), and two independent vectors in $\mathbb{F}_5^2$ form a basis (Theorem 4.1.10(b)). Since $3 \cdot 2 = 6 = 1$, we have $3^{-1} = 2$ and

$$M^{-1} = 2\begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -2 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 2 \end{pmatrix}.$$

So

$$[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = M^{-1}N = \begin{pmatrix} 3 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 9 & 21 \\ 7 & 17 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix}.$$

Check: $4(1,1) + 2(1,4) = (6,12) = (1,2)$ and $1(1,1) + 2(1,4) = (3,9) = (3,4)$.

(b) $\mathbf{v} = (1,2) + (3,4) = (4,6) = (4,1)$, and

$$[\mathbf{v}]_{\mathring{\mathcal{B}}} = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

Check: $0(1,1) + 4(1,4) = (4,16) = (4,1) = \mathbf{v}$.

Solution to Exercise 4.5.20

Uses. Theorem 4.5.17 (Change of basis), Proposition 2.3.16 (The matrix of T_A), Definition 2.3.1 (Matrix of a linear map)
Hint 1. With $P = [\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, the formula is $[T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = P^{-1}AP$.
Hint 2. $P^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$. For the check, write $A(1,1)$ and $A(2,1)$ as combinations of $(1,1)$ and $(2,1)$.
Sketch. $AP = \begin{pmatrix} 2 & 6 \\ 2 & 3 \end{pmatrix}$ and $P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. Indeed $A(1,1) = (2,2) = 2(1,1)$ and $A(2,1) = (6,3) = 3(2,1)$.

By Proposition 2.3.16, $[T_A]_{\mathcal{E}\mathcal{E}} = A$, and $P = [\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ has the vectors of $\mathring{\mathcal{B}}$ as columns. It is invertible over every field, with $P^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$ (multiply to check). By Theorem 4.5.17,

$$[T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = P^{-1}AP = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 6 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Directly: $A(1,1) = (2,2) = 2(1,1) + 0(2,1)$ and $A(2,1) = (6,3) = 0(1,1) + 3(2,1)$, so the columns of $[T_A]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$ are $(2,0)$ and $(0,3)$. In the basis $\mathring{\mathcal{B}}$, T_A just scales each basis vector. Finding such bases is the subject of Section 7.4.

Solution to Exercise 4.5.32

Uses. Proposition 4.5.27 (Similarity invariants), Theorem 4.5.25 (Similar matrices are one operator in two bases), Definition 4.5.23 (Similar matrices)
Techniques. Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. Compare traces first, and remember that only I_2 is similar to I_2.
Hint 2. For (b), $N = M - I_2$ satisfies $N^2 = 0$ and $N \neq 0$. Find $\mathbf{w}$ with $N\mathbf{w} \neq \mathbf{0}$; what is the matrix of T_M in the basis $N\mathbf{w}, \mathbf{w}$?
Sketch. (a) No: $S^{-1}I_2S = I_2 \neq J$. (c) No: $\mathrm{tr} = 3 \neq 2$. (b) Yes: with $\mathbf{w} = \mathbf{e}_1$, $N\mathbf{w} = (1,-1)$; in the basis $(1,-1), (1,0)$, T_M fixes the first vector and sends the second to itself plus the first, so $S = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ works. Check $MS = SJ$.

(a) Not similar: $S^{-1}I_2S = I_2 \neq J$ for every invertible S.

(c) Not similar: $\mathrm{tr}\, M = 3 \neq 2 = \mathrm{tr}\, J$ (Proposition 4.5.27(b)).

(b) Similar. Both matrices have trace 2 and rank 2, which only says that similarity is possible. Let $N = M - I_2 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Then $N^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, while $N\mathbf{e}_1 = (1,-1) \neq \mathbf{0}$. Let $\mathring{\mathcal{B}} = (1,-1), (1,0)$, a basis since the vectors are not parallel. Then $M(1,-1) = (1,-1) + N(1,-1) = (1,-1)$, because $N(1,-1) = N^2\mathbf{e}_1 = \mathbf{0}$, and $M(1,0) = (1,0) + N\mathbf{e}_1 = (1,0) + (1,-1)$. So the columns of $[T_M]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}}$ are $(1,0)$ and $(1,1)$, that is, $[T_M]_{\mathring{\mathcal{B}}\mathring{\mathcal{B}}} = J$. By Theorem 4.5.25 and Theorem 4.5.17, $S = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, whose columns are the vectors of $\mathring{\mathcal{B}}$, satisfies $S^{-1}MS = J$. Check:

$$MS = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}, \quad SJ = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}.$$

Solution to Exercise 4.5.38

Uses. Proposition 4.5.12 (Calculus of change of coordinates), Corollary 3.5.25 (The invertible matrix theorem), Theorem 4.1.10 (Independent lists are short), Theorem 3.3.10 ($GA = \mathrm{rref}(A)$)

Hint 1. $[\mathrm{id}]_{\mathcal{BE}}$ is the inverse of the matrix M whose columns are the vectors of $\mathcal{B}$. Row reduce $(M \mid I_3)$ over $\mathbb{F}_3$, where $2^{-1} = 2$.
Hint 2. For (c), $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = \left([\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}}\right)^{-1}[\mathrm{id}]_{\mathcal{EB}}$. For (d), add the three vectors.
Sketch. $M^{-1} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{pmatrix}$ over $\mathbb{F}_3$ (over $\mathbb{Q}$ it is $\frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$, and $1/2 = 2$ in $\mathbb{F}_3$). Then $[(1,2,0)]_{\mathcal{B}} = (0,2,1)$, and $[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$. Over $\mathbb{F}_2$ the three vectors add to $\mathbf{0}$.

(a) Let $M = [\mathrm{id}]_{\mathcal{EB}} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Row reducing $(M \mid I_3)$ over $\mathbb{F}_3$ (or inverting over $\mathbb{Q}$ and reading $1/2$ as 2, since $2 \cdot 2 = 4 = 1$) gives

$$M^{-1} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{pmatrix},$$

and one checks $MM^{-1} = I_3$ over $\mathbb{F}_3$; for instance, the first row of the product is $(1\cdot 2 + 1\cdot 2, 1\cdot 2 + 1\cdot 1, 1\cdot 1 + 1\cdot 2) = (4,3,3) = (1,0,0)$. So M is invertible, its columns form a basis of $\mathbb{F}_3^3$ (Corollary 3.5.25, Theorem 4.1.10(b)), and $[\mathrm{id}]_{\mathcal{BE}} = M^{-1}$ by Proposition 4.5.12(b).

(b) $[(1,2,0)]_{\mathcal{B}} = M^{-1}(1,2,0) = (2+4, 1+4, 2+2) = (6,5,4) = (0,2,1)$. Check: $2(0,1,1) + (1,0,1) = (1,2,3) = (1,2,0)$.

(c) $[\mathrm{id}]_{\mathcal{E}\mathring{\mathcal{B}}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ has inverse $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, so

$$[\mathrm{id}]_{\mathring{\mathcal{B}}\mathcal{B}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

Check the first column: $0(1,0,0) + 1(1,1,0) + 0(1,1,1) = (1,1,0)$, the first vector of $\mathcal{B}$.

(d) No: over $\mathbb{F}_2$, $(1,1,0) + (0,1,1) + (1,0,1) = (2,2,2) = \mathbf{0}$, a nontrivial dependence.

Solution to Exercise 4.5.39

Uses. Proposition 4.5.12 (Calculus of change of coordinates), Example 4.5.26 (A similarity and a non-similarity), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Remark 4.5.28 (What the invariants cannot do), Theorem 4.5.8 (Dimension of $\mathrm{L}(V,W)$), Theorem 4.5.6 (Classification by dimension)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Two of the six are false, and the counterexamples are in this section.
Hint 2. For (c), compute $\dim \operatorname{im} T$. For (f), compute $S^{-1}(cI_n)S$.
Sketch. (a) True: the inverse of $[\mathrm{id}]_{\mathcal{B}'\mathcal{B}}$ is $[\mathrm{id}]_{\mathcal{B}\mathcal{B}'}$. (b) False: $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ over $\mathbb{Q}$. (c) True by rank–nullity. (d) False: I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. (e) True: $2\dim W = 6$. (f) True: $S^{-1}(cI_n)S = cI_n$.

(a) True, by Proposition 4.5.12(b).

(b) False. Over $\mathbb{Q}$, $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ (Example 4.5.26), and the diagonals $(1,1)$ and $(2,0)$ differ.

(c) True. $\ker T = \{\mathbf{0}\}$, so by Theorem 4.2.11, $\dim \operatorname{im} T = \dim V = \dim W$. Then $\operatorname{im} T = W$ by Proposition 4.1.21(b), so T is bijective, hence an isomorphism.

(d) False. I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ both have rank 2 and trace 2, and they are not similar (Remark 4.5.28).

(e) True. W is finite-dimensional: if $\mathbf{v}_1, \mathbf{v}_2$ is a basis of V, then $T \mapsto T(\mathbf{v}_1)$ is a surjective linear map $\mathrm{L}(V,W) \to W$ (every $\mathbf{w}$ is $T(\mathbf{v}_1)$ for the T with $\mathbf{v}_1 \mapsto \mathbf{w}$ and $\mathbf{v}_2 \mapsto \mathbf{0}$), so $\dim W \le 6$ by Theorem 4.2.11. Then Theorem 4.5.8 gives $2\dim W = 6$, so $\dim W = 3$ and $W \cong F^3$ (Theorem 4.5.6).

(f) True. $S^{-1}(cI_n)S = cS^{-1}S = cI_n$ for every invertible S.

Solution to Exercise 4.5.40

Uses. Corollary 4.1.11 (Spanning lists are long), Definition 4.5.10 (Change-of-coordinates matrix), Proposition 4.5.12 (Calculus of change of coordinates)
Hint 1. Write $x = (x-a)+a$ and expand x^j by the binomial theorem, which holds in $F[x]$ because multiplication of polynomials is commutative.
Hint 2. For (b), expand $(x-a)^j$ in powers of x. For (c), the two matrices of (a) and (b) are inverse to each other; compute the (i,j) entry of their product.
Sketch. $x^j = \sum_i \binom{j}{i} a^{j-i}(x-a)^i$ shows that $\mathcal{B}_a$ spans, hence is a basis, and gives the columns of $[\mathrm{id}]_{\mathcal{B}_a\mathcal{B}}$. Likewise $(x-a)^j = \sum_i \binom{j}{i}(-a)^{j-i}x^i$. The product $[\mathrm{id}]_{\mathcal{B}\mathcal{B}_a}[\mathrm{id}]_{\mathcal{B}_a\mathcal{B}} = I$ has (i,j) entry $a^{j-i}\sum_k (-1)^{k-i}\binom{k}{i}\binom{j}{k}$.

Here an integer such as $\binom{j}{i}$ stands for the corresponding sum of 1's in F.

(a) By the binomial theorem in $F[x]$,

$$x^j = ((x-a)+a)^j = \sum_{i=0}^{j} \binom{j}{i} a^{j-i}(x-a)^i.$$

So $1, x, \ldots, x^n$ lie in $\mathrm{span}(\mathcal{B}_a)$, which is therefore all of $F[x]_{\le n}$. A spanning list of length $n+1 = \dim F[x]_{\le n}$ is a basis (Corollary 4.1.11(b)). The display gives $[x^j]_{\mathcal{B}_a}$, the j-th column of $[\mathrm{id}]_{\mathcal{B}_a\mathcal{B}}$: its i-th entry is $\binom{j}{i}a^{j-i}$ for $i \le j$ and 0 for $i > j$.

(b) Similarly $(x-a)^j = \sum_{i=0}^{j}\binom{j}{i}(-a)^{j-i}x^i$, so the (i,j) entry of $[\mathrm{id}]_{\mathcal{B}\mathcal{B}_a}$ is $\binom{j}{i}(-a)^{j-i}$ for $i \le j$ and 0 otherwise.

(c) By Proposition 4.5.12(b), the two matrices are inverse to each other. The (i,j) entry of $[\mathrm{id}]_{\mathcal{B}\mathcal{B}_a}[\mathrm{id}]_{\mathcal{B}_a\mathcal{B}}$ is, for $i \le j$,

$$\sum_{k=i}^{j} \binom{k}{i}(-a)^{k-i}\binom{j}{k}a^{j-k} = a^{j-i}\sum_{k=i}^{j}(-1)^{k-i}\binom{j}{k}\binom{k}{i},$$

since only $i \le k \le j$ contribute. With $F = \mathbb{Q}$ and $a = 1$, this entry of I_{n+1} is 1 if $i = j$ and 0 if $i < j$, which is the identity. (The case $j > n$ is covered by taking n larger.)

Solution to Exercise 4.5.41

Uses. Theorem 4.1.16 (Extending to a basis), Theorem 4.5.8 (Dimension of $\mathrm{L}(V,W)$), Proposition 4.5.5 (Isomorphisms preserve linear structure), Proposition 1.6.7 (Standard bases)
Techniques. Choose a convenient basis (Technique 4.5.33), Transport along an isomorphism (Technique 4.5.7), Dimension counting (Technique 4.2.24)
Hint 1. Extend $\mathbf{v}_0$ to a basis $\mathcal{B}$ of V with $\mathbf{v}_0$ first. What does $T(\mathbf{v}_0) = \mathbf{0}$ say about $[T]_{\mathcal{C}\mathcal{B}}$?
Hint 2. The condition is that the first column is zero. Transport K to matrices with Theorem 4.5.8.
Sketch. With $\mathbf{v}_0$ the first vector of $\mathcal{B}$, $T \in K$ iff the first column of $[T]_{\mathcal{C}\mathcal{B}}$ is $\mathbf{0}$. The isomorphism $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ carries K onto the matrices with zero first column, a subspace with basis U_{ij}, $j \ge 2$, of dimension $(n-1)m$.

If $m = 0$, then $\mathrm{L}(V,W) = \{0\} = K$, of dimension 0. Let $m \ge 1$. Extend $\mathbf{v}_0$ to a basis $\mathcal{B} = \mathbf{v}_0, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of V (Theorem 4.1.16), and choose a basis $\mathcal{C}$ of W. The first column of $[T]_{\mathcal{C}\mathcal{B}}$ is $[T(\mathbf{v}_0)]_{\mathcal{C}}$, which is $\mathbf{0}$ exactly when $T(\mathbf{v}_0) = \mathbf{0}$. So the isomorphism $\Phi : T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ of Theorem 4.5.8 carries K onto

$$Z = \{A \in F^{m\times n} \mid \text{the first column of } A \text{ is } \mathbf{0}\}.$$

Z consists of the combinations of the matrix units U_{ij} with $j \ge 2$, which are independent (Proposition 1.6.7(b)); so Z is a subspace with $\dim Z = m(n-1)$. Then $K = \mathrm{im}(\Phi^{-1}|_Z)$ is a subspace of $\mathrm{L}(V,W)$ with $\dim K = \dim Z$, by Proposition 4.5.5(b) and (c) applied to the isomorphism Φ^{-1}.

Solution to Exercise 4.5.41

Uses. Definition 2.1.21 (Pointwise operations), Theorem 4.1.16 (Extending to a basis), Theorem 2.2.1 (Linear maps are determined by a basis), Theorem 4.2.11 (Rank–nullity), Theorem 4.5.8 (Dimension of $\mathrm{L}(V,W)$), Proposition 2.6.8 (Kernels and images are subspaces)
Techniques. Dimension counting (Technique 4.2.24), Choose a convenient basis (Technique 4.5.33)
Hint 1. K is the kernel of a linear map $\mathrm{L}(V,W) \to W$. Which one?

Hint 2. Evaluation $\varepsilon(T) = T(\mathbf{v}_0)$ is linear and surjective. Apply rank–nullity in $\mathrm{L}(V, W)$, whose dimension you know.
Sketch. $\varepsilon : T \mapsto T(\mathbf{v}_0)$ is linear with kernel K, and surjective because a map can send $\mathbf{v}_0$ anywhere. So $\dim K = nm - m$.

Let $\varepsilon : \mathrm{L}(V, W) \to W$, $\varepsilon(T) = T(\mathbf{v}_0)$. It is linear because the operations on $\mathrm{L}(V, W)$ are pointwise: $(S + cT)(\mathbf{v}_0) = S(\mathbf{v}_0) + cT(\mathbf{v}_0)$. Its kernel is K, which is therefore a subspace (Proposition 2.6.8). It is surjective: given $\mathbf{w} \in W$, extend $\mathbf{v}_0$ to a basis $\mathbf{v}_0, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of V, and let T be the linear map with $T(\mathbf{v}_0) = \mathbf{w}$ and $T(\mathbf{v}_j) = \mathbf{0}$ for $j \geq 2$ (Theorem 2.2.1). By Theorem 4.2.11(b) and Theorem 4.5.8,

$$\dim K = \dim \mathrm{L}(V, W) - \dim W = nm - m = (n-1)m.$$

Solution to Exercise 4.5.42

Uses. Theorem 4.1.16 (Extending to a basis), Theorem 4.5.8 (Dimension of $\mathrm{L}(V, W)$), Proposition 4.5.5 (Isomorphisms preserve linear structure), Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Proposition 1.6.7 (Standard bases), Proposition 2.6.22 (Testing injectivity and surjectivity on a basis)
Techniques. Choose a convenient basis (Technique 4.5.33), Choose a basis adapted to a preserved subspace (Technique 4.5.36), Transport along an isomorphism (Technique 4.5.7)
Hint 1. Choose a basis of V that starts with a basis of U, and a basis of W that starts with a basis of U'.
Hint 2. In these bases, the condition says that certain entries of $[T]_{\mathcal{C}\mathcal{B}}$ vanish. Which ones, and how many?
Hint 3. Compare the proof of Proposition 4.5.35. Then transport along $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$.
Sketch. With adapted bases, $\mathrm{im}(T\restriction_U) \subseteq U'$ iff $T(\mathbf{u}_j) \in U'$ for $j \leq k$, iff the first k columns of $[T]_{\mathcal{C}\mathcal{B}}$ vanish in rows $k'+1, \ldots, m$. Those $k(m-k')$ entries are forced to be 0 and the other $mn - k(m-k')$ are free, so the set corresponds to the span of the remaining matrix units.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be a basis of U, extended to a basis $\mathcal{B}$ of V, and $\mathbf{u}'_1, \ldots, \mathbf{u}'_{k'}$ a basis of U', extended to a basis $\mathcal{C}$ of W (Theorem 4.1.16). A vector of W lies in U' if and only if its $\mathcal{C}$-coordinates vanish after position k', as in the proof of Proposition 4.5.35. Since $T(\mathbf{u}_1), \ldots, T(\mathbf{u}_k)$ spans $\mathrm{im}(T\restriction_U)$ (Proposition 2.6.22(d)) and U' is a subspace, $\mathrm{im}(T\restriction_U) \subseteq U'$ holds if and only if $T(\mathbf{u}_j) \in U'$ for $j = 1, \ldots, k$, that is, if and only if the entries of $[T]_{\mathcal{C}\mathcal{B}}$ in rows $k'+1, \ldots, m$ and columns $1, \ldots, k$ are all 0.

So the isomorphism $T \mapsto [T]_{\mathcal{C}\mathcal{B}}$ (Theorem 4.5.8) carries our set onto the set Z of matrices whose entries in those $k(m-k')$ positions vanish. Z is the span of the matrix units U_{ij} in the other $mn - k(m-k')$ positions, which are independent (Proposition 1.6.7(b)); so Z is a subspace of that dimension. By Proposition 4.5.5(b), (c), applied to the inverse isomorphism, our set is a subspace of $\mathrm{L}(V, W)$ of the same dimension. For $V = W$ and $U = U'$, take $m = n$ and $k' = k$.

Solution to Exercise 4.5.43

Uses. Exercise 2.5.46, Theorem 4.1.10 (Independent lists are short), Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.5.24 (Similarity is an equivalence relation), Proposition 1.5.14 (Appending a vector)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint 1. (a) Apply the hypothesis to $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_1 + \mathbf{e}_2$.
Hint 2. (b) By (a) there is $\mathbf{v}$ with $A\mathbf{v} \notin \mathrm{span}(\mathbf{v})$. For the matrix, you need $A(A\mathbf{v})$ in terms of $\mathbf{v}$ and $A\mathbf{v}$: use $A^2 - (\mathrm{tr}\, A)A + \delta(A)I_2 = 0$ (Exercise 2.5.46(a)).
Hint 3. (c) Both matrices are similar to the same C. (d) What are the trace and δ of A^{T}, and when is A^{T} scalar?
Sketch. (a) $A\mathbf{e}_1 = \lambda\mathbf{e}_1$, $A\mathbf{e}_2 = \mu\mathbf{e}_2$, and $A(\mathbf{e}_1 + \mathbf{e}_2) = \nu(\mathbf{e}_1 + \mathbf{e}_2)$ force $\lambda = \nu = \mu$. (b) Pick $\mathbf{v}$ with $A\mathbf{v} \notin \mathrm{span}(\mathbf{v})$; then $\mathbf{v}, A\mathbf{v}$ is independent, and $A^2\mathbf{v} = (\mathrm{tr}\, A)A\mathbf{v} - \delta(A)\mathbf{v}$ gives the second column. (c) Both are similar to the same C_A. (d) A^{T} has the same trace and δ, and is scalar iff A is. (e) Both matrices have trace -1 and $\delta = 1$ and are not scalar.

(a) By hypothesis $A\mathbf{e}_1 = \lambda\mathbf{e}_1$, $A\mathbf{e}_2 = \mu\mathbf{e}_2$, and $A(\mathbf{e}_1 + \mathbf{e}_2) = \nu(\mathbf{e}_1 + \mathbf{e}_2)$ for some scalars. The last equation reads $\lambda\mathbf{e}_1 + \mu\mathbf{e}_2 = \nu\mathbf{e}_1 + \nu\mathbf{e}_2$, so $\lambda = \nu = \mu$. The columns of A are then $\lambda\mathbf{e}_1$ and $\lambda\mathbf{e}_2$, that is, $A = \lambda I_2$.

(b) By (a) there is $\mathbf{v}$ with $A\mathbf{v} \notin \mathrm{span}(\mathbf{v})$. Then $\mathbf{v} \neq \mathbf{0}$ (as $A\mathbf{0} = \mathbf{0}$ lies in every span), so the list $\mathbf{v}$ is independent, and $\mathbf{v}, A\mathbf{v}$ is independent by Proposition 1.5.14. Two independent vectors in F^2 form a basis (Theorem 4.1.10(b)). In this basis, T_A sends the first vector $\mathbf{v}$ to the second vector $A\mathbf{v}$, with coordinates $(0, 1)$. By Exercise 2.5.46(a), $A^2 = (\mathrm{tr}\, A)A - \delta(A)I_2$, so

$$A(A\mathbf{v}) = -\delta(A)\mathbf{v} + (\mathrm{tr}\, A)A\mathbf{v},$$

with coordinates $(-\delta(A), \mathrm{tr}\, A)$. So the matrix is C_A.

(c) Let A and B be non-scalar with $\operatorname{tr} A = \operatorname{tr} B$ and $\delta(A) = \delta(B)$. Then $C_A = C_B$. By (b) and Theorem 4.5.25, $A \sim C_A$ and $B \sim C_B = C_A$, so $A \sim B$ (Proposition 4.5.24).

(d) If $A = cI_2$, then $A^{\mathsf{T}} = A$. Otherwise A^{T} is not scalar either, since $(cI_2)^{\mathsf{T}} = cI_2$ and $\left(A^{\mathsf{T}}\right)^{\mathsf{T}} = A$. The diagonal of A^{T} is that of A, and $\delta\left(A^{\mathsf{T}}\right) = ad - cb = \delta(A)$. So $A \sim A^{\mathsf{T}}$ by (c).

(e) R is not scalar, $\operatorname{tr} R = -1$, and $\delta(R) = 0 \cdot (-1) - (-1) \cdot 1 = 1$. The rotation matrix is not scalar, its trace is $2\cos(2\pi/3) = -1$, and its δ is $\cos^2 + \sin^2 = 1$. By (c), the two are similar over $\mathbb{R}$.

Solution to Exercise 4.5.44

Uses. Exercise 4.5.43, Proposition 4.5.27 (Similarity invariants), Definition 4.5.23 (Similar matrices)
Techniques. Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. (a) is a direct expansion; then $\delta\left(S^{-1}\right)\delta(S) = \delta(I_2) = 1$.
Hint 2. (b) Existence is Exercise 4.5.43(b). For uniqueness, which invariants separate the matrices in the list? What are the trace and δ of $\left(\begin{smallmatrix} 0 & -d \\ 1 & t \end{smallmatrix}\right)$?
Sketch. (a) Expanding, $\delta(XY) = (ad - bc)(eh - fg)$ for $X = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $Y = \left(\begin{smallmatrix} e & f \\ g & h \end{smallmatrix}\right)$; so $\delta\left(S^{-1}AS\right) = \delta\left(S^{-1}\right)\delta(A)\delta(S) = \delta(A)$. (b) A scalar matrix is similar only to itself; a non-scalar A is similar to $\left(\begin{smallmatrix} 0 & -\delta(A) \\ 1 & \operatorname{tr} A \end{smallmatrix}\right)$, which is not scalar; and $\left(\begin{smallmatrix} 0 & -d \\ 1 & t \end{smallmatrix}\right)$ has trace t and $\delta = d$, so different (t, d) give non-similar matrices. (c) q scalars plus q^2 pairs (t, d).

(a) For $X = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} e & f \\ g & h \end{smallmatrix}\right)$, $XY = \left(\begin{smallmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{smallmatrix}\right)$, and expanding,

$$\begin{aligned} \delta(XY) &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\ &= adeh + bcfg - adfg - bceh = (ad - bc)(eh - fg), \end{aligned}$$

since the terms $acef$ and $bdgh$ cancel. If $B = S^{-1}AS$, then $\delta\left(S^{-1}\right)\delta(S) = \delta\left(S^{-1}S\right) = \delta(I_2) = 1$, and since scalars commute,

$$\delta(B) = \delta\left(S^{-1}\right)\delta(A)\delta(S) = \delta(A).$$

(b) *Existence.* If A is scalar, it is in the list. Otherwise $A \sim \left(\begin{smallmatrix} 0 & -\delta(A) \\ 1 & \operatorname{tr} A \end{smallmatrix}\right)$ by Exercise 4.5.43(b) and Theorem 4.5.25.

Uniqueness. A scalar matrix cI_2 is similar only to itself (Exercise 4.5.39(f)), so it is not similar to any other matrix of the list; the matrices $\left(\begin{smallmatrix} 0 & -d \\ 1 & t \end{smallmatrix}\right)$ are not scalar, because of the entry 1. Finally, $\left(\begin{smallmatrix} 0 & -d \\ 1 & t \end{smallmatrix}\right)$ has trace t and $\delta = 0 \cdot t - (-d) \cdot 1 = d$. By (a) and Proposition 4.5.27(b), two matrices of this form with different pairs (t, d) are not similar.

(c) Over $\mathbb{F}_q$ the list has q scalar matrices and q^2 matrices $\left(\begin{smallmatrix} 0 & -d \\ 1 & t \end{smallmatrix}\right)$, and by (b) each similarity class contains exactly one of them. So there are $q^2 + q$ classes.

Solution to Exercise 4.5.45

Uses. Definition 4.5.23 (Similar matrices), Corollary 4.5.21 (Matrices of one linear map), Proposition 4.5.27 (Similarity invariants)
Techniques. Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. Compare the last sentence with Definition 4.5.23. Which relation between the left and right factors does similarity require?
Hint 2. For a counterexample, look for two 2×2 matrices of rank 1 that differ in an invariant.
Sketch. Similarity needs $B = S^{-1}AS$, with the left factor the inverse of the right one; nothing makes $G'^{-1}G = S^{-1}$. The argument proves only $B = GAH$, which is Corollary 4.5.21. Counterexample: $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ have rank 1 but traces 0 and 1.

The error is in the last sentence. Similarity requires $B = S^{-1}AS$: the left factor must be the inverse of the right factor. The argument produces $B = G''AS$ with $G'' = G'^{-1}G$ and $S = HH'^{-1}$ both invertible, but nothing forces $G'' = S^{-1}$. What it proves is that A and B are matrices of one linear map with possibly different bases on the domain and the target (Corollary 4.5.21).

The claim is false: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ both have rank 1, but their traces are 0 and 1, so they are not similar (Proposition 4.5.27(b)).

Solution to Exercise 4.5.46

Uses. Definition 4.5.23 (Similar matrices), Proposition 4.5.27 (Similarity invariants), Lemma 2.4.24 (Matrix units)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. Only one statement is true. For counterexamples, use matrix units, and remember that only 0 is similar to 0 and only I_n to I_n.
Hint 2. $U_{11} \sim U_{22}$ (swap the basis vectors). Try $C = U_{11}$ in (a) and (b).
Sketch. (a) False: $U_{11} \sim U_{22}$ and $U_{11} \sim U_{11}$, but $2U_{11}$ (rank 1 or 0) is not similar to $U_{11} + U_{22}$ (rank 2). (b) False: $U_{11}U_{11} = U_{11}$ but $U_{22}U_{11} = 0$. (c) True by Proposition 4.5.27(d). (d) False: $U_{12}^2 = 0 = 0^2$.

We write U_{ij} for the matrix units (here in $F^{n\times n}$). Swapping $\mathbf{e}_1$ and $\mathbf{e}_2$ in the standard basis shows $U_{11} \sim U_{22}$: with S the matrix of the swap, $S^{-1}U_{11}S = U_{22}$.

(a) False. Take $A = C = D = U_{11}$ and $B = U_{22}$. Then $A + C = 2U_{11}$, while $B + D = U_{11} + U_{22}$ has rank 2. If $2 \neq 0$ in F, then $2U_{11}$ has rank 1; if $2 = 0$, it is the zero matrix, of rank 0. Either way the ranks differ, so the sums are not similar.

(b) False. With $A = U_{11}$, $B = U_{22}$, and $C = U_{11}$, we get $AC = U_{11}$ and $BC = U_{22}U_{11} = 0$ (Lemma 2.4.24(c)). Only 0 is similar to 0.

(c) True: apply Proposition 4.5.27(d) with $-c$ in place of c.

(d) False. $A = U_{12}$ and $B = 0$ satisfy $A^2 = 0 = B^2$, but $A \neq 0$ is not similar to 0.

Solution to Exercise 4.5.47

Uses. Lemma 2.4.24 (Matrix units), Proposition 2.4.27 (The trace is linear), Definition 2.4.26 (Trace), Proposition 1.6.7 (Standard bases)
Techniques. Cyclicity of trace (Technique 2.4.29), Check it on a basis (Technique 2.2.7)
Hint 1. Apply the hypothesis to $A = SX$: it says $f(XS) = f(SX)$ for every X and every invertible S.
Hint 2. For $k \neq l$, the matrix $I_n + U_{kl}$ is invertible, with inverse $I_n - U_{kl}$. Deduce $f(XU_{kl}) = f(U_{kl}X)$.
Hint 3. Take $X = U_{ik}$. Then $XU_{kl} = U_{il}$ and $U_{kl}X = \delta_{li}U_{kk}$. What does this say about $f(U_{il})$ for $i \neq l$, and about $f(U_{ll})$?
Sketch. Invariance with $A = SX$ gives $f(XS) = f(SX)$. With $S = I + U_{kl}$ $(k \neq l)$ and linearity, $f(XU_{kl}) = f(U_{kl}X)$. For $X = U_{ik}$ this reads $f(U_{il}) = \delta_{li}f(U_{kk})$. So $f(U_{il}) = 0$ for $i \neq l$ and $f(U_{ll}) = f(U_{kk})$ for all $k \neq l$. With c the common value, $f(A) = \sum a_{ij}f(U_{ij}) = c\,\mathrm{tr}\,A$.

If $n = 1$, then $f(a) = af(1) = f(1)\,\mathrm{tr}(a)$ by linearity. Let $n \geq 2$.

Step 1. For $X \in F^{n\times n}$ and invertible S, the hypothesis with $A = SX$ gives $f\big(S^{-1}SXS\big) = f(SX)$, that is,

$$f(XS) = f(SX).$$

Step 2. Let $k \neq l$. By Lemma 2.4.24(c), $U_{kl}U_{kl} = \delta_{lk}U_{kl} = 0$, so $(I_n + U_{kl})(I_n - U_{kl}) = I_n - U_{kl}^2 = I_n$, and likewise in the other order. So $S = I_n + U_{kl}$ is invertible, and Step 1 gives $f(X + XU_{kl}) = f(X + U_{kl}X)$. By linearity,

$$f(XU_{kl}) = f(U_{kl}X) \quad \text{for all } X \text{ and all } k \neq l.$$

Step 3. Take $X = U_{ik}$. By Lemma 2.4.24(c), $XU_{kl} = U_{il}$ and $U_{kl}X = \delta_{li}U_{kk}$, so

$$f(U_{il}) = \delta_{li}f(U_{kk}) \quad \text{for all } i \text{ and all } k \neq l.$$

If $i \neq l$, choose any $k \neq l$ (possible since $n \geq 2$); then $f(U_{il}) = 0$. If $i = l$, we get $f(U_{ll}) = f(U_{kk})$ for every $k \neq l$. So all $f(U_{kk})$ are equal to a single $c \in F$.

Conclusion. For $A = \sum_{i,j} a_{ij}U_{ij}$ (Proposition 1.6.7(b)), linearity gives

$$f(A) = \sum_{i,j} a_{ij}f\big(U_{ij}\big) = c\sum_i a_{ii} = c\,\mathrm{tr}\,A.$$

Conversely, every $c\,\mathrm{tr}$ is linear and similarity-invariant (Proposition 2.4.27, Proposition 4.5.27(b)).

Solution to Exercise 4.5.48

Uses. Definition 4.5.30 (Trace of an operator), Lemma 2.4.24 (Matrix units), Proposition 1.6.7 (Standard bases), Proposition 2.4.12 (Matrices of sums and scalar multiples), Proposition 2.4.27 (The trace is linear)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. Use the basis of matrix units. The diagonal entry of $[T]$ belonging to U_{ij} is the U_{ij}-coordinate of $T(U_{ij})$, which is the (i,j) entry of $AU_{ij}B$.
Hint 2. Write $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$, so the (i,j) entry of $AU_{ij}B$ is $\left(\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_i\right)\left(\mathbf{e}_j^{\mathsf{T}}B\mathbf{e}_j\right)$.
Sketch. The (i,j) entry of $AU_{ij}B$ is $a_{ii}b_{jj}$, so $\operatorname{tr} T = \sum_{i,j} a_{ii}b_{jj} = (\operatorname{tr} A)(\operatorname{tr} B)$. With $B = I_n$ or $A = I_n$ the traces are $n \operatorname{tr} A$ and $n \operatorname{tr} B$, and $X \mapsto AX - XA$ has trace $n \operatorname{tr} A - n \operatorname{tr} A = 0$.

In the basis of matrix units $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Proposition 1.6.7(b), Lemma 2.4.24), the U_{ij}-coordinate of a matrix is its (i,j) entry. So the diagonal entry of $[T]$ in the position of U_{ij} is the (i,j) entry of $T(U_{ij}) = A\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}B$, which by Lemma 2.4.24(b) is

$$\mathbf{e}_i^{\mathsf{T}}A\mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}B\mathbf{e}_j = a_{ii}b_{jj}.$$

Summing over all i and j (Definition 4.5.30),

$$\operatorname{tr} T = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ii}b_{jj} = (\operatorname{tr} A)(\operatorname{tr} B).$$

With $B = I_n$, the operator $X \mapsto AX$ has trace $(\operatorname{tr} A)(\operatorname{tr} I_n) = n \operatorname{tr} A$; with $A = I_n$, $X \mapsto XB$ has trace $n \operatorname{tr} B$. The trace of an operator is linear in the operator, by Proposition 2.4.12 and Proposition 2.4.27, so $X \mapsto AX - XA$ has trace $n \operatorname{tr} A - n \operatorname{tr} A = 0$.

Solution to Exercise 4.5.49

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 1.5.14 (Appending a vector), Theorem 4.1.10 (Independent lists are short), Definition 4.5.30 (Trace of an operator), Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.5.27 (Similarity invariants), Proposition 4.5.24 (Similarity is an equivalence relation)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint 1. If $\mathbf{v} = P\mathbf{w}$, compute $P\mathbf{v}$.
Hint 2. For independence of the combined list, split a relation into a part in $\operatorname{im} P$ and a part in $\ker P$. Count with rank–nullity.
Hint 3. For (c), both matrices are similar to the same N_r.
Sketch. If $\mathbf{v} = P\mathbf{w}$ then $P\mathbf{v} = P^2\mathbf{w} = \mathbf{v}$, so a vector in $\ker P \cap \operatorname{im} P$ equals $P\mathbf{v} = \mathbf{0}$. A relation among the combined list gives a vector in both subspaces, hence zero, so the list is independent; it has $r + (n-r) = n$ vectors. P fixes the first r and kills the rest, so its matrix is N_r and $\operatorname{tr} P = r$. In (c), $A \sim N_r \sim B$.

(a) If $\mathbf{v} = P\mathbf{w} \in \operatorname{im} P$, then $P\mathbf{v} = P^2\mathbf{w} = P\mathbf{w} = \mathbf{v}$. If also $\mathbf{v} \in \ker P$, then $\mathbf{v} = P\mathbf{v} = \mathbf{0}$. Let $\mathbf{p}_1, \dots, \mathbf{p}_r$ be a basis of $\operatorname{im} P$ and $\mathbf{k}_1, \dots, \mathbf{k}_s$ a basis of $\ker P$, where $s = n - r$ by Theorem 4.2.11. Suppose $\sum_i a_i\mathbf{p}_i + \sum_j b_j\mathbf{k}_j = \mathbf{0}$. The vector $\mathbf{x} = \sum_i a_i\mathbf{p}_i = -\sum_j b_j\mathbf{k}_j$ lies in $\operatorname{im} P \cap \ker P = \{\mathbf{0}\}$, so all a_i and all b_j are 0 by independence of the two bases. So the combined list of n vectors is independent, hence a basis of V (Theorem 4.1.10(b)).

(b) $P\mathbf{p}_i = \mathbf{p}_i$ and $P\mathbf{k}_j = \mathbf{0}$, so the matrix of P in this basis has columns $\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{0}, \dots, \mathbf{0}$; it is N_r. Hence $\operatorname{tr} P = \operatorname{tr} N_r = r \cdot 1$ (Definition 4.5.30). Over a field of characteristic p, this determines r only modulo p.

(c) If $A \sim B$, then $\operatorname{rank} A = \operatorname{rank} B$ (Proposition 4.5.27(a)). Conversely, if both have rank r, apply (b) to T_A and T_B on F^n: each has matrix N_r in a suitable basis, so $A \sim N_r$ and $B \sim N_r$ by Theorem 4.5.25, and $A \sim B$ by Proposition 4.5.24.

Solution to Exercise 4.5.50

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.5.24 (Similarity is an equivalence relation), Proposition 2.4.23 (Transpose of a product)
Techniques. Choose a convenient basis (Technique 4.5.33), Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Choose $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$, and look at $\mathbf{v}, N\mathbf{v}, \dots, N^{n-1}\mathbf{v}$.
Hint 2. These vectors are independent by Proposition 4.1.26. Where does N send each of them?
Hint 3. For (c), check that J_n^{T} satisfies the hypotheses of (b). The transpose of a power is the power of the transpose.

Sketch. Pick $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$; since $N^n\mathbf{v} = \mathbf{0}$, Proposition 4.1.26 makes $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ independent, hence a basis. N sends each vector to the next and the last to $\mathbf{0}$, which is J_n. (b) T_A and T_B both have matrix J_n. (c) $\left(J_n^{\mathsf{T}}\right)^k = \left(J_n^k\right)^{\mathsf{T}}$.

(a) Since $N^{n-1} \neq 0$, there is $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$, and $N^n\mathbf{v} = \mathbf{0}$. By Proposition 4.1.26 with $k = n-1$, the list $\mathcal{B} = \mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is independent, and it has n vectors, so it is a basis (Theorem 4.1.10(b)). N sends the j-th vector $N^{j-1}\mathbf{v}$ to $N^j\mathbf{v}$, the $(j+1)$-th vector, for $j < n$, and the last vector to $N^n\mathbf{v} = \mathbf{0}$. So the columns of $[N]_{\mathcal{B}\mathcal{B}}$ are $\mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{0}$, and $[N]_{\mathcal{B}\mathcal{B}} = J_n$.

(b) T_A and T_B satisfy the hypotheses of (a) on F^n (since $(T_A)^k = T_{A^k}$). So $A \sim J_n$ and $B \sim J_n$ by Theorem 4.5.25, and $A \sim B$ (Proposition 4.5.24).

(c) $J_n\mathbf{e}_j = \mathbf{e}_{j+1}$ for $j < n$ and $J_n\mathbf{e}_n = \mathbf{0}$, so $J_n^{n-1}\mathbf{e}_1 = \mathbf{e}_n \neq \mathbf{0}$ and $J_n^n = 0$. By Proposition 2.4.23, $\left(J_n^{\mathsf{T}}\right)^k = \left(J_n^k\right)^{\mathsf{T}}$, so J_n^{T} satisfies the same two conditions, and $J_n \sim J_n^{\mathsf{T}}$ by (b). Concretely, reversing the order of the basis $\mathcal{B}$ turns J_n into J_n^{T}.

Solution to Exercise 4.5.51

Uses. Proposition 4.3.16 (Rank-one matrices), Proposition 2.4.28 (Cyclicity of the trace), Theorem 4.2.11 (Rank–nullity), Theorem 4.1.16 (Extending to a basis), Proposition 1.5.14 (Appending a vector), Theorem 4.1.10 (Independent lists are short), Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.5.27 (Similarity invariants)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17), Choose a convenient basis (Technique 4.5.33)
Hint 1. Write $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$. Then $A\mathbf{x} = \left(\mathbf{v}^{\mathsf{T}}\mathbf{x}\right)\mathbf{u}$, so $\ker A = \left\{\mathbf{x} \;\middle|\; \mathbf{v}^{\mathsf{T}}\mathbf{x} = 0\right\}$, and $t = \mathbf{v}^{\mathsf{T}}\mathbf{u}$.
Hint 2. If $t \neq 0$, then $\mathbf{u} \notin \ker A$; use $\mathbf{u}$ followed by a basis of $\ker A$. If $t = 0$, then $\mathbf{u} \in \ker A$; pick $\mathbf{w}$ with $A\mathbf{w} = \mathbf{u}$ and use $\mathbf{w}, \mathbf{u}$, followed by more vectors of $\ker A$.
Sketch. $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$, $t = \mathbf{v}^{\mathsf{T}}\mathbf{u}$, $\ker A = \left\{\mathbf{x} \;\middle|\; \mathbf{v}^{\mathsf{T}}\mathbf{x} = 0\right\}$ has dimension $n-1$. If $t \neq 0$: in the basis $\mathbf{u}$, then a basis of $\ker A$, A has matrix tU_{11}. If $t = 0$: choose $\mathbf{w}$ with $\mathbf{v}^{\mathsf{T}}\mathbf{w} = 1$, so $A\mathbf{w} = \mathbf{u}$; extend $\mathbf{u}$ to a basis of $\ker A$; in the basis $\mathbf{w}, \mathbf{u}, \ldots$, the matrix is U_{21}.

By Proposition 4.3.16, $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$. Then $A\mathbf{x} = \mathbf{u}\left(\mathbf{v}^{\mathsf{T}}\mathbf{x}\right) = \left(\mathbf{v}^{\mathsf{T}}\mathbf{x}\right)\mathbf{u}$, so $\ker A = \left\{\mathbf{x} \;\middle|\; \mathbf{v}^{\mathsf{T}}\mathbf{x} = 0\right\}$, which has dimension $n-1$ by Theorem 4.2.11. By Proposition 2.4.28, $t = \operatorname{tr}\left(\mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \operatorname{tr}\left(\mathbf{v}^{\mathsf{T}}\mathbf{u}\right) = \mathbf{v}^{\mathsf{T}}\mathbf{u}$.

Case $t \neq 0$**.** Then $A\mathbf{u} = t\mathbf{u} \neq \mathbf{0}$, so $\mathbf{u} \notin \ker A$. Let $\mathbf{k}_2, \ldots, \mathbf{k}_n$ be a basis of $\ker A$; by Proposition 1.5.14, $\mathbf{k}_2, \ldots, \mathbf{k}_n, \mathbf{u}$ is independent, so (reordering) $\mathbf{u}, \mathbf{k}_2, \ldots, \mathbf{k}_n$ is a basis of F^n (Theorem 4.1.10(b)). T_A sends $\mathbf{u}$ to $t\mathbf{u}$ and each $\mathbf{k}_j$ to $\mathbf{0}$, so its matrix in this basis is tU_{11}.

Case $t = 0$**.** Then $\mathbf{u} \in \ker A$. Some entry v_i of $\mathbf{v}$ is nonzero; let $\mathbf{w} = v_i^{-1}\mathbf{e}_i$, so $\mathbf{v}^{\mathsf{T}}\mathbf{w} = 1$ and $A\mathbf{w} = \mathbf{u}$. Extend $\mathbf{u}$ to a basis $\mathbf{u}, \mathbf{k}_3, \ldots, \mathbf{k}_n$ of $\ker A$ (Theorem 4.1.16). Since $A\mathbf{w} \neq \mathbf{0}$, $\mathbf{w} \notin \ker A$, so as before $\mathbf{w}, \mathbf{u}, \mathbf{k}_3, \ldots, \mathbf{k}_n$ is a basis of F^n. T_A sends $\mathbf{w}$ to $\mathbf{u}$, the second basis vector, and the other basis vectors to $\mathbf{0}$. So its matrix is U_{21}.

In both cases Theorem 4.5.25 gives the similarity. Two matrices of rank 1 with the same trace t are similar to the same matrix, hence to each other; conversely, similar matrices have the same trace (Proposition 4.5.27(b)).

Solution to Exercise 4.5.52

Uses. Theorem 3.5.27 (Rank normal form), Proposition 2.5.17 (Algebra of inverses), Example 2.5.14 (Invertible maps and matrices)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint 1. The property "is a sum of two invertible matrices" survives $A \mapsto GAH$ with G, H invertible. So assume $A = N_r$.
Hint 2. Write $N_r = \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ as the sum of two diagonal matrices with nonzero diagonal entries. You need a scalar c with $c \neq 0$ and $1 - c \neq 0$.
Sketch. If $A = GN_rH$ and $N_r = X + Y$, then $A = GXH + GYH$. Choose $c \in F$ with $c \neq 0, 1$, and split $N_r = \operatorname{diag}(c, \ldots, c, 1, \ldots, 1) + \operatorname{diag}(1-c, \ldots, 1-c, -1, \ldots, -1)$. Over $\mathbb{F}_2$, the only invertible 1×1 matrix is (1), and $(1) + (1) = (0)$.

By Theorem 3.5.27, $G'AH' = N_r$ for some invertible G', H', where $r = \operatorname{rank} A$; so $A = GN_rH$ with $G = G'^{-1}$ and $H = H'^{-1}$ invertible. If $N_r = X + Y$ with X and Y invertible, then $A = GXH + GYH$, and GXH and GYH are invertible (Proposition 2.5.17(b)). So it suffices to treat N_r.

Choose $c \in F$ with $c \neq 0$ and $c \neq 1$, which is possible because F has at least three elements. Let

$$X = \operatorname{diag}(c, \ldots, c, 1, \ldots, 1), \quad Y = \operatorname{diag}(1-c, \ldots, 1-c, -1, \ldots, -1),$$

with r entries c (respectively $1-c$) followed by $n-r$ entries 1 (respectively -1). Then $X+Y=N_r$, and X and Y are invertible because their diagonal entries are nonzero (Example 2.5.14(d)).

Over $\mathbb{F}_2$, the only invertible 1×1 matrix is (1), and $(1)+(1)=(0)\neq(1)$. So (1) is not a sum of two invertible matrices.

Solution to Exercise 4.5.53

Uses. Exercise 4.5.52, Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Proposition 2.5.17 (Algebra of inverses)
Techniques. Choose a convenient basis (Technique 4.5.33), Block decomposition (Technique 2.4.35)
Hint 1. As in Exercise 4.5.52, it suffices to treat N_r. Over $\mathbb{F}_2$ a diagonal matrix with a 0 on the diagonal is not invertible, so the splitting must use non-diagonal blocks.
Hint 2. Block diagonal matrices with invertible blocks are invertible. Split N_r into diagonal blocks and decompose each block separately. Which small blocks can you decompose? Try (0), I_2, $\mathrm{diag}(1,0)$, and I_3.
Hint 3. $(0)=(1)+(1)$, $I_2=\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}0&1\\1&1\end{pmatrix}$, $\mathrm{diag}(1,0)=\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}0&1\\1&0\end{pmatrix}$. For I_3, try X with $X\mathbf{e}_1=\mathbf{e}_3$, $X\mathbf{e}_2=\mathbf{e}_1+\mathbf{e}_3$, $X\mathbf{e}_3=\mathbf{e}_2$, and $Y=I_3+X$.
Sketch. Reduce to N_r. Decompose the blocks $(0)=(1)+(1)$, I_2, $\mathrm{diag}(1,0)$, and I_3 explicitly over $\mathbb{F}_2$. Then cut N_r into diagonal blocks of these types: if $r=0$, use (0)'s; if $r=1$, use $\mathrm{diag}(1,0)$ and (0)'s (here $n\geq 2$ is needed); if $r\geq 2$, write I_r with I_2's and at most one I_3, and the rest with (0)'s. Block diagonal sums of invertible blocks finish the proof.

As in the solution of Exercise 4.5.52, $A=GN_rH$ with G, H invertible, and it suffices to write $N_r=X+Y$ with X, Y invertible. All arithmetic is in $\mathbb{F}_2$, where $1+1=0$.

Block diagonal matrices. If $D=\mathrm{diag}(D_1,\dots,D_s)$ is block diagonal and $D_i=X_i+Y_i$ with X_i, Y_i invertible, then $D=\mathrm{diag}(X_1,\dots,X_s)+\mathrm{diag}(Y_1,\dots,Y_s)$, and $\mathrm{diag}(X_1,\dots,X_s)$ is invertible with inverse $\mathrm{diag}\left(X_1^{-1},\dots,X_s^{-1}\right)$ by Theorem 2.4.34; likewise for the Y_i.

Small blocks. Over $\mathbb{F}_2$:

- $(0)=(1)+(1)$.
- $I_2=\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}0&1\\1&1\end{pmatrix}$, and both summands are invertible: $\begin{pmatrix}1&1\\1&0\end{pmatrix}$ has inverse $\begin{pmatrix}0&1\\1&1\end{pmatrix}$ (multiply to check), and vice versa.
- $\mathrm{diag}(1,0)=\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}0&1\\1&0\end{pmatrix}$, and $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ is its own inverse.
- $I_3=X_3+Y_3$ with

$$X_3=\begin{pmatrix}0&1&0\\0&0&1\\1&1&0\end{pmatrix},\quad Y_3=I_3+X_3=\begin{pmatrix}1&1&0\\0&1&1\\1&1&1\end{pmatrix}.$$

  X_3 is invertible: its columns $\mathbf{e}_3$, $\mathbf{e}_1+\mathbf{e}_3$, $\mathbf{e}_2$ span $\mathbb{F}_2^3$, since $\mathbf{e}_1=(\mathbf{e}_1+\mathbf{e}_3)+\mathbf{e}_3$. Y_3 is invertible: if $Y_3(x,y,z)=(x+y,y+z,x+y+z)=\mathbf{0}$, then $x=y=z$ (from the first two entries) and the third entry is $3x=x=0$.

Cutting N_r into blocks. If $r=0$, then N_r is a block diagonal matrix of n blocks (0). If $r=1$, then since $n\geq 2$, $N_1=\mathrm{diag}(\mathrm{diag}(1,0),(0),\dots,(0))$. If $r\geq 2$, write $r=2a$ or $r=2a+3$ with $a\geq 0$; then I_r is block diagonal with a blocks I_2 and, if r is odd, one block I_3, and $N_r=\mathrm{diag}(I_r,(0),\dots,(0))$. In every case each block has a decomposition from the list above, and the block diagonal step gives $N_r=X+Y$ with X and Y invertible.

Solution to Exercise 4.5.54

Uses. Definition 1.2.13 (Polynomials), Proposition 2.7.7 (Bijective linear maps are invertible), Proposition 4.5.5 (Isomorphisms preserve linear structure), Proposition 4.1.21 (Dimensions of subspaces), Example 2.1.9 (Maps of polynomials)
Hint 1. (a) In terms of coefficient sequences, S shifts every coefficient up by one place.
Hint 2. (b) Compare dimensions. (c) Every polynomial has an antiderivative; which polynomials have derivative 0? (d) Contrapositive of Proposition 4.5.5(c).
Sketch. (a) $S\left(a_0+a_1x+\cdots+a_dx^d\right)=a_0x+\cdots+a_dx^{d+1}$ is linear, injective, and onto $xF[x]$; $1\notin xF[x]$. (b) A subspace isomorphic to V has dimension $\dim V$, so it is V. (c) $D\left(\sum a_kx^{k+1}/(k+1)\right)=\sum a_kx^k$, and $D(1)=0$. (d) If W were finite-dimensional, so would V be.

(a) For $p = a_0 + a_1x + \cdots + a_dx^d$, $S(p) = a_0x + a_1x^2 + \cdots + a_dx^{d+1}$, whose coefficient sequence is that of p shifted by one place. So S is linear, and $S(p) = \mathbf{0}$ forces all $a_k = 0$, so S is injective. Every $q = b_1x + \cdots + b_ex^e \in xF[x]$ is $S\big(b_1 + b_2x + \cdots + b_ex^{e-1}\big)$. So S is a linear bijection onto $xF[x]$, an isomorphism (Proposition 2.7.7). The subspace $xF[x]$ is proper, since $1 \notin xF[x]$; so S is not surjective as a map $F[x] \to F[x]$.

(b) If U is a subspace of a finite-dimensional V and $U \cong V$, then $\dim U = \dim V$ (Proposition 4.5.5(c)), so $U = V$ by Proposition 4.1.21(b).

(c) $D\big(a_0 + a_1x + \cdots + a_dx^d\big) = a_1 + 2a_2x + \cdots + da_dx^{d-1}$ (Example 2.1.9). Given $q = \sum_{k=0}^{e} b_kx^k \in \mathbb{R}[x]$, the polynomial $p = \sum_{k=0}^{e} \frac{b_k}{k+1}x^{k+1}$ satisfies $D(p) = q$, so D is surjective. But $D(1) = 0$ with $1 \neq 0$, so D is not injective. Rank–nullity forbids both phenomena for operators on finite-dimensional spaces (Corollary 4.2.18).

(d) If W were finite-dimensional, then V, being isomorphic to W, would be finite-dimensional by Proposition 4.5.5(c), applied to an isomorphism $W \to V$.

Solution to Exercise 4.5.55

Uses. Theorem 4.5.25 (Similar matrices are one operator in two bases), Definition 4.5.23 (Similar matrices), Proposition 2.5.28 (Inverse of a 2×2 matrix)
Techniques. Choose a convenient basis (Technique 4.5.33)
Hint 1. (a) Look for $\mathbf{v} \in \mathbb{C}^2$ with $R\mathbf{v} = i\mathbf{v}$ and $\mathbf{w}$ with $R\mathbf{w} = -i\mathbf{w}$. Then use them as the columns of S.
Hint 2. (b) If $S^{-1}RS = \operatorname{diag}(p, q)$ with S real, then the first column $\mathbf{s}$ of S satisfies $R\mathbf{s} = p\mathbf{s}$. Write this out.
Hint 3. (c) Try $S = \operatorname{diag}(1, -1)$, a reflection.
Sketch. (a) $R(1, -i) = (i, 1) = i(1, -i)$ and $R(1, i) = -i(1, i)$, so $S = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$ works. (b) $R(x, y) = p(x, y)$ means $-y = px$ and $x = py$, so $\big(1 + p^2\big)x = 0$, giving $x = y = 0$; but columns of an invertible S are nonzero. (c) $\operatorname{diag}(1, -1)R \operatorname{diag}(1, -1) = -R$: conjugating a rotation by a reflection reverses its direction.

(a) $R(x, y) = (-y, x)$. So $R(1, -i) = (i, 1) = i(1, -i)$ and $R(1, i) = (-i, 1) = -i(1, i)$. Let $S = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$; it is invertible, since $1 \cdot i - 1 \cdot (-i) = 2i \neq 0$ (Proposition 2.5.28). In the basis of its columns, T_R has matrix $\operatorname{diag}(i, -i)$, so $S^{-1}RS = \operatorname{diag}(i, -i)$ (Theorem 4.5.25). Equivalently, $RS = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} = S \operatorname{diag}(i, -i)$.

(b) Suppose $S^{-1}RS = \operatorname{diag}(p, q)$ with $S \in \mathbb{R}^{2\times 2}$ invertible and $p, q \in \mathbb{R}$. Then $RS = S \operatorname{diag}(p, q)$, and comparing first columns, $R\mathbf{s} = p\mathbf{s}$ for the first column $\mathbf{s} = (x, y)$ of S. That is, $-y = px$ and $x = py$, so $x = py = -p^2x$ and $\big(1 + p^2\big)x = 0$. Since $1 + p^2 > 0$, $x = 0$, and then $y = -px = 0$. But $\mathbf{s} \neq \mathbf{0}$, because S is invertible. So no such S exists, and R is not similar over $\mathbb{R}$ to a real diagonal matrix.

(c) With $S = \operatorname{diag}(1, -1) = S^{-1}$,

$$S^{-1}RS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -R.$$

Geometrically, $-R$ is the rotation by $-\pi/2$, and the similarity is a change of basis from $\mathbf{e}_1, \mathbf{e}_2$ to $\mathbf{e}_1, -\mathbf{e}_2$, which reverses the orientation of the plane. A counterclockwise quarter turn, viewed in a mirror, is a clockwise one. So over $\mathbb{R}$, a rotation and its inverse are similar, although they are different operators.

Solution to Exercise 4.5.56

Uses. Theorem 4.5.17 (Change of basis), Definition 4.5.30 (Trace of an operator), Proposition 2.5.28 (Inverse of a 2×2 matrix)
Hint 1. In the basis $\mathcal{U} = (1, 2), (1, 1)$, the matrix of T is diagonal. Which diagonal matrix?
Hint 2. $[T]_{\mathcal{E}\mathcal{E}} = Q[T]_{\mathcal{U}\mathcal{U}}Q^{-1}$ with $Q = [\mathrm{id}]_{\mathcal{E}\mathcal{U}} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$.
Sketch. $[T]_{\mathcal{U}\mathcal{U}} = \operatorname{diag}(3, -1)$, $Q^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$, and $[T]_{\mathcal{E}\mathcal{E}} = \begin{pmatrix} -5 & 4 \\ -8 & 7 \end{pmatrix}$. The trace is 2 either way.

$T(1, 2) = 3(1, 2)$ and $T(1, 1) = -(1, 1)$, so in the basis $\mathcal{U} = (1, 2), (1, 1)$ (a basis over every field, since $1 \cdot 1 - 1 \cdot 2 = -1 \neq 0$), $[T]_{\mathcal{U}\mathcal{U}} = \operatorname{diag}(3, -1)$. With $Q = [\mathrm{id}]_{\mathcal{E}\mathcal{U}} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ and $Q^{-1} = [\mathrm{id}]_{\mathcal{U}\mathcal{E}} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$ (Proposition 2.5.28), Theorem 4.5.17 gives

$$[T]_{\mathcal{E}\mathcal{E}} = Q\begin{pmatrix}3 & 0\\ 0 & -1\end{pmatrix}Q^{-1} = \begin{pmatrix}3 & -1\\ 6 & -1\end{pmatrix}\begin{pmatrix}-1 & 1\\ 2 & -1\end{pmatrix} = \begin{pmatrix}-5 & 4\\ -8 & 7\end{pmatrix}.$$

Check: $\begin{pmatrix}-5 & 4\\ -8 & 7\end{pmatrix}\begin{pmatrix}1\\ 2\end{pmatrix} = \begin{pmatrix}3\\ 6\end{pmatrix}$ and $\begin{pmatrix}-5 & 4\\ -8 & 7\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} = \begin{pmatrix}-1\\ -1\end{pmatrix}$. The trace is $-5+7=2$ from the standard matrix and $3+(-1)=2$ from the diagonal one, as Definition 4.5.30 requires.

Section 4.6

Solution to Exercise 4.6.6

Uses. Theorem 1.4.3 (Subspace criterion), Definition 4.6.1 (Product space)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), check the subspace criterion one component at a time.
Hint 2. For (b), try the smallest interesting case, $V = W = F$. Is there a line in $F \times F = F^2$ that is not a product?
Sketch. (a) Componentwise. (b) False: the diagonal $\{(t,t) \mid t \in F\}$ is a subspace of $F \times F$. A product $V' \times W'$ containing $(1,1)$ has $1 \in V'$ and $0 \in W'$, so it contains $(1,0)$, which is not on the diagonal.

(a) $(\mathbf{0},\mathbf{0}) \in V' \times W'$. If $(\mathbf{v},\mathbf{w}), (\mathbf{v}',\mathbf{w}') \in V' \times W'$ and $a \in F$, then $a(\mathbf{v},\mathbf{w}) + (\mathbf{v}',\mathbf{w}') = (a\mathbf{v}+\mathbf{v}', a\mathbf{w}+\mathbf{w}')$, whose components lie in V' and W' because these are subspaces. By Theorem 1.4.3, $V' \times W'$ is a subspace.

(b) The claim is false. Take $V = W = F$ and the diagonal $D = \{(t,t) \mid t \in F\}$, which is a subspace of $F \times F$ (it is $\operatorname{span}((1,1))$). Suppose $D = V' \times W'$. Since $(1,1) \in D$, we get $1 \in V'$ and $0 \in W'$ (as W' is a subspace), so $(1,0) \in V' \times W' = D$. But $(1,0)$ is not of the form (t,t).

Solution to Exercise 4.6.17

Uses. Lemma 4.6.14 (Equal cosets), Definition 1.3.10 (Span)
Hint 1. A general element of U is $a(1,0,1) + b(0,1,1)$. Write it out.
Hint 2. By Lemma 4.6.14, $\mathbf{v} + U = \mathbf{w} + U$ exactly when $\mathbf{v} - \mathbf{w} \in U$, that is, when $\mathbf{v}$ and $\mathbf{w}$ have the same value of $z - x - y$.
Sketch. (a) $a(1,0,1) + b(0,1,1) = (a, b, a+b)$. (b) The values of $z-x-y$ are $0, 1, 1, -1$, so only $(0,0,1) + U = (2,-1,2) + U$; also $(1,2,3) + U = U$.

(a) $a(1,0,1) + b(0,1,1) = (a, b, a+b)$, and every (x,y,z) with $z = x + y$ has this form with $a = x$, $b = y$. So $U = \{(x,y,z) \mid z - x - y = 0\}$. By Lemma 4.6.14, $(x,y,z) \in \mathbf{v} + U$ if and only if $(x,y,z) - \mathbf{v} \in U$, that is, $(z - v_3) - (x - v_1) - (y - v_2) = 0$, which is $z - x - y = v_3 - v_1 - v_2$.

(b) Let $f(x,y,z) = z - x - y$. By (a) and Lemma 4.6.14, two cosets $\mathbf{v} + U$ and $\mathbf{w} + U$ are equal if and only if $f(\mathbf{v}) = f(\mathbf{w})$. The values are

$$f(1,2,3) = 0, \quad f(0,0,1) = 1, \quad f(2,-1,2) = 1, \quad f(1,1,1) = -1.$$

So $(0,0,1) + U = (2,-1,2) + U$, and the other two cosets are different from each other and from it. (The first one is U itself, since $(1,2,3) \in U$.)

Solution to Exercise 4.6.24

Uses. Lemma 4.6.14 (Equal cosets), Definition 4.6.20 (Quotient space)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18)
Hint 1. By Lemma 4.6.14, two representatives of one coset differ by $t(1,1)$. For each formula, compare its values at $\mathbf{v}$ and at $\mathbf{v} + t(1,1)$.
Hint 2. To show that a formula is not well defined, one pair of representatives of the same coset suffices, for example $(0,0)$ and $(1,1)$.
Sketch. (a) No: $(0,0)$ and $(1,1)$ give 0 and 2. (b) Yes, linear. (c) Yes (a function of (b)), not linear. (d) Yes, linear; it equals $-\mathrm{id}$, since $(v_2, v_1) + (v_1, v_2) \in U$. (e) No: $(0,0) + U$ and $(1,1) + U$ go to U and $(1,0) + U \neq U$.

Two representatives of the same coset differ by $t(1,1)$ for some $t \in \mathbb{R}$ (Lemma 4.6.14). Write $\mathbf{v}' = \mathbf{v} + t(1,1)$.

(a) Not well defined: $(0,0) + U = (1,1) + U$, but the formula gives 0 and 2.

(b) Well defined, since $v_1' - v_2' = (v_1 + t) - (v_2 + t) = v_1 - v_2$. It is linear: $f(a(\mathbf{v} + U) + (\mathbf{w} + U)) = f((a\mathbf{v} + \mathbf{w}) + U) = a(v_1 - v_2) + (w_1 - w_2)$.

(c) Well defined, since it is the square of the function in (b). Not linear: $f((1,0)+U)=1$ but $f(2((1,0)+U))=f((2,0)+U)=4\neq 2$.

(d) Well defined: swapping the coordinates of $\mathbf{v}'$ gives $(v_2,v_1)+t(1,1)$, which lies in the same coset as (v_2,v_1). It is linear, by the same computation as in (b). In fact $(v_2,v_1)+(v_1,v_2)=(v_1+v_2)(1,1)\in U$, so $(v_2,v_1)+U=-(\mathbf{v}+U)$, and $S=-\mathrm{id}$.

(e) Not well defined: $(0,0)+U=(1,1)+U$, but the formula gives $(0,0)+U=U$ and $(1,0)+U$, which differ because $(1,0)\notin U$.

Solution to Exercise 4.6.37

Uses. Theorem 4.6.34 (First isomorphism theorem), Lemma 4.6.14 (Equal cosets), Proposition 1.2.18 (Degree and evaluation versus the operations)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. U is the kernel of $E:F[x]\to F^2$, $E(p)=(p(1),p(2))$. Is E surjective? Try polynomials of degree at most 1.
Hint 2. An isomorphism carries bases to bases, so compute $E(1)$ and $E(x)$. For x^2+U, solve $E(x^2)=aE(1)+bE(x)$.
Sketch. E is linear with kernel U, and surjective: $a+(b-a)(x-1)\mapsto(a,b)$. The first isomorphism theorem gives $F[x]/U\cong F^2$, and $p+U=q+U$ iff $p-q\in U$ iff $E(p)=E(q)$. $\tilde{E}$ sends $1+U, x+U$ to $(1,1),(1,2)$, a basis. From $(1,4)=-2(1,1)+3(1,2)$, $x^2+U=-2(1+U)+3(x+U)$. Over $\mathbb{F}_2$ this reads $x^2+U=x+U$.

(a) Let $E:F[x]\to F^2$, $E(p)=(p(1),p(2))$. It is linear by Proposition 1.2.18, and $\ker E=U$. It is surjective: for $(a,b)\in F^2$, the polynomial $p=a+(b-a)(x-1)$ has $p(1)=a$ and $p(2)=a+(b-a)=b$. By Theorem 4.6.34, $\tilde{E}:F[x]/U\to F^2$, $p+U\mapsto(p(1),p(2))$, is an isomorphism. Moreover $p+U=q+U$ if and only if $p-q\in U$ (Lemma 4.6.14), that is, $p(1)-q(1)=0$ and $p(2)-q(2)=0$.

(b) $\tilde{E}(1+U)=(1,1)$ and $\tilde{E}(x+U)=(1,2)$. These are independent, since $(1,2)-(1,1)=(0,1)$, so together with $(1,1)$ they span F^2; hence they form a basis of F^2, and their preimages $1+U, x+U$ form a basis of $F[x]/U$ ($\tilde{E}^{-1}$ is an isomorphism). Next, $\tilde{E}(x^2+U)=(1,4)$, and $a(1,1)+b(1,2)=(1,4)$ gives $a+b=1$, $a+2b=4$, so $b=3$ and $a=-2$. Hence

$$x^2+U=-2(1+U)+3(x+U)=(3x-2)+U.$$

Check: $x^2-(3x-2)=x^2-3x+2$ vanishes at 1 and at 2. Over $\mathbb{F}_2$ we have $3=1$ and $2=0$, so the answer reads $x^2+U=x+U$: there $2=0$, so U is the set of polynomials vanishing at 0 and 1, and x^2-x is one of them.

Solution to Exercise 4.6.38

Uses. Theorem 4.6.5 (Dimension of a product), Proposition 4.6.7 (Maps into and out of a product), Theorem 4.2.11 (Rank–nullity), Theorem 4.1.10 (Independent lists are short)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. For (a), use Theorem 4.6.5 with the basis $1, x, x^2$ of each factor.
Hint 2. For (b), D is the kernel of $(p,q)\mapsto p(0)-q(1)$, a linear map $V\times V\to F$. Find its rank, then guess five independent elements of D.
Sketch. (a) $(1,\mathbf{0}),(x,\mathbf{0}),(x^2,\mathbf{0}),(\mathbf{0},1),(\mathbf{0},x),(\mathbf{0},x^2)$; dimension 6. (b) $D=\ker\varphi$ for the surjective $\varphi(p,q)=p(0)-q(1)$, so $\dim D=5$. A basis: $(x,\mathbf{0}),(x^2,\mathbf{0}),(\mathbf{0},x-1),(\mathbf{0},x^2-1),(1,1)$.

(a) By Theorem 4.6.5, applied to the basis $1, x, x^2$ of V in each factor,

$$(1,\mathbf{0}),\quad (x,\mathbf{0}),\quad (x^2,\mathbf{0}),\quad (\mathbf{0},1),\quad (\mathbf{0},x),\quad (\mathbf{0},x^2)$$

is a basis of $V\times V$, which therefore has dimension 6.

(b) Let $\varphi:V\times V\to F, \varphi(p,q)=p(0)-q(1)$. By Proposition 4.6.7(b), with $S_1=\mathrm{ev}_0$ and $S_2=-\mathrm{ev}_1$, it is linear, and $D=\ker\varphi$ is a subspace. Since $\varphi(1,\mathbf{0})=1$, the image is a nonzero subspace of F, hence all of F. By rank–nullity, $\dim D=6-1=5$. The five pairs

$$(x,\mathbf{0}),\quad (x^2,\mathbf{0}),\quad (\mathbf{0},x-1),\quad (\mathbf{0},x^2-1),\quad (1,1)$$

lie in D (for each, $p(0) = q(1)$). They are independent: if $a(x, \mathbf{0}) + b\big(x^2, \mathbf{0}\big) + c(\mathbf{0}, x - 1) + d\big(\mathbf{0}, x^2 - 1\big) + e(1, 1) = \mathbf{0}$, the first component $e + ax + bx^2$ is the zero polynomial, so $a = b = e = 0$, and then the second component $c(x - 1) + d\big(x^2 - 1\big) = -(c + d) + cx + dx^2$ vanishes, so $c = d = 0$. Five independent vectors in the five-dimensional space D form a basis (Theorem 4.1.10(b)).

Solution to Exercise 4.6.39

Uses. Proposition 4.6.7 (Maps into and out of a product), Proposition 2.6.8 (Kernels and images are subspaces), Theorem 4.4.7 (Criteria for direct sums), Proposition 4.6.29 (Complements and the quotient), Proposition 2.7.7 (Bijective linear maps are invertible)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), the map $\mathbf{v} \mapsto (\mathbf{v}, T(\mathbf{v}))$ is linear by Proposition 4.6.7(a). Its first component makes it injective.
Hint 2. For (b), split $(\mathbf{v}, \mathbf{w})$ as $(\mathbf{v}, T(\mathbf{v})) + (\mathbf{0}, \mathbf{w} - T(\mathbf{v}))$. Then use Proposition 4.6.29.
Hint 3. For (c), the two conditions say that each $\mathbf{v}$ is the first component of exactly one element $(\mathbf{v}, \mathbf{w})$ of G. Define $S(\mathbf{v}) = \mathbf{w}$ and prove it is linear by uniqueness.
Sketch. (a) $\varphi(\mathbf{v}) = (\mathbf{v}, T(\mathbf{v}))$ is linear and injective with image G_T. (b) $(\mathbf{v}, \mathbf{w}) = (\mathbf{v}, T\mathbf{v}) + (\mathbf{0}, \mathbf{w} - T\mathbf{v})$, and $(\mathbf{v}, T\mathbf{v}) = (\mathbf{0}, \mathbf{w})$ forces $\mathbf{v} = \mathbf{0}$, then $\mathbf{w} = \mathbf{0}$. Then $(V \times W)/G_T \cong \{\mathbf{0}\} \times W \cong W$. (c) The conditions give existence and uniqueness of $\mathbf{w}$ with $(\mathbf{v}, \mathbf{w}) \in G$; call it $S(\mathbf{v})$. Since G is a subspace, $(a\mathbf{v} + \mathbf{v}', aS\mathbf{v} + S\mathbf{v}') \in G$, so S is linear, and $G = G_S$.

(a) Let $\varphi : V \to V \times W$, $\varphi(\mathbf{v}) = (\mathbf{v}, T(\mathbf{v}))$. It is linear by Proposition 4.6.7(a) (with the maps id_V and T), its image is G_T, which is therefore a subspace (Proposition 2.6.8), and it is injective because the first component of $\varphi(\mathbf{v})$ is $\mathbf{v}$. So $\varphi : V \to G_T$ is a bijective linear map, hence an isomorphism.

(b) Every $(\mathbf{v}, \mathbf{w}) \in V \times W$ equals $(\mathbf{v}, T(\mathbf{v})) + (\mathbf{0}, \mathbf{w} - T(\mathbf{v}))$, a vector of G_T plus a vector of $\{\mathbf{0}\} \times W$. If $(\mathbf{v}, T(\mathbf{v})) = (\mathbf{0}, \mathbf{w})$ lies in both subspaces, then $\mathbf{v} = \mathbf{0}$, so $\mathbf{w} = T(\mathbf{0}) = \mathbf{0}$. So the sum is direct (Theorem 4.4.7) and $V \times W = G_T \oplus (\{\mathbf{0}\} \times W)$. By Proposition 4.6.29, $(V \times W)/G_T \cong \{\mathbf{0}\} \times W$, and $(\mathbf{0}, \mathbf{w}) \mapsto \mathbf{w}$ is an isomorphism $\{\mathbf{0}\} \times W \to W$.

(c) If $G = G_S$, then $(\mathbf{0}, \mathbf{w}) \in G_S$ forces $\mathbf{w} = S(\mathbf{0}) = \mathbf{0}$, and $\mathbf{v}$ is the first component of $(\mathbf{v}, S(\mathbf{v})) \in G_S$.

Conversely, assume the two conditions. For each $\mathbf{v} \in V$ there is some $\mathbf{w}$ with $(\mathbf{v}, \mathbf{w}) \in G$, and only one: if $(\mathbf{v}, \mathbf{w}), (\mathbf{v}, \mathbf{w}') \in G$, then their difference $(\mathbf{0}, \mathbf{w} - \mathbf{w}')$ lies in $G \cap (\{\mathbf{0}\} \times W) = \{\mathbf{0}\}$, so $\mathbf{w} = \mathbf{w}'$. Define $S(\mathbf{v})$ to be this $\mathbf{w}$. For $\mathbf{v}, \mathbf{v}' \in V$ and $a \in F$, the vector $a(\mathbf{v}, S(\mathbf{v})) + (\mathbf{v}', S(\mathbf{v}')) = (a\mathbf{v} + \mathbf{v}', aS(\mathbf{v}) + S(\mathbf{v}'))$ lies in G, so by uniqueness $S(a\mathbf{v} + \mathbf{v}') = aS(\mathbf{v}) + S(\mathbf{v}')$. Thus S is linear. By construction $G_S \subseteq G$, and every $(\mathbf{v}, \mathbf{w}) \in G$ has $\mathbf{w} = S(\mathbf{v})$, so $G = G_S$.

Solution to Exercise 4.6.40

Uses. Definition 4.6.12 (Coset), Definition 4.6.20 (Quotient space)
Techniques. Double containment (Technique 1.3.14)
Hint 1. Prove each equality of sets by double containment. For $\supseteq$ in (a), write $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = (\mathbf{v} + \mathbf{u}) + (\mathbf{w} + \mathbf{0})$.
Hint 2. For (b), $a\mathbf{v} + \mathbf{u} = a\big(\mathbf{v} + a^{-1}\mathbf{u}\big)$. For (c), how many elements does a coset of U have?
Sketch. (a) $(\mathbf{v} + \mathbf{u}) + (\mathbf{w} + \mathbf{u}') = (\mathbf{v} + \mathbf{w}) + (\mathbf{u} + \mathbf{u}')$, and conversely $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = (\mathbf{v} + \mathbf{u}) + (\mathbf{w} + \mathbf{0})$. (b) Similar, with a^{-1}. (c) $0(\mathbf{v} + U) = \{\mathbf{0}\}$, which is not a coset when $U \neq \{\mathbf{0}\}$, because every coset is in bijection with U.

(a) An element of the left side is $(\mathbf{v} + \mathbf{u}) + (\mathbf{w} + \mathbf{u}') = (\mathbf{v} + \mathbf{w}) + (\mathbf{u} + \mathbf{u}')$ with $\mathbf{u}, \mathbf{u}' \in U$, and $\mathbf{u} + \mathbf{u}' \in U$. Conversely, $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = (\mathbf{v} + \mathbf{u}) + (\mathbf{w} + \mathbf{0})$ is in the left side.

(b) An element of the left side is $a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$ with $a\mathbf{u} \in U$. Conversely, $a\mathbf{v} + \mathbf{u} = a\big(\mathbf{v} + a^{-1}\mathbf{u}\big)$ with $a^{-1}\mathbf{u} \in U$.

(c) $0(\mathbf{v} + U) = \{0\mathbf{x} \mid \mathbf{x} \in \mathbf{v} + U\} = \{\mathbf{0}\}$. A coset $\mathbf{x} + U$ is in bijection with U via $\mathbf{u} \mapsto \mathbf{x} + \mathbf{u}$, so when $U \neq \{\mathbf{0}\}$ it has at least two elements, and $\{\mathbf{0}\}$ is not a coset. So scalar multiplication by 0 leads out of the set of cosets, and the set operations do not even define a scalar multiplication on it. By (a) and (b), the set operations agree with those of Definition 4.6.20 in all other cases. The quotient definition sets $0(\mathbf{v} + U) = 0\mathbf{v} + U = U$ instead.

In particular, the sum in V/U is the sum of the cosets as sets, which gives a second reason why addition is well defined.

Solution to Exercise 4.6.41

Uses. Proposition 4.6.26 (The quotient map), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Lemma 4.6.14 (Equal cosets), Theorem 4.4.7 (Criteria for direct sums), Definition 1.6.1 (Basis)
Hint 1. A relation $\sum_j c_j(\mathbf{v}_j + U) = \mathbf{0} + U$ says exactly that $\sum_j c_j \mathbf{v}_j \in U$.
Hint 2. For (b), the span of the cosets is $\pi(X)$. When is $\pi(X)$ all of V/U?
Sketch. (a) $\sum_j c_j(\mathbf{v}_j + U) = U$ iff $\sum_j c_j \mathbf{v}_j \in X \cap U$. If $X \cap U = \{\mathbf{0}\}$ and the $\mathbf{v}_j$ are independent, then all $c_j = 0$; conversely, a relation among the $\mathbf{v}_j$, or a nonzero vector of $X \cap U$, gives a nontrivial relation among the cosets. (b) The span of the cosets is $\pi(X)$, and $\pi(X) = V/U$ iff every $\mathbf{v}$ is in $X + U$. (c) Combine: $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of X and $V = U \oplus X$.

Throughout, $\pi : V \to V/U$ is the quotient map. For scalars c_j, $\sum_j c_j(\mathbf{v}_j + U) = \left(\sum_j c_j \mathbf{v}_j\right) + U$, which is the zero coset if and only if $\sum_j c_j \mathbf{v}_j \in U$ (Corollary 4.6.15(c)).

(a) Suppose the cosets are independent. If $\sum_j c_j \mathbf{v}_j = \mathbf{0}$, then $\sum_j c_j(\mathbf{v}_j + U) = U$, so all $c_j = 0$; thus $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is independent. If $\mathbf{x} = \sum_j c_j \mathbf{v}_j \in X \cap U$, then $\sum_j c_j(\mathbf{v}_j + U) = U$, so all $c_j = 0$ and $\mathbf{x} = \mathbf{0}$. Conversely, assume both conditions and let $\sum_j c_j(\mathbf{v}_j + U) = U$. Then $\sum_j c_j \mathbf{v}_j \in X \cap U = \{\mathbf{0}\}$, so $\sum_j c_j \mathbf{v}_j = \mathbf{0}$ and all $c_j = 0$ by independence.

(b) The span of $\pi(\mathbf{v}_1), \ldots, \pi(\mathbf{v}_k)$ is $\operatorname{im}(\pi\restriction_X) = \{\mathbf{x} + U \mid \mathbf{x} \in X\}$ (Proposition 2.1.16(a)). This is all of V/U if and only if for every $\mathbf{v} \in V$ there is $\mathbf{x} \in X$ with $\mathbf{v} - \mathbf{x} \in U$ (Lemma 4.6.14), that is, $\mathbf{v} \in X + U$. So the cosets span V/U if and only if $X + U = V$.

(c) By (a) and (b), the cosets form a basis of V/U if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is independent (hence a basis of X), $X \cap U = \{\mathbf{0}\}$, and $X + U = V$. The last two conditions say that X is a complement of U (Theorem 4.4.7). This matches Proposition 4.6.29: $\pi\restriction_X$ is an isomorphism exactly when X is a complement, and isomorphisms carry bases to bases.

Solution to Exercise 4.6.42

Uses. Theorem 4.6.34 (First isomorphism theorem), Proposition 4.6.29 (Complements and the quotient), Theorem 4.4.19 (Dimension of a sum), Lemma 2.4.24 (Matrix units)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. The function $\psi(A) = a_{12} - a_{21}$ is linear on V. What is its kernel?
Hint 2. For (b), compute $S \cap K$ and the dimensions of S and K. For (c), what is $-A$ over $\mathbb{F}_2$?
Sketch. (a) $\psi(A) = a_{12} - a_{21}$ is linear and surjective with kernel S; apply Theorem 4.6.34. (b) $A \in S \cap K$ gives $A = -A$, so $(1+1)A = 0$ and $A = 0$; $\dim S = 3$, $\dim K = 1$. (c) Over $\mathbb{F}_2$, $-A = A$, so $K = S$. $\operatorname{span}(U_{12})$ is a complement, since $\psi(U_{12}) = 1$.

(a) Let $\psi : V \to F$, $\psi(A) = a_{12} - a_{21}$. It is linear, and surjective since $\psi(U_{12}) = 1$ (Lemma 2.4.24). A 2×2 matrix is symmetric exactly when $a_{12} = a_{21}$, so $\ker \psi = S$. By Theorem 4.6.34, $\tilde{\psi} : V/S \to F$, $A + S \mapsto a_{12} - a_{21}$, is well defined and an isomorphism.

(b) If $A \in S \cap K$, then $A = A^{\mathsf{T}} = -A$, so $(1+1)A = 0$, and $A = 0$ after multiplying by $(1+1)^{-1}$. The space S has basis $U_{11}, U_{22}, U_{12} + U_{21}$, so $\dim S = 3$. For $A \in K$, the diagonal entries satisfy $a_{ii} = -a_{ii}$, so they vanish, and $a_{21} = -a_{12}$; hence $K = \operatorname{span}(U_{12} - U_{21})$ has dimension 1. By Theorem 4.4.19, $\dim(S + K) = 3 + 1 - 0 = 4$, so $S + K = V$ and $V = S \oplus K$.

(c) Over $\mathbb{F}_2$, $-A = A$ for every A, so $K = S$ and $S \cap K = S \neq \{0\}$. Let $L = \operatorname{span}(U_{12})$. A nonzero multiple cU_{12} is not symmetric, so $L \cap S = \{0\}$, and $\dim S + \dim L = 4$, so $V = S \oplus L$. Equivalently, by Proposition 4.6.29: $\psi(U_{12}) = 1 \neq 0$, so $\pi\restriction_L$ is an isomorphism onto the one-dimensional space V/S.

The quotient $V/S \cong F$ is described by the same formula over every field; only the familiar complement of skew-symmetric matrices breaks down in characteristic 2.

Solution to Exercise 4.6.43

Uses. Theorem 4.6.30 (Universal property of the quotient), Theorem 4.6.34 (First isomorphism theorem)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18)
Hint 1. Compare the student's implication with the one in Technique 4.6.18. Which way does each go?
Hint 2. Take $T = \mathrm{id}$, so that $\ker T = \{\mathbf{0}\}$ is inside every U.

Sketch. The student proved "equal outputs imply equal inputs", which is injectivity. Well-definedness needs "equal inputs imply equal outputs". With $T = \mathrm{id}_{\mathbb{R}^2}$ and $U = \mathrm{span}(\mathbf{e}_1)$: $\mathbf{e}_1 + U = \mathbf{0} + U$ but $T(\mathbf{e}_1) \neq T(\mathbf{0})$. The right hypothesis is $U \subseteq \ker T$.

The argument proves the wrong implication. It shows that if two cosets have the same value, then they are equal, which is injectivity of S, assuming S exists. Well-definedness is the converse: if $\mathbf{v} + U = \mathbf{v}' + U$, then $T(\mathbf{v}) = T(\mathbf{v}')$. That is, $\mathbf{v} - \mathbf{v}' \in U$ must imply $\mathbf{v} - \mathbf{v}' \in \ker T$, which needs $U \subseteq \ker T$, the opposite of the given containment.

Example: $T = \mathrm{id}_{\mathbb{R}^2}$, so $\ker T = \{\mathbf{0}\} \subseteq U$ for $U = \mathrm{span}(\mathbf{e}_1)$. Then $\mathbf{e}_1 + U = \mathbf{0} + U$, but the formula assigns $T(\mathbf{e}_1) = \mathbf{e}_1$ to the first name and $T(\mathbf{0}) = \mathbf{0}$ to the second.

By Theorem 4.6.30, the formula defines a (linear) map exactly when $U \subseteq \ker T$. Together with the student's hypothesis $\ker T \subseteq U$, that means $U = \ker T$, and then the student's argument is a correct proof that $S = \tilde{T}$ is injective, which is half of Theorem 4.6.34.

Solution to Exercise 4.6.44

Uses. Theorem 4.6.34 (First isomorphism theorem), Theorem 4.6.30 (Universal property of the quotient), Proposition 4.6.26 (The quotient map), Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 1.6.16 (The coordinate map is a bijection)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. A nonzero linear map $f : V \to F$ is surjective, since its image is a nonzero subspace of F. Apply Theorem 4.6.34.
Hint 2. Conversely, if $\dim(V/U) = 1$, compose π with a coordinate map $V/U \to F$.
Hint 3. For (b), both f and g induce isomorphisms $V/U \to F$ with $U = \ker f = \ker g$. What are the linear maps $F \to F$?
Sketch. (a) If $f \neq 0$, then $\mathrm{im}\, f = F$ and $V/\ker f \cong F$. If $\mathbf{v}_0 + U$ is a basis of V/U, let φ be its coordinate map and $f = \varphi \circ \pi$; then $\ker f = \ker \pi = U$ and $f(\mathbf{v}_0) = 1$. (b) $\tilde{g} \circ \tilde{f}^{-1} : F \to F$ is multiplication by some $c \neq 0$, so $\tilde{g} = c\tilde{f}$ and $g = cf$.

(a) Suppose $U = \ker f$ with $f \neq 0$. Pick $\mathbf{v}$ with $f(\mathbf{v}) = a \neq 0$; then $f((t/a)\mathbf{v}) = t$ for every $t \in F$, so f is surjective. By Theorem 4.6.34, $V/U \cong F$, which has dimension 1.

Conversely, suppose $\dim(V/U) = 1$, and let $\mathbf{v}_0 + U$ be a basis of V/U. Its coordinate map $\varphi : V/U \to F$, $c(\mathbf{v}_0 + U) \mapsto c$, is a linear bijection (Proposition 1.6.19, Proposition 1.6.16). Let $f = \varphi \circ \pi : V \to F$. It is linear, and since φ is injective, $\ker f = \ker \pi = U$ (Proposition 4.6.26). It is nonzero because $f(\mathbf{v}_0) = \varphi(\mathbf{v}_0 + U) = 1$.

(b) If $g = cf$ with $c \neq 0$, then $g(\mathbf{v}) = 0$ if and only if $f(\mathbf{v}) = 0$. Conversely, let $U = \ker f = \ker g$. As in (a), f and g are surjective, so Theorem 4.6.34 gives isomorphisms $\tilde{f}, \tilde{g} : V/U \to F$ with $f = \tilde{f} \circ \pi$ and $g = \tilde{g} \circ \pi$. The map $h = \tilde{g} \circ \tilde{f}^{-1} : F \to F$ is linear, so $h(t) = th(1)$ for all t; let $c = h(1)$, which is nonzero because h is bijective. Then $\tilde{g} = h \circ \tilde{f} = c\tilde{f}$, and $g = \tilde{g} \circ \pi = c\tilde{f} \circ \pi = cf$.

Solution to Exercise 4.6.45

Uses. Theorem 4.6.30 (Universal property of the quotient), Theorem 4.6.34 (First isomorphism theorem), Proposition 4.1.19 (Extending linear maps and separating vectors), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. One direction is one line. For the other, T_B kills $\ker T_A$, so it factors through $F^n/\ker T_A$, and that quotient is isomorphic to $\mathrm{im}\, T_A$.
Hint 2. You get a linear map $R_0 : \mathrm{im}\, T_A \to F^p$ with $R_0(A\mathbf{x}) = B\mathbf{x}$. It is only defined on $\mathrm{im}\, T_A$. Extend it to F^m with Proposition 4.1.19(a).
Sketch. If $B = CA$ and $A\mathbf{x} = \mathbf{0}$, then $B\mathbf{x} = \mathbf{0}$. Conversely, Theorem 4.6.30 gives S with $T_B = S \circ \pi$, and $R_0 = S \circ \widetilde{T_A}^{-1} : \mathrm{im}\, T_A \to F^p$ satisfies $R_0(A\mathbf{x}) = B\mathbf{x}$. Extend R_0 to $R \in \mathrm{L}(F^m, F^p)$ and let C be its matrix.

If $B = CA$ and $A\mathbf{x} = \mathbf{0}$, then $B\mathbf{x} = C(A\mathbf{x}) = \mathbf{0}$, so $\ker T_A \subseteq \ker T_B$.

Conversely, suppose $\ker T_A \subseteq \ker T_B$, and let $\pi : F^n \to F^n/\ker T_A$ be the quotient map. By Theorem 4.6.30(a), there is a linear $S : F^n/\ker T_A \to F^p$ with $S(\mathbf{x} + \ker T_A) = B\mathbf{x}$. By Theorem 4.6.34, $\widetilde{T_A} : F^n/\ker T_A \to \mathrm{im}\, T_A$, $\mathbf{x} + \ker T_A \mapsto A\mathbf{x}$, is an isomorphism. Let $R_0 = S \circ \widetilde{T_A}^{-1} : \mathrm{im}\, T_A \to F^p$; it is linear, and

$$R_0(A\mathbf{x}) = S(\mathbf{x} + \ker T_A) = B\mathbf{x} \quad \text{for all } \mathbf{x} \in F^n.$$

By Proposition 4.1.19(a), R_0 extends to a linear map $R : F^m \to F^p$. By Proposition 2.3.17, $R = T_C$ for some $C \in F^{p\times m}$. Then $CA\mathbf{x} = R(A\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x}$; taking $\mathbf{x} = \mathbf{e}_j$ shows that CA and B have the same columns, so $CA = B$.

The extension step is where choices enter: for $p \geq 1$, C is unique exactly when $\operatorname{im} T_A = F^m$, that is, $\operatorname{rank} A = m$.

Solution to Exercise 4.6.46

Uses. Proposition 4.6.7 (Maps into and out of a product), Proposition 4.6.26 (The quotient map), Theorem 4.6.34 (First isomorphism theorem)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. For (b) "⇐": given $(\mathbf{w}_1, \mathbf{w}_2)$, choose preimages $\mathbf{x}_1$ under T_1 and $\mathbf{x}_2$ under T_2. You need one vector that works for both. Split $\mathbf{x}_1 - \mathbf{x}_2$ along $\ker T_1 + \ker T_2$.
Hint 2. For (b) "⇒": to write $\mathbf{v} \in V$ as a sum, find $\mathbf{x}$ with $T(\mathbf{x}) = (T_1(\mathbf{v}), \mathbf{0})$.
Hint 3. For (c), apply (a) and (b) to the two quotient maps.
Sketch. (a) Both components vanish. (b) If $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{k}_1 + \mathbf{k}_2$ with $\mathbf{k}_i \in \ker T_i$, then $\mathbf{v} = \mathbf{x}_1 - \mathbf{k}_1 = \mathbf{x}_2 + \mathbf{k}_2$ maps to $(\mathbf{w}_1, \mathbf{w}_2)$. Conversely, if $T(\mathbf{x}) = (T_1\mathbf{v}, \mathbf{0})$, then $\mathbf{v} = (\mathbf{v} - \mathbf{x}) + \mathbf{x} \in \ker T_1 + \ker T_2$. (c) The quotient maps are surjective with kernels U and W; then use Theorem 4.6.34.

T is linear by Proposition 4.6.7(a).

(a) $T(\mathbf{v}) = (\mathbf{0}, \mathbf{0})$ if and only if $T_1(\mathbf{v}) = \mathbf{0}$ and $T_2(\mathbf{v}) = \mathbf{0}$.

(b) Suppose $\ker T_1 + \ker T_2 = V$, and let $(\mathbf{w}_1, \mathbf{w}_2) \in W_1 \times W_2$. Since T_1, T_2 are surjective, there are $\mathbf{x}_1, \mathbf{x}_2 \in V$ with $T_1(\mathbf{x}_1) = \mathbf{w}_1$ and $T_2(\mathbf{x}_2) = \mathbf{w}_2$. Write $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{k}_1 + \mathbf{k}_2$ with $\mathbf{k}_i \in \ker T_i$, and let

$$\mathbf{v} = \mathbf{x}_1 - \mathbf{k}_1 = \mathbf{x}_2 + \mathbf{k}_2.$$

Then $T_1(\mathbf{v}) = T_1(\mathbf{x}_1) = \mathbf{w}_1$ and $T_2(\mathbf{v}) = T_2(\mathbf{x}_2) = \mathbf{w}_2$, so $T(\mathbf{v}) = (\mathbf{w}_1, \mathbf{w}_2)$.

Conversely, suppose T is surjective, and let $\mathbf{v} \in V$. Choose $\mathbf{x}$ with $T(\mathbf{x}) = (T_1(\mathbf{v}), \mathbf{0})$. Then $T_2(\mathbf{x}) = \mathbf{0}$ and $T_1(\mathbf{v} - \mathbf{x}) = \mathbf{0}$, so $\mathbf{v} = (\mathbf{v} - \mathbf{x}) + \mathbf{x} \in \ker T_1 + \ker T_2$.

(c) The quotient maps $\pi_U : V \to V/U$ and $\pi_W : V \to V/W$ are surjective with kernels U and W (Proposition 4.6.26). The map in question is (π_U, π_W), so by (a) its kernel is $U \cap W$, and by (b) it is surjective if and only if $U + W = V$. In that case Theorem 4.6.34 gives $V/(U \cap W) \cong V/U \times V/W$.

Part (c) is a linear-algebra version of the Chinese remainder theorem: when $U + W = V$, a pair of conditions "$\mathbf{v} \equiv \mathbf{a} (\mathrm{mod}\, U)$" and "$\mathbf{v} \equiv \mathbf{b} (\mathrm{mod}\, W)$" always has a solution, unique modulo $U \cap W$.

Solution to Exercise 4.6.47

Uses. Proposition 4.6.7 (Maps into and out of a product), Theorem 4.6.34 (First isomorphism theorem), Theorem 4.6.5 (Dimension of a product), Proposition 4.1.21 (Dimensions of subspaces), Proposition 4.2.6 (Injective maps preserve dimension), Theorem 4.6.27 (Dimension of a quotient)
Techniques. Pass to a quotient (Technique 4.6.31), Dimension counting (Technique 4.2.24)
Hint 1. Find a linear map out of V whose kernel is $U_1 \cap \cdots \cap U_m$ and whose target is finite-dimensional.
Hint 2. Use $T : V \to V/U_1 \times \cdots \times V/U_m$, $\mathbf{v} \mapsto (\mathbf{v} + U_1, \ldots, \mathbf{v} + U_m)$, and Theorem 4.6.34. The image need not be the whole product.
Sketch. T has kernel $\cap_i U_i$, so $V/\cap_i U_i \cong \operatorname{im} T$, a subspace of a product of dimension $\sum_i \dim(V/U_i)$. (b) Substitute $\dim(V/X) = \dim V - \dim X$. (c) $4 + 4 + 3 - 2 \cdot 5 = 1$.

(a) Let $T : V \to V/U_1 \times \cdots \times V/U_m$, $T(\mathbf{v}) = (\mathbf{v} + U_1, \ldots, \mathbf{v} + U_m)$. It is linear by Proposition 4.6.7(a), and $T(\mathbf{v}) = \mathbf{0}$ if and only if $\mathbf{v} \in U_i$ for every i, so $\ker T = U_1 \cap \cdots \cap U_m =: X$. The product is finite-dimensional of dimension $\sum_i \dim(V/U_i)$ (Theorem 4.6.5), so its subspace $\operatorname{im} T$ is finite-dimensional with $\dim \operatorname{im} T \leq \sum_i \dim(V/U_i)$ (Proposition 4.1.21). By Theorem 4.6.34, $\tilde{T} : V/X \to \operatorname{im} T$ is an isomorphism; its inverse is injective with finite-dimensional domain $\operatorname{im} T$, so by Proposition 4.2.6(a), V/X is finite-dimensional with $\dim(V/X) = \dim \operatorname{im} T \leq \sum_i \dim(V/U_i)$.

(b) By Theorem 4.6.27, (a) reads $\dim V - \dim X \leq \sum_{i=1}^m (\dim V - \dim U_i) = m \dim V - \sum_i \dim U_i$. Rearranging, $\dim X \geq \sum_i \dim U_i - (m-1) \dim V$.

(c) By (b), the intersection has dimension at least $4 + 4 + 3 - 2 \cdot 5 = 1$, so it contains a nonzero vector.

Solution to Exercise 4.6.48

Uses. Proposition 4.6.7 (Maps into and out of a product), Theorem 4.6.5 (Dimension of a product), Theorem 4.2.11 (Rank–nullity), Proposition 2.4.14 (Algebra of matrices)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $A\mathbf{1}$ is the vector of row sums and $A^{\mathsf{T}}\mathbf{1}$ the vector of column sums, so $Z = \ker\Phi$. The image of Φ is not all of $F^n \times F^n$: both vectors have the same sum of entries.
Hint 2. Let $D = \{(\mathbf{r}, \mathbf{c}) \mid r_1 + \cdots + r_n = c_1 + \cdots + c_n\}$, of dimension $2n - 1$. To show $\operatorname{im}\Phi = D$, look for a preimage supported on the first row and first column.
Sketch. $Z = \ker\Phi$. $\operatorname{im}\Phi \subseteq D$, since both sums equal the sum of all entries. For $(\mathbf{r}, \mathbf{c}) \in D$ with common sum s, $A = \mathbf{r}\mathbf{e}_1^{\mathsf{T}} + \mathbf{e}_1\mathbf{c}^{\mathsf{T}} - s\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}$ has $\Phi(A) = (\mathbf{r}, \mathbf{c})$. So $\operatorname{rank}\Phi = 2n - 1$ and $\dim Z = n^2 - 2n + 1$.

Φ is linear (Proposition 4.6.7(a)), and $A\mathbf{1}$ and $A^{\mathsf{T}}\mathbf{1}$ are the vectors of row sums and column sums of A. So $Z = \ker\Phi$, a subspace. Let $\sigma(\mathbf{x}) = x_1 + \cdots + x_n$ and

$$D = \{(\mathbf{r}, \mathbf{c}) \in F^n \times F^n \mid \sigma(\mathbf{r}) = \sigma(\mathbf{c})\}.$$

D is the kernel of $(\mathbf{r}, \mathbf{c}) \mapsto \sigma(\mathbf{r}) - \sigma(\mathbf{c})$, a linear map onto F (it sends $(\mathbf{e}_1, \mathbf{0})$ to 1), so $\dim D = 2n - 1$ by Theorem 4.6.5 and rank–nullity.

Both $\sigma(A\mathbf{1})$ and $\sigma\left(A^{\mathsf{T}}\mathbf{1}\right)$ equal the sum of all entries of A, so $\operatorname{im}\Phi \subseteq D$. Conversely, let $(\mathbf{r}, \mathbf{c}) \in D$ with $s = \sigma(\mathbf{r}) = \sigma(\mathbf{c})$, and let

$$A = \mathbf{r}\mathbf{e}_1^{\mathsf{T}} + \mathbf{e}_1\mathbf{c}^{\mathsf{T}} - s\mathbf{e}_1\mathbf{e}_1^{\mathsf{T}}.$$

Using $\mathbf{e}_1^{\mathsf{T}}\mathbf{1} = 1$, $\mathbf{c}^{\mathsf{T}}\mathbf{1} = s$, and $\mathbf{r}^{\mathsf{T}}\mathbf{1} = s$,

$$A\mathbf{1} = \mathbf{r} + s\mathbf{e}_1 - s\mathbf{e}_1 = \mathbf{r}, \quad A^{\mathsf{T}}\mathbf{1} = \mathbf{e}_1\left(\mathbf{r}^{\mathsf{T}}\mathbf{1}\right) + \mathbf{c} - s\mathbf{e}_1 = \mathbf{c}.$$

So $\operatorname{im}\Phi = D$, and by rank–nullity $\dim Z = n^2 - (2n - 1) = (n - 1)^2$.

Solution to Exercise 4.6.49

Uses. Proposition 4.6.26 (The quotient map), Proposition 4.2.3 (Properties of restriction), Theorem 4.6.34 (First isomorphism theorem), Theorem 4.6.27 (Dimension of a quotient)
Techniques. Pass to a quotient (Technique 4.6.31)
Hint 1. The map is the quotient map $U + W \to (U + W)/U$ restricted to W. Why does every coset have a representative in W?
Hint 2. For (b), take dimensions on both sides of (a).
Sketch. $(\mathbf{u} + \mathbf{w}) + U = \mathbf{w} + U$, so the map is onto, and $\mathbf{w} + U = U$ iff $\mathbf{w} \in U$. By the first isomorphism theorem, $W/(U \cap W) \cong (U + W)/U$. Then $\dim W - \dim(U \cap W) = \dim(U + W) - \dim U$.

(a) Let $\pi : U + W \to (U + W)/U$ be the quotient map of the space $U + W$ by its subspace U, and $\varphi = \pi\restriction_W$, which is linear. Every element of $(U + W)/U$ is $(\mathbf{u} + \mathbf{w}) + U = \mathbf{w} + U = \varphi(\mathbf{w})$ for some $\mathbf{u} \in U$, $\mathbf{w} \in W$, so φ is surjective. Its kernel is $W \cap \ker\pi = W \cap U$ (Proposition 4.2.3(c)). By Theorem 4.6.34, $W/(U \cap W) \cong \operatorname{im}\varphi = (U + W)/U$.

(b) Let U and W be finite-dimensional. Then $U + W$ is finite-dimensional (it is spanned by bases of U and W), and isomorphic finite-dimensional spaces have equal dimensions. By (a) and Theorem 4.6.27, applied in W and in $U + W$,

$$\dim W - \dim(U \cap W) = \dim(U + W) - \dim U,$$

which rearranges to $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

Solution to Exercise 4.6.50

Uses. Proposition 4.6.26 (The quotient map), Proposition 2.6.8 (Kernels and images are subspaces), Theorem 4.6.30 (Universal property of the quotient), Theorem 4.6.34 (First isomorphism theorem), Lemma 4.6.14 (Equal cosets), Theorem 4.2.11 (Rank–nullity)
Techniques. Pass to a quotient (Technique 4.6.31)

Hint 1. (a) $X/U = \operatorname{im}(\pi\restriction_X)$. For $\pi^{-1}(X/U) \subseteq X$: if $\mathbf{v}+U = \mathbf{x}+U$ with $\mathbf{x} \in X$, then $\mathbf{v}-\mathbf{x} \in U \subseteq X$.
Hint 2. (b) Given Y, take $X = \pi^{-1}(Y)$. Check that it is a subspace containing U and that $\pi(X) = Y$. Uniqueness comes from (a).
Hint 3. (c) Induce a map $V/U \to V/X$ from the quotient map $V \to V/X$ with Theorem 4.6.30, and compute its kernel.
Sketch. (a) $X/U = \pi(X)$ is a subspace, and $\mathbf{v}+U \in X/U$ forces $\mathbf{v} \in X+U = X$. (b) $X = \pi^{-1}(Y)$ is a subspace containing $\ker\pi = U$, with $\pi(X) = Y$ since π is onto; if also $Y = X'/U$, then $X' = \pi^{-1}(Y) = X$ by (a). (c) π_X kills U, so it induces $\rho : V/U \to V/X$, $\mathbf{v}+U \mapsto \mathbf{v}+X$, onto, with kernel X/U. (d) $\pi\restriction_X$ has kernel U and image X/U.

(a) $X/U = \operatorname{im}(\pi\restriction_X)$, the image of a linear map, so it is a subspace (Proposition 2.6.8). Clearly $X \subseteq \pi^{-1}(X/U)$. Conversely, if $\pi(\mathbf{v}) \in X/U$, then $\mathbf{v}+U = \mathbf{x}+U$ for some $\mathbf{x} \in X$, so $\mathbf{v}-\mathbf{x} \in U \subseteq X$ (Lemma 4.6.14), and $\mathbf{v} = \mathbf{x}+(\mathbf{v}-\mathbf{x}) \in X$.

(b) Let Y be a subspace of V/U and $X = \pi^{-1}(Y)$. It is a subspace of V: it contains $\mathbf{0}$, and if $\mathbf{v}, \mathbf{v}' \in X$ and $a \in F$, then $\pi(a\mathbf{v}+\mathbf{v}') = a\pi(\mathbf{v})+\pi(\mathbf{v}') \in Y$. It contains U, because $\pi(U) = \{\mathbf{0}\} \subseteq Y$. And $X/U = \pi(X) = Y$: the containment $\subseteq$ is the definition of X, and every $\mathbf{y} \in Y$ is $\pi(\mathbf{v})$ for some $\mathbf{v}$ (as π is surjective), which then lies in X. For uniqueness, if $Y = X'/U$ with $U \subseteq X'$, then $X' = \pi^{-1}(X'/U) = \pi^{-1}(Y) = X$ by (a).

(c) Let $\pi_X : V \to V/X$ be the quotient map. Its kernel X contains U, so by Theorem 4.6.30(a) it induces a linear $\rho : V/U \to V/X$, $\rho(\mathbf{v}+U) = \mathbf{v}+X$, with image $\operatorname{im}\pi_X = V/X$ and kernel $\{\mathbf{v}+U \mid \mathbf{v} \in X\} = X/U$. By Theorem 4.6.34, $(V/U)/(X/U) \cong V/X$.

(d) $\pi\restriction_X : X \to V/U$ is linear with kernel $X \cap U = U$ and image X/U, so rank–nullity gives $\dim X = \dim U + \dim(X/U)$.

So the subspaces of V/U correspond to the subspaces of V that contain U, with dimensions shifted down by $\dim U$. The set X/U is also literally the quotient of the space X by its subspace U, which explains the notation.

Solution to Exercise 4.6.51

Uses. Theorem 4.6.27 (Dimension of a quotient), Theorem 4.5.6 (Classification by dimension), Example 4.6.36 (The first isomorphism theorem at work), Example 4.6.23 (Quotient spaces), Theorem 4.6.34 (First isomorphism theorem), Example 2.1.9 (Maps of polynomials)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. In finite dimensions, everything is decided by dimensions (Theorem 4.5.6 and Theorem 4.6.27).
Hint 2. In general, look in $F[x]$. The backward shift of Example 4.6.36(d) gives $F[x]/\operatorname{span}(1) \cong F[x]$, and multiplication by x maps $F[x]$ isomorphically onto the polynomials with constant term 0.
Sketch. Finite dimensions: all three are true, by comparing $\dim V - \dim U$ and $\dim V - \dim W$. In general all three are false in $V = F[x]$: (a) $U = \operatorname{span}(1)$; (b) $U = \operatorname{span}(1)$, $W = \{\mathbf{0}\}$; (c) $U = F[x]$, $W = \operatorname{im} M = \{p \mid p(0) = 0\}$, where $V/U = \{\mathbf{0}\}$ but $V/W \cong F$.

Finite-dimensional V. All three statements are true. By Theorem 4.6.27, $\dim(V/U) = \dim V - \dim U$ and $\dim(V/W) = \dim V - \dim W$, and finite-dimensional spaces are isomorphic exactly when their dimensions agree (Theorem 4.5.6).

(a) $\dim(V/U) = \dim V - \dim U < \dim V$ when $U \neq \{\mathbf{0}\}$.

(b) $\dim V - \dim U = \dim V - \dim W$ gives $\dim U = \dim W$.

(c) $\dim U = \dim W$ gives $\dim(V/U) = \dim(V/W)$.

General V. All three fail for $V = F[x]$.

(a) By Example 4.6.36(d), $F[x]/\operatorname{span}(1) \cong F[x]$, although $\operatorname{span}(1) \neq \{\mathbf{0}\}$.

(b) Take $U = \operatorname{span}(1)$ and $W = \{\mathbf{0}\}$. Then $V/U \cong V \cong V/W$ (Example 4.6.23(c)), but U is one-dimensional and W is zero, so $U \not\cong W$.

(c) Take $U = F[x]$ and $W = \{p \mid p(0) = 0\}$. Multiplication by x, $M(p) = xp$, is linear and injective (Example 2.1.9), and its image is W (a polynomial with constant term 0 is x times the polynomial with its

coefficients shifted down). So $U = F[x] \cong W$. But $V/U = \{\mathbf{0}\}$, while $V/W \cong F$ by Theorem 4.6.34 applied to ev_0, which is surjective with kernel W.

The finite-dimensional arguments all rest on subtracting dimensions, which is exactly what is unavailable in infinite dimensions.

Solution to Exercise 4.6.52

> *Uses.* Proposition 4.6.32 (Operators on a quotient), Proposition 4.2.3 (Properties of restriction), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces)
> *Techniques.* Restrict to a subspace (Technique 4.2.5), Pass to a quotient (Technique 4.6.31)
> *Hint 1.* Apply rank–nullity to $\pi|_{\operatorname{im} T}$, where $\pi : V \to V/U$.
> *Hint 2.* Its kernel is $\operatorname{im} T \cap U$, which contains $T(U) = \operatorname{im} T_U$. Its image is $\pi(\operatorname{im} T)$. Compare it with $\operatorname{im}(T/U)$.
> *Sketch.* $\pi|_{\operatorname{im} T}$ has kernel $\operatorname{im} T \cap U \supseteq T(U)$ and image $\{T\mathbf{v} + U\} = \operatorname{im}(T/U)$. So $\operatorname{rank} T = \dim(\operatorname{im} T \cap U) + \operatorname{rank}(T/U) \geq \operatorname{rank} T_U + \operatorname{rank}(T/U)$, with equality iff $\operatorname{im} T \cap U = T(U)$. Example: the 2×2 shift with $U = \operatorname{span}(\mathbf{e}_1)$.

(a) Let $X = \operatorname{im} T$ and $\pi : V \to V/U$. The restriction $\pi|_X$ is linear with kernel $X \cap U$ (Proposition 4.2.3(c)) and image $\{T(\mathbf{v}) + U \mid \mathbf{v} \in V\} = \{(T/U)(\mathbf{v} + U)\} = \operatorname{im}(T/U)$. By rank–nullity,

$$\operatorname{rank} T = \dim X = \dim(X \cap U) + \operatorname{rank}(T/U).$$

Since $T(U) \subseteq U$ and $T(U) \subseteq X$, we have $\operatorname{im} T_U = T(U) \subseteq X \cap U$, so $\dim(X \cap U) \geq \operatorname{rank} T_U$ (Proposition 4.1.21), and (a) follows.

(b) By the computation in (a), equality holds if and only if $\dim(X \cap U) = \dim T(U)$, which, as $T(U) \subseteq X \cap U$, holds if and only if $T(U) = X \cap U$ (Proposition 4.1.21(b)).

(c) Let $N \in \mathrm{L}(F^2)$ with $N(\mathbf{e}_1) = \mathbf{0}$, $N(\mathbf{e}_2) = \mathbf{e}_1$, and $U = \operatorname{span}(\mathbf{e}_1)$. Then $N_U = 0$, and $(N/U)(\mathbf{e}_2 + U) = \mathbf{e}_1 + U = \mathbf{0} + U$, so $N/U = 0$. Both ranks are 0, but $\operatorname{rank} N = 1$. Here $\operatorname{im} N \cap U = U$, while $N(U) = \{\mathbf{0}\}$.

In matrix terms (Proposition 4.6.32(b)): if $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ with square diagonal blocks, then $\operatorname{rank} M \geq \operatorname{rank} A + \operatorname{rank} C$, and the example $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ shows that the off-diagonal block can add rank.

Solution to Exercise 4.6.53

> *Uses.* Theorem 4.6.30 (Universal property of the quotient), Theorem 4.6.27 (Dimension of a quotient), Theorem 4.2.7 (The image of T_A), Theorem 4.2.11 (Rank–nullity)
> *Techniques.* Pass to a quotient (Technique 4.6.31)
> *Hint 1.* Check the containments $\operatorname{im} T_{BC} \subseteq \operatorname{im} T_B$ and $\operatorname{im} T_{ABC} \subseteq \operatorname{im} T_{AB}$, and that T_A maps $\operatorname{im} T_B$ into $\operatorname{im} T_{AB}$.
> *Hint 2.* Define $\varphi : \operatorname{im} T_B \to \operatorname{im} T_{AB}/\operatorname{im} T_{ABC}$ by $\varphi(\mathbf{x}) = A\mathbf{x} + \operatorname{im} T_{ABC}$. Why does it vanish on $\operatorname{im} T_{BC}$? Why is the induced map surjective?
> *Hint 3.* A surjective linear map cannot increase dimension. Translate both quotient dimensions with Theorem 4.6.27.
> *Sketch.* Let $X = \operatorname{im} T_B \supseteq Y = \operatorname{im} T_{BC}$ and $X' = \operatorname{im} T_{AB} \supseteq Y' = \operatorname{im} T_{ABC}$. The map $\mathbf{x} \mapsto A\mathbf{x} + Y'$ on X kills Y (as $ABC\mathbf{z} \in Y'$), so it induces $X/Y \to X'/Y'$, which is onto since $AB\mathbf{z} + Y'$ is the image of $B\mathbf{z} + Y$. Hence $\operatorname{rank}(AB) - \operatorname{rank}(ABC) \leq \operatorname{rank} B - \operatorname{rank}(BC)$.

Let $X = \operatorname{im} T_B$, $Y = \operatorname{im} T_{BC}$, $X' = \operatorname{im} T_{AB}$, and $Y' = \operatorname{im} T_{ABC}$. Since $BC\mathbf{z} = B(C\mathbf{z})$ and $ABC\mathbf{z} = AB(C\mathbf{z})$, we have $Y \subseteq X$ and $Y' \subseteq X'$, and $AX \subseteq X'$ since $A(B\mathbf{z}) = (AB)\mathbf{z}$.

Define $\varphi : X \to X'/Y'$ by $\varphi(\mathbf{x}) = A\mathbf{x} + Y'$. It is linear, being the composite of $T_A|_X : X \to X'$ with the quotient map $X' \to X'/Y'$. For $\mathbf{x} = BC\mathbf{z} \in Y$, $A\mathbf{x} = ABC\mathbf{z} \in Y'$, so $\varphi(\mathbf{x}) = \mathbf{0}$. By Theorem 4.6.30, φ induces a linear map $\tilde{\varphi} : X/Y \to X'/Y'$ with $\tilde{\varphi}(\mathbf{x} + Y) = A\mathbf{x} + Y'$. It is surjective: every element of X'/Y' is $AB\mathbf{z} + Y' = \tilde{\varphi}(B\mathbf{z} + Y)$.

By rank–nullity, $\dim(X'/Y') = \dim \operatorname{im} \tilde{\varphi} \leq \dim(X/Y)$. By Theorem 4.6.27 and Theorem 4.2.7,

$$\operatorname{rank}(AB) - \operatorname{rank}(ABC) \leq \operatorname{rank} B - \operatorname{rank}(BC),$$

which rearranges to the inequality.

The quotient proof makes the meaning visible: T_A maps "what B reaches beyond BC" onto "what AB reaches beyond ABC", and a linear map cannot make a space bigger.

Solution to Exercise 4.6.54

Uses. Theorem 4.4.7 (Criteria for direct sums), Proposition 4.4.16 (Complements exist), Proposition 4.6.29 (Complements and the quotient), Theorem 4.5.8 (Dimension of $\mathrm{L}(V,W)$), Proposition 4.1.34 (Size of a vector space over a finite field), Proposition 4.2.6 (Injective maps preserve dimension)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Dimension counting (Technique 4.2.24)
Hint 1. A complement is a choice of one vector in each coset of U, made linearly (Proposition 4.6.29). Starting from W_0, each $\mathbf{w} \in W_0$ can be moved within its coset by a vector of U, namely $S(\mathbf{w})$.
Hint 2. For (b), given a complement W, decompose each $\mathbf{w} \in W_0$ along $V = U \oplus W$ as $\mathbf{w} = \mathbf{u} + \mathbf{x}$, and set $S(\mathbf{w}) = -\mathbf{u}$.
Hint 3. For (c), $\mathrm{L}(W_0, U)$ has dimension $(n-k)k$ (Theorem 4.5.8); count its elements.
Sketch. (a) W_S is the image of $\mathbf{w} \mapsto \mathbf{w} + S\mathbf{w}$. If $\mathbf{w} + S\mathbf{w} \in U$, then $\mathbf{w} \in U \cap W_0 = \{\mathbf{0}\}$; and $U + W_S \supseteq U + W_0 = V$. (b) If $V = U \oplus W$, let $-S(\mathbf{w})$ be the U-component of $\mathbf{w} \in W_0$; S is linear by uniqueness of components, $W_S \subseteq W$, and dimensions agree. If $W_S = W_{S'}$, comparing $\mathbf{w} + S\mathbf{w} = \mathbf{w}' + S'\mathbf{w}'$ in $W_0 \oplus U$ gives $S = S'$. (c) $|\mathrm{L}(W_0, U)| = q^{k(n-k)}$. In $\mathbb{F}_3^2$: 3.

By Proposition 4.4.16(b), $\dim W_0 = n - k$.

(a) W_S is the image of the linear map $\mathbf{w} \mapsto \mathbf{w} + S(\mathbf{w})$, $W_0 \to V$, so it is a subspace. If $\mathbf{w} + S(\mathbf{w}) \in U$, then $\mathbf{w} \in U$ (as $S(\mathbf{w}) \in U$), so $\mathbf{w} \in U \cap W_0 = \{\mathbf{0}\}$ and the vector is $\mathbf{0}$. So $U \cap W_S = \{\mathbf{0}\}$. Every $\mathbf{w} \in W_0$ equals $(\mathbf{w} + S(\mathbf{w})) - S(\mathbf{w}) \in W_S + U$, so $U + W_S \supseteq U + W_0 = V$. Hence $V = U \oplus W_S$ (Theorem 4.4.7).

(b) Let W be a complement of U. Each $\mathbf{w} \in W_0$ has a unique decomposition $\mathbf{w} = \mathbf{u} + \mathbf{x}$ with $\mathbf{u} \in U$, $\mathbf{x} \in W$; set $S(\mathbf{w}) = -\mathbf{u}$. If $\mathbf{w} = \mathbf{u} + \mathbf{x}$ and $\mathbf{w}' = \mathbf{u}' + \mathbf{x}'$, then $a\mathbf{w} + \mathbf{w}' = (a\mathbf{u} + \mathbf{u}') + (a\mathbf{x} + \mathbf{x}')$ is the decomposition of $a\mathbf{w} + \mathbf{w}'$, so S is linear, with values in U. Now $\mathbf{w} + S(\mathbf{w}) = \mathbf{x} \in W$, so $W_S \subseteq W$. Both have dimension $n - k$: W by Proposition 4.4.16(b), and W_S because $\mathbf{w} \mapsto \mathbf{w} + S(\mathbf{w})$ is injective (if $\mathbf{w} + S(\mathbf{w}) = \mathbf{0}$, then $\mathbf{w} = -S(\mathbf{w}) \in U \cap W_0 = \{\mathbf{0}\}$) and Proposition 4.2.6(a). So $W_S = W$.

For uniqueness, suppose $W_S = W_{S'}$. For $\mathbf{w} \in W_0$, the vector $\mathbf{w} + S(\mathbf{w})$ lies in $W_{S'}$, so $\mathbf{w} + S(\mathbf{w}) = \mathbf{w}' + S'(\mathbf{w}')$ for some $\mathbf{w}' \in W_0$. Then $\mathbf{w} - \mathbf{w}' = S'(\mathbf{w}') - S(\mathbf{w}) \in W_0 \cap U = \{\mathbf{0}\}$, so $\mathbf{w}' = \mathbf{w}$ and $S(\mathbf{w}) = S'(\mathbf{w})$. Hence $S = S'$.

(c) By (a) and (b), $S \mapsto W_S$ is a bijection from $\mathrm{L}(W_0, U)$ onto the set of complements of U. The space $\mathrm{L}(W_0, U)$ has dimension $(n-k)k$ (Theorem 4.5.8), so it has $q^{k(n-k)}$ elements (Proposition 4.1.34). For a line U in $\mathbb{F}_3^2$, $n = 2$ and $k = 1$, so there are $3^1 = 3$ complements. Indeed, $\mathbb{F}_3^2$ has $(9-1)/2 = 4$ lines, and every line other than U meets U only in $\mathbf{0}$.

Solution to Exercise 4.6.55

Uses. Proposition 4.6.29 (Complements and the quotient), Proposition 4.6.32 (Operators on a quotient), Example 2.1.9 (Maps of polynomials), Proposition 1.2.18 (Degree and evaluation versus the operations), Definition 2.3.1 (Matrix of a linear map)
Techniques. Pass to a quotient (Technique 4.6.31), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Scale a nonzero element of degree d to make it monic. By minimality of d, which polynomials of degree less than d lie in U?
Hint 2. To show $F[x]_{\le d-1} + U = F[x]$, induct on the degree of q: subtract a multiple of $x^{e-d}p \in U$ to kill the leading term of q.
Hint 3. For (b), apply Proposition 4.6.32 to $M(q) = xq$. The last basis vector goes to $x^d + U$; rewrite x^d modulo p.
Sketch. (a) A nonzero element of degree d divided by its leading coefficient is p. Nonzero elements of U have degree $\ge d$, so $U \cap F[x]_{\le d-1} = \{\mathbf{0}\}$. If $\deg q = e \ge d$ with leading coefficient b, then $x^{e-d}p \in U$ and $\deg\left(q - bx^{e-d}p\right) < e$, so by induction $q \in F[x]_{\le d-1} + U$. By Proposition 4.6.29, $F[x]/U \cong F[x]_{\le d-1}$. (b) $M(U) \subseteq U$ gives $X = M/U$, and $X\left(x^j + U\right) = x^{j+1} + U$ for $j < d-1$, $X\left(x^{d-1} + U\right) = x^d + U = -\left(c_0 + \cdots + c_{d-1}x^{d-1}\right) + U$: the companion matrix of p.

(a) Let $p_0 \in U$ be nonzero of degree d with leading coefficient $a \neq 0$. Then $p = a^{-1}p_0 \in U$ has degree d and leading coefficient 1, so it has the stated form. By the hypothesis applied j times, $x^j p \in U$ for every $j \ge 0$, and $x^j p$ has degree $d + j$ and leading coefficient 1.

Intersection. A nonzero element of U has degree at least d, by the choice of d, so $U \cap F[x]_{\le d-1} = \{\mathbf{0}\}$.

Sum. We show by strong induction on e that every q of degree e lies in $F[x]_{\le d-1} + U$; the zero polynomial lies there trivially. If $e \le d-1$, then $q \in F[x]_{\le d-1}$. If $e \ge d$, let b be the leading coefficient of q. Then $q - bx^{e-d}p$

is zero or has degree less than e, because the coefficients of x^e cancel (Proposition 1.2.18). By induction it lies in $F[x]_{\le d-1} + U$, and since $bx^{e-d}p \in U$, so does q.

Hence $F[x] = F[x]_{\le d-1} \oplus U$, and by Proposition 4.6.29, $q \mapsto q + U$ restricts to an isomorphism $F[x]_{\le d-1} \to F[x]/U$. So $\dim(F[x]/U) = d$, and $1+U, x+U, \dots, x^{d-1}+U$ is a basis of $F[x]/U$ (the image of the basis $1, x, \dots, x^{d-1}$). If $d = 0$, then $p = 1$ and $U = F[x]$.

(b) Multiplication by x, $M(q) = xq$, is linear (Example 2.1.9), and $M(U) \subseteq U$ by hypothesis. By Proposition 4.6.32(a), $X = M/U$ is an operator on $F[x]/U$ with $X(q+U) = xq + U$. On the basis,

$$X\big(x^j + U\big) = x^{j+1} + U \quad (0 \le j \le d-2),$$

and, since $p \in U$,

$$X\big(x^{d-1} + U\big) = x^d + U = \big(x^d - p\big) + U = -c_0(1+U) - c_1(x+U) - \cdots - c_{d-1}\big(x^{d-1}+U\big).$$

So the matrix of X has 1s just below the diagonal, last column $(-c_0, -c_1, \dots, -c_{d-1})$, and zeros elsewhere:

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -c_{d-1} \end{pmatrix}.$$

This is the companion matrix of p, which reappears in Section 7.3.

One can show that U is exactly the set of polynomial multiples of p (polynomial multiplication and division come in Section 7.1). So the quotients of $F[x]$ by such subspaces are finite-dimensional, and multiplication by x on them is the basic example of an operator with a cyclic vector.

Section 4.7

Solution to Exercise 4.7.24

Uses. Proposition 4.7.20 (Affine independence through differences)
Hint 1. Subtract the first point from the others and test the differences for linear independence.
Hint 2. In (b), the relation $a(-1,1,0) + b(-1,0,1) + c(0,1,1) = \mathbf{0}$ leads to $-2a = 0$. What happens to this step over $\mathbb{F}_2$?
Sketch. (a) The differences $(1,1,1)$ and $(3,3,3)$ are dependent; $2\mathbf{v}_0 - 3\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$. (b) The differences are independent over $\mathbb{R}$, so the points are the vertices of a tetrahedron. (c) Over $\mathbb{F}_2$, the four points add up to $\mathbf{0}$, and $1+1+1+1 = 0$.

(a) The differences from $(1,2,3)$ are $(1,1,1)$ and $(3,3,3) = 3(1,1,1)$, which are linearly dependent, so the list is affinely dependent (Proposition 4.7.20). Translating $3(1,1,1) - (3,3,3) = \mathbf{0}$ back as in the proof of the proposition gives the affine dependence

$$2(1,2,3) - 3(2,3,4) + (4,5,6) = (2-6+4, 4-9+5, 6-12+6) = \mathbf{0}, \quad 2 - 3 + 1 = 0.$$

(b) The differences from $(1,0,0)$ are $(-1,1,0)$, $(-1,0,1)$, $(0,1,1)$. If $a(-1,1,0) + b(-1,0,1) + c(0,1,1) = \mathbf{0}$, then $-a - b = 0$, $a + c = 0$, and $b + c = 0$. So $b = -a$ and $c = -a$, and the last equation gives $-2a = 0$, so $a = 0$ over $\mathbb{R}$, and then $b = c = 0$. The differences are independent, and the four points are affinely independent.

(c) Over $\mathbb{F}_2$ the step $-2a = 0 \Rightarrow a = 0$ fails, and $a = b = c = 1$ is a relation: $(1,1,0) + (1,0,1) + (0,1,1) = (2,2,2) = \mathbf{0}$. Correspondingly,

$$(1,0,0) + (0,1,0) + (0,0,1) + (1,1,1) = (2,2,2) = \mathbf{0} \quad \text{and} \quad 1+1+1+1 = 0$$

in $\mathbb{F}_2$, an affine dependence. The four points lie in the plane $x_1 + x_2 + x_3 = 1$ of $\mathbb{F}_2^3$.

Solution to Exercise 4.7.43

Uses. Theorem 2.4.34 (Block multiplication), Theorem 2.5.10 (Left and right inverses agree)
Hint 1. Multiply the two block matrices in both orders.

Hint 2. For the 3×3 matrix, A is the rotation $\begin{pmatrix}0&-1\\1&0\end{pmatrix}$, whose inverse is $\begin{pmatrix}0&1\\-1&0\end{pmatrix}$.
Sketch. Both products are $\begin{pmatrix}I_n&\mathbf{0}\\0&1\end{pmatrix}=I_{n+1}$. The translation by $\mathbf{a}$ inverts to the translation by $-\mathbf{a}$. For the 3×3 matrix, $-A^{-1}(3,1)=-(1,-3)=(-1,3)$, giving $\begin{pmatrix}0&1&-1\\-1&0&3\\0&0&1\end{pmatrix}$.

By Theorem 2.4.34,

$$\begin{pmatrix}A&\mathbf{b}\\0&1\end{pmatrix}\begin{pmatrix}A^{-1}&-A^{-1}\mathbf{b}\\0&1\end{pmatrix}=\begin{pmatrix}AA^{-1}&-AA^{-1}\mathbf{b}+\mathbf{b}\\0&1\end{pmatrix}=\begin{pmatrix}I_n&\mathbf{0}\\0&1\end{pmatrix}=I_{n+1},$$

and

$$\begin{pmatrix}A^{-1}&-A^{-1}\mathbf{b}\\0&1\end{pmatrix}\begin{pmatrix}A&\mathbf{b}\\0&1\end{pmatrix}=\begin{pmatrix}A^{-1}A&A^{-1}\mathbf{b}-A^{-1}\mathbf{b}\\0&1\end{pmatrix}=I_{n+1}.$$

For the translation, $A=I_n$ and $\mathbf{b}=\mathbf{a}$, so the inverse is $\begin{pmatrix}I_n&-\mathbf{a}\\0&1\end{pmatrix}$, the matrix of the translation by $-\mathbf{a}$.

For the 3×3 matrix, $A=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ and $\mathbf{b}=(3,1)$. Since $A\begin{pmatrix}0&1\\-1&0\end{pmatrix}=I_2$, $A^{-1}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$, and $-A^{-1}\mathbf{b}=-(1,-3)=(-1,3)$. So the inverse is $\begin{pmatrix}0&1&-1\\-1&0&3\\0&0&1\end{pmatrix}$. Check: the product of the first row of the inverse with the third column of the original is $0\cdot 3+1\cdot 1+(-1)\cdot 1=0$, as it should be.

Solution to Exercise 4.7.46

Uses. Theorem 4.7.8 (Affine subspaces of F^n are solution sets), Theorem 4.7.14 (Affine subspaces are the sets closed under affine combinations), Proposition 4.7.3 (Direction and base points), Corollary 2.6.31 (Fibers are translates of the kernel)
Hint 1. (a) and (b) are solution sets of linear equations, in $\mathbb{R}^3$ and in $\mathbb{R}[x]$. Find one solution and the homogeneous solutions.
Hint 2. For (c) and (d), look for an affine combination of three points of the set, such as $\mathbf{u}+\mathbf{v}-\mathbf{w}$, that leaves the set.
Sketch. (a) $(2,-1,0)+\operatorname{span}((1,-2,1))$, a line. (b) The fiber of $p\mapsto(p(0),p(1))$ over $(1,3)$: $1+2x+\{p\mid p(0)=p(1)=0\}$, an affine subspace with infinite-dimensional direction. (c) No: $(1,0)+(0,1)-(0,0)=(1,1)$. (d) No: the set is $\{(0,0),(1,1),(2,1)\}$, and $(1,1)+(2,1)-(0,0)=(0,2)$ is missing.

(a) Yes. The homogeneous equations $x_1+x_2+x_3=0$, $x_1=x_3$ give $x_2=-2x_3$, so their solutions are $\operatorname{span}((1,-2,1))$. One solution of the system is $(2,-1,0)$: $2-1+0=1$ and $2-0=2$. By Theorem 4.7.8(a), the set is $(2,-1,0)+\operatorname{span}((1,-2,1))$, a line, with direction $\operatorname{span}((1,-2,1))$ and dimension 1.

(b) Yes. The map $T:\mathbb{R}[x]\to\mathbb{R}^2$, $T(p)=(p(0),p(1))$, is linear, and the set is the fiber $T^{-1}((1,3))$, which contains $1+2x$. By Corollary 2.6.31 it is $1+2x+\ker T$, an affine subspace with direction $\ker T=\{p\mid p(0)=p(1)=0\}$. This direction contains $x^j(x-1)$ for every $j\geq 1$, and these polynomials have distinct degrees, so $\ker T$ is infinite-dimensional and the dimension is not defined.

(c) No. The points $(1,0)$, $(0,1)$, $(0,0)$ lie in the set, and the affine combination $(1,0)+(0,1)-(0,0)=(1,1)$ does not, since $1\cdot 1\neq 0$. By Theorem 4.7.14 the set is not affine.

(d) No. In $\mathbb{F}_3$, $0^2=0$, $1^2=1$, and $2^2=4=1$, so the set is $\{(0,0),(1,1),(2,1)\}$. The affine combination $(1,1)+(2,1)-(0,0)=(3,2)=(0,2)$ is not in it, so it is not affine by Theorem 4.7.14.

Solution to Exercise 4.7.47

Uses. Remark 4.7.26 (Frames exist, and coordinates are coordinates), Proposition 4.7.32 (Barycentric coordinates on a simplex)
Hint 1. Since $\mathbf{v}_0=\mathbf{0}$, (t_1,t_2) are the coordinates of the point in the basis $(4,0),(1,3)$, and $t_0=1-t_1-t_2$.
Hint 2. Solve $4t_1+t_2=x$ and $3t_2=y$. A negative t_i puts the point beyond the edge opposite $\mathbf{v}_i$.
Sketch. $t_2=y/3$ and $t_1=(x-y/3)/4$. This gives $\mathbf{p}:\left(\frac{1}{4},\frac{5}{12},\frac{1}{3}\right)$, inside; $\mathbf{q}:\left(-\frac{1}{4},\frac{7}{12},\frac{2}{3}\right)$, beyond the edge $\mathbf{v}_1\mathbf{v}_2$; $\mathbf{r}:\left(1,-\frac{1}{3},\frac{1}{3}\right)$, beyond the edge $\mathbf{v}_0\mathbf{v}_2$.

The differences $(4,0)$ and $(1,3)$ are independent, so the points form an affine frame of $\mathbb{R}^2$. By Remark 4.7.26, since $\mathbf{v}_0 = \mathbf{0}$, a point (x,y) has $(x,y) = t_1(4,0) + t_2(1,3)$, that is, $4t_1 + t_2 = x$ and $3t_2 = y$, and $t_0 = 1 - t_1 - t_2$. So $t_2 = y/3$ and $t_1 = (x - y/3)/4$.

For $\mathbf{p} = (2,1)$: $t_2 = \frac{1}{3}$, $t_1 = \left(2 - \frac{1}{3}\right)/4 = \frac{5}{12}$, $t_0 = 1 - \frac{5}{12} - \frac{4}{12} = \frac{1}{4}$. All are positive, so $\mathbf{p}$ lies in the triangle (Proposition 4.7.32). Check: $\frac{5}{12}(4,0) + \frac{1}{3}(1,3) = \left(\frac{5}{3} + \frac{1}{3}, 1\right) = (2,1)$.

For $\mathbf{q} = (3,2)$: $t_2 = \frac{2}{3}$, $t_1 = \left(3 - \frac{2}{3}\right)/4 = \frac{7}{12}$, $t_0 = 1 - \frac{7}{12} - \frac{8}{12} = -\frac{1}{4}$. So $\mathbf{q}$ is outside, beyond the edge $\mathbf{v}_1\mathbf{v}_2$ opposite $\mathbf{v}_0$. Check: $\frac{7}{12}(4,0) + \frac{2}{3}(1,3) = \left(\frac{7}{3} + \frac{2}{3}, 2\right) = (3,2)$.

For $\mathbf{r} = (-1,1)$: $t_2 = \frac{1}{3}$, $t_1 = \left(-1 - \frac{1}{3}\right)/4 = -\frac{1}{3}$, $t_0 = 1 + \frac{1}{3} - \frac{1}{3} = 1$. So $\mathbf{r}$ is outside, beyond the edge $\mathbf{v}_0\mathbf{v}_2$ opposite $\mathbf{v}_1$. Check: $-\frac{1}{3}(4,0) + \frac{1}{3}(1,3) = (-1,1)$.

Solution to Exercise 4.7.48

Uses. Theorem 4.7.42 (Affine maps as matrices), Example 4.7.36 (Affine maps), Definition 4.7.41 (Homogeneous coordinates)
Hint 1. Each matrix is $\begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix}$ with A the linear part and $\mathbf{b}$ the image of $\mathbf{0}$. The scaling about $(1,1)$ is $\mathbf{v} \mapsto 3\mathbf{v} - 2(1,1)$.
Hint 2. By Theorem 4.7.42(c), $M_T M_R$ is the matrix of $T \circ R$, which rotates first.
Sketch. $M_T = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, $M_R = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $M_S = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{pmatrix}$. $M_T M_R = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ is $\mathbf{v} \mapsto R\mathbf{v} + (2,-1)$, and $M_R M_T = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ is $\mathbf{v} \mapsto R\mathbf{v} + (1,2)$. (c) $(2,-1)$, and $(4,-2,2)$.

(a) By Theorem 4.7.42, the matrix of $\mathbf{v} \mapsto A\mathbf{v} + \mathbf{b}$ is $\begin{pmatrix} A & \mathbf{b} \\ 0 & 1 \end{pmatrix}$. The scaling about $(1,1)$ is $\mathbf{v} \mapsto (1,1) + 3(\mathbf{v} - (1,1)) = 3\mathbf{v} - (2,2)$ (Example 4.7.36). So

$$M_T = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_R = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_S = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) Multiplying,

$$M_T M_R = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_R M_T = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

By Theorem 4.7.42(c), the first is the matrix of $T \circ R$, which rotates and then translates: $\mathbf{v} \mapsto R\mathbf{v} + (2,-1)$. The second is $R \circ T$, which translates and then rotates: $\mathbf{v} \mapsto R(\mathbf{v} + (2,-1)) = R\mathbf{v} + (1,2)$, since $R(2,-1) = (1,2)$. The two maps have the same linear part but differ by a translation, so the order matters.

(c) By Definition 4.7.41, $(6,-3,3)$ represents $3^{-1}(6,-3) = (2,-1)$. Multiplying $\widehat{(2,-1)} = (2,-1,1)$ by 2 gives $(4,-2,2)$.

Solution to Exercise 4.7.49

Uses. Proposition 4.7.45 (Perspective projection in homogeneous coordinates), Theorem 4.7.42 (Affine maps as matrices), Definition 4.7.41 (Homogeneous coordinates)
Hint 1. Parametrize the line as $\mathbf{c} + s(\mathbf{v} - \mathbf{c})$ and solve for the s that makes the third coordinate 0.
Hint 2. For (b), compute $\tau_{\mathbf{a}}(\pi(\tau_{-\mathbf{a}}(\mathbf{v})))$ with the formula for π; note that $\tau_{-\mathbf{a}}$ does not change z.
Hint 3. In (c), the translation matrices are $\begin{pmatrix} I_3 & \pm\mathbf{a} \\ 0 & 1 \end{pmatrix}$. Multiply $PM_{\tau_{-\mathbf{a}}}$ first.
Sketch. (a) $s = d/(d-z)$, and the first coordinate is $c_1 + s(x - c_1) = (dx - c_1 z)/(d - z)$. (b) $\pi(x - c_1, y - c_2, z) + \mathbf{a}$ has first coordinate $d(x - c_1)/(d - z) + c_1$, the same. (c) The product is $\begin{pmatrix} 1 & 0 & -c_1/d & 0 \\ 0 & 1 & -c_2/d & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{pmatrix}$, which sends $\hat{\mathbf{v}}$ to $(x - c_1 z/d, y - c_2 z/d, 0, 1 - z/d)$.

(a) The line through $\mathbf{c}$ and $\mathbf{v}$ consists of the points $\mathbf{c} + s(\mathbf{v} - \mathbf{c})$, whose third coordinate is $d + s(z - d)$. It vanishes exactly for $s = d/(d - z)$, which exists since $z \neq d$. The first coordinate is then

$$c_1 + \frac{d}{d-z}(x - c_1) = \frac{c_1(d-z) + dx - dc_1}{d-z} = \frac{dx - c_1 z}{d-z},$$

and the second is $(dy - c_2 z)/(d-z)$ in the same way.

(b) $\tau_{-\mathbf{a}}(\mathbf{v}) = (x - c_1, y - c_2, z)$ has third coordinate $z \neq d$, so by the formula for π (the case $c_1 = c_2 = 0$ of (a)),

$$\pi(\tau_{-\mathbf{a}}(\mathbf{v})) = \left(\frac{d(x-c_1)}{d-z}, \frac{d(y-c_2)}{d-z}, 0\right).$$

Adding $\mathbf{a}$, the first coordinate becomes $(dx - dc_1 + c_1(d-z))/(d-z) = (dx - c_1 z)/(d-z)$, and likewise the second. By (a), this is $\pi_{\mathbf{c}}(\mathbf{v})$.

(c) With $M_{\tau_{\pm\mathbf{a}}} = \begin{pmatrix} I_3 & \pm\mathbf{a} \\ 0 & 1 \end{pmatrix}$,

$$PM_{\tau_{-\mathbf{a}}} = \begin{pmatrix} 1 & 0 & 0 & -c_1 \\ 0 & 1 & 0 & -c_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{pmatrix}, \quad M_{\tau_{\mathbf{a}}} PM_{\tau_{-\mathbf{a}}} = \begin{pmatrix} 1 & 0 & -c_1/d & 0 \\ 0 & 1 & -c_2/d & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{pmatrix}.$$

(In the second product, row 1 is row 1 plus c_1 times row 4 of $PM_{\tau_{-\mathbf{a}}}$, and similarly for row 2.) This matrix sends $\hat{\mathbf{v}} = (x, y, z, 1)$ to $(x - c_1 z/d, y - c_2 z/d, 0, 1 - z/d)$. The last entry is nonzero, and dividing by it gives first coordinate $(x - c_1 z/d)/(1 - z/d) = (dx - c_1 z)/(d-z)$, and similarly for the second, which is $\pi_{\mathbf{c}}(\mathbf{v})$ by (a).

Solution to Exercise 4.7.50

Uses. Proposition 4.7.45 (Perspective projection in homogeneous coordinates), Definition 4.7.41 (Homogeneous coordinates), Proposition 4.7.17 (The affine hull)

Hint 1. P is linear and $\widehat{\mathbf{v} + s\mathbf{w}} = \hat{\mathbf{v}} + s\begin{pmatrix}\mathbf{w}\\0\end{pmatrix}$, so $P\widehat{\mathbf{v} + s\mathbf{w}} = \mathbf{h}_1 + s\mathbf{h}_2$ with $\mathbf{h}_1 = P\hat{\mathbf{v}}$ and $\mathbf{h}_2 = P\begin{pmatrix}\mathbf{w}\\0\end{pmatrix}$.

Hint 2. Let α and β be the last entries of $\mathbf{h}_1$ and $\mathbf{h}_2$. Dividing $\mathbf{h}_1 + s\mathbf{h}_2$ by $\alpha + s\beta$ writes the image point as an affine combination of $\pi(\mathbf{v})$ and $\mathbf{q}_{\mathbf{w}}$ with coefficients $\alpha/(\alpha + s\beta)$ and $s\beta/(\alpha + s\beta)$.

Sketch. $\mathbf{h}_2 = (w_1, w_2, 0, -w_3/d)$ represents $\mathbf{q}_{\mathbf{w}}$. With $\alpha = 1 - v_3/d$ and $\beta = -w_3/d$, the point $\pi(\mathbf{v} + s\mathbf{w})$ is the first three entries of $\mathbf{h}_1 + s\mathbf{h}_2$ divided by $\alpha + s\beta$, which equals $\alpha/(\alpha + s\beta) \cdot \pi(\mathbf{v}) + s\beta/(\alpha + s\beta) \cdot \mathbf{q}_{\mathbf{w}}$, an affine combination. In (b), z is constant along the line, so $\pi(\mathbf{v} + s\mathbf{w}) = \pi(\mathbf{v}) + (s/\alpha)(w_1, w_2, 0)$.

(a) $P\begin{pmatrix}\mathbf{w}\\0\end{pmatrix} = (w_1, w_2, 0, -w_3/d) = \mathbf{h}_2$, whose last entry $\beta = -w_3/d$ is nonzero, so it represents $\beta^{-1}(w_1, w_2, 0) = \mathbf{q}_{\mathbf{w}}$. Let $\mathbf{h}_1 = P\hat{\mathbf{v}} = (v_1, v_2, 0, \alpha)$ with $\alpha = 1 - v_3/d \neq 0$; it represents $\pi(\mathbf{v})$ by Proposition 4.7.45. Since $\widehat{\mathbf{v} + s\mathbf{w}} = \hat{\mathbf{v}} + s\begin{pmatrix}\mathbf{w}\\0\end{pmatrix}$ and P is linear, $P\widehat{\mathbf{v} + s\mathbf{w}} = \mathbf{h}_1 + s\mathbf{h}_2$. If $v_3 + sw_3 \neq d$, its last entry $\alpha + s\beta = 1 - (v_3 + sw_3)/d$ is nonzero, and by Proposition 4.7.45,

$$\pi(\mathbf{v} + s\mathbf{w}) = \frac{1}{\alpha + s\beta}((v_1, v_2, 0) + s(w_1, w_2, 0)) = \frac{\alpha}{\alpha + s\beta}\pi(\mathbf{v}) + \frac{s\beta}{\alpha + s\beta}\mathbf{q}_{\mathbf{w}},$$

using $(v_1, v_2, 0) = \alpha\pi(\mathbf{v})$ and $(w_1, w_2, 0) = \beta\mathbf{q}_{\mathbf{w}}$. The coefficients add up to 1, so the point lies in $\operatorname{aff}(\pi(\mathbf{v}), \mathbf{q}_{\mathbf{w}})$ (Proposition 4.7.17), which is the line through $\pi(\mathbf{v})$ and $\mathbf{q}_{\mathbf{w}}$ if these differ, and the single point $\mathbf{q}_{\mathbf{w}}$ otherwise. The point $\mathbf{q}_{\mathbf{w}}$ depends only on $\mathbf{w}$, so this holds for every line $\mathbf{v} + \operatorname{span}(\mathbf{w})$ with $v_3 \neq d$. (A line in this direction with $v_3 = d$ contains a point $\mathbf{v} + s\mathbf{w}$ with $v_3 + sw_3 \neq d$, and we may start from that point instead.)

(b) If $w_3 = 0$, then $(w_1, w_2) \neq (0, 0)$, and every point of the line has third coordinate $v_3 \neq d$. So

$$\pi(\mathbf{v} + s\mathbf{w}) = \frac{1}{\alpha}(v_1 + sw_1, v_2 + sw_2, 0) = \pi(\mathbf{v}) + \frac{s}{\alpha}(w_1, w_2, 0).$$

As s runs over $\mathbb{R}$, so does s/α, and the image is the line $\pi(\mathbf{v}) + \operatorname{span}((w_1, w_2, 0))$.

Solution to Exercise 4.7.51

Uses. Theorem 4.7.14 (Affine subspaces are the sets closed under affine combinations), Proposition 4.7.17 (The affine hull)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Try the smallest field first.
Hint 2. If F has an element $t \notin \{0,1\}$, take $\mathbf{a} \in E_1 \setminus E_2$ and $\mathbf{b} \in E_2 \setminus E_1$, and ask where $(1-t)\mathbf{a} + t\mathbf{b}$ lies.
Hint 3. If $\mathbf{c} = (1-t)\mathbf{a} + t\mathbf{b} \in E_1$, then $\mathbf{b}$ is an affine combination of $\mathbf{a}$ and $\mathbf{c}$.
Sketch. Over $\mathbb{F}_2$ it is false: in $V = \mathbb{F}_2$, $\{0\} \cup \{1\} = \mathbb{F}_2$. If $|F| \geq 3$ it is true: with $\mathbf{a}, \mathbf{b}$ as in the hint and $t \notin \{0,1\}$, $\mathbf{c} = (1-t)\mathbf{a} + t\mathbf{b} \in E_1 \cup E_2$. If $\mathbf{c} \in E_1$, then $\mathbf{b} = (1-1/t)\mathbf{a} + (1/t)\mathbf{c} \in E_1$; if $\mathbf{c} \in E_2$, then $\mathbf{a} = (1/(1-t))\mathbf{c} - (t/(1-t))\mathbf{b} \in E_2$. Both are contradictions.

The answer depends on F. Over $\mathbb{F}_2$ the statement is false: in $V = \mathbb{F}_2$, the points $\{0\}$ and $\{1\}$ are affine subspaces, neither contains the other, and their union $\mathbb{F}_2$ is a subspace.

Now suppose F has more than two elements, and choose $t \in F$ with $t \neq 0$ and $t \neq 1$. Suppose $E_1 \cup E_2$ is affine but neither of E_1, E_2 contains the other. Pick $\mathbf{a} \in E_1$ with $\mathbf{a} \notin E_2$, and $\mathbf{b} \in E_2$ with $\mathbf{b} \notin E_1$. The point $\mathbf{c} = (1-t)\mathbf{a} + t\mathbf{b}$ is an affine combination of elements of $E_1 \cup E_2$, so it lies in $E_1 \cup E_2$ by Theorem 4.7.14.

If $\mathbf{c} \in E_1$, then

$$\left(1 - \frac{1}{t}\right)\mathbf{a} + \frac{1}{t}\mathbf{c} = \left(1 - \frac{1}{t} + \frac{1-t}{t}\right)\mathbf{a} + \mathbf{b} = \mathbf{b},$$

an affine combination of $\mathbf{a}, \mathbf{c} \in E_1$, so $\mathbf{b} \in E_1$, a contradiction. If $\mathbf{c} \in E_2$, then

$$\frac{1}{1-t}\mathbf{c} - \frac{t}{1-t}\mathbf{b} = \mathbf{a} + \frac{t}{1-t}\mathbf{b} - \frac{t}{1-t}\mathbf{b} = \mathbf{a},$$

an affine combination of $\mathbf{c}, \mathbf{b} \in E_2$ (the coefficients add up to $(1-t)/(1-t) = 1$), so $\mathbf{a} \in E_2$, a contradiction. Hence, when $|F| \geq 3$, the statement is true. The argument breaks down over $\mathbb{F}_2$ exactly because the line through $\mathbf{a}$ and $\mathbf{b}$ has no third point.

Solution to Exercise 4.7.52

Uses. Proposition 4.7.10 (Intersections of affine subspaces), Theorem 4.4.19 (Dimension of a sum), Proposition 4.1.21 (Dimensions of subspaces), Corollary 4.7.6 (Parallel affine subspaces are equal or disjoint)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Let $U = \operatorname{dir} H$ and $U' = \operatorname{dir} H'$. If $U \neq U'$, what is $U + U'$?
Hint 2. To find a common point, write $\mathbf{h}' - \mathbf{h} = \mathbf{u} + \mathbf{u}'$ with $\mathbf{h} \in H$, $\mathbf{h}' \in H'$, $\mathbf{u} \in U$, $\mathbf{u}' \in U'$.
Sketch. If $U \neq U'$, then $U' \not\subseteq U$ (equal dimensions), so $U + U'$ strictly contains U and equals V. Writing $\mathbf{h}' - \mathbf{h} = \mathbf{u} + \mathbf{u}'$ gives the common point $\mathbf{h} + \mathbf{u} = \mathbf{h}' - \mathbf{u}'$. By Proposition 4.7.10 the intersection has direction $U \cap U'$, of dimension $2(n-1) - n = n - 2$.

Let $U = \operatorname{dir} H$ and $U' = \operatorname{dir} H'$, both of dimension $n-1$, and suppose H and H' are not parallel, so $U \neq U'$. If $U' \subseteq U$, then $U' = U$ by Proposition 4.1.21(b), which is not the case. So some vector of U' lies outside U, and $U \subsetneq U + U'$. Then $n - 1 < \dim(U + U') \leq n$, so $U + U' = V$ by Proposition 4.1.21(b).

Pick $\mathbf{h} \in H$ and $\mathbf{h}' \in H'$, and write $\mathbf{h}' - \mathbf{h} = \mathbf{u} + \mathbf{u}'$ with $\mathbf{u} \in U$ and $\mathbf{u}' \in U'$. Then $\mathbf{h} + \mathbf{u} = \mathbf{h}' - \mathbf{u}'$ lies in H and in H'. By Proposition 4.7.10, $H \cap H'$ is an affine subspace with direction $U \cap U'$, and by Theorem 4.4.19,

$$\dim(U \cap U') = \dim U + \dim U' - \dim(U + U') = 2(n-1) - n = n - 2.$$

For $n = 2$, the hyperplanes are lines. Two lines in a plane are either parallel, or they meet in an affine subspace of dimension 0, a single point.

Solution to Exercise 4.7.53

Uses. Proposition 1.5.14 (Appending a vector), Theorem 4.1.10 (Independent lists are short), Definition 4.7.5 (Direction, dimension, parallel)
Hint 1. The lines meet if and only if $\mathbf{p} + a\mathbf{u} = \mathbf{q} + b\mathbf{w}$ for some scalars. Rewrite this as a statement about $\mathbf{q} - \mathbf{p}$.
Hint 2. The lines are parallel if and only if $\operatorname{span}(\mathbf{u}) = \operatorname{span}(\mathbf{w})$, that is, if and only if $\mathbf{u}, \mathbf{w}$ is dependent. Then use Proposition 1.5.14.

Sketch. $\ell_1 \cap \ell_2 \neq \emptyset$ if and only if $\mathbf{q}-\mathbf{p} \in \operatorname{span}(\mathbf{u},\mathbf{w})$, and the lines are parallel if and only if $\mathbf{u},\mathbf{w}$ is dependent. So "disjoint and not parallel" means "$\mathbf{u},\mathbf{w}$ independent and $\mathbf{q}-\mathbf{p} \notin \operatorname{span}(\mathbf{u},\mathbf{w})$", which is the independence of $\mathbf{u},\mathbf{w},\mathbf{q}-\mathbf{p}$ by Proposition 1.5.14. In dimension 2, three vectors are dependent.

The lines meet if and only if $\mathbf{p}+a\mathbf{u} = \mathbf{q}+b\mathbf{w}$ for some $a,b \in F$, that is, if and only if $\mathbf{q}-\mathbf{p} = a\mathbf{u}-b\mathbf{w}$ for some a,b, that is, if and only if $\mathbf{q}-\mathbf{p} \in \operatorname{span}(\mathbf{u},\mathbf{w})$.

The lines are parallel if and only if $\operatorname{span}(\mathbf{u}) = \operatorname{span}(\mathbf{w})$. Since $\mathbf{u},\mathbf{w} \neq \mathbf{0}$, this holds if and only if $\mathbf{w} = c\mathbf{u}$ for some c (necessarily nonzero), that is, if and only if $\mathbf{u},\mathbf{w}$ is linearly dependent: a nontrivial relation $a\mathbf{u}+b\mathbf{w}=\mathbf{0}$ must have $a,b \neq 0$, as both vectors are nonzero.

So ℓ_1,ℓ_2 are disjoint and not parallel if and only if $\mathbf{u},\mathbf{w}$ is independent and $\mathbf{q}-\mathbf{p} \notin \operatorname{span}(\mathbf{u},\mathbf{w})$. By Proposition 1.5.14, this is equivalent to the independence of $\mathbf{u},\mathbf{w},\mathbf{q}-\mathbf{p}$. (If $\mathbf{u},\mathbf{w}$ is dependent, so is the longer list, and both sides fail.)

If $\dim V = 2$, no list of three vectors is independent (Theorem 4.1.10), so two lines are never disjoint and non-parallel.

Solution to Exercise 4.7.54

Uses. Proposition 4.7.17 (The affine hull), Theorem 4.7.8 (Affine subspaces of F^n are solution sets), Proposition 4.1.21 (Dimensions of subspaces)
Hint 1. The direction is spanned by $(1,1,-1,0)$ and $(-1,1,1,1)$. Find the dimension, then the number of equations from Theorem 4.7.8(b).
Hint 2. An equation $a_1x_1+\cdots+a_4x_4=c$ works for the direction if $(a_1,\ldots,a_4)$ kills both direction vectors. Find two independent such rows, then compute the right-hand sides at $(1,0,2,1)$.
Sketch. The two differences are independent, so the hull is a plane and needs $4-2=2$ equations. The rows $(1,0,1,0)$ and $(1,-1,0,2)$ kill both differences, and the system $x_1+x_3=3$, $x_1-x_2+2x_4=3$ works.

By Proposition 4.7.17(b), the hull is $(1,0,2,1)+\operatorname{span}(\mathbf{d}_1,\mathbf{d}_2)$ with $\mathbf{d}_1=(1,1,-1,0)$ and $\mathbf{d}_2=(-1,1,1,1)$. These are independent (the last entries force the coefficient of $\mathbf{d}_2$ to be 0, and then that of $\mathbf{d}_1$), so the hull is a plane, and by Theorem 4.7.8(b) it takes $4-2=2$ equations, and no fewer.

A row $\mathbf{a}=(a_1,a_2,a_3,a_4)$ with $A\mathbf{d}_1 = A\mathbf{d}_2 = 0$ satisfies $a_1+a_2-a_3=0$ and $-a_1+a_2+a_3+a_4=0$. Taking $(a_3,a_4)=(1,0)$ gives $(1,0,1,0)$, and $(a_3,a_4)=(0,2)$ gives $(1,-1,0,2)$. With $A=\begin{pmatrix}1&0&1&0\\1&-1&0&2\end{pmatrix}$, of rank 2, and $A(1,0,2,1)=(3,3)$, the system is

$$x_1+x_3=3, \quad x_1-x_2+2x_4=3.$$

Its solution set is an affine subspace containing $(1,0,2,1)$ with direction $\ker T_A$ of dimension $4-2=2$, which contains $\mathbf{d}_1,\mathbf{d}_2$; so the direction is $\operatorname{span}(\mathbf{d}_1,\mathbf{d}_2)$ by Proposition 4.1.21(b), and the solution set is the hull. Check the other points: $2+1=3$, $2-1+2=3$, and $0+3=3$, $0-1+4=3$.

Solution to Exercise 4.7.55

Uses. Theorem 2.1.3 (Checking linearity), Definition 4.7.35 (Affine map)
Hint 1. Replace f by $g=f-f(\mathbf{0})$. It satisfies the same identity and $g(\mathbf{0})=\mathbf{0}$.
Hint 2. Take $\mathbf{v}=\mathbf{0}$ to get $g\left(\frac{1}{2}\mathbf{u}\right)=\frac{1}{2}g(\mathbf{u})$. Then $g(\mathbf{u}+\mathbf{v})=2g\left(\frac{1}{2}(\mathbf{u}+\mathbf{v})\right)$.
Hint 3. Additivity gives $g(k\mathbf{u})=kg(\mathbf{u})$ for integers k. For $r=k/m$, compute $mg(r\mathbf{u})$.
Sketch. $g=f-f(\mathbf{0})$ satisfies the midpoint identity and $g(\mathbf{0})=\mathbf{0}$, so $g(\mathbf{u}/2)=g(\mathbf{u})/2$ and $g(\mathbf{u}+\mathbf{v})=2g((\mathbf{u}+\mathbf{v})/2)=g(\mathbf{u})+g(\mathbf{v})$. By induction $g(k\mathbf{u})=kg(\mathbf{u})$ for $k\geq 0$, also for $k<0$ since $g(-\mathbf{u})=-g(\mathbf{u})$, and $mg((k/m)\mathbf{u})=g(k\mathbf{u})=kg(\mathbf{u})$. So g is linear over $\mathbb{Q}$ and $f=g+f(\mathbf{0})$.

Let $g(\mathbf{x})=f(\mathbf{x})-f(\mathbf{0})$. Then $g(\mathbf{0})=\mathbf{0}$, and subtracting $f(\mathbf{0})=\frac{1}{2}f(\mathbf{0})+\frac{1}{2}f(\mathbf{0})$ from the hypothesis gives

$$g\left(\frac{1}{2}\mathbf{u}+\frac{1}{2}\mathbf{v}\right)=\frac{1}{2}g(\mathbf{u})+\frac{1}{2}g(\mathbf{v}) \quad \text{for all } \mathbf{u},\mathbf{v}.$$

With $\mathbf{v}=\mathbf{0}$ this says $g\left(\frac{1}{2}\mathbf{u}\right)=\frac{1}{2}g(\mathbf{u})$. Applying that to $\mathbf{u}+\mathbf{v}$,

$$g(\mathbf{u}+\mathbf{v})=2g\left(\frac{1}{2}(\mathbf{u}+\mathbf{v})\right)=2\left(\frac{1}{2}g(\mathbf{u})+\frac{1}{2}g(\mathbf{v})\right)=g(\mathbf{u})+g(\mathbf{v}).$$

So g is additive. By induction on k, $g(k\mathbf{u}) = kg(\mathbf{u})$ for integers $k \geq 0$ (the case $k = 0$ is $g(\mathbf{0}) = \mathbf{0}$, and $g((k+1)\mathbf{u}) = g(k\mathbf{u}) + g(\mathbf{u})$). Since $g(\mathbf{u}) + g(-\mathbf{u}) = g(\mathbf{0}) = \mathbf{0}$, the same holds for negative k. For $r = k/m$ with $k, m \in \mathbb{Z}$ and $m > 0$,

$$mg(r\mathbf{u}) = g(mr\mathbf{u}) = g(k\mathbf{u}) = kg(\mathbf{u}),$$

so $g(r\mathbf{u}) = rg(\mathbf{u})$. Hence g is additive and homogeneous over $\mathbb{Q}$, so it is linear by Theorem 2.1.3, and $f = g + f(\mathbf{0})$ is affine.

Solution to Exercise 4.7.56

Uses. Definition 4.7.18 (Affine independence), Example 4.7.19 (Affine independence)
Hint 1. Changing a coefficient changes the vector that the combination represents. Is the new combination still $\mathbf{0}$?
Hint 2. Test the claim on a list containing $\mathbf{0}$.
Sketch. Replacing c_0 by $c_0 - s$ changes the combination to $-s\mathbf{v}_0$, which need not be $\mathbf{0}$, so the new coefficients need not form an affine dependence. The forward half is true. The converse is false: $\mathbf{0}, \mathbf{e}_1$ is affinely independent but linearly dependent.

The first half is correct: an affine dependence is a nontrivial linear relation, so a linearly independent list has none.

The error is in the converse. After c_0 is replaced by $c_0 - s$, the combination represents

$$(c_0 - s)\mathbf{v}_0 + c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0} - s\mathbf{v}_0 = -s\mathbf{v}_0,$$

which is not $\mathbf{0}$ unless $s = 0$ or $\mathbf{v}_0 = \mathbf{0}$. So the new coefficients are not an affine dependence, and affine independence says nothing about them. (The last sentence of the argument is also confused, but the claim fails before it.)

The converse is false. The list $\mathbf{0}, \mathbf{e}_1$ in $\mathbb{R}^2$ consists of two distinct points, so it is affinely independent (Example 4.7.19(a)), but it contains $\mathbf{0}$, so it is linearly dependent. Likewise $(0,0), (1,0), (0,1)$ is affinely independent and linearly dependent (Example 4.7.19(b)).

Solution to Exercise 4.7.57

Uses. Theorem 4.7.22 (Characterizations of affine independence), Proposition 4.7.17 (The affine hull), Corollary 4.7.6 (Parallel affine subspaces are equal or disjoint), Proposition 4.7.3 (Direction and base points), Proposition 4.1.34 (Size of a vector space over a finite field)
Hint 1. For (a), write $\mathbf{g}$ as an affine combination of $\mathbf{v}_i$ and $\mathbf{m}_i$. To see that two medians differ, show that if the median from $\mathbf{v}_0$ contained $\mathbf{v}_1$, then $\mathbf{v}_2$ would lie in $\operatorname{aff}(\mathbf{v}_0, \mathbf{v}_1)$.
Hint 2. For (b), in $\mathbb{F}_3$ we have $1/2 = 2 = -1$, so $\mathbf{m}_i = -(\mathbf{v}_j + \mathbf{v}_k)$. Compute $\mathbf{m}_i - \mathbf{v}_i$.
Hint 3. Let $\mathbf{s} = \mathbf{v}_0 + \mathbf{v}_1 + \mathbf{v}_2$. Show $\mathbf{s} \neq \mathbf{0}$, and that $\mathbf{v}_1 - \mathbf{v}_0 \in \operatorname{span}(\mathbf{s})$ would give an affine dependence, using $3 = 0$.
Sketch. (a) $\mathbf{g} = \frac{1}{3}\mathbf{v}_i + \frac{2}{3}\mathbf{m}_i$ lies on every median. If two medians were equal, the median from $\mathbf{v}_0$ would contain $\mathbf{v}_1$, so $\mathbf{m}_0 \in \operatorname{aff}(\mathbf{v}_0, \mathbf{v}_1)$ and $\mathbf{v}_2 = 2\mathbf{m}_0 - \mathbf{v}_1$ would be in it too. Distinct lines share at most one point. (b) $\mathbf{m}_i - \mathbf{v}_i = -\mathbf{s}$ for every i, and $\mathbf{s} \neq \mathbf{0}$, so each median is $\mathbf{v}_i + \operatorname{span}(\mathbf{s})$. If $\mathbf{v}_j - \mathbf{v}_i = c\mathbf{s}$, the coefficients of the resulting relation add up to $-3c = 0$, giving an affine dependence. So the medians are three distinct parallel lines, disjoint by Corollary 4.7.6, with 3 points each, covering all 9 points.

(a) Write $\mathbf{m}_0 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$, and similarly $\mathbf{m}_1, \mathbf{m}_2$. Then

$$\frac{1}{3}\mathbf{v}_0 + \frac{2}{3}\mathbf{m}_0 = \frac{1}{3}(\mathbf{v}_0 + \mathbf{v}_1 + \mathbf{v}_2) = \mathbf{g},$$

an affine combination of $\mathbf{v}_0$ and $\mathbf{m}_0$, so $\mathbf{g}$ lies on the median from $\mathbf{v}_0$ (Proposition 4.7.17), and likewise on the other two. The median from $\mathbf{v}_0$ is a line, since $\mathbf{m}_0 \neq \mathbf{v}_0$: otherwise $\mathbf{v}_0$ would be an affine combination of $\mathbf{v}_1, \mathbf{v}_2$, contradicting Theorem 4.7.22(V).

Suppose two medians coincide, say those from $\mathbf{v}_0$ and $\mathbf{v}_1$ (the other cases are the same up to renaming). Then $\mathbf{v}_1$ lies on the median from $\mathbf{v}_0$, so this median is $\operatorname{aff}(\mathbf{v}_0, \mathbf{v}_1)$, and it contains $\mathbf{m}_0$. Then $\mathbf{v}_2 = 2\mathbf{m}_0 - \mathbf{v}_1$ is an affine combination of points of $\operatorname{aff}(\mathbf{v}_0, \mathbf{v}_1)$, so it lies there, contradicting Theorem 4.7.22(V). So the medians are three distinct lines. Two distinct lines meet in at most one point: their intersection is empty or

an affine subspace of dimension at most 1 (Proposition 4.7.10), and if it had dimension 1 it would equal both lines (Proposition 4.1.21(b) applied to the directions, and Proposition 4.7.3). So $\mathbf{g}$ is the only common point.

(b) In $\mathbb{F}_3$, $2 \cdot 2 = 4 = 1$, so $1/2 = 2 = -1$, and $\mathbf{m}_i = -(\mathbf{v}_j + \mathbf{v}_k)$, where $\{i,j,k\} = \{0,1,2\}$. Let $\mathbf{s} = \mathbf{v}_0 + \mathbf{v}_1 + \mathbf{v}_2$. Then $\mathbf{m}_i - \mathbf{v}_i = -\mathbf{s}$ for every i. If $\mathbf{s} = \mathbf{0}$, then $\mathbf{v}_2 = -\mathbf{v}_0 - \mathbf{v}_1 = 2\mathbf{v}_0 + 2\mathbf{v}_1$ with $2 + 2 = 4 = 1$, an affine combination of $\mathbf{v}_0, \mathbf{v}_1$, contradicting affine independence. So $\mathbf{s} \neq \mathbf{0}$, and the median from $\mathbf{v}_i$ is $\mathbf{v}_i + \operatorname{span}(\mathbf{s})$. All three medians have direction $\operatorname{span}(\mathbf{s})$, so they are parallel.

Suppose two of them coincide, say $\mathbf{v}_1 - \mathbf{v}_0 = c\mathbf{s}$ for some $c \in \mathbb{F}_3$ (by Proposition 4.7.3(c)). Then

$$(-1-c)\mathbf{v}_0 + (1-c)\mathbf{v}_1 - c\mathbf{v}_2 = \mathbf{0},$$

with coefficient sum $-3c = 0$, and the coefficients are not all 0 (that would need $c = -1$, $c = 1$, and $c = 0$). This is an affine dependence, a contradiction. So the three medians are distinct parallel lines, hence pairwise disjoint (Corollary 4.7.6). Each has $|\operatorname{span}(\mathbf{s})| = 3$ points, so together they have $9 = \left|\mathbb{F}_3^2\right|$ points (Proposition 4.1.34), and they cover $\mathbb{F}_3^2$.

Solution to Exercise 4.7.58

Uses. Theorem 4.7.39 (Affine maps are determined by a frame), Theorem 4.7.42 (Affine maps as matrices), Proposition 4.7.20 (Affine independence through differences), Corollary 3.5.25 (The invertible matrix theorem), Proposition 2.4.8 (Four views of the product)
Techniques. Homogenize: add a coordinate 1 (Technique 4.7.21)
Hint 1. In (a), $f(\mathbf{0})$ is the translation part, and $f(\mathbf{e}_j) - f(\mathbf{0})$ is column j of the linear part.
Hint 2. In (b), compose: if h sends $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2$ to $(1,1), (3,1), (1,2)$, then $g = f \circ h^{-1}$.
Hint 3. In (c), $M\hat{\mathbf{a}}_i = \hat{\mathbf{b}}_i$ for all i says $M\hat{A} = \hat{B}$. The columns of $\hat{A}$ are independent by the second approach to Proposition 4.7.20.
Sketch. (a) $f(\mathbf{x}) = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\mathbf{x} + (2,1)$. (b) $h(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} + (1,1)$, and $g = f \circ h^{-1}$ is $g(\mathbf{y}) = \begin{pmatrix} 1 & -1 \\ \frac{1}{2} & 2 \end{pmatrix}\mathbf{y} + \left(2, -\frac{3}{2}\right)$. (c) $M_g\hat{A} = \hat{B}$ column by column, and $\hat{A}$ has independent columns, so it is invertible.

(a) By Theorem 4.7.39 there is exactly one such map. Its translation part is $f(\mathbf{0}) = (2,1)$, and its linear part L has $L(\mathbf{e}_1) = (4,2) - (2,1) = (2,1)$ and $L(\mathbf{e}_2) = (1,3) - (2,1) = (-1,2)$. So

$$f(\mathbf{x}) = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad M_f = \begin{pmatrix} 2 & -1 & 2 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(b) The affine map $h(\mathbf{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} + (1,1)$ sends $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2$ to $(1,1), (3,1), (1,2)$, and it is bijective with $h^{-1}(\mathbf{y}) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}(\mathbf{y} - (1,1))$. The map $g = f \circ h^{-1}$ is affine and sends $(1,1), (3,1), (1,2)$ to $(2,1), (4,2), (1,3)$, so it is the unique such map (Theorem 4.7.39). Explicitly,

$$g(\mathbf{y}) = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\left(\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\left(\mathbf{y} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) + \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & -1 \\ \frac{1}{2} & 2 \end{pmatrix}\mathbf{y} + \begin{pmatrix} 2 \\ -\frac{3}{2} \end{pmatrix}.$$

Check: $g(1,1) = \left(0, \frac{5}{2}\right) + \left(2, -\frac{3}{2}\right) = (2,1)$, $g(3,1) = \left(2, \frac{7}{2}\right) + \left(2, -\frac{3}{2}\right) = (4,2)$, $g(1,2) = \left(-1, \frac{9}{2}\right) + \left(2, -\frac{3}{2}\right) = (1,3)$.

(c) The columns of $\hat{A}$ are the lifts $\hat{\mathbf{a}}_i$, which are linearly independent by the second approach to Proposition 4.7.20. So $\hat{A}$ is invertible (Corollary 3.5.25). Let g be the affine map with $g(\mathbf{a}_i) = \mathbf{b}_i$ (Theorem 4.7.39). By Theorem 4.7.42(a), $M_g\hat{\mathbf{a}}_i = \hat{\mathbf{b}}_i$ for each i, which says that $M_g\hat{A} = \hat{B}$ column by column (Proposition 2.4.8). Hence $M_g = \hat{B}\hat{A}^{-1}$. For (b),

$$\hat{A} = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}, \quad \hat{A}^{-1} = \begin{pmatrix} -\frac{1}{2} & -1 & \frac{5}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & -1 \end{pmatrix}, \quad \hat{B}\hat{A}^{-1} = \begin{pmatrix} 2 & 4 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}\hat{A}^{-1} = \begin{pmatrix} 1 & -1 & 2 \\ \frac{1}{2} & 2 & -\frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix},$$

in agreement with (b).

Solution to Exercise 4.7.59

Uses. Corollary 4.7.23 (Affinely independent lists are short), Definition 4.7.28 (Convex combination, convex set, convex hull)
Techniques. Slide along a dependence (Technique 4.7.33)
Hint 1. The $n+2$ points are affinely dependent. Take an affine dependence $\sum_i d_i \mathbf{v}_i = \mathbf{0}$, $\sum_i d_i = 0$.
Hint 2. Split the indices by the sign of d_i, and move the negative terms to the other side.
Sketch. Let $\sum_i d_i \mathbf{v}_i = \mathbf{0}$, $\sum_i d_i = 0$, not all $d_i = 0$. Put $I = \{i \mid d_i > 0\}$ and J the rest; both are nonempty, and $s = \sum_{i\in I} d_i = -\sum_{j\in J} d_j > 0$. Then $\sum_{i\in I}(d_i/s)\mathbf{v}_i = \sum_{j\in J}(-d_j/s)\mathbf{v}_j$ is a common point. In the examples, $(1,1)$ is the midpoint of $(2,0)$ and $(0,2)$, and lies inside the triangle $(0,0),(4,0),(0,4)$.

By Corollary 4.7.23, the $n+2$ points are affinely dependent: there are $d_1, \dots, d_{n+2}$, not all 0, with $\sum_i d_i \mathbf{v}_i = \mathbf{0}$ and $\sum_i d_i = 0$. Let $I = \{i \mid d_i > 0\}$ and $J = \{j \mid d_j \le 0\}$. Since the d_i add up to 0 and are not all 0, some d_i is positive and some is negative, so I and J are nonempty. Let $s = \sum_{i\in I} d_i > 0$; then $\sum_{j\in J} d_j = -s$. Moving the terms with $j \in J$ to the other side of the dependence and dividing by s,

$$\mathbf{p} = \sum_{i\in I} \frac{d_i}{s}\mathbf{v}_i = \sum_{j\in J} \frac{-d_j}{s}\mathbf{v}_j.$$

On the left, the coefficients are positive and add up to 1; on the right, they are nonnegative and add up to $(-(-s))/s = 1$. So $\mathbf{p}$ lies in both convex hulls.

For $(0,0),(2,0),(0,2),(1,1)$: the dependence $(2,0)+(0,2)-2(1,1) = \mathbf{0}$ (coefficients $0,1,1,-2$, sum 0) gives $I = \{2,3\}$ and $J = \{1,4\}$, with common point $(1,1) = \frac{1}{2}(2,0) + \frac{1}{2}(0,2)$, which is also the endpoint $(1,1)$ of the segment $\operatorname{conv}\{(0,0),(1,1)\}$. For $(0,0),(4,0),(0,4),(1,1)$: $2(0,0)+(4,0)+(0,4)-4(1,1) = \mathbf{0}$, with coefficients $2,1,1,-4$ adding up to 0, gives $I = \{1,2,3\}$ and $J = \{4\}$: the point $(1,1) = \frac{1}{2}(0,0) + \frac{1}{4}(4,0) + \frac{1}{4}(0,4)$ lies in the triangle.

Solution to Exercise 4.7.60

Uses. Theorem 4.7.37 (Affine maps preserve affine combinations)
Hint 1. The point $\left(\frac{1}{2},\frac{1}{2}\right)$ is the midpoint of $(0,0)$ and $(1,1)$, and also of $(1,0)$ and $(0,1)$.
Hint 2. For (b), try $h'(\mathbf{y}) = c - y_1 - y_2 + 2y_3$ and choose c.
Sketch. (a) Affine maps preserve midpoints, so $h\left(\frac{1}{2},\frac{1}{2}\right)$ is the average of two positive numbers and also of two negative numbers. (b) $h'(\mathbf{y}) = \frac{1}{2} - y_1 - y_2 + 2y_3$ gives $\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}$.

(a) Suppose such an affine h exists. The point $\mathbf{m} = \left(\frac{1}{2},\frac{1}{2}\right)$ is the affine combination $\frac{1}{2}(0,0) + \frac{1}{2}(1,1)$, and also $\frac{1}{2}(1,0) + \frac{1}{2}(0,1)$. By Theorem 4.7.37(II),

$$h(\mathbf{m}) = \frac{1}{2}h(0,0) + \frac{1}{2}h(1,1) > 0 \quad \text{and} \quad h(\mathbf{m}) = \frac{1}{2}h(1,0) + \frac{1}{2}h(0,1) < 0,$$

a contradiction. In words, the segment between the two positive points crosses the segment between the two negative points, and at the crossing point an affine function would have to be both positive and negative.

(b) The four points become $\varphi(0,0) = (0,0,0)$, $\varphi(1,1) = (1,1,1)$, $\varphi(1,0) = (1,0,0)$, and $\varphi(0,1) = (0,1,0)$. Let $h'(\mathbf{y}) = \frac{1}{2} - y_1 - y_2 + 2y_3$. Then h' takes the values $\frac{1}{2}, \frac{1}{2} - 2 + 2 = \frac{1}{2}, -\frac{1}{2}$, and $-\frac{1}{2}$ at these points, as required. The midpoint argument of (a) no longer applies, because φ is not affine: the midpoint of $\varphi(0,0)$ and $\varphi(1,1)$ is $\left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$, while that of $\varphi(1,0)$ and $\varphi(0,1)$ is $\left(\frac{1}{2},\frac{1}{2},0\right)$.

Solution to Exercise 4.7.61

Uses. Theorem 4.7.14 (Affine subspaces are the sets closed under affine combinations)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. Over $\mathbb{F}_2$, an affine combination of $\mathbf{s}_1, \dots, \mathbf{s}_k$ is the sum of the $\mathbf{s}_i$ whose coefficient is 1, and there is an odd number of them.

Hint 2. Show by induction that the sum of any odd number of elements of S lies in S: replace the first three summands by their sum.
Sketch. ($\Rightarrow$) $\mathbf{u}+\mathbf{v}+\mathbf{w}$ is an affine combination, as $1+1+1=1$ in $\mathbb{F}_2$. ($\Leftarrow$) An affine combination over $\mathbb{F}_2$ is a sum of an odd number $2j+1$ of elements of S; replacing three of them by their sum, which is in S, lowers j, so by induction the sum lies in S. Then Theorem 4.7.14 applies. The first set is the subspace $x_1+x_2+x_3=0$; the second is the plane $x_1+x_2+x_3=1$, and both pass the test.

($\Rightarrow$) If S is affine, then $\mathbf{u}+\mathbf{v}+\mathbf{w}=1\mathbf{u}+1\mathbf{v}+1\mathbf{w}$ is an affine combination, since $1+1+1=1$ in $\mathbb{F}_2$, so it lies in S by Theorem 4.7.14.

($\Leftarrow$) Every coefficient in $\mathbb{F}_2$ is 0 or 1, so an affine combination $\sum_i c_i\mathbf{s}_i$ of elements of S is the sum of those $\mathbf{s}_i$ with $c_i=1$, and their number is odd, because the coefficients add up to 1 in $\mathbb{F}_2$. We show by induction on $j\geq 0$ that a sum of $2j+1$ elements of S (repetitions allowed) lies in S. For $j=0$ it is a single element of S. For $j\geq 1$, the sum $\mathbf{s}_1+\mathbf{s}_2+\mathbf{s}_3+\mathbf{s}_4+\cdots+\mathbf{s}_{2j+1}$ equals $\mathbf{s}'+\mathbf{s}_4+\cdots+\mathbf{s}_{2j+1}$ with $\mathbf{s}'=\mathbf{s}_1+\mathbf{s}_2+\mathbf{s}_3\in S$ by hypothesis, a sum of $2(j-1)+1$ elements of S, which lies in S by induction. So S is closed under affine combinations, and it is affine by Theorem 4.7.14.

The first set consists of the vectors with $x_1+x_2+x_3=0$, and each of its triples sums to a vector with the same property, so it passes the test; it is in fact a subspace. The second consists of the vectors with $x_1+x_2+x_3=1$, and the sum of three such vectors has coordinate sum $1+1+1=1$, so it passes too: it is the affine plane $(1,0,0)+\{\mathbf{x}\mid x_1+x_2+x_3=0\}$. (It is the plane of Exercise 4.7.24(c).)

Solution to Exercise 4.7.62

Uses. Theorem 4.7.42 (Affine maps as matrices), Example 4.7.36 (Affine maps), Proposition 2.5.28 (Inverse of a 2×2 matrix), Proposition 3.1.19 (Solution sets are translates of the kernel)
Hint 1. The rotation about $\mathbf{a}$ has matrix $\begin{pmatrix} R & \mathbf{a}-R\mathbf{a} \\ 0 & 1\end{pmatrix}$. Multiply in the right order.
Hint 2. A fixed point solves $(I_2-R_\theta)\mathbf{c}=\mathbf{b}$. Compute $ad-bc$ for I_2-R_θ.
Hint 3. If $\mathbf{c}=R_\theta\mathbf{c}+\mathbf{b}$, then $\mathbf{b}=\mathbf{c}-R_\theta\mathbf{c}$; substitute into f.
Sketch. (a) $\begin{pmatrix}0&-1&1\\1&0&-1\\0&0&1\end{pmatrix}\begin{pmatrix}0&-1&0\\1&0&0\\0&0&1\end{pmatrix}=\begin{pmatrix}-1&0&1\\0&-1&-1\\0&0&1\end{pmatrix}$, the map $\mathbf{v}\mapsto-\mathbf{v}+(1,-1)$, with fixed point $\left(\frac{1}{2},-\frac{1}{2}\right)$; it is $\mathbf{v}\mapsto\mathbf{c}-(\mathbf{v}-\mathbf{c})$.
(b) I_2-R_θ has $ad-bc=(1-\cos\theta)^2+\sin^2\theta=2-2\cos\theta\neq 0$, so the fixed point exists and is unique, and $f(\mathbf{v})=R_\theta\mathbf{v}+\mathbf{c}-R_\theta\mathbf{c}=\mathbf{c}+R_\theta(\mathbf{v}-\mathbf{c})$.

(a) Let $R=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$. The rotation about $(0,0)$ has matrix $M_1=\begin{pmatrix}R&\mathbf{0}\\0&1\end{pmatrix}$. The rotation about $\mathbf{a}=(1,0)$ is $\mathbf{v}\mapsto R\mathbf{v}+(\mathbf{a}-R\mathbf{a})$ (Example 4.7.36), with $\mathbf{a}-R\mathbf{a}=(1,0)-(0,1)=(1,-1)$. The composite applies M_1 first, so by Theorem 4.7.42(c) its matrix is

$$\begin{pmatrix}0&-1&1\\1&0&-1\\0&0&1\end{pmatrix}\begin{pmatrix}0&-1&0\\1&0&0\\0&0&1\end{pmatrix}=\begin{pmatrix}-1&0&1\\0&-1&-1\\0&0&1\end{pmatrix},$$

the map $g(\mathbf{v})=-\mathbf{v}+(1,-1)$. Its fixed points solve $2\mathbf{v}=(1,-1)$, so $\mathbf{c}=\left(\frac{1}{2},-\frac{1}{2}\right)$, and $g(\mathbf{v})=-\mathbf{v}+2\mathbf{c}=\mathbf{c}-(\mathbf{v}-\mathbf{c})$. This is the rotation by 180° about $\mathbf{c}$, since $R_\pi=-I_2$.

(b) A fixed point is a solution of $R_\theta\mathbf{c}+\mathbf{b}=\mathbf{c}$, that is, of $(I_2-R_\theta)\mathbf{c}=\mathbf{b}$, where

$$I_2-R_\theta=\begin{pmatrix}1-\cos\theta&\sin\theta\\-\sin\theta&1-\cos\theta\end{pmatrix}.$$

Here $ad-bc=(1-\cos\theta)^2+\sin^2\theta=2-2\cos\theta$, which is nonzero because $\cos\theta\neq 1$. By Proposition 2.5.28, I_2-R_θ is invertible, so the fixed point $\mathbf{c}=(I_2-R_\theta)^{-1}\mathbf{b}$ exists and is unique. Then $\mathbf{b}=\mathbf{c}-R_\theta\mathbf{c}$, and

$$f(\mathbf{v})=R_\theta\mathbf{v}+\mathbf{c}-R_\theta\mathbf{c}=\mathbf{c}+R_\theta(\mathbf{v}-\mathbf{c}),$$

the rotation by θ about $\mathbf{c}$. Part (a) is the case $\theta=\pi$, $\mathbf{b}=(1,-1)$.

Solution to Exercise 4.7.63

Uses. Proposition 4.7.3 (Direction and base points), Corollary 4.7.6 (Parallel affine subspaces are equal or disjoint), Proposition 4.1.34 (Size of a vector space over a finite field), Proposition 4.1.37 (Counting subspaces)
Techniques. Double counting (Technique 4.1.36)
Hint 1. Group the affine subspaces by their direction. How many translates does a k-dimensional subspace U have?
Hint 2. The translates of U are disjoint and cover F^n, and each has $|U| = q^k$ elements.
Sketch. Each k-dimensional affine subspace has exactly one direction, a k-dimensional subspace U. The distinct translates of U partition F^n (Corollary 4.7.6), and each has q^k elements, so there are $q^n/q^k = q^{n-k}$ of them. (b) $\mathbb{F}_3^2$ has 4 one-dimensional subspaces, so $4 \cdot 3 = 12$ lines; $\mathbb{F}_2^3$ has 7 two-dimensional subspaces, so $7 \cdot 2 = 14$ planes.

(a) By Proposition 4.7.3(a), each affine subspace has exactly one direction, so the k-dimensional affine subspaces are grouped by their directions, which are the k-dimensional subspaces U of F^n. Fix such a U. By Corollary 4.7.6, the distinct translates of U are pairwise disjoint and every vector lies in one of them. Each translate $\mathbf{v} + U$ is in bijection with U via $\mathbf{u} \mapsto \mathbf{v} + \mathbf{u}$, so it has $|U| = q^k$ elements (Proposition 4.1.34). Since $|F^n| = q^n$, there are exactly $q^n/q^k = q^{n-k}$ translates of U. Summing over U gives the claim.

(b) By Proposition 4.1.37, $\mathbb{F}_3^2$ has $(9-1)/(3-1) = 4$ one-dimensional subspaces, so it has $4 \cdot 3^{2-1} = 12$ lines. And $\mathbb{F}_2^3$ has $((8-1)(8-2))/((4-1)(4-2)) = 42/6 = 7$ two-dimensional subspaces, so it has $7 \cdot 2^{3-2} = 14$ planes.

Solution to Exercise 4.7.64

Uses. Exercise 2.1.39, Definition 4.7.35 (Affine map)
Hint 1. First extend the identity to all $t \in \mathbb{R}$. For $t > 1$, the point $\mathbf{v}$ lies on the segment from $\mathbf{u}$ to $\mathbf{w} = (1-t)\mathbf{u} + t\mathbf{v}$.
Hint 2. For $t < 0$, swap the roles of $\mathbf{u}$ and $\mathbf{v}$.
Hint 3. Now $g = f - f(\mathbf{0})$ respects all affine combinations of two points and fixes $\mathbf{0}$. Use Exercise 2.1.39(a).
Sketch. For $t > 1$, $\mathbf{v} = (1-1/t)\mathbf{u} + (1/t)\mathbf{w}$ with $1/t \in (0,1)$, so $f(\mathbf{v}) = (1-1/t)f(\mathbf{u}) + (1/t)f(\mathbf{w})$, which solves to $f(\mathbf{w}) = (1-t)f(\mathbf{u}) + tf(\mathbf{v})$. For $t < 0$, write the point as $(1-t')\mathbf{v} + t'\mathbf{u}$ with $t' = 1 - t > 1$. So $g = f - f(\mathbf{0})$ satisfies $g(a\mathbf{u} + b\mathbf{v}) = ag(\mathbf{u}) + bg(\mathbf{v})$ whenever $a + b = 1$, and $g(\mathbf{0}) = \mathbf{0}$; since $1 + 1 \neq 0$ in $\mathbb{R}$, g is linear.

We first show that the identity holds for all $t \in \mathbb{R}$. Let $t > 1$ and $\mathbf{w} = (1-t)\mathbf{u} + t\mathbf{v}$. Then

$$\left(1 - \frac{1}{t}\right)\mathbf{u} + \frac{1}{t}\mathbf{w} = \left(1 - \frac{1}{t} + \frac{1-t}{t}\right)\mathbf{u} + \mathbf{v} = \mathbf{v},$$

and $1/t \in (0,1)$, so the hypothesis gives $f(\mathbf{v}) = \left(1 - \frac{1}{t}\right)f(\mathbf{u}) + \frac{1}{t}f(\mathbf{w})$. Multiplying by t and solving,

$$f(\mathbf{w}) = tf(\mathbf{v}) - (t-1)f(\mathbf{u}) = (1-t)f(\mathbf{u}) + tf(\mathbf{v}).$$

If $t < 0$, then $(1-t)\mathbf{u} + t\mathbf{v} = (1-t')\mathbf{v} + t'\mathbf{u}$ with $t' = 1 - t > 1$, and the case just proved, with $\mathbf{u}$ and $\mathbf{v}$ exchanged, gives $f((1-t')\mathbf{v} + t'\mathbf{u}) = (1-t')f(\mathbf{v}) + t'f(\mathbf{u}) = (1-t)f(\mathbf{u}) + tf(\mathbf{v})$.

Now let $g(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$. Then $g(\mathbf{0}) = \mathbf{0}$, and for $a + b = 1$,

$$g(a\mathbf{u} + b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v}) - (a+b)f(\mathbf{0}) = ag(\mathbf{u}) + bg(\mathbf{v}).$$

Since $1 + 1 \neq 0$ in $\mathbb{R}$, Exercise 2.1.39(a) shows that g is linear. Hence $f = g + f(\mathbf{0})$ is affine.

Solution to Exercise 4.7.65

Uses. Proposition 4.7.32 (Barycentric coordinates on a simplex), Theorem 4.7.22 (Characterizations of affine independence)
Hint 1. Work in barycentric coordinates (Proposition 4.7.32). A vertex $\mathbf{v}_0$ has coordinates $(1, 0, \ldots, 0)$.
Hint 2. If $\mathbf{v}_0 = (1-t)\mathbf{a} + t\mathbf{b}$, compare coordinates $i \geq 1$: a sum of two nonnegative numbers is 0.
Hint 3. If $\mathbf{p}$ is not a vertex, two of its coordinates are positive, say $t_0, t_1 > 0$. Move $\mathbf{p}$ a little in the direction $\pm(\mathbf{v}_0 - \mathbf{v}_1)$.
Sketch. If $\mathbf{v}_0 = (1-t)\mathbf{a} + t\mathbf{b}$, the barycentric coordinates satisfy $0 = (1-t)\alpha_i + t\beta_i$ for $i \geq 1$ with both terms nonnegative, so $\alpha_i = \beta_i = 0$ and $\mathbf{a} = \mathbf{b} = \mathbf{v}_0$. If $\mathbf{p}$ has $t_0, t_1 > 0$, then with $\delta = \min(t_0, t_1)$, the points $\mathbf{p} \pm \delta(\mathbf{v}_0 - \mathbf{v}_1)$ lie in the simplex, differ, and have midpoint $\mathbf{p}$.

Let σ have vertices $\mathbf{v}_0, \ldots, \mathbf{v}_k$, and write points of σ in barycentric coordinates (Proposition 4.7.32). The vertex $\mathbf{v}_j$ has coordinates $\mathbf{e}_j$, with 1 in place j and 0 elsewhere.

Vertices are extreme. Suppose $\mathbf{v}_j = (1-t)\mathbf{a} + t\mathbf{b}$ with $\mathbf{a}, \mathbf{b} \in \sigma$ and $t \in (0,1)$, and let $\mathbf{a}, \mathbf{b}$ have barycentric coordinates (α_i) and (β_i). Then $\sum_i ((1-t)\alpha_i + t\beta_i)\mathbf{v}_i$ is an affine combination representing $\mathbf{v}_j$, so by unique-

ness (Theorem 4.7.22(IV)), $(1-t)\alpha_i + t\beta_i = 0$ for $i \neq j$. Both terms are nonnegative and $1-t, t > 0$, so $\alpha_i = \beta_i = 0$ for $i \neq j$, and then $\alpha_j = \beta_j = 1$. Hence $\mathbf{a} = \mathbf{b} = \mathbf{v}_j$.

Other points are not extreme. Let $\mathbf{p} \in \sigma$ have coordinates (t_i) and not be a vertex. The t_i are nonnegative and add up to 1, and they are not a standard basis vector, so at least two of them are positive, say $t_r, t_s > 0$ with $r \neq s$. Let $\delta = \min(t_r, t_s) > 0$, and $\mathbf{a} = \mathbf{p} + \delta(\mathbf{v}_r - \mathbf{v}_s)$, $\mathbf{b} = \mathbf{p} - \delta(\mathbf{v}_r - \mathbf{v}_s)$. Their coordinates are those of $\mathbf{p}$ with t_r, t_s replaced by $t_r + \delta, t_s - \delta$ and by $t_r - \delta, t_s + \delta$; these are nonnegative and add up to 1, so $\mathbf{a}, \mathbf{b} \in \sigma$. They differ, since $\mathbf{v}_r \neq \mathbf{v}_s$, and $\mathbf{p} = \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b}$. So $\mathbf{p}$ is not extreme.

The extreme points are defined in terms of the set σ alone, so σ determines its vertex set. In particular, a k-simplex is not a k'-simplex for $k' \neq k$.

Section 4.8

Solution to Exercise 4.8.4

Uses. Proposition 2.4.36 (Products of upper triangular matrices), Lemma 2.4.24 (Matrix units), Theorem 4.1.10 (Independent lists are short), Theorem 4.3.3 (Rank of a product), Proposition 4.2.26 (Subspaces that must meet), Exercise 4.3.30, Proposition 4.5.27 (Similarity invariants), Proposition 2.4.27 (The trace is linear)
Techniques. Too many vectors force a dependence (Technique 4.1.30), Dimension counting (Technique 4.2.24), Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. (a) In which subspace of $F^{2\times 2}$ do the powers of A lie, and what is its dimension? (b) Compare ranks. (c) Use Proposition 4.2.26.
Hint 2. (d) Iterate Sylvester's inequality. (e) Traces differ by 2. What happens when $2 = 0$?
Sketch. (a) True: I_2, A, A^2, A^3 lie in the 3-dimensional space of upper triangular matrices (Technique 4.1.30). (b) False: $\operatorname{rank}(BA) \leq \operatorname{rank} A \leq 2$. (c) True: $\dim(U \cap W) \geq 4 + 4 - 7 = 1$. (d) True: $0 = \operatorname{rank} T^3 \geq 3 \operatorname{rank} T - 10$; the bound is attained by a shift of size 3 next to one of size 2. (e) True exactly when $1 + 1 \neq 0$ in F: then $\operatorname{tr}(A + I_2) = \operatorname{tr} A + 2 \neq \operatorname{tr} A$. When $1 + 1 = 0$, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ are similar.

(a) True. The upper triangular matrices form a subspace of $F^{2\times 2}$ of dimension 3, with basis the matrix units U_{11}, U_{12}, U_{22} (Lemma 2.4.24), and products of upper triangular matrices are upper triangular (Proposition 2.4.36). So I_2, A, A^2, A^3 are four vectors in a 3-dimensional space, and they are dependent (Theorem 4.1.10(a)). The move is Technique 4.1.30 in the smallest space available; in all of $F^{2\times 2}$ it would only give a relation among $I_2, A, \ldots, A^4$.

(b) False. By Theorem 4.3.3, $\operatorname{rank}(BA) \leq \operatorname{rank} A \leq 2$, while $\operatorname{rank} I_3 = 3$.

(c) True. By Proposition 4.2.26, $\dim(U \cap W) \geq 4 + 4 - 7 = 1$.

(d) True. Let A be the standard matrix of T, so $A^3 = 0$. By Exercise 4.3.30(c) with $A_1 = A_2 = A_3 = A$,

$$0 = \operatorname{rank} A^3 \geq 3 \operatorname{rank} A - 2 \cdot 5,$$

so $\operatorname{rank} A \leq 10/3$, and $\operatorname{rank} A \leq 3$. The bound is attained: the operator that sends $\mathbf{e}_3 \mapsto \mathbf{e}_2 \mapsto \mathbf{e}_1 \mapsto \mathbf{0}$ and $\mathbf{e}_5 \mapsto \mathbf{e}_4 \mapsto \mathbf{0}$ has rank 3 and cube 0.

(e) True exactly when $1 + 1 \neq 0$ in F, as over $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{F}_p$ for odd p. Similar matrices have the same trace (Proposition 4.5.27(b)), and $\operatorname{tr}(A + I_2) = \operatorname{tr} A + 2$ (Proposition 2.4.27). If $2 \neq 0$, the traces differ, so $A \not\sim A + I_2$. If $2 = 0$, as over $\mathbb{F}_2$, the traces agree, and for $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ we get $A + I_2 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, which is $S^{-1}AS$ for the swap $S = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = S^{-1}$. The move, comparing an invariant, can only prove that two matrices are not similar; to prove similarity, a matrix S has to be found (Technique 4.5.29).

Solution to Exercise 4.8.5

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Proposition 1.5.15 (Short lists, zeros, and repeats), Theorem 4.1.10 (Independent lists are short)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Why should $T^{k+1}\mathbf{v}$ be $\mathbf{0}$?

Hint 2. Try $T = \mathrm{id}_V$.
Sketch. The step "$T^{k+j}\mathbf{v} = \mathbf{0}$ for $j \geq 1$" needs $T^{k+1}\mathbf{v} = \mathbf{0}$, which is not assumed. For $T = \mathrm{id}_V$, $\mathbf{v} \neq \mathbf{0}$, and $k = 1$, the list $\mathbf{v}, \mathbf{v}$ is dependent. With the missing hypothesis the argument is the proof of Proposition 4.1.26.

The error. The argument uses $T^{k+j}\mathbf{v} = \mathbf{0}$ for $j \geq 1$, which follows from $T^{k+1}\mathbf{v} = \mathbf{0}$ but not from anything in the hypotheses. Only $T^k\mathbf{v} \neq \mathbf{0}$ is assumed, and that says nothing about later powers.

The claim is false. Let $T = \mathrm{id}_V$ and $\mathbf{v} \neq \mathbf{0}$, and $k = 1$. Then $T\mathbf{v} = \mathbf{v} \neq \mathbf{0}$, but the list $\mathbf{v}, T\mathbf{v} = \mathbf{v}, \mathbf{v}$ has a repeated vector and is dependent (Proposition 1.5.15). A less trivial example is the rotation $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ on $\mathbb{R}^2$, with $\mathbf{v} = \mathbf{e}_1$ and $k = 2$: $R^2\mathbf{e}_1 = -\mathbf{e}_1 \neq \mathbf{0}$, and $\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1$ is dependent. In fact, whenever $k \geq \dim V$, the list has more than $\dim V$ entries and is dependent (Theorem 4.1.10(a)), whatever T is.

With the hypothesis $T^{k+1}\mathbf{v} = \mathbf{0}$ added, every step of the argument is correct, and it is the proof of Proposition 4.1.26.

Solution to Exercise 4.8.6

Uses. Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Example 2.6.4 (Injective, surjective, both, neither), Exercise 4.5.54, Corollary 4.2.18 (Injective if and only if surjective)
Techniques. Dimension counting (Technique 4.2.24), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Which results does the argument use, and what are their hypotheses?
Hint 2. Look at multiplication by x on $F[x]$.
Sketch. Rank–nullity and Proposition 4.1.21(b) assume V finite-dimensional; for infinite-dimensional V the equation $\dim \operatorname{im} T = \dim V - 0$ has no meaning as a computation with numbers. Multiplication by x on $F[x]$ is injective, and 1 is not in its image. For finite-dimensional V the argument is correct (Corollary 4.2.18).

The error. Both steps need V to be finite-dimensional. Rank–nullity (Theorem 4.2.11(b)) is stated for finite-dimensional domains, and "a subspace of the same dimension is everything" (Proposition 4.1.21(b)) is stated for finite-dimensional V. If V is infinite-dimensional, "$\dim V$" is not a number, and subtracting $\dim \ker T = 0$ from it carries no information. Even the conclusion that $\operatorname{im} T$ and V are isomorphic, which is true for an injective T, does not make them equal: $F[x]$ is isomorphic to its proper subspace $xF[x]$ (Exercise 4.5.54).

The claim is false. Let $M \in \mathrm{L}(F[x])$ be multiplication by x. If $xp = \mathbf{0}$, then $p = \mathbf{0}$, so M is injective. Every value xp has constant coefficient 0, so the constant polynomial 1 is not a value, and M is not surjective (Example 2.6.4).

For finite-dimensional V the argument is correct, and it is the proof of Corollary 4.2.18.

Solution to Exercise 4.8.7

Uses. Exercise 4.2.44, Proposition 4.2.26 (Subspaces that must meet), Proposition 4.1.21 (Dimensions of subspaces), Theorem 4.2.11 (Rank–nullity), Proposition 4.2.3 (Properties of restriction), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 4.4.2 (Sums are the smallest containing subspaces), Theorem 4.3.8 (Row rank equals column rank), Proposition 1.5.14 (Appending a vector)
Techniques. Restrict to a subspace (Technique 4.2.5), Dimension counting (Technique 4.2.24)
Hint 1. (a) $\operatorname{rank}(T\restriction_H) = \dim H - \dim(H \cap \ker T)$. How small can $H \cap \ker T$ be inside $\ker T$?
Hint 2. (b) $A'\mathbf{x}' = A\mathbf{x}$, where $\mathbf{x}$ is $\mathbf{x}'$ with a 0 inserted in position j. So A' is T_A restricted to the hyperplane $x_j = 0$. Directly: $\operatorname{im} A = \operatorname{im} A' + \operatorname{span}(\mathbf{a}_j)$.
Sketch. (a) $\operatorname{rank}(T\restriction_H) = (n-1) - \dim(H \cap \ker T)$, and $\dim \ker T - 1 \leq \dim(H \cap \ker T) \leq \dim \ker T$ by Proposition 4.2.26; the upper value is attained exactly when $\ker T \subseteq H$. (b) $\operatorname{im} A' = \operatorname{im}\left(T_A\restriction_{H_j}\right)$ with $H_j = \{\mathbf{x} \mid x_j = 0\}$, and $\ker A \subseteq H_j$ means no solution of $A\mathbf{x} = \mathbf{0}$ has $x_j \neq 0$, which means that $\mathbf{a}_j$ is not a combination of the others. Directly, $\operatorname{im} A = \operatorname{im} A' + \operatorname{span}(\mathbf{a}_j)$. (c) One column at a time, and rows by transposing.

(a) By Exercise 4.2.44(a), $\operatorname{rank}(T\restriction_H) = (n-1) - \dim(H \cap \ker T)$. Since $H \cap \ker T \subseteq \ker T$, its dimension is at most $\dim \ker T$, with equality if and only if $H \cap \ker T = \ker T$, that is, $\ker T \subseteq H$ (Proposition 4.1.21). By Proposition 4.2.26, $\dim(H \cap \ker T) \geq (n-1) + \dim \ker T - n = \dim \ker T - 1$. So $\dim(H \cap \ker T)$ is $\dim \ker T$ or $\dim \ker T - 1$, and with $\operatorname{rank} T = n - \dim \ker T$ (Theorem 4.2.11(b)),

$$\operatorname{rank}(T{\restriction_H}) = \begin{cases} \operatorname{rank} T - 1 & \text{if } \ker T \subseteq H \\ \operatorname{rank} T & \text{otherwise.} \end{cases}$$

(b) *From (a).* Let $H_j = \{\mathbf{x} \in F^n \mid x_j = 0\}$, a subspace of dimension $n-1$. For $\mathbf{x}' \in F^{n-1}$, let $\mathbf{x} \in H_j$ be $\mathbf{x}'$ with a 0 inserted in position j; this is a bijection $F^{n-1} \to H_j$, and $A'\mathbf{x}' = A\mathbf{x}$. So $\operatorname{im} A' = \operatorname{im}\left(T_A{\restriction_{H_j}}\right)$, and by (a) the rank drops if and only if $\ker A \subseteq H_j$, that is, if and only if every solution of $A\mathbf{x} = \mathbf{0}$ has $x_j = 0$. If $\mathbf{a}_j = \sum_{i \neq j} c_i \mathbf{a}_i$, then $\mathbf{x} = \mathbf{e}_j - \sum_{i \neq j} c_i \mathbf{e}_i$ solves $A\mathbf{x} = \mathbf{0}$ with $x_j = 1$. Conversely, if $A\mathbf{x} = \mathbf{0}$ with $x_j \neq 0$, then $\mathbf{a}_j = -x_j^{-1} \sum_{i \neq j} x_i \mathbf{a}_i$. So $\ker A \subseteq H_j$ exactly when $\mathbf{a}_j$ is not a combination of the other columns.

Directly. By Proposition 2.6.18, $\operatorname{im} A'$ is spanned by the columns other than $\mathbf{a}_j$, and $\operatorname{im} A = \operatorname{im} A' + \operatorname{span}(\mathbf{a}_j)$ (Proposition 4.4.2). If $\mathbf{a}_j \in \operatorname{im} A'$, the two images are equal. Otherwise a basis of $\operatorname{im} A'$ followed by $\mathbf{a}_j$ is independent (Proposition 1.5.14) and spans $\operatorname{im} A$, so $\operatorname{rank} A = \operatorname{rank} A' + 1$. The direct proof is shorter here, because deleting a column has an obvious effect on the column space; the restriction proof is the one that also applies to an arbitrary hyperplane, where there are no columns to delete.

(c) By (b), deleting one column lowers the rank by at most 1. Deleting a row of A deletes a column of A^{T}, and $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$ (Theorem 4.3.8), so it also lowers the rank by at most 1. Deleting k rows and l columns one at a time, each time from a matrix with at least two rows or two columns respectively, lowers the rank by at most $k + l$.

Solution to Exercise 4.8.8

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Proposition 1.5.15 (Short lists, zeros, and repeats), Proposition 1.2.35 (Zero products), Proposition 4.1.34 (Size of a vector space over a finite field)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Dimension counting (Technique 4.2.24)
Hint 1. Compute $T^k\mathbf{w}$ and $T^{k+1}\mathbf{w}$. Only one term of $\mathbf{w}$ survives T^k.
Hint 2. If $a_0 \neq 0$, the chain of $\mathbf{w}$ is independent by Proposition 4.1.26, and it has the right length. If $a_0 = 0$, what is $T^k\mathbf{w}$?
Sketch. (a) $T^j\mathbf{w} \in W$, $T^k\mathbf{w} = a_0 T^k\mathbf{v}$, and $T^{k+1}\mathbf{w} = \mathbf{0}$. If $a_0 \neq 0$, the chain of $\mathbf{w}$ is independent of length $k+1 = \dim W$, so it is a basis. If $a_0 = 0$, its last vector is $\mathbf{0}$. (b) The coordinates $(a_0, \ldots, a_k)$ of $\mathbf{w}$ are unique, with $a_0 \neq 0$: $(q-1)q^k$ vectors.

(a) By Proposition 4.1.26, $\mathbf{v}, \ldots, T^k\mathbf{v}$ is a basis of W, so $\dim W = k+1$. For $j \geq 0$,

$$T^j\mathbf{w} = \sum_{i=0}^{k} a_i T^{i+j}\mathbf{v} \in W,$$

and since $T^l\mathbf{v} = \mathbf{0}$ for $l \geq k+1$, only the terms with $i + j \leq k$ survive. In particular $T^k\mathbf{w} = a_0 T^k\mathbf{v}$ and $T^{k+1}\mathbf{w} = \mathbf{0}$.

If $a_0 \neq 0$, then $T^k\mathbf{w} \neq \mathbf{0}$ (Proposition 1.2.35), so $\mathbf{w}, T\mathbf{w}, \ldots, T^k\mathbf{w}$ is independent by Proposition 4.1.26. It is a list of $k+1 = \dim W$ vectors in W, so it is a basis of W (Theorem 4.1.10(b)). If $a_0 = 0$, then $T^k\mathbf{w} = \mathbf{0}$ is an entry of the list, which is therefore dependent (Proposition 1.5.15) and not a basis.

(b) If $\operatorname{span}\left(\mathbf{w}, \ldots, T^k\mathbf{w}\right) = W$, then this spanning list of length $\dim W$ is a basis (Corollary 4.1.11(b)), so by (a) the condition is again $a_0 \neq 0$. The vectors of W correspond bijectively to their coordinate vectors $(a_0, \ldots, a_k) \in F^{k+1}$, and those with $a_0 \neq 0$ number $(q-1)q^k$. So exactly $(q-1)q^k$ of the q^{k+1} vectors of W (Proposition 4.1.34) generate it, that is, all except those in the hyperplane $\operatorname{span}\left(T\mathbf{v}, \ldots, T^k\mathbf{v}\right) = \operatorname{im}(T{\restriction_W})$.

Solution to Exercise 4.8.9

Uses. Theorem 4.1.10 (Independent lists are short), Proposition 2.4.14 (Algebra of matrices), Corollary 4.2.18 (Injective if and only if surjective), Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 2.4.36 (Products of upper triangular matrices), Exercise 4.1.56
Techniques. Too many vectors force a dependence (Technique 4.1.30), Dimension counting (Technique 4.2.24)
Hint 1. (a) The $d+1$ matrices $I_n, A, \ldots, A^d$ lie in S. In a relation among them, look at the smallest index with a nonzero coefficient, and multiply by a power of A^{-1}.
Hint 2. (b) $X \mapsto AX$ maps S into S and is injective. Which $X \in S$ does surjectivity produce?

Sketch. (a) $\sum_{i\geq j} c_i A^i = 0$ with $c_j \neq 0$; multiplying by A^{-j} gives $c_j I_n = -A\big(c_{j+1}I_n + \cdots + c_d A^{d-j-1}\big)$, so A^{-1} is a polynomial in A, hence in S. (b) $L(X) = AX$ is injective on S, hence surjective, so $AX = I_n$ for some $X \in S$, and $X = A^{-1}$. (c) Upper triangular matrices form such an S.

By induction, $A^i \in S$ for every $i \geq 0$: $A^0 = I_n \in S$, and $A^{i+1} = A^i A$ is a product of two elements of S.

(a) The $d+1$ matrices $I_n, A, \ldots, A^d$ lie in the d-dimensional space S, so they satisfy a relation $\sum_{i=0}^{d} c_i A^i = 0$ with not all $c_i = 0$ (Theorem 4.1.10(a)). Let j be the smallest index with $c_j \neq 0$. Multiplying $\sum_{i=j}^{d} c_i A^i = 0$ by $\big(A^{-1}\big)^j$ gives

$$c_j I_n + c_{j+1} A + \cdots + c_d A^{d-j} = 0.$$

If $j = d$ this says $c_d I_n = 0$, which is impossible since $n \geq 1$. So $j < d$, and

$$A\big(c_{j+1}I_n + c_{j+2}A + \cdots + c_d A^{d-j-1}\big) = -c_j I_n.$$

The matrix $B = -c_j^{-1}\big(c_{j+1}I_n + \cdots + c_d A^{d-j-1}\big)$ lies in S and satisfies $AB = I_n$, so $B = A^{-1}$ (multiply by A^{-1} on the left). This is the argument of Exercise 4.1.56, run inside S.

(b) Let $L : S \to S$, $L(X) = AX$. It maps S into S by hypothesis, and it is linear by Proposition 2.4.14. It is injective: if $AX = 0$, then $X = A^{-1}AX = 0$. Since S is finite-dimensional, L is surjective (Corollary 4.2.18), so $AX = I_n$ for some $X \in S$, and then $X = A^{-1}$ (Theorem 3.5.23).

The second proof is shorter and gives no formula; the first expresses A^{-1} as a polynomial in A of degree less than $\dim S$, and so shows that A^{-1} lies in every such S that contains A.

(c) The upper triangular matrices form a subspace of $F^{n\times n}$ that contains I_n and is closed under products (Proposition 2.4.36). By (a) or (b), it contains the inverse of each of its invertible elements.

Solution to Exercise 4.8.10

Uses. Proposition 4.3.5 (Rank of a product, exactly), Theorem 4.2.11 (Rank–nullity), Proposition 4.1.21 (Dimensions of subspaces), Proposition 2.6.15 (Kernels and images of composites)
Techniques. Restrict to a subspace (Technique 4.2.5)
Hint 1. Write both sides with (4.3.1): $\operatorname{rank}(AB) = \operatorname{rank} B - \dim(\operatorname{im} B \cap \ker A)$ and $\operatorname{rank} A + \operatorname{rank} B - n = \operatorname{rank} B - \dim \ker A$.
Hint 2. When does a subspace of $\ker A$ have the dimension of $\ker A$?
Sketch. (a) The difference of the two sides is $\dim \ker A - \dim(\operatorname{im} B \cap \ker A)$, which is 0 exactly when $\operatorname{im} B \cap \ker A = \ker A$. (b) $AB = 0$ means $\operatorname{im} B \subseteq \ker A$, so $\operatorname{rank} B \leq n - \operatorname{rank} A$, with equality exactly when the containment is an equality. (c) (a) with $A = B = N$; for the shift with $n \geq 2$, $\ker N = \operatorname{span}(\mathbf{e}_1) \subseteq \operatorname{im} N$ and $\operatorname{rank} N^2 = n - 2$.

(a) By (4.3.1) and Theorem 4.2.11(a),

$$\operatorname{rank}(AB) - (\operatorname{rank} A + \operatorname{rank} B - n) = \dim \ker A - \dim(\operatorname{im} B \cap \ker A).$$

Since $\operatorname{im} B \cap \ker A \subseteq \ker A$, this is 0 if and only if $\operatorname{im} B \cap \ker A = \ker A$ (Proposition 4.1.21(b)), that is, if and only if $\ker A \subseteq \operatorname{im} B$.

(b) $AB = 0$ means $A(B\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$, that is, $\operatorname{im} B \subseteq \ker A$. So $\operatorname{rank} B \leq \dim \ker A = n - \operatorname{rank} A$, with equality if and only if $\operatorname{im} B = \ker A$ (Proposition 4.1.21(b)). (This also follows from (a), since $\operatorname{rank}(AB) = 0$.)

(c) Take $A = B = N$ in (a). For the shift, $\ker N = \operatorname{span}(\mathbf{e}_1)$ and $\operatorname{im} N = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1})$, so $\ker N \subseteq \operatorname{im} N$ when $n \geq 2$. Indeed N^2 sends $\mathbf{e}_j$ to $\mathbf{e}_{j-2}$ for $j \geq 3$ and kills $\mathbf{e}_1, \mathbf{e}_2$, so $\operatorname{rank} N^2 = n - 2 = 2(n-1) - n$. For $n = 1$ the shift is 0, and $\ker N = F^1$ is not contained in $\operatorname{im} N = \{\mathbf{0}\}$; indeed $\operatorname{rank} N^2 = 0 \neq 2 \cdot 0 - 1$.

Solution to Exercise 4.8.11

Uses. Theorem 4.4.19 (Dimension of a sum), Theorem 4.2.11 (Rank–nullity), Proposition 4.2.3 (Properties of restriction), Example 4.8.3 (Sylvester's inequality, three ways)
Techniques. Restrict to a subspace (Technique 4.2.5), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Does $A(B\mathbf{x}) = \mathbf{0}$ say anything about $A\mathbf{x}$?
Hint 2. Try $A = B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. For the repair, restrict T_B to $\ker(AB)$.

Sketch. "A kills $B\mathbf{x}$" is not "A kills $\mathbf{x}$", and $\ker(AB) \subseteq \ker A + \ker B$ is false: for $A = B = N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $\ker N^2 = F^2$ but $\ker N + \ker N = \operatorname{span}(\mathbf{e}_1)$. The claim is true (Sylvester): T_B maps $\ker(AB)$ into $\ker A$ with kernel $\ker B$, so $\dim \ker(AB) \leq \dim \ker B + \dim \ker A$.

The error. From $A(B\mathbf{x}) = \mathbf{0}$ we learn that $B\mathbf{x} \in \ker A$, which says nothing about $A\mathbf{x}$. The step treats $AB\mathbf{x} = \mathbf{0}$ like a product of numbers that vanishes only if a factor does. The resulting containment $\ker(AB) \subseteq \ker A + \ker B$ is false. For $A = B = N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, we have $N^2 = 0$, so $\ker(AB) = F^2$, while $\ker N + \ker N = \operatorname{span}(\mathbf{e}_1)$.

The claim is true. It is Sylvester's inequality for square matrices. A repair is the kernel proof of Example 4.8.3. The vector $B\mathbf{x}$ lies in $\ker A$ for every $\mathbf{x} \in \ker(AB)$, so T_B restricts to a linear map $R : \ker(AB) \to \ker A$ (Proposition 4.2.3), and $\ker R = \ker(AB) \cap \ker B = \ker B$. Rank–nullity for R gives

$$\dim \ker(AB) = \dim \ker B + \dim \operatorname{im} R \leq \dim \ker B + \dim \ker A.$$

The repaired argument uses B to move $\ker(AB)$ into $\ker A$, where the false argument tried to place $\ker(AB)$ inside $\ker A + \ker B$ directly. In the example, R is N restricted to F^2, which maps F^2 onto $\ker N$ with kernel $\ker N$, and the inequality $2 \leq 1 + 1$ is an equality.

Solution to Exercise 4.8.12

Uses. Theorem 3.5.27 (Rank normal form), Theorem 2.4.34 (Block multiplication), Corollary 2.6.31 (Fibers are translates of the kernel), Proposition 4.5.5 (Isomorphisms preserve linear structure), Proposition 4.1.34 (Size of a vector space over a finite field)
Techniques. Choose a convenient basis (Technique 4.5.33), Transport along an isomorphism (Technique 4.5.7)
Hint 1. The condition $AXA = A$ survives $A \mapsto GAH$: if $A = G'NH'$ with G', H' invertible, substitute and set $Y = H'XG'$.
Hint 2. For $N = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \in F^{m\times n}$ and $Y \in F^{n\times m}$ in blocks, $NYN = \left(\begin{smallmatrix} Y_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right)$.
Sketch. By Theorem 3.5.27, $A = G'NH'$ with $G' = G^{-1}$ and $H' = H^{-1}$ invertible. Then $AXA = G'(NYN)H'$ with $Y = H'XG'$, and $X \mapsto H'XG'$ is an isomorphism of $F^{n\times m}$. $NYN = \left(\begin{smallmatrix} Y_{11} & 0 \\ 0 & 0 \end{smallmatrix}\right)$, so K corresponds to $Y_{11} = 0$, of dimension $nm - r^2$, and $X_0 = H'^{-1}N^{\mathsf{T}}G'^{-1}$ solves $AXA = A$. Over $\mathbb{F}_q$: q^{mn-r^2}. (c) $AXA = sA$ with s the sum of the entries of X, so Σ_A is the set of X whose entries add up to 1.

(a) By Theorem 3.5.27 there are invertible G and H with $GAH = N$, where $N = \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \in F^{m\times n}$. Let $G' = G^{-1}$ and $H' = H^{-1}$, so $A = G'NH'$. For $X \in F^{n\times m}$, put $Y = H'XG' \in F^{n\times m}$. Then

$$AXA = G'NH'XG'NH' = G'(NYN)H'.$$

Since G' and H' are invertible, $AXA = 0$ if and only if $NYN = 0$, and $AXA = A$ if and only if $NYN = N$. Write $Y = \left(\begin{smallmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{smallmatrix}\right)$ with $Y_{11} \in F^{r\times r}$. By Theorem 2.4.34,

$$NYN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} Y_{11} & 0 \\ 0 & 0 \end{pmatrix}.$$

So $NYN = 0$ exactly when $Y_{11} = 0$, and $NYN = N$ exactly when $Y_{11} = I_r$. The map $\Phi(X) = H'XG'$ is linear and invertible on $F^{n\times m}$, with inverse $Y \mapsto HYG$. It carries K onto $K_0 = \{Y \mid Y_{11} = 0\}$, the matrices whose top left $r \times r$ block is zero, a subspace of dimension $nm - r^2$. By Proposition 4.5.5, K is a subspace and $\dim K = mn - r^2$.

The matrix $Y_0 = N^{\mathsf{T}} \in F^{n\times m}$ has $Y_{11} = I_r$, so $X_0 = \Phi^{-1}(Y_0) = HN^{\mathsf{T}}G$ lies in Σ_A. Finally, Σ_A is the set of solutions of $\Psi(X) = A$ for the linear map $\Psi(X) = AXA$, so it is $X_0 + \ker \Psi = X_0 + K$ for any solution X_0 (Corollary 2.6.31).

(b) $X \mapsto X - X_0$ is a bijection from Σ_A onto K, which has q^{mn-r^2} elements (Proposition 4.1.34). So there are q^{mn-r^2} solutions. For an invertible $A \in F^{n\times n}$, this is $q^0 = 1$: the only solution is A^{-1}.

(c) Here $A = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ with $\mathbf{1} = (1, 1)$, so $AXA = \mathbf{1}\big(\mathbf{1}^{\mathsf{T}}X\mathbf{1}\big)\mathbf{1}^{\mathsf{T}} = sA$, where $s = \mathbf{1}^{\mathsf{T}}X\mathbf{1}$ is the sum of the four entries of X. So $AXA = A$ if and only if the entries of X add up to 1. This is a translate of the 3-dimensional subspace of matrices whose entries add up to 0, in agreement with $mn - r^2 = 4 - 1$.

Solution to Exercise 4.8.13

Uses. Theorem 2.4.34 (Block multiplication), Proposition 4.3.1 (Invertible factors preserve rank), Lemma 4.8.2 (Ranks of block triangular matrices), Example 4.8.3 (Sylvester's inequality, three ways), Exercise 4.3.31
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. The left factor subtracts A times the second block row from the first, and the right factor subtracts the first block column times C from the second.
Hint 2. The result is $\left(\begin{smallmatrix} 0 & -ABC \\ B & 0 \end{smallmatrix}\right)$. Compare its rank with that of M, using Lemma 4.8.2.
Sketch. (a) $\left(\begin{smallmatrix} 0 & -ABC \\ B & 0 \end{smallmatrix}\right)$. (b) The outer factors are invertible, so $\operatorname{rank} M = \operatorname{rank} B + \operatorname{rank}(ABC)$ (swap the block columns and use the block diagonal case of the lemma), while $\operatorname{rank} M \geq \operatorname{rank}(AB) + \operatorname{rank}(BC)$ by the lower triangular case. (c) $B = I_n$.

(a) By Theorem 2.4.34,

$$\begin{pmatrix} I_m & -A \\ 0 & I_n \end{pmatrix}\begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix} = \begin{pmatrix} AB - AB & -ABC \\ B & BC \end{pmatrix} = \begin{pmatrix} 0 & -ABC \\ B & BC \end{pmatrix},$$

and

$$\begin{pmatrix} 0 & -ABC \\ B & BC \end{pmatrix}\begin{pmatrix} I_p & -C \\ 0 & I_q \end{pmatrix} = \begin{pmatrix} 0 & -ABC \\ B & -BC + BC \end{pmatrix} = \begin{pmatrix} 0 & -ABC \\ B & 0 \end{pmatrix}.$$

(b) The matrices $\left(\begin{smallmatrix} I_m & -A \\ 0 & I_n \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} I_p & -C \\ 0 & I_q \end{smallmatrix}\right)$ are invertible, with inverses $\left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} I_p & C \\ 0 & I_q \end{smallmatrix}\right)$, as block multiplication checks. So $\operatorname{rank} M = \operatorname{rank}\left(\begin{smallmatrix} 0 & -ABC \\ B & 0 \end{smallmatrix}\right)$ by Proposition 4.3.1. Exchanging the two block columns of the last matrix (multiplication on the right by an invertible permutation matrix) gives $\left(\begin{smallmatrix} -ABC & 0 \\ 0 & B \end{smallmatrix}\right)$, whose rank is $\operatorname{rank}(ABC) + \operatorname{rank} B$ by Lemma 4.8.2(b). On the other hand, M is block lower triangular with diagonal blocks AB and BC, so $\operatorname{rank} M \geq \operatorname{rank}(AB) + \operatorname{rank}(BC)$ by Lemma 4.8.2(a). Together,

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leq \operatorname{rank} M = \operatorname{rank} B + \operatorname{rank}(ABC).$$

(c) Take $n = p$ and $B = I_n$, with $A \in F^{m\times n}$ and $C \in F^{n\times q}$. Then (b) reads $\operatorname{rank} A + \operatorname{rank} C \leq n + \operatorname{rank}(AC)$, which is Sylvester's inequality for the product AC, and the computation in (a) becomes the one in Example 4.8.3 (up to the order of the blocks). Exercise 4.3.31 proves the same inequality by restriction.

Solution to Exercise 4.8.14

Uses. Proposition 4.6.32 (Operators on a quotient), Proposition 4.1.26 (Nilpotent chains are independent), Corollary 4.6.15 (Cosets partition V), Theorem 4.6.27 (Dimension of a quotient), Theorem 4.1.10 (Independent lists are short), Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps)
Techniques. Pass to a quotient (Technique 4.6.31), Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Apply Proposition 4.1.26 to the induced operator T/U on V/U and the vector $\mathbf{v} + U$.
Hint 2. For (b), let k be the largest index with $T^k\mathbf{v} \notin U$, and count in V/U.
Sketch. (a) $(T/U)^j(\mathbf{v} + U) = T^j\mathbf{v} + U$, which is nonzero for $j = k$ and zero for $j = k + 1$; Proposition 4.1.26 in V/U gives independence. A combination of the $T^i\mathbf{v}$ that lies in U gives a relation among the cosets, so all coefficients vanish. (b) For $\mathbf{v}$ with $T^k\mathbf{v} \notin U$, $T^{k+1}\mathbf{v} \in U$, (a) gives $k + 1 \leq \dim V/U = d$; if $\mathbf{v} \in U$ already, every $T^j\mathbf{v} \in U$. The case $U = \{\mathbf{0}\}$ is Corollary 4.1.27.

(a) By Proposition 4.6.32(a), there is an operator T/U on V/U with $(T/U)(\mathbf{x} + U) = T\mathbf{x} + U$, and by induction on j, $(T/U)^j(\mathbf{v} + U) = T^j\mathbf{v} + U$. By Corollary 4.6.15(c), $T^k\mathbf{v} + U$ is not the zero coset, since $T^k\mathbf{v} \notin U$, and $T^{k+1}\mathbf{v} + U$ is the zero coset. By Proposition 4.1.26, applied to T/U and $\mathbf{v} + U$, the cosets $\mathbf{v} + U, \ldots, T^k\mathbf{v} + U$ are independent.

Now suppose $\mathbf{y} = \sum_{i=0}^{k} a_i T^i\mathbf{v} \in U$. Then $\sum_i a_i\left(T^i\mathbf{v} + U\right) = \mathbf{y} + U = U$, the zero of V/U, so every $a_i = 0$ and $\mathbf{y} = \mathbf{0}$. Taking $\mathbf{y} = \mathbf{0}$ shows that the list is independent, and in general this shows that its span meets U only in $\mathbf{0}$.

(b) Let $\mathbf{v} \in V$. If $\mathbf{v} \in U$, then $T^j\mathbf{v} \in U$ for every j, because T maps U into itself. Otherwise let k be the largest index with $T^k\mathbf{v} \notin U$; it exists and $k < m$, since $T^m\mathbf{v} \in U$ and, once a vector $T^j\mathbf{v}$ lies in U, so do all later ones.

By (a), V/U contains $k+1$ independent cosets, and $\dim V/U = \dim V - \dim U = d$ (Theorem 4.6.27), so $k+1 \le d$ (Theorem 4.1.10(a)). Hence $T^d\mathbf{v} = T^{d-k-1}\big(T^{k+1}\mathbf{v}\big) \in U$. So $T^d(V) \subseteq U$.

For $U = \{\mathbf{0}\}$, the hypothesis says that T is nilpotent, and the conclusion is $T^{\dim V} = 0$, which is Corollary 4.1.27. In general T is only nilpotent "modulo U", and the bound is the dimension of what is left, V/U.

Problems for Chapter 4

Solution to Problem 4.1

> *Uses.* Theorem 4.2.11 (Rank–nullity), Theorem 4.1.16 (Extending to a basis), Proposition 4.1.21 (Dimensions of subspaces), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Theorem 1.5.5 (Independence is decided at zero), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
> *Techniques.* Extend a basis of a subspace (Technique 4.1.18), Dimension counting (Technique 4.2.24)
> *Hint 1.* (b) Extend a basis $\mathbf{u}_1, \dots, \mathbf{u}_k$ of $\ker T$ to $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$, and show that $T(\mathbf{w}_1), \dots, T(\mathbf{w}_r)$ is a basis of $\operatorname{im} T$.
> *Hint 2.* (c) For spanning, write $T(\mathbf{v}) = \sum c_j T(\mathbf{v}_j)$; then $\mathbf{v} - \sum c_j \mathbf{v}_j \in \ker T$. For independence, apply T to a relation. Notice which step of (b) needed $\dim V < \infty$ and which step of (c) does not.
> *Sketch.* (b) The values $T(\mathbf{w}_j)$ span $\operatorname{im} T$ because the $\mathbf{u}_i$ are killed, and a relation $\sum b_j T(\mathbf{w}_j) = \mathbf{0}$ puts $\sum b_j \mathbf{w}_j$ in $\ker T$, so it is a combination of the $\mathbf{u}_i$, and independence of the basis forces $b = \mathbf{0}$. Count: $n = k + r$. (c) $\mathbf{v} - \sum c_j \mathbf{v}_j \in \ker T$ gives spanning; applying T to a relation kills the kernel part and leaves a relation among the basis $T(\mathbf{v}_j)$. So V has a finite basis of length $\dim \ker T + \dim \operatorname{im} T$.

(a) Let V be finite-dimensional and $T \in \mathrm{L}(V, W)$. Then $\operatorname{im} T$ is finite-dimensional, and $\dim \ker T + \dim \operatorname{im} T = \dim V$.

(b) Since $\ker T$ is a subspace of the finite-dimensional space V, it is finite-dimensional (Proposition 4.1.21(a)). Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be a basis of $\ker T$ and extend it to a basis $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$ of V (Theorem 4.1.16), so $\dim V = k + r$. We show that $T(\mathbf{w}_1), \dots, T(\mathbf{w}_r)$ is a basis of $\operatorname{im} T$; then $\operatorname{im} T$ is finite-dimensional with $\dim \operatorname{im} T = r = \dim V - \dim \ker T$.

Spanning. By Proposition 2.1.16, $\operatorname{im} T$ is spanned by the values of T on the basis, $\mathbf{0}, \dots, \mathbf{0}, T(\mathbf{w}_1), \dots, T(\mathbf{w}_r)$, and the zeros can be dropped.

Independence. If $\sum_j b_j T(\mathbf{w}_j) = \mathbf{0}$, then $T\big(\sum_j b_j \mathbf{w}_j\big) = \mathbf{0}$, so $\sum_j b_j \mathbf{w}_j \in \ker T$ and $\sum_j b_j \mathbf{w}_j = \sum_i a_i \mathbf{u}_i$ for some scalars a_i. The relation $\sum_i a_i \mathbf{u}_i - \sum_j b_j \mathbf{w}_j = \mathbf{0}$ among the vectors of a basis forces every $b_j = 0$ (Theorem 1.5.5).

(c) Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be a basis of $\ker T$ and $\mathbf{v}_1, \dots, \mathbf{v}_r$ as described; here V is any vector space, and we only assume that $\ker T$ and $\operatorname{im} T$ are finite-dimensional. Let $\mathcal{L} = \mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r$.

$\mathcal{L}$ spans V. Let $\mathbf{v} \in V$. Then $T(\mathbf{v}) \in \operatorname{im} T$, so $T(\mathbf{v}) = \sum_j c_j T(\mathbf{v}_j)$ for some scalars c_j, and $T\big(\mathbf{v} - \sum_j c_j \mathbf{v}_j\big) = \mathbf{0}$. So $\mathbf{v} - \sum_j c_j \mathbf{v}_j \in \ker T$ is a combination of the $\mathbf{u}_i$, and $\mathbf{v}$ is a combination of $\mathcal{L}$.

$\mathcal{L}$ is independent. If $\sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{v}_j = \mathbf{0}$, apply T: the first sum is killed, and $\sum_j b_j T(\mathbf{v}_j) = \mathbf{0}$ (Proposition 2.1.4). The $T(\mathbf{v}_j)$ are independent, so every $b_j = 0$, and then $\sum_i a_i \mathbf{u}_i = \mathbf{0}$ forces every $a_i = 0$.

So $\mathcal{L}$ is a basis of V, V is finite-dimensional, and $\dim V = k + r = \dim \ker T + \dim \operatorname{im} T$, which is rank–nullity again. The proof in (b) needs $\dim V < \infty$ before it starts, to extend a basis of the kernel to a basis of V. The proof in (c) only needs finite spanning lists for the kernel and the image, and it produces one for V. Contrapositively, if V is infinite-dimensional, then $\ker T$ or $\operatorname{im} T$ is. For instance multiplication by x on $F[x]$ has kernel $\{\mathbf{0}\}$, so its image $xF[x]$ is infinite-dimensional, and differentiation over $\mathbb{Q}$ has a 1-dimensional kernel and the infinite-dimensional image $\mathbb{Q}[x]$.

Solution to Problem 4.2

> *Uses.* Theorem 4.3.8 (Row rank equals column rank), Theorem 4.3.21 (Column–row factorization), Proposition 4.3.23 (Factorizations through F^k), Proposition 2.4.23 (Transpose of a product), Theorem 4.3.3 (Rank of a product), Theorem 4.1.10 (Independent lists are short), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 2.4.8 (Four views of the product), Theorem 3.5.14 (Third Fundamental Theorem, completed)
> *Techniques.* Transpose to trade rows for columns (Technique 3.2.27), Restrict to a subspace (Technique 4.2.5)

Hint 1. (b) A factorization $A = CR$ with inner size $r = \operatorname{rank} A$ transposes to $A^{\mathsf{T}} = R^{\mathsf{T}}C^{\mathsf{T}}$.
Hint 2. (c) The columns in J form a basis of $\operatorname{im} A$, so $A = CX$ where C consists of these columns. Restrict to the rows in I: $A_{I\cdot} = C_{I\cdot}X$, and $C_{I\cdot} = A_{IJ}$.
Sketch. (b) $A = CR$ with $C \in F^{m\times r}$, $R \in F^{r\times n}$ (Theorem 4.3.21), so $A^{\mathsf{T}} = R^{\mathsf{T}}C^{\mathsf{T}}$ has rank at most r; applied to A^{T}, $\operatorname{rank} A \le \operatorname{rank} A^{\mathsf{T}}$. (c) Let $C =$ the columns in J, a basis of $\operatorname{im} A$, so $A = CX$; taking the rows in I, $A_{I\cdot} = A_{IJ}X$ has rank r, so $\operatorname{rank} A_{IJ} \ge r$ and the $r \times r$ matrix A_{IJ} is invertible. Conversely an invertible $k \times k$ submatrix A_{IJ} makes the columns J of A independent, so $k \le \operatorname{rank} A$.

(a) For every $A \in F^{m\times n}$, $\operatorname{rank} A^{\mathsf{T}} = \operatorname{rank} A$: the rows of A span a subspace of F^n of the same dimension as the subspace of F^m spanned by its columns.

(b) Let $r = \operatorname{rank} A$. If $r = 0$, then $A = 0$ and $A^{\mathsf{T}} = 0$. Otherwise, by Theorem 4.3.21, $A = CR$ with $C \in F^{m\times r}$ and $R \in F^{r\times n}$. By Proposition 2.4.23, $A^{\mathsf{T}} = R^{\mathsf{T}}C^{\mathsf{T}}$ with inner size r, so $\operatorname{rank} A^{\mathsf{T}} \le r = \operatorname{rank} A$ by Proposition 4.3.23(a). Applying this to A^{T}, whose transpose is A, gives $\operatorname{rank} A \le \operatorname{rank} A^{\mathsf{T}}$.

(c) Let $C \in F^{m\times r}$ consist of the columns of A with indices in J. They are r independent vectors in the r-dimensional space $\operatorname{im} A$ (Proposition 2.6.18), so they form a basis of $\operatorname{im} A$ (Theorem 4.1.10(b)). Every column of A is therefore a combination of the columns of C, so $A = CX$ for some $X \in F^{r\times n}$ (Proposition 2.4.8). Taking the rows with indices in I on both sides, $A_{I\cdot} = C_{I\cdot}X$, where $A_{I\cdot} \in F^{r\times n}$ consists of the rows of A in I and $C_{I\cdot} = A_{IJ}$. The r rows of $A_{I\cdot}$ are independent, so $\operatorname{rank} A_{I\cdot} = r$ by (a). By Theorem 4.3.3,

$$r = \operatorname{rank}\bigl(A_{IJ}X\bigr) \le \operatorname{rank} A_{IJ} \le r,$$

so the $r \times r$ matrix A_{IJ} has rank r and is invertible (Theorem 3.5.14).

The largest invertible submatrix. By (a) there are r independent rows and r independent columns, so by what we just proved, A has an invertible $r \times r$ submatrix. Conversely, let A_{IJ} be an invertible $k \times k$ submatrix. A relation among the columns of A with indices in J gives, in the rows with indices in I, the same relation among the columns of A_{IJ}, which are independent. So the k columns of A in J are independent, and $k \le \operatorname{rank} A$.

The point of (c) is that the rows and the columns can be chosen separately: any r independent rows meet any r independent columns in an invertible submatrix. In Section 5.5 the same characterization of the rank is phrased with determinants.

Solution to Problem 4.3

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Theorem 1.5.5 (Independence is decided at zero), Proposition 1.2.35 (Zero products), Theorem 4.1.10 (Independent lists are short)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. (b) Apply T^k to a relation $\sum_i a_i T^i\mathbf{v} = \mathbf{0}$: only $a_0T^k\mathbf{v}$ survives. Continue with T^{k-1}, and so on. For the corollary, let k be the last index with $T^k\mathbf{v} \ne \mathbf{0}$.
Hint 2. (c) Apply T^j to $\mathbf{w} = \sum_i a_iT^i\mathbf{v}$. Which terms survive, and why are the surviving vectors independent?
Sketch. (b) After $a_0, \dots, a_{i-1}$ are known to vanish, T^{k-i} applied to the relation leaves $a_iT^k\mathbf{v} = \mathbf{0}$. For the corollary, the chain of a nonzero $\mathbf{v}$ has at most n entries, so $T^n\mathbf{v} = \mathbf{0}$. (c) $T^j\mathbf{w} = \sum_{i\le k-j} a_iT^{i+j}\mathbf{v}$, and $T^j\mathbf{v}, \dots, T^k\mathbf{v}$ are part of the independent chain, so $T^j\mathbf{w} = \mathbf{0}$ forces $a_0 = \cdots = a_{k-j} = 0$.

(a) Proposition 4.1.26: let $T \in \mathrm{L}(V)$, $\mathbf{v} \in V$, and $k \ge 0$ with $T^k\mathbf{v} \ne \mathbf{0}$ and $T^{k+1}\mathbf{v} = \mathbf{0}$. Then $\mathbf{v}, T\mathbf{v}, \dots, T^k\mathbf{v}$ is linearly independent. Corollary 4.1.27: if $\dim V = n$ and $T^m = 0$ for some $m \ge 1$, then $T^n = 0$.

(b) *The proposition.* For $l \ge k+1$, $T^l\mathbf{v} = T^{l-k-1}\bigl(T^{k+1}\mathbf{v}\bigr) = \mathbf{0}$. Let $\sum_{i=0}^{k} a_iT^i\mathbf{v} = \mathbf{0}$. We show $a_i = 0$ by induction on i. Suppose $a_0 = \cdots = a_{i-1} = 0$ (no assumption for $i = 0$). Applying T^{k-i} to $\sum_{l=i}^{k} a_lT^l\mathbf{v} = \mathbf{0}$ gives

$$a_iT^k\mathbf{v} + \sum_{l=i+1}^{k} a_lT^{k+l-i}\mathbf{v} = \mathbf{0},$$

and every exponent in the sum is at least $k+1$, so $a_iT^k\mathbf{v} = \mathbf{0}$. Since $T^k\mathbf{v} \ne \mathbf{0}$, $a_i = 0$ (Proposition 1.2.35). So the list is independent (Theorem 1.5.5).

The corollary. Let $\mathbf{v} \neq \mathbf{0}$. Since $T^m\mathbf{v} = \mathbf{0}$, there is a largest k with $T^k\mathbf{v} \neq \mathbf{0}$, and $T^{k+1}\mathbf{v} = \mathbf{0}$. By the proposition, $\mathbf{v}, \ldots, T^k\mathbf{v}$ is an independent list of length $k+1$ in V, so $k+1 \leq n$ (Theorem 4.1.10(a)), and $T^n\mathbf{v} = T^{n-k-1}\left(T^{k+1}\mathbf{v}\right) = \mathbf{0}$. Also $T^n\mathbf{0} = \mathbf{0}$. So $T^n = 0$.

(c) Fix j with $1 \leq j \leq k+1$. For $i \geq k-j+1$, $T^j\left(T^i\mathbf{v}\right) = T^{i+j}\mathbf{v} = \mathbf{0}$, since $i+j \geq k+1$. So the right side is contained in $\ker T^j \cap W$. Conversely, let $\mathbf{w} = \sum_{i=0}^{k} a_i T^i\mathbf{v} \in W$ with $T^j\mathbf{w} = \mathbf{0}$. The terms with $i+j \geq k+1$ vanish, so

$$\mathbf{0} = T^j\mathbf{w} = \sum_{i=0}^{k-j} a_i T^{i+j}\mathbf{v}.$$

The vectors $T^j\mathbf{v}, \ldots, T^k\mathbf{v}$ are entries of the independent chain, so $a_0 = \cdots = a_{k-j} = 0$, and $\mathbf{w} \in \operatorname{span}\left(T^{k-j+1}\mathbf{v}, \ldots, T^k\mathbf{v}\right)$. These j vectors are independent, being entries of the chain, so $\dim\left(\ker T^j \cap W\right) = j$.

So on W the kernels of the powers grow by exactly one dimension at each step, until $\ker T^{k+1} \cap W = W$. In particular $\ker T \cap W$ is the line spanned by the last vector $T^k\mathbf{v}$, the vector that becomes an eigenvector when the chain is shifted in Proposition 7.2.37.

Solution to Problem 4.4

Uses. Theorem 4.4.27 (Idempotents are projections), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Theorem 4.5.17 (Change of basis), Theorem 2.4.34 (Block multiplication), Theorem 4.2.11 (Rank–nullity), Proposition 4.5.5 (Isomorphisms preserve linear structure), Proposition 4.4.21 (Two out of three)
Techniques. Choose a basis adapted to a preserved subspace (Technique 4.5.36), Study a linear map on a space of matrices (Technique 4.2.29), Split the space with an idempotent (Technique 4.4.28)
Hint 1. A basis of $\operatorname{im} P$ followed by a basis of $\ker P$ gives an invertible S with $S^{-1}PS = D = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. The map $X \mapsto S^{-1}XS$ is an isomorphism of $F^{n\times n}$ that turns δ into $Y \mapsto DY - YD$.
Hint 2. Compute $DY - YD$ for $Y = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}$.
Sketch. With $D = S^{-1}PS$ and $Y = S^{-1}XS$, $\delta(X) = S(DY - YD)S^{-1}$ and $DY - YD = \begin{pmatrix} 0 & Y_{12} \\ -Y_{21} & 0 \end{pmatrix}$. So $\ker\delta$ corresponds to block diagonal Y (dimension $r^2 + (n-r)^2$) and $\operatorname{im}\delta$ to block off-diagonal Y (dimension $2r(n-r)$). DZD and $(I-D)Z(I-D)$ pick out the diagonal blocks, so they vanish exactly on the off-diagonal matrices, and the conditions transport. Block diagonal plus block off-diagonal is a direct sum, and the diagonal part of Y is $DYD + (I-D)Y(I-D)$.

If $r = 0$ or $r = n$, then $P = 0$ or $P = I_n$ (an idempotent of full rank is invertible, and $P^2 = P$ then gives $P = I_n$), so $\delta = 0$, and all three statements hold with $\ker\delta = F^{n\times n}$ and $\operatorname{im}\delta = \{0\}$. So let $1 \leq r \leq n-1$. By Theorem 4.4.27, $F^n = \operatorname{im} P \oplus \ker P$, and P fixes $\operatorname{im} P$ pointwise and kills $\ker P$. Let $S \in F^{n\times n}$ have as columns a basis of $\operatorname{im} P$ followed by a basis of $\ker P$; together they form a basis of F^n (Theorem 4.4.13), so S is invertible. In this basis T_P has the matrix $D = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, so $S^{-1}PS = D$ (Theorem 4.5.17). For $X \in F^{n\times n}$ write $Y = S^{-1}XS$. Then

$$S^{-1}\delta(X)S = S^{-1}PS\,S^{-1}XS - S^{-1}XS\,S^{-1}PS = DY - YD.$$

The map $X \mapsto S^{-1}XS$ is linear and invertible, with inverse $Y \mapsto SYS^{-1}$, so it carries $\ker\delta$ onto $\ker\delta_D$ and $\operatorname{im}\delta$ onto $\operatorname{im}\delta_D$, where $\delta_D(Y) = DY - YD$ (Proposition 4.5.5). In blocks, with $Y_{11} \in F^{r\times r}$, Theorem 2.4.34 gives

$$DY - YD = \begin{pmatrix} Y_{11} & Y_{12} \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} Y_{11} & 0 \\ Y_{21} & 0 \end{pmatrix} = \begin{pmatrix} 0 & Y_{12} \\ -Y_{21} & 0 \end{pmatrix}.$$

(a) $\delta_D(Y) = 0$ if and only if $Y_{12} = 0$ and $Y_{21} = 0$. So $\ker\delta_D$ is the space of block diagonal matrices, of dimension $r^2 + (n-r)^2$, and $\ker\delta$ has the same dimension.

(b) By rank–nullity, $\operatorname{rank}\delta = n^2 - r^2 - (n-r)^2 = 2r(n-r)$. The image of δ_D consists of the block off-diagonal matrices $\begin{pmatrix} 0 & Z_{12} \\ Z_{21} & 0 \end{pmatrix}$, since Y_{12} and $-Y_{21}$ can be anything. For $Z' = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$,

$$DZ'D = \begin{pmatrix} Z_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad (I_n - D)Z'(I_n - D) = \begin{pmatrix} 0 & 0 \\ 0 & Z_{22} \end{pmatrix},$$

so $Z' \in \operatorname{im}\delta_D$ if and only if $DZ'D = 0$ and $(I_n - D)Z'(I_n - D) = 0$. For $Z = SZ'S^{-1}$, we have $PZP = S(DZ'D)S^{-1}$ and $(I_n - P)Z(I_n - P) = S((I_n - D)Z'(I_n - D))S^{-1}$, using $I_n - P = S(I_n - D)S^{-1}$. So $Z \in \operatorname{im}\delta$ if and only if $PZP = 0$ and $(I_n - P)Z(I_n - P) = 0$.

(c) By (a) and (b), $\dim\ker\delta + \dim\operatorname{im}\delta = n^2$. A matrix in both is, after conjugation by S, both block diagonal and block off-diagonal, hence 0. So $F^{n\times n} = \ker\delta \oplus \operatorname{im}\delta$ by Proposition 4.4.21. Let $X_0 = PXP + (I_n - P)X(I_n - P)$. It commutes with P: since $P^2 = P$ and $P(I_n - P) = 0 = (I_n - P)P$, both PX_0 and X_0P equal PXP. And $X - X_0 = PX(I_n - P) + (I_n - P)XP$ satisfies the two conditions of (b), because $P(X - X_0)P = 0$ and $(I_n - P)(X - X_0)(I_n - P) = 0$ by the same two identities. So $X = X_0 + (X - X_0)$ is the decomposition.

Choosing the basis made the computation a matter of reading blocks. The formula in (c) needs no basis at all, and could have been guessed from the blocks: PXP and $(I_n - P)X(I_n - P)$ are the two diagonal blocks of X seen through the decomposition $F^n = \operatorname{im}P \oplus \ker P$.

Solution to Problem 4.5

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Corollary 4.1.11 (Spanning lists are long), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 2.1.24 (Algebra of composition), Exercise 2.1.32
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Dimension counting (Technique 4.2.24)
Hint 1. Δ lowers the degree by exactly one when the leading coefficient is multiplied by a nonzero number. Compute the degree and leading coefficient of $\Delta^k p$, and apply Proposition 4.1.26 to Δ and p.
Hint 2. $\Delta^k = (S - \mathrm{id})^k$ is a combination of $\mathrm{id}, S, \ldots, S^k$. For (c), $(x+1)^p = x^p + 1$ over $\mathbb{F}_p$.
Sketch. (a) For $\deg q = d \geq 1$ with leading coefficient c and $dc \neq 0$, Δq has degree $d-1$ and leading coefficient dc; so $\Delta^k p$ has degree $n-k$ and leading coefficient $n(n-1)\cdots(n-k+1)c \neq 0$ for $k \leq n$, and $\Delta^{n+1}p = \mathbf{0}$. By Proposition 4.1.26, $p, \Delta p, \ldots, \Delta^n p$ is a basis of $F[x]_{\leq n}$. Each $\Delta^k p = \sum_j \binom{k}{j}(-1)^{k-j}p(x+j)$ lies in the span of the translates, so they span, and $n+1$ spanning vectors form a basis. (b) $\Delta^2 x^2 = 2$, so $1 = \frac{1}{2}\left((x+2)^2 - 2(x+1)^2 + x^2\right)$. (c) $(x+j)^p = x^p + j$.

(a) All the maps involved send $F[x]_{\leq n}$ into itself, since $p(x+j)$ has the same degree as p. By the binomial theorem, for $j \geq 1$ the polynomial $\Delta(x^j) = (x+1)^j - x^j$ has degree at most $j-1$, and its coefficient of x^{j-1} is $j * 1$; and $\Delta(1) = \mathbf{0}$. Let $q = q_0 + q_1 x + \cdots + q_d x^d$ with $d \geq 1$ and $c = q_d \neq 0$. By linearity, $\Delta q = \sum_{j=0}^{d} q_j \Delta(x^j)$, and only the term $j = d$ can contribute to the coefficient of x^{d-1}, which is therefore dc. If $dc \neq 0$, then Δq has degree exactly $d-1$ and leading coefficient dc; and Δ kills constants. Let p have degree n and leading coefficient c. By induction on k, for $k \leq n$, $\Delta^k p$ has degree $n-k$ and leading coefficient $n(n-1)\cdots(n-k+1)c$. This coefficient is nonzero: its integer factors $n, n-1, \ldots, n-k+1$ are factors of $n!$, which is nonzero in F, so none of them is 0 in F, and a product of nonzero elements of a field is nonzero. In particular $\Delta^n p = n!c$ is a nonzero constant, and $\Delta^{n+1}p = \mathbf{0}$.

By Proposition 4.1.26, applied to Δ on $F[x]_{\leq n}$, the list $p, \Delta p, \ldots, \Delta^n p$ is independent. It has $n+1 = \dim F[x]_{\leq n}$ entries, so it is a basis (Theorem 4.1.10(b)). Since S commutes with id, the binomial theorem for commuting operators (which has the same proof as for numbers, using Proposition 2.1.24) gives

$$\Delta^k p = (S - \mathrm{id})^k p = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} S^j p = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} p(x+j).$$

So every $\Delta^k p$ with $k \leq n$ lies in the span of $p(x), p(x+1), \ldots, p(x+n)$, and this span is all of $F[x]_{\leq n}$. A spanning list of length $n+1 = \dim F[x]_{\leq n}$ is a basis (Corollary 4.1.11(b)).

(b) For $p = x^2$, $\Delta p = 2x+1$ and $\Delta^2 p = 2$. By the formula above, $2 = \Delta^2 p = (x+2)^2 - 2(x+1)^2 + x^2$, so

$$1 = \frac{1}{2}(x+2)^2 - (x+1)^2 + \frac{1}{2}x^2.$$

Check: $\frac{1}{2}(x^2 + 4x + 4) - (x^2 + 2x + 1) + \frac{1}{2}x^2 = 2 - 1 = 1$.

(c) Over $\mathbb{F}_p$, the binomial coefficients $\binom{p}{i}$ with $0 < i < p$ are 0, so $S(x^p) = (x+1)^p = x^p + 1$, and $S(1) = 1$. By induction on j, $S^j(x^p) = x^p + j * 1$. So the translates of x^p are the polynomials $x^p + c$ with $c \in \mathbb{F}_p$; they lie in $\operatorname{span}(x^p, 1)$ and include x^p and $x^p + 1$, so they span exactly this 2-dimensional space. For $n = p$ the factor p of $n!$ is 0 in $\mathbb{F}_p$, the chain $x^p, \Delta x^p = 1, \Delta^2 x^p = \mathbf{0}$ stops after two steps, and the argument of (a) breaks exactly there.

Solution to Problem 4.6

Uses. Proposition 4.3.13 (Dimensions of the four subspaces), Theorem 3.3.12 (Matrices in RREF), Theorem 4.2.11 (Rank–nullity), Theorem 4.1.10 (Independent lists are short), Proposition 4.4.2 (Sums are the smallest containing subspaces), Theorem 4.4.19 (Dimension of a sum)
Techniques. Dimension counting (Technique 4.2.24), Reduce to RREF (Technique 3.3.22)
Hint 1. (a) A combination of the rows is $\left(\sum a_i \mathbf{u}_i + \sum b_j \mathbf{w}_j, \sum a_i \mathbf{u}_i\right)$. For the kernel, $\mathbf{u} + \mathbf{w} = \mathbf{0}$ means $\mathbf{u} = -\mathbf{w} \in U \cap W$.
Hint 2. (b) Row reduction keeps the row space. In R, the rows with a pivot in the left half come first. Count: if s rows have a nonzero left half and d others are nonzero, the kernel in (a) has dimension $(s + d) - s$.
Sketch. (a) Combinations of the rows give exactly the pairs $(\mathbf{u} + \mathbf{w}, \mathbf{u})$; projecting to the left half gives $U + W$, and $\mathbf{u} + \mathbf{w} = \mathbf{0}$ forces $\mathbf{u} \in U \cap W$. (b) The nonzero rows of R are a basis of Z. The s rows with a pivot in the left half have independent left halves (pivot columns), which span the image $U + W$. The other d nonzero rows are $(\mathbf{0}, \mathbf{c})$, independent, in the kernel, whose dimension is $(s + d) - s = d$; so the $\mathbf{c}$'s are a basis of $U \cap W$. (c) Over $\mathbb{F}_2$: $U + W$ has basis $(1,0,0,1), (0,1,0,1), (0,0,1,1)$ and $U \cap W = \operatorname{span}((1,0,1,0))$. Over $\mathbb{Q}$, $U + W = \mathbb{Q}^4$ and $U \cap W = \{\mathbf{0}\}$.

(a) A combination of the rows of M with coefficients a_i and b_j is

$$\left(\sum_i a_i \mathbf{u}_i + \sum_j b_j \mathbf{w}_j, \sum_i a_i \mathbf{u}_i\right) = (\mathbf{u} + \mathbf{w}, \mathbf{u}),$$

with $\mathbf{u} = \sum_i a_i \mathbf{u}_i \in U$ and $\mathbf{w} = \sum_j b_j \mathbf{w}_j \in W$, and every such pair arises. So the row space is Z. The map $\pi(\mathbf{a}, \mathbf{b}) = \mathbf{a}$ is linear, and $\pi(Z) = \{\mathbf{u} + \mathbf{w}\} = U + W$. A pair $(\mathbf{u} + \mathbf{w}, \mathbf{u})$ lies in $\ker \pi$ exactly when $\mathbf{u} = -\mathbf{w}$, that is, $\mathbf{u} \in U \cap W$; and for $\mathbf{u} \in U \cap W$, $(\mathbf{0}, \mathbf{u}) = (\mathbf{u} + (-\mathbf{u}), \mathbf{u}) \in Z$. So $\ker(\pi\restriction_Z) = \{(\mathbf{0}, \mathbf{u}) \mid \mathbf{u} \in U \cap W\}$, which is isomorphic to $U \cap W$ by $(\mathbf{0}, \mathbf{u}) \mapsto \mathbf{u}$.

(b) Row operations do not change the row space, and the nonzero rows of R form a basis of it (Proposition 4.3.13(b)), so they form a basis of Z. In R the pivots move to the right from one row to the next (Theorem 3.3.12), so the nonzero rows split into s rows $\mathbf{r}_1, \ldots, \mathbf{r}_s$ with their pivot in the left half, followed by d rows of the form $(\mathbf{0}, \mathbf{c}_1), \ldots, (\mathbf{0}, \mathbf{c}_d)$.

The left halves of the $\mathbf{r}_i$ form a basis of $U + W$. They span $\pi(Z) = U + W$, since π of the other basis vectors of Z is $\mathbf{0}$. They are independent: in a relation $\sum_i c_i \pi(\mathbf{r}_i) = \mathbf{0}$, look at the pivot column of $\mathbf{r}_i$, which lies in the left half and in which every other row of R has a 0; it gives $c_i = 0$.

The $\mathbf{c}_j$ form a basis of $U \cap W$. The vectors $(\mathbf{0}, \mathbf{c}_j)$ are independent, being nonzero rows of R, and they lie in $\ker(\pi\restriction_Z)$. By rank–nullity for $\pi\restriction_Z$, this kernel has dimension $\dim Z - \dim(U + W) = (s + d) - s = d$. So they form a basis of it (Theorem 4.1.10(b)), and by (a) the $\mathbf{c}_j$ form a basis of $U \cap W$. In passing, $\dim(U + W) + \dim(U \cap W) = s + d = \dim Z = \dim U + \dim W$ when the $\mathbf{u}_i$ and the $\mathbf{w}_j$ are bases, which is Theorem 4.4.19 once more.

(c) The matrix is

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Over $\mathbb{F}_2$, where $-1 = 1$, Gauss–Jordan elimination gives

$$R = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

So $U + W$ has basis $(1,0,0,1), (0,1,0,1), (0,0,1,1)$, the hyperplane $x_1 + x_2 + x_3 + x_4 = 0$, and $U \cap W = \operatorname{span}((1,0,1,0))$. Indeed $(1,0,1,0) = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{w}_1 + \mathbf{w}_2$ over $\mathbb{F}_2$.

Over $\mathbb{Q}$, the same elimination ends with a pivot in each of the first four columns, since

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \end{pmatrix},$$

so $U + W = \mathbb{Q}^4$ and $U \cap W = \{\mathbf{0}\}$. Directly: $U = \{(a, a+b, b, 0)\}$ and $W = \{(d, 0, c, c-d)\}$, and a common vector has $a = b = c = d$ and $2a = 0$. Over $\mathbb{Q}$ this forces $a = 0$, and over $\mathbb{F}_2$ it does not.

The column method of Example 4.4.20 finds the intersection as a relation among the given vectors; this method finds bases of both spaces in one reduction, with no relation to interpret.

Solution to Problem 4.7

Uses. Definition 4.5.23 (Similar matrices), Proposition 4.5.27 (Similarity invariants), Proposition 4.3.1 (Invertible factors preserve rank), Proposition 2.4.14 (Algebra of matrices), Theorem 4.4.19 (Dimension of a sum), Proposition 4.4.2 (Sums are the smallest containing subspaces), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Compare components (Technique 4.4.8)
Hint 1. Compute $A(I_n - A - B)$ and $(I_n - A - B)B$.
Hint 2. (c) Show $\operatorname{im}(A+B) = \operatorname{im} A + \operatorname{im} B$ and $\operatorname{im} A \cap \operatorname{im} B = \{\mathbf{0}\}$. For the first, $(A+B)A = A$.
Sketch. (a) $AG = -AB = GB$ for $G = I_n - A - B$, so $B = G^{-1}AG$. (b) $A = I_n$, $B = 0$. (c) $(A+B)^2 = A + AB + BA + B = A + B$. $(A+B)A\mathbf{x} = A\mathbf{x}$ gives $\operatorname{im} A \subseteq \operatorname{im}(A+B)$, likewise for B, so $\operatorname{im}(A+B) = \operatorname{im} A + \operatorname{im} B$; and $\mathbf{y} = A\mathbf{x} = B\mathbf{z}$ gives $\mathbf{y} = A\mathbf{y} = AB\mathbf{z} = \mathbf{0}$.

(a) Let $G = I_n - A - B$. Using $A^2 = A$ and $B^2 = B$,

$$AG = A - A^2 - AB = -AB, \quad GB = B - AB - B^2 = -AB.$$

So $AG = GB$, and since G is invertible, $B = G^{-1}AG$. So $A \sim B$, and $\operatorname{rank} A = \operatorname{rank} B$ (Proposition 4.5.27(a)). The rank statement alone also follows from Proposition 4.3.1: $\operatorname{rank} A = \operatorname{rank}(AG) = \operatorname{rank}(GB) = \operatorname{rank} B$.

(b) $A = I_n$ and $B = 0$ are idempotent of ranks n and 0, and $I_n - A - B = 0$ is not invertible.

(c) $(A+B)^2 = A^2 + AB + BA + B^2 = A + B$ (Proposition 2.4.14). For every $\mathbf{x}$, $(A+B)A\mathbf{x} = A^2\mathbf{x} + BA\mathbf{x} = A\mathbf{x}$, so $\operatorname{im} A \subseteq \operatorname{im}(A+B)$; similarly $(A+B)B\mathbf{x} = B\mathbf{x}$ gives $\operatorname{im} B \subseteq \operatorname{im}(A+B)$. Hence $\operatorname{im} A + \operatorname{im} B \subseteq \operatorname{im}(A+B)$ (Proposition 4.4.2), and the converse containment holds because $(A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x}$. So $\operatorname{im}(A+B) = \operatorname{im} A + \operatorname{im} B$. If $\mathbf{y} \in \operatorname{im} A \cap \operatorname{im} B$, say $\mathbf{y} = A\mathbf{x} = B\mathbf{z}$, then $\mathbf{y} = A^2\mathbf{x} = A\mathbf{y} = AB\mathbf{z} = \mathbf{0}$. By Theorem 4.4.19,

$$\operatorname{rank}(A+B) = \dim(\operatorname{im} A + \operatorname{im} B) = \operatorname{rank} A + \operatorname{rank} B.$$

Over $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, (c) also follows from tr = rank for idempotents (Exercise 4.3.43), since the trace is additive. The argument above works over every field, where the trace of an idempotent only determines its rank modulo the characteristic.

Solution to Problem 4.8

Uses. Exercise 4.3.30, Exercise 4.2.52, Proposition 4.2.15 (Rank through any matrix), Theorem 4.2.11 (Rank–nullity), Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Exercise 4.5.50
Techniques. Dimension counting (Technique 4.2.24), Apply an operator to a relation to kill terms (Technique 4.1.25)

Hint 1. (a) Sylvester with k factors: $0 = \operatorname{rank} N^k \geq k \operatorname{rank} N - (k-1)n$.
Hint 2. (b) The k jumps $d_1, \ldots, d_k$ add up to $\dim\ker N^k = n$, and none exceeds $d_1 = n/k$. (c) Let m be the least exponent with $N^m = 0$, and compare $n = d_1 + \cdots + d_m$ with md_1.
Sketch. (a) $\operatorname{rank} N \leq (k-1)n/k$. (b) $d_1 + \cdots + d_k = n$ with every $d_j \leq n/k$ forces $d_j = n/k$, so $\dim\ker N^j = jn/k$, and $\dim\ker N^{k-1} < n$. Example: n/k shifts of size k. (c) With $N^m = 0$ minimal, $n = d_1 + \cdots + d_m \leq m$, so $m = n$ and $N^{n-1} \neq 0$; a vector with $N^{n-1}\mathbf{v} \neq \mathbf{0}$ has a chain of length n, which is a basis.

(a) Let A be the matrix of N in some basis, so $A^k = 0$ and $\operatorname{rank} A = \operatorname{rank} N$ (Proposition 4.2.15). By Exercise 4.3.30(c) with all k factors equal to A,

$$0 = \operatorname{rank} A^k \geq k \operatorname{rank} A - (k-1)n,$$

so $\operatorname{rank} N \leq (k-1)n/k$ and $\dim\ker N = n - \operatorname{rank} N \geq n/k$. The kernel proof of Exercise 4.2.52 and this one bound the same quantity; the kernel proof is the kernel proof of Sylvester's inequality (Example 4.8.3), applied k times.

(b) Since $N^k = 0$, $\sum_{j=1}^{k} d_j = \dim\ker N^k - \dim\ker N^0 = n$. By Exercise 4.2.52(b), $d_j \leq d_1 = n/k$ for every j. A sum of k numbers, each at most n/k, equals n only if every term equals n/k. So $\dim\ker N^j = d_1 + \cdots + d_j = jn/k$ for $0 \leq j \leq k$. In particular $\dim\ker N^{k-1} = n - n/k < n$, so $N^{k-1} \neq 0$: a nilpotent operator whose kernel is as small as the bound allows takes the full k steps to die.

For an example, let $m = n/k$ and choose a basis $\mathbf{b}_{i,l}$ of V, indexed by $1 \leq i \leq m$ and $1 \leq l \leq k$. Define N by $N\mathbf{b}_{i,1} = \mathbf{0}$ and $N\mathbf{b}_{i,l} = \mathbf{b}_{i,l-1}$ for $l \geq 2$. Then N^k kills every basis vector, so $N^k = 0$, and N sends the basis vectors $\mathbf{b}_{i,l}$ with $l \geq 2$ to distinct basis vectors and the others to $\mathbf{0}$. So $\operatorname{im} N$ is spanned by the $m(k-1)$ independent vectors $\mathbf{b}_{i,l}$ with $l \leq k-1$, $\operatorname{rank} N = m(k-1)$, and $\dim\ker N = n - m(k-1) = m$.

(c) For $n = 1$ there is nothing to prove, since $N^0 = \mathrm{id} \neq 0$. Let $n \geq 2$, and let m be the least exponent with $N^m = 0$; it exists by Corollary 4.1.27, with $m \leq n$. With the jumps as above, $n = d_1 + \cdots + d_m \leq md_1 = m$, so $m = n$ and $N^{n-1} \neq 0$. (Part (a) with $k = n-1$ gives the same: $N^{n-1} = 0$ would force $1 \geq n/(n-1)$.) So there is $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$ and $N^n\mathbf{v} = \mathbf{0}$. By Proposition 4.1.26, $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is independent, and it has n entries, so it is a basis (Theorem 4.1.10(b)). In this basis N is the shift (Exercise 4.5.50).

Solution to Problem 4.9

Uses. Theorem 4.2.11 (Rank–nullity), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.1.16 (Linear maps preserve spans and linear relations), Corollary 4.2.18 (Injective if and only if surjective)
Techniques. Dimension counting (Technique 4.2.24), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. Let $\rho_i = \operatorname{rank} T_i$, and $\rho_0 = \rho_k = 0$. Show $\dim V_i = \rho_{i-1} + \rho_i$ with rank–nullity. Then the alternating sum telescopes.
Hint 2. (c) In an exact sequence, $\rho_i = n_i - \rho_{i-1}$, so $\rho_i = r_i$. Conversely, build $T_i : F^{n_i} \to F^{n_{i+1}}$ that kills the first r_{i-1} standard basis vectors and sends the next r_i to $\mathbf{e}_1, \ldots, \mathbf{e}_{r_i}$.
Sketch. (a) $\dim V_i = \dim\ker T_i + \operatorname{rank} T_i = \rho_{i-1} + \rho_i$ (with $\ker T_1 = \{\mathbf{0}\}$ and, at V_k, $\dim V_k = \rho_{k-1}$ by surjectivity), so $\sum (-1)^i(\rho_{i-1} + \rho_i) = -\rho_0 + (-1)^k\rho_k = 0$. (b) $k = 2$: an isomorphism, $n_1 = n_2$. $k = 3$: $\dim V_2 = \dim V_1 + \dim V_3$, rank–nullity for T_2. (c) Necessity: $\rho_i = r_i$ by induction, and ranks are nonnegative, $\rho_k = 0$. Sufficiency: the maps above have $\operatorname{im} T_i = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{r_i}) = \ker T_{i+1}$.

Write $\rho_i = \operatorname{rank} T_i$ for $1 \leq i \leq k-1$, and $\rho_0 = \rho_k = 0$.

(a) We claim $\dim V_i = \rho_{i-1} + \rho_i$ for every i. For $i = 1$: T_1 is injective, so rank–nullity gives $\dim V_1 = 0 + \rho_1$. For $2 \leq i \leq k-1$: $\dim\ker T_i = \dim\operatorname{im} T_{i-1} = \rho_{i-1}$ by exactness, and rank–nullity for T_i gives $\dim V_i = \rho_{i-1} + \rho_i$. For $i = k$: T_{k-1} is surjective, so $\dim V_k = \rho_{k-1} = \rho_{k-1} + \rho_k$. Hence

$$\sum_{i=1}^{k} (-1)^i \dim V_i = \sum_{i=1}^{k} (-1)^i \rho_{i-1} + \sum_{i=1}^{k} (-1)^i \rho_i = \sum_{i=0}^{k-1} (-1)^{i+1} \rho_i + \sum_{i=1}^{k} (-1)^i \rho_i,$$

and the terms with $1 \leq i \leq k-1$ cancel in pairs, leaving $-\rho_0 + (-1)^k\rho_k = 0$.

(b) For $k = 2$, T_1 is injective and surjective, an isomorphism, and (a) says $\dim V_1 = \dim V_2$. For $k = 3$, T_1 is injective, T_2 is surjective, and $\operatorname{im} T_1 = \ker T_2$; (a) says $\dim V_2 = \dim V_1 + \dim V_3 = \dim\ker T_2 + \dim\operatorname{im} T_2$,

which is rank–nullity for T_2. For example, $\ker T \xrightarrow{\iota} V \xrightarrow{T} \operatorname{im} T$, with ι the inclusion and T viewed as a map onto its image, is exact. So (a) is rank–nullity applied along a chain of maps.

(c) *Necessity.* By the claim in (a), $\rho_i = n_i - \rho_{i-1}$ with $\rho_0 = 0$, so by induction $\rho_i = r_i$ for $1 \le i \le k-1$. Ranks are nonnegative, so $r_i \ge 0$ for these i, and $r_k = n_k - \rho_{k-1} = 0$ by the claim at $i = k$.

Sufficiency. Suppose all $r_i \ge 0$ and $r_k = 0$. Let $V_i = F^{n_i}$. Then $r_{i-1} + r_i = n_i$, so we may define T_i, for $1 \le i \le k-1$, on the standard basis of F^{n_i} by

$$T_i(\mathbf{e}_j) = \mathbf{0} \quad (j \le r_{i-1}), \quad T_i\left(\mathbf{e}_{r_{i-1}+j}\right) = \mathbf{e}_j \quad (1 \le j \le r_i)$$

(Theorem 2.2.1). This makes sense because $r_i \le r_i + r_{i+1} = n_{i+1}$. The values on the basis are $\mathbf{0}$ and the independent vectors $\mathbf{e}_1, \dots, \mathbf{e}_{r_i}$, so $\operatorname{im} T_i = \operatorname{span}\left(\mathbf{e}_1, \dots, \mathbf{e}_{r_i}\right)$ (Proposition 2.1.16) and $\operatorname{rank} T_i = r_i$. A vector $\sum_j x_j \mathbf{e}_j$ is killed by T_i exactly when $x_j = 0$ for $j > r_{i-1}$, so $\ker T_i = \operatorname{span}\left(\mathbf{e}_1, \dots, \mathbf{e}_{r_{i-1}}\right)$. Hence $\operatorname{im} T_{i-1} = \ker T_i$ for $2 \le i \le k-1$; $\ker T_1 = \{\mathbf{0}\}$ since $r_0 = 0$; and $\operatorname{im} T_{k-1} = \operatorname{span}\left(\mathbf{e}_1, \dots, \mathbf{e}_{r_{k-1}}\right) = F^{n_k}$ since $r_{k-1} = n_k - r_k = n_k$. The sequence is exact.

The condition $r_k = 0$ alone is the vanishing of the alternating sum in (a), since $r_k = \sum_{i=1}^{k} (-1)^{k-i} n_i$. It is not sufficient: $(n_1, n_2, n_3, n_4) = (1, 0, 0, 1)$ has alternating sum 0, but $r_2 = -1$, and indeed $F \to \{\mathbf{0}\}$ cannot be injective.

Solution to Problem 4.10

Uses. Proposition 4.4.16 (Complements exist), Proposition 4.4.31 (Defining maps summand by summand), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 4.5.5 (Isomorphisms preserve linear structure), Corollary 4.1.35 (Counting independent lists and bases), Theorem 4.4.27 (Idempotents are projections), Proposition 4.4.26 (Properties of projections), Proposition 4.1.37 (Counting subspaces), Exercise 4.4.49, Exercise 4.1.54, Theorem 4.2.11 (Rank–nullity)
Techniques. Double counting (Technique 4.1.36), Construct a map by choosing it on a basis (Technique 2.2.4)
Hint 1. (a) Fix a complement X of U. A map with kernel U and image W is determined by its restriction to X, which must be an isomorphism $X \to W$. Count isomorphisms by the image of a fixed basis of X.
Hint 2. (b) An idempotent is the projection onto its image along its kernel. Count pairs (image, complement). (c) Count triples (kernel, image, isomorphism).
Sketch. (a) $T \mapsto T{\restriction_X}$ is a bijection onto the isomorphisms $X \to W$ (inverse: $\varphi \mapsto \varphi \circ \Pr_{X,U}$), and isomorphisms correspond to bases of W: $N_q(r,r)$. (b) $P \mapsto (\operatorname{im} P, \ker P)$ is a bijection onto pairs of complementary subspaces; there are $N_q(n,r)/N_q(r,r)$ choices of image and $q^{r(n-r)}$ complements of each. For $n = 2$: $1 + (q+1)q + 1 = q^2 + q + 2$. (c) Choose the kernel ($N_q(n,n-r)/N_q(n-r,n-r)$ ways, equal to the number of r-dimensional subspaces), the image, and then (a). For $n = q = 2$: $1 + 9 + 6 = 16$.

(a) By Proposition 4.4.16, U has a complement X, with $\dim X = n - \dim U = r$. Let T have kernel U and image W. Then $T{\restriction_X}$ is injective, since $X \cap \ker T = X \cap U = \{\mathbf{0}\}$, and its image is $T(U + X) = T(V) = W$. So $T{\restriction_X}$ is an isomorphism $X \to W$, and T is determined by it, since T is 0 on U (Proposition 4.4.31). Conversely, for an isomorphism $\varphi : X \to W$, the map T that is 0 on U and φ on X (Proposition 4.4.31) has image $\varphi(X) = W$, and $T(\mathbf{u} + \mathbf{x}) = \varphi(\mathbf{x}) = \mathbf{0}$ exactly when $\mathbf{x} = \mathbf{0}$, so its kernel is U. So the operators in question correspond bijectively to the isomorphisms $X \to W$. Fix a basis $\mathbf{x}_1, \dots, \mathbf{x}_r$ of X. A linear map $X \to W$ is determined by the list of values $\varphi(\mathbf{x}_1), \dots, \varphi(\mathbf{x}_r)$, which can be any list (Theorem 2.2.1), and it is an isomorphism exactly when this list is a basis of W (Proposition 4.5.5). There are $N_q(r,r)$ bases of W.

(b) By Theorem 4.4.27, an idempotent P satisfies $V = \operatorname{im} P \oplus \ker P$ and $P = \Pr_{\operatorname{im} P, \ker P}$. Conversely, for every pair (Y, Z) with $V = Y \oplus Z$, the projection $\Pr_{Y,Z}$ is idempotent with image Y and kernel Z (Proposition 4.4.26). So idempotents of rank r correspond bijectively to pairs (Y, Z) of complementary subspaces with $\dim Y = r$. There are $N_q(n,r)/N_q(r,r)$ subspaces Y of dimension r (Proposition 4.1.37), and each has $q^{r(n-r)}$ complements (Exercise 4.4.49(c)). Multiplying gives the formula.

For $n = 2$: $r = 0$ and $r = 2$ give 1 each (0 and I_2). For $r = 1$, $N_q(2,1)/N_q(1,1) = \left(q^2 - 1\right)/(q-1) = q + 1$, times $q^{1 \cdot 1} = q$. The total is $q^2 + q + 2$, as in Problem 2.8(b) for $q = p$ prime.

(c) An operator of rank r has a kernel of dimension $n-r$ and an image of dimension r (Theorem 4.2.11), and by (a) each such pair is realized by $N_q(r,r)$ operators. The number of $(n-r)$-dimensional subspaces equals the number of r-dimensional ones (Exercise 4.1.54(b)), namely $N_q(n,r)/N_q(r,r)$. So the count is

$$\left(N_q \frac{n,r}{N_q}(r,r)\right)^2 \cdot N_q(r,r) = N_q \frac{(n,r)^2}{N_q}(r,r).$$

For $n=q=2$: $r=0$ gives 1; $r=1$ gives $N_2(2,1)^2/N_2(1,1) = 3^2/1 = 9$; $r=2$ gives $N_2(2,2) = 3\cdot 2 = 6$. The total $16 = 2^4$ is the number of 2×2 matrices over $\mathbb{F}_2$.

Solution to Problem 4.11

Uses. Theorem 4.3.3 (Rank of a product), Exercise 4.3.30, Theorem 2.4.34 (Block multiplication), Proposition 2.4.14 (Algebra of matrices)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. $(AB)^2 = A(BA)B$. Bound $\operatorname{rank}\big((AB)^2\big)$ from above by $\operatorname{rank}(BA)$ and from below by Sylvester's inequality.
Hint 2. For (c), take $n=2m$ and block matrices built from I_m and 0, with $AB\neq 0$ and $BA=0$.
Sketch. (a) $\operatorname{rank}(BA) \geq \operatorname{rank}(A(BA)B) = \operatorname{rank}\big((AB)^2\big) \geq 2\operatorname{rank}(AB) - n$. (b) With $a = \operatorname{rank}(AB)$, $a - \operatorname{rank}(BA) \leq \min(a, n-a) \leq n/2$, and symmetrically. (c) $A = \left(\begin{smallmatrix} 0 & I_m \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 0 \\ 0 & I_m \end{smallmatrix}\right)$ give $AB = A$ of rank m and $BA = 0$; for odd n add a zero row and column.

(a) By associativity, $(AB)^2 = A(BA)B$, so $\operatorname{rank}\big((AB)^2\big) \leq \operatorname{rank}(BA)$ by Theorem 4.3.3 (applied twice). By Sylvester's inequality (Exercise 4.3.30(a)) for the product of AB with itself, $\operatorname{rank}\big((AB)^2\big) \geq 2\operatorname{rank}(AB) - n$. Together, $\operatorname{rank}(BA) \geq 2\operatorname{rank}(AB) - n$.

(b) Let $a = \operatorname{rank}(AB)$. By (a), $a - \operatorname{rank}(BA) \leq n - a$, and trivially $a - \operatorname{rank}(BA) \leq a$. So $a - \operatorname{rank}(BA) \leq \min(a, n-a) \leq n/2$, and since the left side is an integer, it is at most $\lfloor n/2 \rfloor$. Exchanging the roles of A and B gives $\operatorname{rank}(BA) - \operatorname{rank}(AB) \leq \lfloor n/2 \rfloor$.

(c) Let $m = \lfloor n/2 \rfloor$; for $m = 0$ there is nothing to show. If $n = 2m$, let

$$A = \begin{pmatrix} 0 & I_m \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & I_m \end{pmatrix}.$$

By Theorem 2.4.34, $AB = \left(\begin{smallmatrix} 0 & I_m \\ 0 & 0 \end{smallmatrix}\right) = A$, of rank m, and $BA = \left(\begin{smallmatrix} 0 & 0 \\ 0 & I_m \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & I_m \\ 0 & 0 \end{smallmatrix}\right) = 0$. If $n = 2m+1$, use $\left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} B & 0 \\ 0 & 0 \end{smallmatrix}\right)$, whose products are $\left(\begin{smallmatrix} AB & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} BA & 0 \\ 0 & 0 \end{smallmatrix}\right)$. In both cases $\operatorname{rank}(AB) - \operatorname{rank}(BA) = m$.

For $n=2$ this is Example 4.3.4. The proof of (a) is the move of Problem 2.9: products in the other order are sandwiched inside powers, $(AB)^{k+1} = A(BA)^k B$, and a rank inequality controls the powers.

Solution to Problem 4.12

Uses. Theorem 4.3.3 (Rank of a product), Exercise 4.2.50, Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 4.5.17 (Change of basis), Theorem 2.4.34 (Block multiplication), Theorem 3.5.23 (One-sided inverses of square matrices), Proposition 4.3.1 (Invertible factors preserve rank), Lemma 4.8.2 (Ranks of block triangular matrices)
Techniques. Choose a basis adapted to a preserved subspace (Technique 4.5.36), Block decomposition (Technique 2.4.35)
Hint 1. (a) $\operatorname{rank} A = \operatorname{rank}\big(A^2 B\big) \leq \operatorname{rank} A^2 \leq \operatorname{rank} A$. Then use Exercise 4.2.50.
Hint 2. (b) In a basis of $\operatorname{im} A$ followed by a basis of $\ker A$, A becomes $\left(\begin{smallmatrix} A_1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ with A_1 invertible. Write B in the same basis in blocks and translate $A^2 B = A$.
Sketch. (a) The ranks are squeezed, and Exercise 4.2.50 turns $\operatorname{rank} A^2 = \operatorname{rank} A$ into $\operatorname{im} A \cap \ker A = \{\mathbf{0}\}$ and $\operatorname{im} A + \ker A = F^n$. (b) Conjugating by the adapted basis matrix S: $A' = \left(\begin{smallmatrix} A_1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, A_1 invertible, $B' = \left(\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right)$. $A'^2 B' = A'$ gives $A_1 B_{11} = I_r$ and $B_{12} = 0$, so $A'B' - B'A' = \left(\begin{smallmatrix} 0 & 0 \\ -B_{21}A_1 & 0 \end{smallmatrix}\right)$, whose square is 0. (c) Its rank is $\operatorname{rank} B_{21} \leq \min(r, n-r)$; take $A = \left(\begin{smallmatrix} I_k & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} I_k & 0 \\ C & 0 \end{smallmatrix}\right)$ with C of rank k.

(a) By Theorem 4.3.3, $\operatorname{rank} A = \operatorname{rank}\big(A^2 B\big) \leq \operatorname{rank} A^2 \leq \operatorname{rank} A$, so $\operatorname{rank} A^2 = \operatorname{rank} A$. By Exercise 4.2.50, this implies $\ker A \cap \operatorname{im} A = \{\mathbf{0}\}$ and $\ker A + \operatorname{im} A = F^n$, that is, $F^n = \operatorname{im} A \oplus \ker A$.

(b) If $r = 0$, then $A = 0$ and $AB - BA = 0$. If $r = n$, then A is invertible, $A^2B = A$ gives $AB = I_n$, so $B = A^{-1}$ and $AB - BA = 0$. So let $1 \le r \le n-1$, and let S have as columns a basis of $\operatorname{im} A$ followed by a basis of $\ker A$; by (a) and Theorem 4.4.13 these form a basis of F^n, so S is invertible. T_A maps $\operatorname{im} A$ into itself and kills $\ker A$, so by Proposition 4.5.35 and Theorem 4.5.17,

$$A' = S^{-1}AS = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix},$$

with $A_1 \in F^{r\times r}$ the matrix of T_A restricted to $\operatorname{im} A$. This restriction is injective, since its kernel is $\operatorname{im} A \cap \ker A = \{\mathbf{0}\}$, so A_1 is invertible. Write $B' = S^{-1}BS = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$. Conjugating $A^2B = A$ by S gives $A'^2B' = A'$, and by Theorem 2.4.34,

$$A'^2B' = \begin{pmatrix} A_1^2B_{11} & A_1^2B_{12} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since A_1 is invertible, $A_1B_{11} = I_r$ and $B_{12} = 0$. Then also $B_{11}A_1 = I_r$ (Theorem 3.5.23), and

$$A'B' = \begin{pmatrix} A_1B_{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \quad B'A' = \begin{pmatrix} B_{11}A_1 & 0 \\ B_{21}A_1 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ B_{21}A_1 & 0 \end{pmatrix}.$$

So $C' = A'B' - B'A' = \begin{pmatrix} 0 & 0 \\ -B_{21}A_1 & 0 \end{pmatrix}$, and $C'^2 = 0$ by block multiplication. Since $AB - BA = SC'S^{-1}$, we get $(AB - BA)^2 = SC'^2S^{-1} = 0$.

(c) For $r = 0$ and $r = n$, $AB - BA = 0$ by (b). Otherwise, by Proposition 4.3.1, $\operatorname{rank}(AB - BA) = \operatorname{rank} C'$. Exchanging the two block rows of C' (an invertible row operation) gives $\begin{pmatrix} -B_{21}A_1 & 0 \\ 0 & 0 \end{pmatrix}$, so $\operatorname{rank} C' = \operatorname{rank}(B_{21}A_1)$ by Lemma 4.8.2(b), and $\operatorname{rank}(B_{21}A_1) = \operatorname{rank} B_{21}$ because A_1 is invertible. Since $B_{21} \in F^{(n-r)\times r}$, its rank is at most $\min(r, n-r)$.

For the examples, let $2k \le n$, let $A = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$, and let $B = \begin{pmatrix} I_k & 0 \\ C & 0 \end{pmatrix}$, where $C \in F^{(n-k)\times k}$ has I_k in its top k rows and zeros below, so that $\operatorname{rank} C = k$; this needs $n - k \ge k$. Then $A^2B = AB = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} = A$, and $BA = \begin{pmatrix} I_k & 0 \\ C & 0 \end{pmatrix}$, so $AB - BA = \begin{pmatrix} 0 & 0 \\ -C & 0 \end{pmatrix}$ has rank k.

Everything followed from choosing the basis adapted to $F^n = \operatorname{im} A \oplus \ker A$. In it, A is invertible on one summand and 0 on the other, and the hypothesis $A^2B = A$ pins down two of the four blocks of B.

Solution to Problem 4.13

Uses. Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Theorem 4.1.16 (Extending to a basis), Theorem 4.5.17 (Change of basis), Proposition 4.5.27 (Similarity invariants), Lemma 2.4.24 (Matrix units), Proposition 2.4.27 (The trace is linear), Exercise 4.1.43, Proposition 1.4.14 (The span is the smallest subspace containing a list)
Techniques. Extend a basis of a subspace (Technique 4.1.18), Choose a basis adapted to a preserved subspace (Technique 4.5.36), Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. (a) N maps $\ker N^j$ into $\ker N^{j-1}$. Build a basis by taking a basis of $\ker N$, extending it to one of $\ker N^2$, and so on up to $\ker N^n = F^n$.
Hint 2. (b) U_{ij} for $i \ne j$ is nilpotent. Check that $U_{ii} - U_{jj} + U_{ij} - U_{ji}$ has square 0. (c) U_{ii} and $U_{ii} + U_{ij}$.
Sketch. (a) In a basis adapted to $\ker N \subseteq \ker N^2 \subseteq \cdots \subseteq \ker N^n = F^n$, each basis vector from level j is sent into $\ker N^{j-1}$, which is spanned by earlier basis vectors; so the matrix is strictly upper triangular, and $\operatorname{tr} N$ equals its trace, 0. (b) Nilpotents lie in the trace-zero space by (a). Conversely U_{ij} $(i \ne j)$ and $U_{ii} - U_{jj} = E - U_{ij} + U_{ji}$, with E of square 0, are combinations of nilpotents, and these matrices span the trace-zero space. (c) U_{ii} and $U_{ii} + U_{ij}$ are idempotent, so their span contains every matrix unit.

(a) Let $K_j = \ker N^j$, so $K_0 = \{\mathbf{0}\} \subseteq K_1 \subseteq \cdots \subseteq K_n$, and $K_n = F^n$ by Corollary 4.1.27. If $N^j\mathbf{x} = \mathbf{0}$, then $N^{j-1}(N\mathbf{x}) = \mathbf{0}$, so N maps K_j into K_{j-1}. Choose a basis of K_1, extend it to a basis of K_2 (Theorem 4.1.16, applied in K_2), extend that to a basis of K_3, and so on. The result is a basis $\mathcal{B} = \mathbf{b}_1, \ldots, \mathbf{b}_n$ of F^n such that, for each j, the first $\dim K_j$ vectors form a basis of K_j. If $\mathbf{b}_l$ was added at level j, that is, $\mathbf{b}_l \in K_j$ and $l > \dim K_{j-1}$, then $N\mathbf{b}_l \in K_{j-1}$ is a combination of $\mathbf{b}_1, \ldots, \mathbf{b}_{\dim K_{j-1}}$, all of which come before $\mathbf{b}_l$. So column l of

$[T_N]_{\mathcal{B}\mathcal{B}}$ has nonzero entries only in rows $< l$: the matrix is strictly upper triangular, and by Theorem 4.5.17 it equals $S^{-1}NS$ for the invertible S with columns $\mathcal{B}$. Its diagonal is zero, so $\operatorname{tr} N = \operatorname{tr}\left(S^{-1}NS\right) = 0$ (Proposition 4.5.27(b)).

(b) Let W be the trace-zero subspace. By (a) and Proposition 2.4.27, every nilpotent matrix lies in W, and so does their span. Conversely, for $i \neq j$, the matrix unit U_{ij} satisfies $U_{ij}^2 = 0$ (Lemma 2.4.24(c)), so it is nilpotent. Let $E = U_{ii} - U_{jj} + U_{ij} - U_{ji}$. By Lemma 2.4.24(c), E maps $\mathbf{e}_i$ to $\mathbf{e}_i - \mathbf{e}_j$ and $\mathbf{e}_j$ to $\mathbf{e}_i - \mathbf{e}_j$, and kills the other standard basis vectors. So $E\left(\mathbf{e}_i - \mathbf{e}_j\right) = \mathbf{0}$, $E^2\mathbf{e}_i = E^2\mathbf{e}_j = \mathbf{0}$, and $E^2 = 0$. (In rows and columns i, j, E is $\left(\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right)$.) Hence

$$U_{ii} - U_{jj} = E - U_{ij} + U_{ji}$$

is a combination of nilpotent matrices. The matrices U_{ij} with $i \neq j$ and $U_{ii} - U_{nn}$ with $i < n$ lie in W, number $n^2 - 1 = \dim W$ (Exercise 4.1.43), and are independent, since their entries sit in different positions except for the last diagonal entry. So they form a basis of W, and W is contained in the span of the nilpotent matrices.

(c) $U_{ii}^2 = U_{ii}$, and for $i \neq j$, $\left(U_{ii} + U_{ij}\right)^2 = U_{ii} + U_{ij} + U_{ij}U_{ii} + U_{ij}^2 = U_{ii} + U_{ij}$, because $U_{ij}U_{ii} = 0$ and $U_{ij}U_{ij} = 0$ when $i \neq j$. So every matrix unit $U_{ij} = \left(U_{ii} + U_{ij}\right) - U_{ii}$ is a combination of idempotents, and the matrix units span $F^{n\times n}$.

Nilpotent matrices are trapped in a hyperplane because their trace is 0; idempotents are not, since $\operatorname{tr} P = \operatorname{rank} P \cdot 1$ can be anything. Neither set is a subspace: a sum of two nilpotent matrices can be invertible (Problem 2.10(b)).

Solution to Problem 4.14

> *Uses.* Proposition 4.4.12 (Concatenating independent lists), Proposition 4.1.34 (Size of a vector space over a finite field), Corollary 4.1.35 (Counting independent lists and bases), Proposition 4.1.37 (Counting subspaces), Exercise 4.4.49, Proposition 1.5.14 (Appending a vector)
> *Techniques.* Double counting (Technique 4.1.36)
> *Hint 1.* Count pairs (W, basis of W). A list $\mathbf{w}_1, \ldots, \mathbf{w}_j$ is a basis of such a W exactly when a basis of U followed by the $\mathbf{w}$'s is independent.
> *Hint 2.* Choose $\mathbf{w}_i$ outside $\operatorname{span}(U, \mathbf{w}_1, \ldots, \mathbf{w}_{i-1})$, which has q^{k+i-1} elements.
> *Sketch.* The lists $\mathbf{w}_1, \ldots, \mathbf{w}_j$ for which U's basis followed by them is independent number $\prod_{i=0}^{j-1}\left(q^n - q^{k+i}\right) = q^{jk}N_q(n-k, j)$. Each such list spans a W with $W \cap U = \{\mathbf{0}\}$, every basis of such a W is such a list, and each W has $N_q(j, j)$ bases. Divide. $j = n - k$: $q^{k(n-k)}$ complements. $j = 1$: $q^k\left(q^{n-k} - 1\right)/(q-1)$, all lines not in U. In $\mathbb{F}_2^3$ with $k = 1$: 6 lines and 4 planes.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ be a basis of U. For a list $\mathcal{W} = \mathbf{w}_1, \ldots, \mathbf{w}_j$, the concatenation $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathcal{W}$ is independent if and only if $\mathcal{W}$ is independent and $\operatorname{span}(\mathcal{W}) \cap U = \{\mathbf{0}\}$ (Proposition 4.4.12). So the lists $\mathcal{W}$ with this property are exactly the bases of the j-dimensional subspaces W with $W \cap U = \{\mathbf{0}\}$, and each such W has $N_q(j, j)$ bases (Corollary 4.1.35). Let L be the number of such lists; then the number of subspaces is $L/N_q(j, j)$.

To count the lists, choose $\mathbf{w}_1, \mathbf{w}_2, \ldots$ in turn. Given $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$ with $\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$ independent, the list stays independent after appending $\mathbf{w}_i$ exactly when $\mathbf{w}_i$ is not in the span of the list so far (Proposition 1.5.14), a subspace of dimension $k + i - 1$ with q^{k+i-1} elements (Proposition 4.1.34). So

$$L = \prod_{i=1}^{j}\left(q^n - q^{k+i-1}\right) = \prod_{i=0}^{j-1} q^k\left(q^{n-k} - q^i\right) = q^{jk}N_q(n-k, j),$$

and the number of subspaces is $q^{jk}N_q(n-k, j)/N_q(j, j)$.

Checks. For $j = n - k$, the factor $N_q(n-k, n-k)/N_q(n-k, n-k)$ is 1, and the count $q^{k(n-k)}$ is the number of complements of U (Exercise 4.4.49(c)). For $j = 1$ the count is $q^k\left(q^{n-k} - 1\right)/(q-1) = \left(q^n - q^k\right)/(q-1)$: the $q^n - q^k$ vectors outside U, grouped into lines of $q - 1$ nonzero vectors each, which is every line not

contained in U. In $\mathbb{F}_2^3$ with $k=1$: $j=1$ gives $2\cdot 3/1=6$ lines, which are the 7 lines minus U itself; $j=2$ gives $2^2\cdot 1=4$ planes, which are the 7 planes minus the 3 planes that contain U.

Solution to Problem 4.15

Uses. Corollary 4.7.23 (Affinely independent lists are short), Exercise 4.7.59, Proposition 4.7.30 (Convex sets and convex hulls), Proposition 4.7.32 (Barycentric coordinates on a simplex), Definition 4.7.28 (Convex combination, convex set, convex hull)
Techniques. Too many vectors force a dependence (Technique 4.1.30)
Hint 1. Induct on m. For $m\ge n+2$, the induction hypothesis gives, for each i, a point $\mathbf{p}_i$ in every C_j with $j\neq i$.
Hint 2. The $m\ge n+2$ points $\mathbf{p}_i$ are affinely dependent. Split an affine dependence by the signs of its coefficients, as in Radon's theorem, and show that the resulting point lies in every C_l.
Sketch. (a) Induction on m, the case $m=n+1$ being the hypothesis. For $m\ge n+2$, pick $\mathbf{p}_i\in\bigcap_{j\neq i}C_j$. An affine dependence $\sum d_i\mathbf{p}_i=\mathbf{0}$, $\sum d_i=0$, gives $\mathbf{q}=\sum_{i\in I}(d_i/s)\mathbf{p}_i=\sum_{j\in J}(-d_j/s)\mathbf{p}_j$ with $I=\{d_i>0\}$. For each l, l misses I or J, and the points indexed by the other set all lie in C_l, so $\mathbf{q}\in C_l$. (b) The $n+1$ facets of an n-simplex: any n share a vertex, all $n+1$ share nothing.

(a) We induct on m. For $m=n+1$ the hypothesis says exactly that all the sets have a common point. Let $m\ge n+2$, and assume the statement for $m-1$ sets. For each i, the $m-1\ge n+1$ sets C_j with $j\neq i$ satisfy the hypothesis, so by induction there is a point $\mathbf{p}_i$ in every C_j with $j\neq i$.

The $m\ge n+2$ points $\mathbf{p}_1,\dots,\mathbf{p}_m$ are affinely dependent (Corollary 4.7.23): there are $d_1,\dots,d_m$, not all 0, with $\sum_i d_i\mathbf{p}_i=\mathbf{0}$ and $\sum_i d_i=0$. As in the solution of Exercise 4.7.59, let $I=\{i\mid d_i>0\}$ and $J=\{j\mid d_j\le 0\}$; both are nonempty, $s=\sum_{i\in I}d_i>0$, and

$$\mathbf{q}=\sum_{i\in I}\frac{d_i}{s}\mathbf{p}_i=\sum_{j\in J}\frac{-d_j}{s}\mathbf{p}_j,$$

where both sides are convex combinations. Now fix l. If $l\notin I$, then every $\mathbf{p}_i$ with $i\in I$ lies in C_l (because $i\neq l$), so the left side is a convex combination of points of the convex set C_l, and $\mathbf{q}\in C_l$ (Proposition 4.7.30(a)). If $l\in I$, then $l\notin J$, and the right side shows $\mathbf{q}\in C_l$ in the same way. So $\mathbf{q}$ lies in all m sets.

(b) Let $\mathbf{v}_0,\dots,\mathbf{v}_n$ be affinely independent (an affine frame of V) and $C_i=\operatorname{conv}\{\mathbf{v}_j\mid j\neq i\}$ for $i=0,\dots,n$, the $n+1$ facets of the simplex. Any n of them, say all except C_l, contain $\mathbf{v}_l$. But a point in all $n+1$ would lie in the simplex with every barycentric coordinate 0 (Proposition 4.7.32: a point of C_i has ith coordinate 0), while the coordinates add up to 1.

The proof is "too many vectors" in affine form: as soon as there are more points than an affine frame can hold, they satisfy an affine dependence, and the dependence, split by signs, produces the common point.

Solution to Problem 4.16

Uses. Theorem 4.3.3 (Rank of a product), Theorem 3.5.14 (Third Fundamental Theorem, completed), Theorem 4.3.21 (Column–row factorization), Proposition 2.4.14 (Algebra of matrices)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Compute M^2 and rank M.
Hint 2. $(BA)^3=B(AB)^2A$. Show that BA is invertible by comparing ranks, then cancel.
Sketch. (a) $M^2=9M$ and rank $M=2$. Then $\operatorname{rank}(BA)\ge\operatorname{rank}(A(BA)B)=\operatorname{rank}(M^2)=2$, so $BA\in\mathbb{R}^{2\times 2}$ is invertible, and $(BA)^3=BM^2A=9(BA)^2$ gives $BA=9I_2$. (b) The column–row factorization $A=\begin{pmatrix}8&2\\2&5\\-2&4\end{pmatrix}$, $B=\begin{pmatrix}1&0&-1/2\\0&1&1\end{pmatrix}$.

(a) A direct computation gives

$$M^2=\begin{pmatrix}72&18&-18\\18&45&36\\-18&36&45\end{pmatrix}=9M.$$

The first two columns of M are not proportional, and the third is $(-2,4,5)=-\frac{1}{2}(8,2,-2)+(2,5,4)$, so $\operatorname{rank}M=2$. By associativity, $M^2=A(BA)B$, so by Theorem 4.3.3,

$$2 = \operatorname{rank}(9M) = \operatorname{rank}\left(M^2\right) \leq \operatorname{rank}(BA) \leq 2.$$

So $BA \in \mathbb{R}^{2\times 2}$ has rank 2 and is invertible (Theorem 3.5.14). Now

$$(BA)^3 = B(AB)(AB)A = BM^2A = 9BMA = 9BABA = 9(BA)^2,$$

and multiplying by $\left((BA)^{-1}\right)^2$ gives $BA = 9I_2$.

(b) By Theorem 4.3.21, $M = CR$ with C the pivot columns of M and R the nonzero rows of $\operatorname{rref}(M)$:

$$A = \begin{pmatrix} 8 & 2 \\ 2 & 5 \\ -2 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 1 \end{pmatrix}.$$

Indeed AB has columns $(8,2,-2)$, $(2,5,4)$, and $-\frac{1}{2}(8,2,-2) + (2,5,4) = (-2,4,5)$, and

$$BA = \begin{pmatrix} 8+1 & 2-2 \\ 2-2 & 5+4 \end{pmatrix} = 9I_2.$$

The identity $M^2 = 9M$ says that $M/9$ is idempotent of rank 2, and $BA/9$ is the corresponding identity on the 2-dimensional space through which M factors. This is Problem B6 of the 1969 Putnam Competition.

Solution to Problem 4.17

> *Uses.* Theorem 4.2.11 (Rank–nullity), Theorem 3.5.14 (Third Fundamental Theorem, completed), Theorem 4.1.10 (Independent lists are short), Proposition 2.6.18 (The column space is spanned by the columns), Notation 1.1.25 ($\mathbb{F}_p$)
> *Techniques.* Try small cases and small fields (Technique 1.2.38), Dimension counting (Technique 4.2.24)
> *Hint 1.* Show that every row of M_n has n entries 1 and n entries -1, so that $M_n\mathbf{1} = \mathbf{0}$ for $\mathbf{1} = (1,\ldots,1)$.
> *Hint 2.* For the lower bound, reduce the top left $2n \times 2n$ block modulo 2: it becomes $J - I$, where J is the all-ones matrix. Compute $(J-I)^2$ over $\mathbb{F}_2$. Then show that integer columns that are independent modulo 2 are independent over $\mathbb{Q}$.
> *Sketch.* a_{ij} depends only on $(i-j) \bmod (2n+1)$: it is 1 for residues $1,\ldots,n$ and -1 for $n+1,\ldots,2n$. So each row has n entries of each sign, $M_n\mathbf{1} = \mathbf{0}$, and $\operatorname{rank} M_n \leq 2n$. The top left $2n \times 2n$ block M' reduces modulo 2 to $J - I$, and $(J-I)^2 = (2n-2)J + I = I$ over $\mathbb{F}_2$. A rational relation among the columns of M' can be scaled to an integer relation with some odd coefficient, which reduces to a nontrivial relation over $\mathbb{F}_2$; impossible. So M' is invertible over $\mathbb{Q}$, the first $2n$ columns of M_n are independent, and $\operatorname{rank} M_n = 2n$.

Index rows and columns by $1,\ldots,2n+1$. For $i > j$, the entry a_{ij} lies on the $(i-j)$th diagonal below the main one, so $a_{ij} = 1$ if $i - j \leq n$ and $a_{ij} = -1$ if $i - j \geq n+1$. For $i < j$, $a_{ij} = -a_{ji}$ is -1 if $j - i \leq n$ and 1 if $j - i \geq n+1$. In both cases, with ρ the remainder of $i - j$ modulo $2n+1$ (so $\rho \in \{1,\ldots,2n\}$ for $i \neq j$; for $i < j$, $\rho = 2n+1-(j-i)$),

$$a_{ij} = 1 \text{ if } 1 \leq \rho \leq n, \quad a_{ij} = -1 \text{ if } n+1 \leq \rho \leq 2n.$$

Upper bound. In row i, as j runs over the $2n$ indices different from i, the remainder ρ runs over $1,\ldots,2n$, each once. So row i has n entries equal to 1 and n equal to -1, and $M_n\mathbf{1} = \mathbf{0}$ for $\mathbf{1} = (1,\ldots,1)$. Hence $\dim\ker M_n \geq 1$ and $\operatorname{rank} M_n \leq 2n$ (Theorem 4.2.11).

Lower bound. Let M' be the top left $2n \times 2n$ block of M_n. Its entries are integers, 0 on the diagonal and ± 1 elsewhere, so its reduction modulo 2 is $J - I$ over $\mathbb{F}_2$, where J is the all-ones matrix of size $2n$. Since $J^2 = (2n)J$, over $\mathbb{F}_2$

$$(J-I)^2 = J^2 - 2J + I = (2n-2)J + I = I,$$

so $J - I$ is invertible over $\mathbb{F}_2$ and its columns are independent there (Theorem 3.5.14). Now suppose the columns $\mathbf{c}_1,\ldots,\mathbf{c}_{2n}$ of M' satisfy $\sum_j t_j\mathbf{c}_j = \mathbf{0}$ with $t_j \in \mathbb{Q}$, not all 0. Multiplying by a common denominator, we may take the t_j to be integers, and dividing by the largest power of 2 that divides all of them, we may assume that some t_j is odd. Reducing the integer equation $\sum_j t_j\mathbf{c}_j = \mathbf{0}$ modulo 2 gives a relation among the columns of $J - I$ over $\mathbb{F}_2$ with a coefficient 1, which is impossible. So the columns of M' are independent

over $\mathbb{Q}$. A relation among the first $2n$ columns of M_n is, in its first $2n$ rows, a relation among the columns of M', so the first $2n$ columns of M_n are independent, and $\operatorname{rank} M_n \geq 2n$ (Proposition 2.6.18, Theorem 4.1.10).

So $\operatorname{rank} M_n = 2n$, over $\mathbb{Q}$ and therefore also over $\mathbb{R}$ (Gauss–Jordan elimination never leaves $\mathbb{Q}$). The same argument shows that M_n has rank $2n$ over $\mathbb{F}_2$ as well. Reducing modulo a prime can only lower the rank of an integer matrix, so a lower bound modulo p is a lower bound over $\mathbb{Q}$; this is the reason to look at the matrix modulo 2, where it becomes the very simple $J - I$. This is Problem B5 of the 1988 Putnam Competition.

Solution to Problem 4.18

Uses. Theorem 4.2.11 (Rank–nullity), Theorem 4.6.27 (Dimension of a quotient), Theorem 4.6.34 (First isomorphism theorem), Exercise 4.6.49, Exercise 4.6.50, Proposition 4.6.28 (Lifting a basis of a quotient), Proposition 4.2.3 (Properties of restriction), Example 2.6.27 (Differentiation on $F[x]$), Example 4.8.3 (Sylvester's inequality, three ways)
Techniques. Pass to a quotient (Technique 4.6.31), Restrict to a subspace (Technique 4.2.5)
Hint 1. (a) $\dim(W/\operatorname{im} T) = \dim W - \operatorname{rank} T$. For (c), compute $\dim\ker(ST)$ by restricting T to $\ker(ST)$: it maps onto $\ker S \cap \operatorname{im} T$ with kernel $\ker T$.
Hint 2. For the cokernel, use $\operatorname{im}(ST) \subseteq \operatorname{im} S \subseteq W$ and show $\operatorname{im} S/\operatorname{im}(ST) \cong V/(\operatorname{im} T + \ker S)$. For nested $A \subseteq B \subseteq X$, $\dim(X/A) = \dim(X/B) + \dim(B/A)$ by the third isomorphism theorem.
Sketch. (a) Rank–nullity and $\dim(W/\operatorname{im} T) = \dim W - \operatorname{rank} T$. (b) $\ker M = \{\mathbf{0}\}$, $\mathbb{Q}[x]/x\mathbb{Q}[x] \cong \mathbb{Q}$ by ev_0: $\operatorname{ind} M = -1$. $\ker D$ = constants and D is onto: $\operatorname{ind} D = 1$. (c) With $k = \dim(\ker S \cap \operatorname{im} T)$: $\dim\ker(ST) = \dim\ker T + k$. And $\dim(W/\operatorname{im} ST) = \dim(W/\operatorname{im} S) + \dim(V/\operatorname{im} T) - \dim\ker S + k$, using $\operatorname{im} S/\operatorname{im}(ST) \cong V/(\operatorname{im} T + \ker S)$ and $(\operatorname{im} T + \ker S)/\operatorname{im} T \cong \ker S/(\ker S \cap \operatorname{im} T)$. Subtract. $DM = \mathrm{id} + MD$ is bijective (index 0), and MD has kernel and cokernel of dimension 1.

We use one fact about nested subspaces $A \subseteq B \subseteq X$: if X/B and B/A are finite-dimensional, so is X/A, and

$$\dim(X/A) = \dim(X/B) + \dim(B/A). \tag{C.1.7}$$

Indeed B/A is a subspace of X/A with quotient $(X/A)/(B/A) \cong X/B$ (Exercise 4.6.50(c)), so Proposition 4.6.28 applies. Also, a subspace or a quotient of a finite-dimensional space is finite-dimensional.

(a) $\ker T$ is finite-dimensional, and $\dim(W/\operatorname{im} T) = \dim W - \operatorname{rank} T$ (Theorem 4.6.27). By rank–nullity, $\operatorname{ind} T = \dim\ker T - \dim W + \operatorname{rank} T = \dim V - \dim W$.

(b) M is injective, so $\ker M = \{\mathbf{0}\}$. Its image is $x\mathbb{Q}[x]$, the polynomials with constant coefficient 0, which is the kernel of the surjective map $\mathrm{ev}_0 : \mathbb{Q}[x] \to \mathbb{Q}$, $p \mapsto p(0)$. By the first isomorphism theorem (Theorem 4.6.34), $\mathbb{Q}[x]/x\mathbb{Q}[x] \cong \mathbb{Q}$. So M is Fredholm with $\operatorname{ind} M = 0 - 1 = -1$. Over $\mathbb{Q}$, $\ker D$ is the space of constants (Example 2.6.27), and D is surjective, since $x^k = D\big(x^{k+1}/(k+1)\big)$. So $\operatorname{ind} D = 1 - 0 = 1$.

(c) Let $k = \dim(\ker S \cap \operatorname{im} T)$, which is finite since $\ker S$ is finite-dimensional.

The kernel. $\ker(ST) = \{\mathbf{u} \mid T\mathbf{u} \in \ker S\}$. So T restricts to a linear map $R : \ker(ST) \to \ker S \cap \operatorname{im} T$, which is onto: if $\mathbf{y} = T\mathbf{u} \in \ker S$, then $\mathbf{u} \in \ker(ST)$. Its kernel is $\ker T$, which is contained in $\ker(ST)$. By the first isomorphism theorem, $\ker(ST)/\ker T \cong \ker S \cap \operatorname{im} T$, and by Proposition 4.6.28, $\ker(ST)$ is finite-dimensional with

$$\dim\ker(ST) = \dim\ker T + k.$$

The cokernel. We have $\operatorname{im}(ST) \subseteq \operatorname{im} S \subseteq W$. The map $\sigma : V \to \operatorname{im} S/\operatorname{im}(ST)$, $\mathbf{v} \mapsto S\mathbf{v} + \operatorname{im}(ST)$, is linear and surjective. Its kernel is the set of $\mathbf{v}$ with $S\mathbf{v} = ST\mathbf{u}$ for some $\mathbf{u}$, that is, with $\mathbf{v} - T\mathbf{u} \in \ker S$; so $\ker\sigma = \operatorname{im} T + \ker S$, and $\operatorname{im} S/\operatorname{im}(ST) \cong V/(\operatorname{im} T + \ker S)$. Apply (C.1.7) to $\operatorname{im} T \subseteq \operatorname{im} T + \ker S \subseteq V$. The space $V/\operatorname{im} T$ is finite-dimensional, hence so are its subspace $(\operatorname{im} T + \ker S)/\operatorname{im} T$ and its quotient $V/(\operatorname{im} T + \ker S)$, and

$$\dim(V/(\operatorname{im} T + \ker S)) = \dim(V/\operatorname{im} T) - \dim((\operatorname{im} T + \ker S)/\operatorname{im} T).$$

By the second isomorphism theorem (Exercise 4.6.49(a)), $(\operatorname{im} T + \ker S)/\operatorname{im} T \cong \ker S/(\ker S \cap \operatorname{im} T)$, of dimension $\dim\ker S - k$ (Theorem 4.6.27). So

$$\dim(\operatorname{im} S/\operatorname{im}(ST)) = \dim(V/\operatorname{im} T) - \dim\ker S + k.$$

Applying (C.1.7) to $\operatorname{im}(ST) \subseteq \operatorname{im} S \subseteq W$ shows that $W/\operatorname{im}(ST)$ is finite-dimensional, with

$$\dim(W/\operatorname{im}(ST)) = \dim(W/\operatorname{im} S) + \dim(V/\operatorname{im} T) - \dim\ker S + k.$$

The index. Subtracting,

$$\operatorname{ind}(ST) = \dim\ker T + k - \dim(W/\operatorname{im} S) - \dim(V/\operatorname{im} T) + \dim\ker S - k = \operatorname{ind} T + \operatorname{ind} S.$$

The check. $DM(p) = D(xp) = p + xD(p)$, so $DM\left(x^j\right) = (j+1)x^j$. Over $\mathbb{Q}$ this is a bijection on the basis of monomials up to nonzero factors, so DM is invertible and $\operatorname{ind}(DM) = 0 = \operatorname{ind} D + \operatorname{ind} M$. For MD, the kernel is $\ker D$ (constants, dimension 1), and the image is $M(D(\mathbb{Q}[x])) = M(\mathbb{Q}[x]) = x\mathbb{Q}[x]$, with 1-dimensional quotient. So $\operatorname{ind}(MD) = 1 - 1 = 0$, as the formula predicts.

For operators on a finite-dimensional space the index is 0 by (a), and that is the content of "injective if and only if surjective". On $\mathbb{Q}[x]$ the index of M is -1, which measures exactly how far M is from that statement: it is injective but misses a 1-dimensional space. The kernel computation in (c) is the kernel proof of Sylvester's inequality (Example 4.8.3), done with equality.

Section 5.1

Solution to Exercise 5.1.5

Uses. Definition 5.1.1 (Determinant function), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Hint 1. The columns of cA are $c\mathbf{a}_1, \ldots, c\mathbf{a}_n$. Pull the scalars out one column at a time.
Hint 2. For the zero column, a linear map sends $\mathbf{0}$ to 0. Alternatively, $\mathbf{0} = 0 \cdot \mathbf{0}$.
Sketch. Linearity in column j pulls out one factor c, for each of the n columns. A zero column $\mathbf{0} = 0 \cdot \mathbf{0}$ pulls out a factor 0.

The columns of cA are $c\mathbf{a}_1, \ldots, c\mathbf{a}_n$. By linearity in the first column, then the second, and so on,

$$D(c\mathbf{a}_1, c\mathbf{a}_2, \ldots, c\mathbf{a}_n) = cD(\mathbf{a}_1, c\mathbf{a}_2, \ldots, c\mathbf{a}_n) = \cdots = c^n D(\mathbf{a}_1, \ldots, \mathbf{a}_n).$$

Taking $c = -1$ gives $D(-A) = (-1)^n D(A)$. If the j-th column of A is $\mathbf{0}$, then $\mathbf{0} = 0 \cdot \mathbf{0}$, and linearity in column j gives $D(A) = 0 \cdot D(A) = 0$. (Equivalently, the linear map $\mathbf{x} \mapsto D(\ldots, \mathbf{x}, \ldots)$ sends $\mathbf{0}$ to 0 by Proposition 2.1.4.)

Solution to Exercise 5.1.16

Uses. Definition 5.1.14 (The 2×2 determinant), Definition 1.1.19 (Integers modulo n), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$)
Hint 1. Over $\mathbb{Q}$, $\det A = 3 \cdot 2 - 5 \cdot 4$.
Hint 2. Reducing the entries modulo p and then computing gives the same result as computing over $\mathbb{Z}$ and then reducing, by Lemma 1.1.21.
Sketch. Over $\mathbb{Q}$, $\det A = -14$. Over $\mathbb{F}_p$ it is $-14 \bmod p$, which is 0 exactly when $p \mid 14$, that is, $p = 2$ or $p = 7$.

Over $\mathbb{Q}$, $\det A = 3 \cdot 2 - 5 \cdot 4 = -14$. By Lemma 1.1.21, computing $ad - bc$ in $\mathbb{F}_p$ from the remainders of the entries gives the remainder of -14 modulo p. In $\mathbb{F}_7$ this is 0, since $7 \mid 14$. In $\mathbb{F}_2$ the matrix is $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, with determinant $1 \cdot 0 - 1 \cdot 0 = 0$, as it should be since $2 \mid 14$. In general $\det A = 0$ over $\mathbb{F}_p$ exactly when $p \mid 14$, that is, for $p = 2$ and $p = 7$.

Solution to Exercise 5.1.19

Uses. Definition 5.1.1 (Determinant function), Theorem 5.1.13 (The 2×2 determinant), Proposition 2.1.4 (Linear maps preserve $\mathbf{0}$ and negatives)
Hint 1. Normalization and alternation are a single evaluation and a substitution $b = a$, $d = c$. For multilinearity, fix one column and ask whether the value is a linear function of the entries of the other column.
Hint 2. A nonzero constant term, or a product of two entries from the same column, prevents linearity in that column.
Hint 3. Over $\mathbb{F}_2$, $-1 = 1$. Which of the four functions is affected?
Sketch. Over $\mathbb{Q}$: (a) fails all three axioms; (b) is multilinear only; (c) is $-\det$, which is multilinear and alternating but not normalized; (d) is normalized and alternating, but contains the product ac of two entries of the first column, so it is not multilinear. Over $\mathbb{F}_2$ only (c) changes: there $-\det = \det$ is a determinant function.

(a) $D(I_2) = 2 \neq 1$ and $D\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1 \neq 0$, so D is neither normalized nor alternating. For the fixed second column $(0, 1)$, the map $(a, c) \mapsto a + 1$ sends $\mathbf{0}$ to 1, so it is not linear (Proposition 2.1.4), and D is not multilinear.

(b) For a fixed second column (b, d), $(a, c) \mapsto ba - dc$ is linear, and for a fixed first column (a, c), $(b, d) \mapsto ab - cd$ is linear, so D is multilinear. But $D(I_2) = 1 \cdot 0 - 0 \cdot 1 = 0$ and $D\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1$, so D is neither normalized nor alternating.

(c) $D = -\det$. By Theorem 5.1.13, det is multilinear and alternating, and negating both sides of each defining identity shows that $-\det$ is too. But $D(I_2) = -1 \neq 1$.

(d) $D(I_2) = (1 - 0)(1 - 0) = 1$, and if the columns are equal, then $a = b$ and $D(A) = 0$. So D is normalized and alternating. For the fixed second column $(0, 0)$, $D(A) = a(-c) = -ac$. The map $(a, c) \mapsto -ac$ is not linear: it sends $(1, 1)$ to -1 and $(2, 2) = 2(1, 1)$ to $-4 \neq 2 \cdot (-1)$. So D is not multilinear.

Over $\mathbb{F}_2$. In (a), $D(I_2) = 1 + 1 = 0$, and the other two failures are unchanged. In (b) nothing changes. In (d), the map $(a, c) \mapsto -ac = ac$ sends $(1, 0)$ and $(0, 1)$ to 0 but their sum $(1, 1)$ to 1, so it is still not linear. In (c), $-1 = 1$, so $D = -\det = \det$ is a determinant function. So only the answer to (c) changes.

Solution to Exercise 5.1.20

Uses. Definition 5.1.1 (Determinant function), Proposition 5.1.10 (Alternating implies antisymmetric), Lemma 5.1.8 (Column replacement)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. In (a) and (b), every column is a multiple of a standard basis vector. Pull out the scalars, then sort the columns into the order $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ by interchanges.
Hint 2. In (c), use Lemma 5.1.8 to clear the entries above the diagonal, working from the left.
Sketch. (a) $D = 5 \cdot 3 \cdot 2 \cdot D(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) = -30$, one interchange. (b) $D(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) = 1$, two interchanges. (c) Column replacements turn A into I_3, so $D(A) = 1$.

(a) The columns are $5\mathbf{e}_3$, $3\mathbf{e}_2$, and $2\mathbf{e}_1$. By linearity in each column and then Proposition 5.1.10 (interchanging columns 1 and 3),

$$D(A) = 30D(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) = -30D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = -30.$$

(b) The columns are $\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2$. Interchanging columns 1 and 2, then columns 2 and 3,

$$D(A) = D(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) = -D(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2) = D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1.$$

(c) The columns are $\mathbf{e}_1$, $2\mathbf{e}_1 + \mathbf{e}_2$, and $3\mathbf{e}_1 + 4\mathbf{e}_2 + \mathbf{e}_3$. By Lemma 5.1.8, subtracting 2 times column 1 from column 2, then 3 times column 1 from column 3, does not change D, and leaves the columns $\mathbf{e}_1$, $\mathbf{e}_2$, $4\mathbf{e}_2 + \mathbf{e}_3$. Subtracting 4 times the new column 2 from column 3 leaves $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. So $D(A) = D(I_3) = 1$.

Solution to Exercise 5.1.21

Uses. Definition 5.1.1 (Determinant function), Exercise 5.1.5, Remark 5.1.4
Hint 1. Linearity in column j splits column j while the other columns stay fixed. Carry out the splitting honestly for $n = 2$.
Hint 2. Test the claim with $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.
Sketch. For $n \geq 2$, splitting each column separately produces 2^n terms, not 2; the proof keeps only two of them. For $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $\det(A + B) = 1$ but $\det A + \det B = 0$.

The middle equality is wrong. Linearity in column j applies when the other columns are held fixed, so it splits one column at a time. For $n = 2$, splitting column 1 and then column 2 gives four terms:

$$D(A + B) = D(\mathbf{a}_1, \mathbf{a}_2) + D(\mathbf{a}_1, \mathbf{b}_2) + D(\mathbf{b}_1, \mathbf{a}_2) + D(\mathbf{b}_1, \mathbf{b}_2),$$

and in general there are 2^n terms. The argument silently drops all of them except $D(A)$ and $D(B)$. The claim is false for every $n \geq 2$. For $n = 2$, take $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$: then $A + B = I_2$, so $D(A + B) = 1$, while $D(A) = D(B) = 0$ because each has a zero column (Exercise 5.1.5). (For $n = 1$ the claim is true, since $D(a) = a$ by Example 5.1.2.)

Solution to Exercise 5.1.22

Uses. Proposition 5.1.17 (Dependent columns), Theorem 5.1.13 (The 2×2 determinant)

Hint 1. Compare the logical form of the proposition with the way it is used.
Hint 2. Is there a multilinear alternating D that vanishes at a matrix with independent columns?
Sketch. The argument uses the converse of Proposition 5.1.17. The claim is false for $D = 0$ and $A = I_n$.

Proposition 5.1.17 says "dependent columns $\Rightarrow D(A) = 0$". The argument uses the converse, "$D(A) = 0 \Rightarrow$ dependent columns", which does not follow. In fact the claim is false: $D = 0$ is multilinear and alternating, and $D(I_n) = 0$ although the columns of I_n are independent.

For a determinant function, the conclusion is true, but the proof must use normalization: for $n = 2$ see Exercise 5.1.25, and for general n see Theorem 5.4.7.

Solution to Exercise 5.1.23

Uses. Definition 5.1.14 (The 2×2 determinant), Example 5.1.15 (det is signed area)
Hint 1. The pieces are: two right triangles with legs a and c, two with legs b and d, and two rectangles with sides b and c.
Hint 2. For (b), compare the slopes c/a and d/b of the two edges.
Sketch. The area of P is $(a+b)(c+d) - ac - bd - 2bc = ad - bc$. The inequality $ad > bc$ says $c/a < d/b$, so (b, d) is steeper than (a, c) and lies counterclockwise from it.

(a) Because $ad > bc$, the vector (a, c) has smaller slope than (b, d) (see (b)), so P sits in the rectangle as in Figure 5.1.2, and P together with the six pieces below tiles the rectangle without overlaps. (If $ad \leq bc$, the pieces would overlap and the computation below would not apply.) The rectangle has area $(a+b)(c+d)$. Outside P it contains two right triangles with legs a and c (below the edge from $\mathbf{0}$ to (a, c) and above the opposite edge), two right triangles with legs b and d (left of the edge from $\mathbf{0}$ to (b, d) and right of the opposite edge), and two rectangles with sides b and c (in the lower right and upper left corners). So

$$\operatorname{area}(P) = (a+b)(c+d) - 2 \cdot \frac{ac}{2} - 2 \cdot \frac{bd}{2} - 2bc = ac + ad + bc + bd - ac - bd - 2bc = ad - bc.$$

(b) Since $a, b > 0$, dividing $ad > bc$ by ab gives $d/b > c/a$. Both vectors lie in the open first quadrant, and (b, d) has the larger slope, so its polar angle is larger: turning from (a, c) to (b, d) is a counterclockwise turn through an angle in $(0, \pi/2)$. So the signed area of Example 5.1.15 is $+\operatorname{area}(P) = ad - bc = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, consistent with (a).

Solution to Exercise 5.1.24

Uses. Definition 5.1.1 (Determinant function), Definition 1.1.28 (Characteristic)
Hint 1. For (b), only alternation is in doubt. Compute per at a matrix with two equal columns (a, c).
Hint 2. You get $ac + ca = (1+1)ac$. When does this vanish for all a, c?
Sketch. (a) is a direct check. (b) With equal columns, per $A = (1+1)ac$, which vanishes for all a, c exactly when $1 + 1 = 0$. So per is a determinant function exactly when $\operatorname{char} F = 2$, and then per $=$ det.

(a) $\operatorname{per}(I_2) = 1 \cdot 1 + 0 \cdot 0 = 1$. For a fixed second column (b, d), the map $(a, c) \mapsto da + bc$ is linear, and for a fixed first column (a, c), the map $(b, d) \mapsto cb + ad$ is linear, so per is multilinear. Interchanging the columns gives $\begin{pmatrix} b & a \\ d & c \end{pmatrix}$, whose permanent is $bc + ad = \operatorname{per} A$.

(b) By (a), per is a determinant function if and only if it is alternating. A matrix with equal columns is $\begin{pmatrix} a & a \\ c & c \end{pmatrix}$, with permanent $ac + ac = (1+1)ac$. If $1 + 1 = 0$, this is 0 for all a, c. If $1 + 1 \neq 0$, then $a = c = 1$ gives $1 + 1 \neq 0$. So per is a determinant function on $F^{2\times 2}$ if and only if $1 + 1 = 0$, that is, $\operatorname{char} F = 2$ (Definition 1.1.28). In that case $-bc = bc$, so per $=$ det, as Theorem 5.1.13 requires.

Solution to Exercise 5.1.25

Uses. Proposition 5.1.17 (Dependent columns), Definition 1.5.1 (Linear independence), Corollary 3.5.25 (The invertible matrix theorem), Theorem 3.5.12 (First Fundamental Theorem, completed), Definition 2.4.1 (Matrix product), Definition 2.5.12 (Invertible)
Hint 1. One direction is Proposition 5.1.17. For the other, you need an explicit dependence relation built from a, b, c, d.
Hint 2. Compute $d(a, c) - c(b, d)$ and $-b(a, c) + a(b, d)$.
Hint 3. Both are multiples of $\det A$. When are their coefficients not both zero?

Sketch. $d\mathbf{a}_1 - c\mathbf{a}_2 = (\det A, 0)$ and $-b\mathbf{a}_1 + a\mathbf{a}_2 = (0, \det A)$. If $\det A = 0$, one of these is a nontrivial dependence unless $A = 0$. For (b), combine with the Fundamental Theorems, and multiply out to check the inverse.

Write $\mathbf{a}_1 = (a, c)$ and $\mathbf{a}_2 = (b, d)$.

(a) If the columns are dependent, then $\det A = 0$ by Proposition 5.1.17, since det is multilinear and alternating (Theorem 5.1.13). Conversely, suppose $\det A = 0$. We compute

$$d\mathbf{a}_1 - c\mathbf{a}_2 = (da - cb, dc - cd) = (\det A, 0) = \mathbf{0}, \quad -b\mathbf{a}_1 + a\mathbf{a}_2 = (-ba + ab, -bc + ad) = (0, \det A) = \mathbf{0}.$$

If $(c, d) \neq (0, 0)$, the first relation has coefficients $d, -c$ not both zero, so the columns are dependent. If $c = d = 0$ but $(a, b) \neq (0, 0)$, the second relation has coefficients $-b, a$ not both zero. If $a = b = c = d = 0$, then $\mathbf{a}_1 = \mathbf{0}$, and $1 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 = \mathbf{0}$ is a dependence. In every case the columns are dependent.

(b) By Corollary 3.5.25 and Theorem 3.5.12 (as in the proof of Corollary 5.1.18), a square matrix is invertible if and only if its columns are linearly independent. By (a), this happens if and only if $\det A \neq 0$. Now let $C = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Multiplying out,

$$AC = \begin{pmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{pmatrix} = (\det A) I_2, \quad CA = \begin{pmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{pmatrix} = (\det A) I_2.$$

If $\det A \neq 0$, it follows that $(\det A)^{-1} C$ is a two-sided inverse of A.

Solution to Exercise 5.1.26

Uses. Definition 5.1.14 (The 2×2 determinant), Exercise 5.1.25, Exercise 1.1.44
Hint 1. Use $i^2 = -1$.
Hint 2. By Exercise 5.1.25, the columns are dependent exactly when the determinant is 0. Compute it as a polynomial in z.
Sketch. (a) $(1 + i)(1 - i) - 2i = 2 - 2i$. (b) The determinant is $z^2 - i^2 = z^2 + 1$, which vanishes exactly for $z = \pm i$. For $z = i$ the columns are equal; for $z = -i$ they are negatives of each other.

(a) $(1 + i)(1 - i) - 2 \cdot i = 1 - i^2 - 2i = 2 - 2i$.

(b) The determinant is $z \cdot z - i \cdot i = z^2 + 1$. By Exercise 5.1.25, the columns are dependent exactly when $z^2 + 1 = 0$, that is (Exercise 1.1.44), when $z = i$ or $z = -i$. For $z = i$ both columns are (i, i), so $\mathbf{a}_1 - \mathbf{a}_2 = \mathbf{0}$. For $z = -i$ the columns are $(-i, i)$ and $(i, -i)$, so $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{0}$.

Solution to Exercise 5.1.27

Uses. Theorem 5.1.13 (The 2×2 determinant), Definition 2.4.1 (Matrix product), Definition 2.3.15 (The linear map of a matrix), Proposition 2.1.17 (Compositions of linear maps are linear)
Hint 1. Fix A and regard $D(B) = \det(AB)$ as a function of the columns of B.
Hint 2. The columns of AB are $A\mathbf{b}_1$ and $A\mathbf{b}_2$. Show that D is multilinear and alternating, and apply part (a) of Theorem 5.1.13.
Sketch. $D(\mathbf{b}_1, \mathbf{b}_2) = \det(A\mathbf{b}_1, A\mathbf{b}_2)$ is multilinear, because $\mathbf{x} \mapsto A\mathbf{x}$ is linear, and alternating, because equal columns of B give equal columns of AB. So $D(B) = \det(B) D(I_2) = \det B \det A$.

Fix A, and define $D : F^{2\times 2} \to F$ by $D(B) = \det(AB)$. The columns of AB are $A\mathbf{b}_1$ and $A\mathbf{b}_2$ (Definition 2.4.1), so $D(\mathbf{b}_1, \mathbf{b}_2) = \det(A\mathbf{b}_1, A\mathbf{b}_2)$.

Multilinear. For fixed $\mathbf{b}_2$, the map $\mathbf{x} \mapsto \det(A\mathbf{x}, A\mathbf{b}_2)$ is the composite of the linear map T_A with the linear map $\mathbf{y} \mapsto \det(\mathbf{y}, A\mathbf{b}_2)$, so it is linear (Proposition 2.1.17). The same holds in the second column.

Alternating. If $\mathbf{b}_1 = \mathbf{b}_2$, then $A\mathbf{b}_1 = A\mathbf{b}_2$, so $D(B) = 0$.

By part (a) of Theorem 5.1.13, $D(B) = \det(B) D(I_2) = \det(B) \det(A)$, as desired.

Solution to Exercise 5.1.27

Uses. Definition 5.1.14 (The 2×2 determinant), Definition 2.4.1 (Matrix product)
Hint 1. Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ and compute AB.
Hint 2. Expand $\det(AB)$; four of the eight products cancel in pairs.

Sketch. $AB = \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix}$. Expanding the determinant, the terms $acpq$ and $bdrs$ cancel, and the rest factor as $(ad-bc)(ps-qr)$.

With $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$, we have $AB = \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix}$, so

$$\begin{aligned} \det(AB) &= (ap+br)(cq+ds) - (aq+bs)(cp+dr) \\ &= acpq + adps + bcqr + bdrs - acpq - adqr - bcps - bdrs \\ &= ad(ps-qr) - bc(ps-qr) = (ad-bc)(ps-qr) = \det A \det B. \end{aligned}$$

The first approach is the one that generalizes: Theorem 5.4.10 proves $\det(AB) = \det A \det B$ for all n, and the direct computation would be hopeless for large n.

Solution to Exercise 5.1.28

Uses. Exercise 5.1.27, Lemma 5.1.8 (Column replacement), Example 5.1.15 (det is signed area)
Hint 1. Use Exercise 5.1.27 and compute $\det R_\theta$ and $\det S_t$.
Hint 2. For the interpretation: the columns of $R_\theta A$ are the columns of A rotated by θ, and the columns of AS_t are $\mathbf{a}_1$ and $\mathbf{a}_2 + t\mathbf{a}_1$.
Sketch. $\det R_\theta = \cos^2\theta + \sin^2\theta = 1$ and $\det S_t = 1$, so both equalities follow from multiplicativity. Rotating a parallelogram, or sliding one edge parallel to the other, preserves its signed area.

By Exercise 5.1.27, $\det(R_\theta A) = \det(R_\theta)\det A$ and $\det(AS_t) = \det A \det S_t$. Since $\det R_\theta = \cos^2\theta + \sin^2\theta = 1$ and $\det S_t = 1 \cdot 1 - t \cdot 0 = 1$, both equalities hold. (The second also follows from Lemma 5.1.8, since AS_t has columns $\mathbf{a}_1$ and $\mathbf{a}_2 + t\mathbf{a}_1$.)

The columns of $R_\theta A$ are $R_\theta \mathbf{a}_1$ and $R_\theta \mathbf{a}_2$, so $P(R_\theta \mathbf{a}_1, R_\theta \mathbf{a}_2)$ is $P(\mathbf{a}_1, \mathbf{a}_2)$ rotated by θ, and the first equality says that rotation preserves signed area (area and orientation). The second says that sliding the edge $\mathbf{a}_2$ parallel to the edge $\mathbf{a}_1$ preserves signed area. These are the two facts used informally in Intuition (Signed area is linear in each column); with Example 5.1.15 they are now theorems.

Solution to Exercise 5.1.29

Uses. Theorem 5.1.13 (The 2×2 determinant), Definition 5.1.14 (The 2×2 determinant)
Hint 1. By Theorem 5.1.13, every multilinear alternating D is $c\det$ for $c = D(I_2)$. Conversely, which multiples of det are multilinear and alternating?
Hint 2. So the condition is $c\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 6$, that is, $-2c = 6$. Solve it in each field.
Sketch. The candidates are exactly $D = c\det$ with $c \in F$, and the condition is $-2c = 6$. Over $\mathbb{Q}$, $c = -3$. Over $\mathbb{F}_2$, it reads $0 = 0$, so $D = 0$ and $D = \det$ both work. Over $\mathbb{F}_3$, it reads $c = 0$, so only $D = 0$.

By part (a) of Theorem 5.1.13, every multilinear alternating D equals $c\det$ with $c = D(I_2)$. Conversely, for every $c \in F$ the function $c\det$ is multilinear and alternating, since multiplying the identities defining these properties by c preserves them. So the functions in question are the $c\det$ with $c\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 6$. Since $\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 4 - 6 = -2$, the condition is $-2c = 6$.

(a) Over $\mathbb{Q}$, $c = -3$, so $D = -3\det$ is the only one.

(b) Over $\mathbb{F}_2$, $-2 = 0$ and $6 = 0$, so the condition holds for both $c = 0$ and $c = 1$: the answers are $D = 0$ and $D = \det$.

(c) Over $\mathbb{F}_3$, $-2 = 1$ and $6 = 0$, so $c = 0$, and $D = 0$ is the only one.

Solution to Exercise 5.1.30

Uses. Definition 5.1.1 (Determinant function), Theorem 5.1.13 (The 2×2 determinant), Exercise 1.1.18
Hint 1. Normalization and alternation hold over every field. Test multilinearity on $(c\mathbf{e}_1, \mathbf{e}_2)$.
Hint 2. Multilinearity would force $c^2 = c$ for every $c \in F$. Which fields satisfy this? (See Exercise 1.1.18.)
Sketch. $(\det I_2)^2 = 1$, and equal columns give $0^2 = 0$. Linearity in the first column forces $c^2 = (\det(c\mathbf{e}_1, \mathbf{e}_2))^2 = c$ for all c, so $F = \{0, 1\}$, which is $\mathbb{F}_2$. Over $\mathbb{F}_2$, $x^2 = x$ for every x, so $(\det A)^2 = \det A$.

Let $E(A) = (\det A)^2$. Then $E(I_2) = 1$, and if the columns of A are equal, then $\det A = 0$ and $E(A) = 0$. So E is a determinant function if and only if it is multilinear.

Suppose E is multilinear. For every $c \in F$, linearity in the first column gives $E(c\mathbf{e}_1, \mathbf{e}_2) = cE(\mathbf{e}_1, \mathbf{e}_2) = c$, while directly $E(c\mathbf{e}_1, \mathbf{e}_2) = \left(\det\begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}\right)^2 = c^2$. So $c^2 = c$ for every $c \in F$, and by Exercise 1.1.18 every element of F is 0 or 1. Thus $F = \{0, 1\}$. In such a field $1 + 1 \neq 1$ (else $1 = 0$), so $1 + 1 = 0$ and F has the tables of $\mathbb{F}_2$.

Conversely, in $\mathbb{F}_2$ we have $0^2 = 0$ and $1^2 = 1$, so $E(A) = \det A$ for every A, and E is a determinant function by Theorem 5.1.13. So the answer is: exactly when $F = \mathbb{F}_2$.

Solution to Exercise 5.1.31

Uses. Theorem 5.1.13 (The 2×2 determinant), Exercise 5.1.25, Proposition 5.1.10 (Alternating implies antisymmetric), Non-example 5.1.3 (One axiom at a time)
Hint 1. (a) Use part (a) of Theorem 5.1.13 and Exercise 5.1.25.
Hint 2. (b) The hypothesis $D(cA) = c^2 D(A)$ is weaker than multilinearity. Look back at Non-example 5.1.3.
Hint 3. (c) Count the interchanges needed to sort $\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_1$.
Sketch. (a) True: $0 = D(A) = \det(A)D(I_2)$ with $\det A \neq 0$, so $D(I_2) = 0$ and $D = 0$. (b) False: $D = |\det|$. (c) True: two interchanges.

(a) True. By part (a) of Theorem 5.1.13, $D = D(I_2)\det$. Since A is invertible, $\det A \neq 0$ by Exercise 5.1.25, so $0 = D(A) = D(I_2)\det A$ forces $D(I_2) = 0$, and then $D = 0$.

(b) False. Let $D(A) = |\det A|$. It is alternating (Non-example 5.1.3), and $D(cA) = \left|c^2 \det A\right| = c^2 D(A)$. If $D = \lambda \det$ for some $\lambda \in \mathbb{R}$, then evaluating at I_2 gives $\lambda = 1$, and evaluating at $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ gives $1 = -\lambda$. These are incompatible.

(c) True. Interchanging columns 1 and 3 and then columns 2 and 3, Proposition 5.1.10 gives

$$D(\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_1) = -D(\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_2) = D(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3).$$

Solution to Exercise 5.1.32

Uses. Definition 5.1.1 (Determinant function), Theorem 5.1.13 (The 2×2 determinant)
Hint 1. Each coordinate of $\mathbf{x} \times \mathbf{y}$ is a 2×2 determinant built from two coordinates of $\mathbf{x}$ and of $\mathbf{y}$. Use this for multilinearity.
Hint 2. For alternation there are three cases: $\mathbf{x} = \mathbf{y}$, $\mathbf{y} = \mathbf{z}$, and $\mathbf{x} = \mathbf{z}$. In the last two, expand and watch the six terms cancel in pairs.
Sketch. Each $(\mathbf{x} \times \mathbf{y})_k$ is bilinear in $\mathbf{x}, \mathbf{y}$, so each term $(\mathbf{x} \times \mathbf{y})_k z_k$ is linear in each of $\mathbf{x}, \mathbf{y}, \mathbf{z}$. $\mathbf{x} \times \mathbf{x} = \mathbf{0}$, and $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y}$ and $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x}$ expand to six terms cancelling in pairs. $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$.

Multilinear. The coordinate $(\mathbf{x} \times \mathbf{y})_3 = x_1 y_2 - x_2 y_1$ is the determinant of $\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$, and similarly for the other two coordinates, so by Theorem 5.1.13 each coordinate of $\mathbf{x} \times \mathbf{y}$ is linear in $\mathbf{x}$ for fixed $\mathbf{y}$, and linear in $\mathbf{y}$ for fixed $\mathbf{x}$. Hence each term $(\mathbf{x} \times \mathbf{y})_k z_k$ is linear in each of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ when the other two are fixed, and so is their sum D.

Alternating. If $\mathbf{x} = \mathbf{y}$, every coordinate of $\mathbf{x} \times \mathbf{x}$ has the form $x_i x_j - x_j x_i = 0$, so $D = 0$. If $\mathbf{y} = \mathbf{z}$,

$$D = (x_2 y_3 - x_3 y_2) y_1 + (x_3 y_1 - x_1 y_3) y_2 + (x_1 y_2 - x_2 y_1) y_3,$$

and the six terms cancel in pairs: $x_2 y_1 y_3$ with $-x_2 y_1 y_3$, $-x_3 y_1 y_2$ with $x_3 y_1 y_2$, and $-x_1 y_2 y_3$ with $x_1 y_2 y_3$. If $\mathbf{x} = \mathbf{z}$,

$$D = (x_2 y_3 - x_3 y_2) x_1 + (x_3 y_1 - x_1 y_3) x_2 + (x_1 y_2 - x_2 y_1) x_3,$$

and again $x_1 x_2 y_3$, $x_1 x_3 y_2$, and $x_2 x_3 y_1$ each appear once with each sign.

Normalized. $\mathbf{e}_1 \times \mathbf{e}_2 = (0 \cdot 0 - 0 \cdot 1, 0 \cdot 0 - 1 \cdot 0, 1 \cdot 1 - 0 \cdot 0) = \mathbf{e}_3$, so $D(I_3) = D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1$.

Solution to Exercise 5.1.33

Uses. Proposition 5.1.10 (Alternating implies antisymmetric), Proposition 5.1.17 (Dependent columns), Corollary 5.1.18 (Non-invertible matrices), Exercise 5.1.32
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. The columns are $\mathbf{e}_1 + 3\mathbf{e}_2$, $\mathbf{e}_2 + 4\mathbf{e}_3$, and $2\mathbf{e}_1 + \mathbf{e}_3$. Expanding gives eight terms. Discard those with a repeated basis vector.

Hint 2. Exactly two terms survive. One of them needs $D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$; compute it with interchanges.
Hint 3. For (b), solve $A\mathbf{x} = \mathbf{0}$ over $\mathbb{F}_5$, starting from the first row.
Sketch. The surviving terms are $D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and $3 \cdot 4 \cdot 2 \cdot D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) = 24D(I_3)$, so $D(A) = 25D(I_3)$. Over $\mathbb{F}_5$, $3\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = (5, 10, 5) = \mathbf{0}$, and $25 = 0$. Over $\mathbb{Q}$, a determinant function gives $D(A) = 25 \neq 0$.

(a) The columns are $\mathbf{a}_1 = \mathbf{e}_1 + 3\mathbf{e}_2$, $\mathbf{a}_2 = \mathbf{e}_2 + 4\mathbf{e}_3$, and $\mathbf{a}_3 = 2\mathbf{e}_1 + \mathbf{e}_3$. Expanding each column by multilinearity writes $D(A)$ as a sum of eight terms, one for each choice of a basis vector from each column. A term in which some $\mathbf{e}_k$ is chosen twice is 0 because D is alternating. If $\mathbf{e}_1$ is chosen from $\mathbf{a}_1$, then $\mathbf{a}_3$ must contribute $\mathbf{e}_3$, so $\mathbf{a}_2$ must contribute $\mathbf{e}_2$. If $\mathbf{e}_2$ is chosen from $\mathbf{a}_1$, then $\mathbf{a}_2$ must contribute $\mathbf{e}_3$, so $\mathbf{a}_3$ must contribute $\mathbf{e}_1$. So only two terms survive:

$$D(A) = 1 \cdot 1 \cdot 1 \cdot D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + 3 \cdot 4 \cdot 2 \cdot D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1).$$

By two interchanges (Proposition 5.1.10), $D(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) = -D(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2) = D(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Hence $D(A) = (1 + 24)D(I_3) = 25D(I_3)$.

(b) Over $\mathbb{F}_5$, solve $A\mathbf{x} = \mathbf{0}$: the first row gives $x_1 = -2x_3 = 3x_3$, the second gives $x_2 = -3x_1 = -9x_3 = x_3$, and the third, $4x_2 + x_3 = 5x_3 = 0$, then holds automatically. Taking $x_3 = 1$,

$$3\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = (3 + 0 + 2, 9 + 1 + 0, 0 + 4 + 1) = (5, 10, 5) = \mathbf{0} \quad \text{in } \mathbb{F}_5^3.$$

So the columns are dependent over $\mathbb{F}_5$, and Proposition 5.1.17 gives $D(A) = 0$. This agrees with (a), because $25 = 0$ in $\mathbb{F}_5$.

(c) Let D be a determinant function on $\mathbb{Q}^{3\times 3}$. By (a), $D(A) = 25 \neq 0$, so A is invertible by Corollary 5.1.18.

Solution to Exercise 5.1.34

Uses. Lemma 2.1.22 (Functions into a vector space), Definition 1.4.1 (Subspace), Theorem 1.4.3 (Subspace criterion), Proposition 2.1.23 ($\mathrm{L}(V, W)$ is a vector space), Theorem 5.1.13 (The 2×2 determinant), Definition 4.1.5 (Dimension)
Hint 1. For (a), check that $0 \in \mathrm{Alt}_n$ and that $D + sD'$ is multilinear and alternating whenever D and D' are.
Hint 2. For (b), part (a) of Theorem 5.1.13 says every element of Alt_2 is a multiple of det.
Sketch. Sums and scalar multiples of linear maps are linear, and $0 + 0 = 0$, so Alt_n is closed under linear combinations. By Theorem 5.1.13, Alt_2 consists exactly of the multiples of det $\neq 0$, so det alone is a basis.

(a) The zero function is multilinear and alternating. Let $D, D' \in \mathrm{Alt}_n$ and $s \in F$, and let $E = D + sD'$. Fix a column index j and the other columns. The map $\mathbf{x} \mapsto E(\ldots, \mathbf{x}, \ldots)$ is $\mathbf{x} \mapsto D(\ldots, \mathbf{x}, \ldots) + sD'(\ldots, \mathbf{x}, \ldots)$, a linear combination of two linear maps $F^n \to F$, hence linear (Proposition 2.1.23). If A has two equal columns, then $E(A) = D(A) + sD'(A) = 0 + s \cdot 0 = 0$. So $E \in \mathrm{Alt}_n$, and Alt_n is a subspace by Theorem 1.4.3.

(b) By part (b) of Theorem 5.1.13, $\det \in \mathrm{Alt}_2$, so $\mathrm{span}(\det) \subseteq \mathrm{Alt}_2$ by (a). By part (a) of Theorem 5.1.13, every $D \in \mathrm{Alt}_2$ equals $D(I_2) \det \in \mathrm{span}(\det)$. Hence $\mathrm{Alt}_2 = \mathrm{span}(\det)$. Since $\det(I_2) = 1$, det is not the zero function, so the one-element list det is linearly independent and is a basis of Alt_2. Thus $\dim \mathrm{Alt}_2 = 1$.

The expansion in Intuition (What the proof did, and where it leads) suggests that $\dim \mathrm{Alt}_n = 1$ for every n; this follows from Section 5.3.

Solution to Exercise 5.1.35

Uses. Proposition 5.1.10 (Alternating implies antisymmetric), Definition 5.1.1 (Determinant function)
Hint 1. First show that interchanging two adjacent columns negates D. Which hypothesis did the first proof of Proposition 5.1.10 actually use?
Hint 2. If $\mathbf{a}_i = \mathbf{a}_j$ with $j > i + 1$, move column j to position $i + 1$ by interchanging adjacent columns.
Sketch. The expansion proof of Proposition 5.1.10, with $j = i + 1$, uses alternation only for columns i and $i + 1$, so adjacent interchanges negate D. If $\mathbf{a}_i = \mathbf{a}_j$ with $i < j$, then $j - i - 1$ adjacent interchanges bring column j next to column i, giving a matrix A' with $D(A') = 0$ and $D(A) = \pm D(A')$.

Step 1. Fix i and the columns other than i and $i + 1$, and let $f(\mathbf{x}, \mathbf{y})$ be the value of D with $\mathbf{x}$ in column i and $\mathbf{y}$ in column $i + 1$. By hypothesis $f(\mathbf{x}, \mathbf{x}) = 0$ for every $\mathbf{x}$. The first proof of Proposition 5.1.10 used nothing else besides multilinearity, so it gives $f(\mathbf{y}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{y})$. That is, interchanging two adjacent columns negates D.

Step 2. Let A have $\mathbf{a}_i = \mathbf{a}_j$ with $i < j$. If $j = i + 1$, then $D(A) = 0$ by hypothesis. Otherwise, interchange columns $j - 1$ and j, then columns $j - 2$ and $j - 1$, and so on, ending with columns $i + 1$ and $i + 2$. These $j - i - 1$ adjacent interchanges move the vector $\mathbf{a}_j$ from position j to position $i + 1$, and never touch column i. The resulting matrix A' has $\mathbf{a}_i$ in column i and $\mathbf{a}_j = \mathbf{a}_i$ in column $i + 1$, so $D(A') = 0$. By Step 1, $D(A) = (-1)^{j-i-1} D(A') = 0$. Hence D is alternating.

Solution to Exercise 5.1.36

Uses. Proposition 5.1.17 (Dependent columns), Theorem 4.1.10 (Independent lists are short), Remark 5.1.6
Hint 1. What can you say about any k vectors in F^n when $k > n$?
Hint 2. They are dependent. Now repeat the first proof of Proposition 5.1.17 for f.
Sketch. Any $k > n = \dim F^n$ vectors are dependent, so one of them is a combination of the others. Expanding f in that slot gives terms with a repeated argument.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in F^n$. Since $\dim F^n = n < k$, the list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is linearly dependent (Theorem 4.1.10). As in the first proof of Proposition 5.1.17, some $\mathbf{v}_m$ equals $\sum_{j \neq m} b_j \mathbf{v}_j$, and linearity in slot m gives

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \sum_{j \neq m} b_j f\big(\mathbf{v}_1, \ldots, \mathbf{v}_j, \ldots, \mathbf{v}_k\big),$$

where in the j-th term $\mathbf{v}_j$ occupies both slot j and slot m. Each term is 0 because f is alternating. So $f(\mathbf{v}_1, \ldots, \mathbf{v}_k) = 0$ for all inputs, that is, $f = 0$.

For $k = 3$ and $n = 2$, f would measure a signed volume of the parallelepiped spanned by three vectors in the plane. Such a parallelepiped is flat, so its volume is 0, and the result says that no alternating multilinear function can do better.

Solution to Exercise 5.1.37

Uses. Theorem 5.1.13 (The 2×2 determinant), Proposition 5.1.10 (Alternating implies antisymmetric), Remark 5.1.6, Definition 1.5.1 (Linear independence)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. $f_{ij}(\mathbf{x}, \mathbf{y})$ is the 2×2 determinant of rows i and j of the 3×2 matrix with columns $\mathbf{x}, \mathbf{y}$.
Hint 2. For (b), evaluate a vanishing combination at $(\mathbf{e}_1, \mathbf{e}_2)$, $(\mathbf{e}_1, \mathbf{e}_3)$, and $(\mathbf{e}_2, \mathbf{e}_3)$.
Hint 3. For (c), write $\mathbf{x} = \sum_i x_i \mathbf{e}_i$ and $\mathbf{y} = \sum_j y_j \mathbf{e}_j$ and expand. Group the terms (i, j) and (j, i).
Sketch. (a) Each f_{ij} is a 2×2 determinant of coordinates. (b) $f_{ij}(\mathbf{e}_k, \mathbf{e}_l)$ is 1 if $(k, l) = (i, j)$ and 0 for the other pairs $k < l$. (c) Expanding gives $f(\mathbf{x}, \mathbf{y}) = \sum_{i,j} x_i y_j f(\mathbf{e}_i, \mathbf{e}_j)$; diagonal terms vanish, and antisymmetry pairs (i, j) with (j, i) into $(x_i y_j - x_j y_i) f(\mathbf{e}_i, \mathbf{e}_j)$.

(a) $f_{ij}(\mathbf{x}, \mathbf{y}) = \det\begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix}$, and the coordinates x_i, x_j depend linearly on $\mathbf{x}$. By Theorem 5.1.13, f_{ij} is linear in $\mathbf{x}$ for fixed $\mathbf{y}$ and in $\mathbf{y}$ for fixed $\mathbf{x}$, and $f_{ij}(\mathbf{x}, \mathbf{x}) = x_i x_j - x_j x_i = 0$.

(b) For $k < l$, $f_{ij}(\mathbf{e}_k, \mathbf{e}_l) = 1$ if $(k, l) = (i, j)$, and 0 otherwise. If $c_{12} f_{12} + c_{13} f_{13} + c_{23} f_{23} = 0$, evaluating at $(\mathbf{e}_1, \mathbf{e}_2)$, $(\mathbf{e}_1, \mathbf{e}_3)$, and $(\mathbf{e}_2, \mathbf{e}_3)$ gives $c_{12} = c_{13} = c_{23} = 0$.

(c) Expanding $\mathbf{x} = \sum_i x_i \mathbf{e}_i$ and $\mathbf{y} = \sum_j y_j \mathbf{e}_j$,

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{3} \sum_{j=1}^{3} x_i y_j f\big(\mathbf{e}_i, \mathbf{e}_j\big).$$

The terms with $i = j$ vanish since f is alternating. For $i < j$, $f\big(\mathbf{e}_j, \mathbf{e}_i\big) = -f\big(\mathbf{e}_i, \mathbf{e}_j\big)$ by Proposition 5.1.10 (which holds for f by Remark 5.1.6), so the terms (i, j) and (j, i) combine to $\big(x_i y_j - x_j y_i\big) f\big(\mathbf{e}_i, \mathbf{e}_j\big)$. Hence

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i<j} f\big(\mathbf{e}_i, \mathbf{e}_j\big) f_{ij}(\mathbf{x}, \mathbf{y}),$$

as desired. With (a) and (b), the f_{ij} form a basis of the space of multilinear alternating functions of two vectors in F^3, which therefore has dimension 3, compared with dimension 1 for two vectors in F^2 (Exercise 5.1.34).

Solution to Exercise 5.1.38

Uses. Theorem 5.1.13 (The 2×2 determinant), Definition 2.4.26 (Trace), Exercise 5.1.25, Exercise 1.1.44
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. The columns of $A + tI_2$ are $\mathbf{a}_1 + t\mathbf{e}_1$ and $\mathbf{a}_2 + t\mathbf{e}_2$. Expand into four terms.
Hint 2. The mixed terms are $t\det(\mathbf{a}_1, \mathbf{e}_2)$ and $t\det(\mathbf{e}_1, \mathbf{a}_2)$. Compute them from the formula.
Hint 3. For (b), $A + tI_2$ is invertible exactly when its determinant is nonzero (Exercise 5.1.25).
Sketch. Expanding, $\det(A + tI_2) = \det A + t\det(\mathbf{a}_1, \mathbf{e}_2) + t\det(\mathbf{e}_1, \mathbf{a}_2) + t^2$, and the middle terms are ta and td. For the rotation, $\det(A + tI_2) = t^2 + 1$, which vanishes for $t = \pm i$ and for no real t.

(a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The columns of $A + tI_2$ are $\mathbf{a}_1 + t\mathbf{e}_1$ and $\mathbf{a}_2 + t\mathbf{e}_2$. Expanding by multilinearity,

$$\det(A + tI_2) = \det(\mathbf{a}_1, \mathbf{a}_2) + t\det(\mathbf{a}_1, \mathbf{e}_2) + t\det(\mathbf{e}_1, \mathbf{a}_2) + t^2\det(\mathbf{e}_1, \mathbf{e}_2).$$

Here $\det(\mathbf{a}_1, \mathbf{e}_2) = \det\begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix} = a$, $\det(\mathbf{e}_1, \mathbf{a}_2) = \det\begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix} = d$, and $\det(\mathbf{e}_1, \mathbf{e}_2) = 1$. So $\det(A + tI_2) = \det A + (a + d)t + t^2$, and $a + d = \operatorname{tr} A$.

(b) Here $\operatorname{tr} A = 0$ and $\det A = 0 \cdot 0 - (-1) \cdot 1 = 1$, so $\det(A + tI_2) = t^2 + 1$. By Exercise 5.1.25, $A + tI_2$ fails to be invertible exactly when $t^2 + 1 = 0$. Over $\mathbb{C}$ this means $t = i$ or $t = -i$ (Exercise 1.1.44). Over $\mathbb{R}$ there is no such t, since $t^2 + 1 \geq 1$.

So $A\mathbf{x} = -t\mathbf{x}$ has a nonzero solution $\mathbf{x}$ for $t = \pm i$ over $\mathbb{C}$ but for no real t: the rotation by $\pi/2$ moves every nonzero real vector off its own line. This is the first appearance of eigenvalues, which Section 7.3 finds through exactly this kind of determinant.

Solution to Exercise 5.1.39

Uses. Exercise 5.1.25, Definition 1.5.1 (Linear independence), Proposition 5.1.10 (Alternating implies antisymmetric), Proposition 1.2.35 (Zero products)
Hint 1. By Exercise 5.1.25, count the pairs of columns that are linearly independent. Choose the first column, then the second.
Hint 2. Two vectors are independent exactly when the first is nonzero and the second is not a multiple of the first. How many multiples does a nonzero vector of $\mathbb{F}_p^2$ have?
Hint 3. For (b), interchanging the columns is a bijection on $\mathbb{F}_3^{2\times 2}$. What does it do to the determinant?
Sketch. (a) $(p^2 - 1)(p^2 - p)$: a nonzero first column, then a second column outside its p multiples. (b) For $p = 3$ that is 48; interchanging the columns negates det and so swaps det $= 1$ with det $= 2 = -1$. (c) $(4 - 1)(4 - 2) = 6$, all with det $= 1$.

(a) By Exercise 5.1.25, $\det A \neq 0$ exactly when the columns $\mathbf{a}_1, \mathbf{a}_2$ are linearly independent, which happens exactly when $\mathbf{a}_1 \neq \mathbf{0}$ and $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$. (If $\mathbf{a}_1 = \mathbf{0}$, then $1 \cdot \mathbf{a}_1 = \mathbf{0}$; if $\mathbf{a}_2 = c\mathbf{a}_1$, then $c\mathbf{a}_1 - \mathbf{a}_2 = \mathbf{0}$. Conversely, if $x\mathbf{a}_1 + y\mathbf{a}_2 = \mathbf{0}$ with $y \neq 0$, then $\mathbf{a}_2 = -y^{-1}x\mathbf{a}_1$, and if $y = 0$ then $x\mathbf{a}_1 = \mathbf{0}$ with $x \neq 0$, so $\mathbf{a}_1 = \mathbf{0}$.) There are $p^2 - 1$ nonzero vectors in $\mathbb{F}_p^2$. For $\mathbf{a}_1 \neq \mathbf{0}$, the multiples $c\mathbf{a}_1$ with $c \in \mathbb{F}_p$ are pairwise distinct, because $c\mathbf{a}_1 = c'\mathbf{a}_1$ gives $(c - c')\mathbf{a}_1 = \mathbf{0}$ and hence $c = c'$ (Proposition 1.2.35). So there are $p^2 - p$ choices for $\mathbf{a}_2$, and $(p^2 - 1)(p^2 - p)$ matrices in all.

(b) For $p = 3$ the count is $8 \cdot 6 = 48$. The nonzero elements of $\mathbb{F}_3$ are 1 and $2 = -1$. Interchanging the two columns is a bijection from $\mathbb{F}_3^{2\times 2}$ to itself (it is its own inverse), and it negates the determinant (Proposition 5.1.10). So it maps the matrices with $\det A = 1$ bijectively onto those with $\det A = -1 = 2$. These two sets have the same size and together contain all 48 matrices, so each has 24 elements.

(c) For $p = 2$ the count is $(4 - 1)(4 - 2) = 6$, and the only nonzero element of $\mathbb{F}_2$ is 1, so all six have $\det A = 1$.

Solution to Exercise 5.1.40

Uses. Proposition 5.1.10 (Alternating implies antisymmetric), Proposition 5.1.17 (Dependent columns)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Expand every column. A term corresponds to choosing $\mathbf{a}_k$ or $\mathbf{a}_{k+1}$ in each column k, so there are 2^n terms.
Hint 2. Show that if the choices are not all "first" or all "second", then some vector is chosen twice. Look for a column k where the second option is chosen, followed by a column $k + 1$ where the first is.
Hint 3. The all-second term is $D(\mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{a}_1)$. Move $\mathbf{a}_1$ back to the front by adjacent interchanges.

Sketch. Only the two constant choices survive: a switch from "second" in column k to "first" in column $k+1$ puts $\mathbf{a}_{k+1}$ in both columns. So the value is $D(\mathbf{a}_1, ..., \mathbf{a}_n) + D(\mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{a}_1)$, and $n-1$ adjacent interchanges give the second term as $(-1)^{n-1} D(\mathbf{a}_1, ..., \mathbf{a}_n)$. For even n, the alternating sum of the columns telescopes to $\mathbf{0}$.

We prove (b); (a) is the case $n = 3$, where $1 + (-1)^2 = 2$. Indices are read modulo n, so column n is $\mathbf{a}_n + \mathbf{a}_1$.

(b) Expanding each column k into its two terms writes the left side as a sum of 2^n terms, one for each $\varepsilon = (\varepsilon_1, ..., \varepsilon_n) \in \{0, 1\}^n$:

$$T_\varepsilon = D\big(\mathbf{a}_{1+\varepsilon_1}, \mathbf{a}_{2+\varepsilon_2}, ..., \mathbf{a}_{n+\varepsilon_n}\big).$$

Suppose ε is not constant. Choose k with $\varepsilon_k = 1$, and walk forward cyclically from k. Since some entry of ε is 0, there is a first index l after k (cyclically) with $\varepsilon_l = 0$, and then $\varepsilon_{l-1} = 1$. So column $l-1$ of T_ε holds $\mathbf{a}_l$, and column l also holds $\mathbf{a}_l$. These are different columns because $n \geq 2$, so $T_\varepsilon = 0$ since D is alternating. Only the two constant choices remain:

$$D(\mathbf{a}_1 + \mathbf{a}_2, ..., \mathbf{a}_n + \mathbf{a}_1) = D(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) + D(\mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{a}_1).$$

In the second term, interchange columns $n-1$ and n, then columns $n-2$ and $n-1$, and so on, ending with columns 1 and 2. These $n-1$ interchanges move $\mathbf{a}_1$ from the last position to the first and keep the others in order, and each negates D (Proposition 5.1.10). So the second term is $(-1)^{n-1} D(\mathbf{a}_1, ..., \mathbf{a}_n)$, which proves (b).

(c) Let n be even and $\mathbf{c}_k = \mathbf{a}_k + \mathbf{a}_{k+1}$. Then the sum telescopes:

$$\mathbf{c}_1 - \mathbf{c}_2 + \mathbf{c}_3 - \cdots - \mathbf{c}_n = \mathbf{a}_1 + (\mathbf{a}_2 - \mathbf{a}_2) + \cdots + (\mathbf{a}_n - \mathbf{a}_n) - \mathbf{a}_{n+1} = \mathbf{a}_1 - \mathbf{a}_1 = \mathbf{0}.$$

The coefficients are $\pm 1 \neq 0$, so the columns are dependent, and the left side of (b) is 0 by Proposition 5.1.17, matching $1 + (-1)^{n-1} = 0$.

Solution to Exercise 5.1.41

Uses. Theorem 4.1.16 (Extending to a basis), Definition 4.1.5 (Dimension), Proposition 1.6.19 (Coordinates respect linear combinations), Definition 1.5.1 (Linear independence)
Hint 1. Alternation: if two columns are equal, use (ii) to turn one of them into $\mathbf{0}$, then use (i) with $c = 0$.
Hint 2. Additivity in column k: fix the other $n-1$ columns. If they are dependent, show D vanishes for every choice of column k.
Hint 3. If they are independent, extend them to a basis of F^n by one vector $\mathbf{w}$. Write $\mathbf{x} = s\mathbf{w} + (...)$ and use (ii) to strip off the other part.
Sketch. Subtracting one of two equal columns from the other makes a zero column, and (i) with $c = 0$ gives $D = 0$. Fix all columns but the k-th. If they are dependent, (ii) produces a zero column among them, so $D(..., \mathbf{x}, ...) = 0$ for all $\mathbf{x}$. Otherwise extend them by one vector $\mathbf{w}$ to a basis; writing $\mathbf{x} = s(\mathbf{x})\mathbf{w} + \sum_{j \neq k} c_j \mathbf{a}_j$, (ii) and (i) give $D(..., \mathbf{x}, ...) = s(\mathbf{x}) D(..., \mathbf{w}, ...)$, where s is a coordinate function and hence linear.

A zero column. If column k of A is $\mathbf{0}$, then by (i) with $c = 0$, $D(A) = D(..., 0 \cdot \mathbf{0}, ...) = 0 \cdot D(A) = 0$.

Alternating. Suppose $\mathbf{a}_i = \mathbf{a}_j$ with $i \neq j$. Adding -1 times column j to column i does not change D, by (ii), and produces a matrix whose column i is $\mathbf{0}$. So $D(A) = 0$.

Multilinear. Fix k and the columns $\mathbf{a}_j$ for $j \neq k$, and let $g(\mathbf{x})$ be the value of D with $\mathbf{x}$ in column k. We show that g is linear.

Suppose first that the list $(\mathbf{a}_j)_{j \neq k}$ is linearly dependent (so $n \geq 2$). Then some $\mathbf{a}_m$ with $m \neq k$ is a linear combination $\sum b_j \mathbf{a}_j$ of the others, the sum running over $j \notin \{k, m\}$. For any $\mathbf{x}$, adding $-b_j$ times column j to column m for each such j does not change D, by (ii), and turns column m into $\mathbf{0}$. So $g(\mathbf{x}) = 0$ for every $\mathbf{x}$, and g is linear.

Suppose instead that $(\mathbf{a}_j)_{j \neq k}$ is linearly independent. It has $n-1$ vectors, so by Theorem 4.1.16 it extends to a basis of F^n, and since every basis of F^n has n vectors (Definition 4.1.5), exactly one vector $\mathbf{w}$ is added. Let $s(\mathbf{x})$ be the coefficient of $\mathbf{w}$ when $\mathbf{x}$ is written in this basis, so $\mathbf{x} = s(\mathbf{x})\mathbf{w} + \sum_{j \neq k} c_j \mathbf{a}_j$ for some scalars c_j.

The function s is a coordinate of the coordinate map, so it is linear (Proposition 1.6.19). Adding $-c_j$ times column j to column k for each $j \neq k$ does not change D, by (ii), so by (i)

$$g(\mathbf{x}) = g(s(\mathbf{x})\mathbf{w}) = s(\mathbf{x})g(\mathbf{w}).$$

So g is the linear function s times the constant $g(\mathbf{w})$, hence linear. In both cases D is linear in column k, and k was arbitrary.

Solution to Exercise 5.1.42

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Definition 3.2.2 (Elementary matrices), Exercise 5.1.27, Exercise 5.1.25

Hint 1. First show $D(I_2) = 1$, that $D(G) \neq 0$ for invertible G, and that $D\big(GXG^{-1}\big) = D(X)$.

Hint 2. For the shear $S_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, check that $\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} S_1 \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}^{-1} = S_t$ and that $S_1^2 = S_2$. Deduce $D(S_1) = 1$.

Hint 3. Conjugating by the swap $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ handles lower shears and $\begin{pmatrix} 1 & 0 \\ 0 & t \end{pmatrix}$. For P itself, check that $S_1 \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} S_1 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = P$.

Hint 4. Now $D(E) = \det E$ for every elementary matrix. Use Theorem 3.5.19 for invertible A. If A is not invertible, write a zero column with a shear and use $D\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0$.

Sketch. D is invariant under conjugation by invertible matrices. Conjugating S_1 by $\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}$ gives S_t, so $D(S_1) = D(S_2) = D(S_1)^2$, and $D(S_1) = 1$. Conjugating by P gives the lower shears and the second scaling, and P is three shears times $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, so $D(P) = -1$. So $D = \det$ on elementary matrices, hence on invertible matrices. A non-invertible A satisfies $AS = AS\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ or $A = A\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ for a shear S, so $D(A) = 0$.

Write $\operatorname{diag}(s,t) = \begin{pmatrix} s & 0 \\ 0 & t \end{pmatrix}$, $S_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $L_t = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$, and $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Step 0. $D(I_2) = D(\operatorname{diag}(1,1)) = 1$. If G is invertible, then $D(G)D\big(G^{-1}\big) = D(I_2) = 1$, so $D(G) \neq 0$ and $D\big(G^{-1}\big) = D(G)^{-1}$. Hence $D\big(GXG^{-1}\big) = D(G)D(X)D(G)^{-1} = D(X)$ for every X.

Step 1: shears. For $t \neq 0$, a direct multiplication gives $\operatorname{diag}(t,1)S_1 \operatorname{diag}\big(t^{-1},1\big) = S_t$, so $D(S_t) = D(S_1)$ by Step 0. Also $S_1S_1 = S_2$, so $D(S_1)^2 = D(S_2) = D(S_1)$. Since S_1 is invertible, $D(S_1) \neq 0$, and therefore $D(S_1) = 1$. So $D(S_t) = 1$ for all t (for $t = 0$, $S_0 = I_2$). Next, $P = P^{-1}$ and $PS_tP = L_t$, so $D(L_t) = 1$ as well.

Step 2: scalings and the swap. $P \operatorname{diag}(t,1)P = \operatorname{diag}(1,t)$, so $D(\operatorname{diag}(1,t)) = t$. Multiplying out,

$$S_1L_{-1}S_1 = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} S_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \operatorname{diag}(-1,1) = P,$$

so $D(P) = 1 \cdot 1 \cdot 1 \cdot (-1) = -1$.

Step 3: invertible matrices. The elementary matrices in $\mathbb{R}^{2\times 2}$ are P, the scalings $\operatorname{diag}(t,1)$ and $\operatorname{diag}(1,t)$ with $t \neq 0$, and the shears S_t and L_t. By Steps 1 and 2, $D(E) = \det E$ for each of them ($\det P = -1$, $\det \operatorname{diag}(t,1) = \det \operatorname{diag}(1,t) = t$, and $\det S_t = \det L_t = 1$). If A is invertible, then $A = E_1 \cdots E_k$ for some elementary matrices E_i (Theorem 3.5.19), and by multiplicativity of D and of det (Exercise 5.1.27),

$$D(A) = D(E_1) \cdots D(E_k) = \det(E_1) \cdots \det(E_k) = \det A.$$

Step 4: non-invertible matrices. Let A be non-invertible, so $\det A = 0$ and its columns are dependent (Exercise 5.1.25). Note $D(\operatorname{diag}(0,1)) = 0$, and $D(\operatorname{diag}(1,0)) = D(P \operatorname{diag}(0,1)P) = 0$ by Step 0. If $\mathbf{a}_1 = \mathbf{0}$, then $A = A \operatorname{diag}(0,1)$, so $D(A) = D(A) \cdot 0 = 0$. If $\mathbf{a}_1 \neq \mathbf{0}$, a dependence $x\mathbf{a}_1 + y\mathbf{a}_2 = \mathbf{0}$ must have $y \neq 0$, so $\mathbf{a}_2 = c\mathbf{a}_1$ with $c = -y^{-1}x$. Then AS_{-c} has columns $\mathbf{a}_1$ and $\mathbf{0}$, so $AS_{-c} = AS_{-c} \operatorname{diag}(1,0)$, and

$$D(A) = D(A)D(S_{-c}) = D(AS_{-c}) = D(AS_{-c})D(\operatorname{diag}(1,0)) = 0.$$

In both cases $D(A) = 0 = \det A$. Hence $D = \det$.

Section 5.2

Solution to Exercise 5.2.9

Uses. Proposition 5.2.5 (Composition and inverses)
Hint 1. $\sigma^2(i) = \sigma(\sigma(i))$. Apply σ to each entry of the bottom row of σ.
Hint 2. Compare σ^3 with σ^{-1}: what is σ^4?
Sketch. $\sigma^2 = \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$, $\sigma^3 = \begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix} = \sigma^{-1}$, and $\sigma^4 = \mathrm{id}$. Since σ, σ^2, σ^3 are not the identity, $m = 4$.

Applying σ to each entry $2, 4, 1, 3$ of the bottom row of σ gives $\sigma(2), \sigma(4), \sigma(1), \sigma(3) = 4, 3, 2, 1$, so $\sigma^2 = \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$. Applying σ to $4, 3, 2, 1$ gives $3, 1, 4, 2$, so $\sigma^3 = \begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix}$, and applying it once more gives $1, 2, 3, 4$, so $\sigma^4 = \mathrm{id}$. Reading σ from the bottom up, σ^{-1} sends $2, 4, 1, 3$ to $1, 2, 3, 4$, so $\sigma^{-1} = \begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix} = \sigma^3$, as it must be since $\sigma^3\sigma = \sigma^4 = \mathrm{id}$. None of σ, σ^2, σ^3 is the identity, so the least m is 4.

Solution to Exercise 5.2.28

Uses. Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.26 (Sign of a cycle)
Hint 1. Apply sgn and use Proposition 5.2.25. Values of sgn are numbers, so they commute.
Hint 2. For (b), compute the signs of the two given permutations.
Sketch. $\mathrm{sgn}(\sigma^2) = \mathrm{sgn}(\sigma)^2 = 1$, and $\mathrm{sgn}(\sigma\rho\sigma^{-1}\rho^{-1}) = \mathrm{sgn}(\sigma)^2\mathrm{sgn}(\rho)^2 = 1$. A transposition is odd, and $(1\ 2)(3\ 4\ 5)$ has sign $(-1)(+1) = -1$, so neither is a square.

(a) By Proposition 5.2.25, $\mathrm{sgn}(\sigma^2) = \mathrm{sgn}(\sigma)^2 = 1$, and

$$\mathrm{sgn}(\sigma\rho\sigma^{-1}\rho^{-1}) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\rho)\,\mathrm{sgn}(\sigma^{-1})\,\mathrm{sgn}(\rho^{-1}) = \mathrm{sgn}(\sigma)^2\mathrm{sgn}(\rho)^2 = 1,$$

where we used $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ and that the integers $\mathrm{sgn}(\sigma), \mathrm{sgn}(\rho)$ commute.

(b) $\mathrm{sgn}((1\ 2)) = -1$, and by Proposition 5.2.26 and Proposition 5.2.25, $\mathrm{sgn}((1\ 2)(3\ 4\ 5)) = (-1)^1(-1)^2 = -1$. By (a) every square has sign 1, so neither permutation is a square.

Solution to Exercise 5.2.33

Uses. Proposition 5.2.16 (Cycle decomposition), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign)
Hint 1. Write τ in two-line notation first. For $\sigma\tau$, apply τ first.
Hint 2. For the signs, use the cycle notation and Proposition 5.2.26, and check the answers against Proposition 5.2.25.
Sketch. $\sigma = (1\ 2\ 5\ 3\ 6\ 4)$ and $\tau = \begin{pmatrix}1&2&3&4&5&6\\4&6&3&1&5&2\end{pmatrix}$. Then $\sigma\tau = \begin{pmatrix}1&2&3&4&5&6\\1&4&6&2&3&5\end{pmatrix} = (2\ 4)(3\ 6\ 5)$, $\tau\sigma = \begin{pmatrix}1&2&3&4&5&6\\6&5&2&4&3&1\end{pmatrix} = (1\ 6)(2\ 5\ 3)$, and $\sigma^{-1} = \begin{pmatrix}1&2&3&4&5&6\\4&1&5&6&2&3\end{pmatrix} = (1\ 4\ 6\ 3\ 5\ 2)$. The signs are $-1, 1, -1, -1, -1$.

Following $1 \mapsto 2 \mapsto 5 \mapsto 3 \mapsto 6 \mapsto 4 \mapsto 1$, we get $\sigma = (1\ 2\ 5\ 3\ 6\ 4)$, a 6-cycle. The permutation τ interchanges $1, 4$ and $2, 6$, so $\tau = \begin{pmatrix}1&2&3&4&5&6\\4&6&3&1&5&2\end{pmatrix}$.

$\sigma\tau$. Applying σ to the bottom row $4, 6, 3, 1, 5, 2$ of τ gives $\sigma(4), \sigma(6), \sigma(3), \sigma(1), \sigma(5), \sigma(2) = 1, 4, 6, 2, 3, 5$. So $\sigma\tau = \begin{pmatrix}1&2&3&4&5&6\\1&4&6&2&3&5\end{pmatrix} = (2\ 4)(3\ 6\ 5)$.

$\tau\sigma$. Applying τ to the bottom row $2, 5, 6, 1, 3, 4$ of σ gives $6, 5, 2, 4, 3, 1$. So $\tau\sigma = \begin{pmatrix}1&2&3&4&5&6\\6&5&2&4&3&1\end{pmatrix} = (1\ 6)(2\ 5\ 3)$.

σ^{-1}. Running the cycle of σ backwards, $\sigma^{-1} = (4\ 6\ 3\ 5\ 2\ 1) = (1\ 4\ 6\ 3\ 5\ 2)$, which in two-line notation is $\begin{pmatrix}1&2&3&4&5&6\\4&1&5&6&2&3\end{pmatrix}$.

Signs. By Proposition 5.2.26, a 6-cycle has sign $(-1)^5 = -1$, so $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1}) = -1$. Also $\mathrm{sgn}(\tau) = (-1)(-1) = 1$, $\mathrm{sgn}(\sigma\tau) = (-1)\cdot 1 = -1$ (a transposition times a 3-cycle), and $\mathrm{sgn}(\tau\sigma) = -1$ in the same way. These agree with Proposition 5.2.25: $\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\tau\sigma) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau) = -1$.

Solution to Exercise 5.2.34

Uses. Definition 5.2.21 (Inversion), Corollary 5.2.24 (Formulas for the sign), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign)
Hint 1. In (a), the bottom row of $(k\ l)$ is $1, \ldots, k-1, l, k+1, \ldots, l-1, k, l+1, \ldots, n$. Only pairs involving position k or position l can be inverted.
Hint 2. In (b), the bottom row is $2, 3, \ldots, n, 1$. In (c), every pair is inverted, and ω is a product of the disjoint transpositions $(i\ n+1-i)$ for $i < (n+1)/2$.

Sketch. (a) Position k (entry l) is inverted with each of the $l-k$ later positions up to l, and position l (entry k) with each of the $l-k-1$ positions strictly between. (b) The inversions are (i,n) for $i<n$, so $\mathrm{inv}=n-1$ and the sign is $(-1)^{n-1}$. (c) $\mathrm{inv}(\omega)=n(n-1)/2$, while ω is a product of $\lfloor n/2\rfloor$ disjoint transpositions. So ω is even exactly when $n\equiv 0$ or 1 modulo 4.

(a) The bottom row of $(k\ l)$ has l in position k, k in position l, and i in position i otherwise. A pair of positions avoiding k and l has entries $i<j$ in order. Let $m\notin\{k,l\}$. If $m<k$, its entry m is smaller than both l and k, so neither pair (m,k) nor (m,l) is inverted. If $m>l$, its entry m is larger than both, so again neither is inverted. If $k<m<l$, then (k,m) is inverted since $l>m$, and (m,l) is inverted since $m>k$. Finally, (k,l) is inverted since $l>k$. So the inversions are (k,m) and (m,l) for the $l-k-1$ values of m strictly between k and l, together with (k,l): in total $2(l-k-1)+1=2(l-k)-1$.

(b) The bottom row of $(1\ 2\ \ldots\ n)$ is $2,3,\ldots,n,1$. The entries $2,\ldots,n$ are in increasing order, and each is larger than the final entry 1. So the inversions are (i,n) for $1\le i\le n-1$, and $\mathrm{inv}=n-1$. By Corollary 5.2.24 the sign is $(-1)^{n-1}$, which is the sign of an n-cycle in Proposition 5.2.26.

(c) The bottom row of ω is $n,n-1,\ldots,1$, so every pair $i<j$ is an inversion and $\mathrm{inv}(\omega)=n(n-1)/2$. Hence $\mathrm{sgn}(\omega)=(-1)^{n(n-1)/2}$ by Corollary 5.2.24. On the other hand, ω interchanges i and $n+1-i$, so it is the product of the disjoint transpositions $(i\ n+1-i)$ for $1\le i<(n+1)/2$; there are $\lfloor n/2\rfloor$ of them (the middle point is fixed when n is odd). So $\mathrm{sgn}(\omega)=(-1)^{\lfloor n/2\rfloor}$ by Proposition 5.2.25, and the two expressions agree. Now $\lfloor n/2\rfloor$ is even exactly when $n\equiv 0$ or 1 modulo 4, so ω is even exactly for these n. For example ω is odd in S_2 and S_3, even in S_4 and S_5, and odd in S_6 and S_7.

Solution to Exercise 5.2.35

Uses. Lemma 5.2.22 (A transposition changes inv by an odd number), Exercise 5.2.34
Hint 1. Test the claim with $\sigma=\mathrm{id}$ and $\tau=(1\ 3)$.
Hint 2. Look at an entry that sits between the two interchanged positions. Is it still on the same side of each interchanged entry?
Sketch. The claim fails for $\sigma=\mathrm{id},\tau=(1\ 3)$: $\mathrm{inv}((1\ 3))=3$. The error is "nothing else": every entry between the two interchanged positions also changes its order relative to both interchanged entries. The correct statement is Lemma 5.2.22: the change is odd, and it is ±1 for adjacent transpositions.

The claim is false: $\mathrm{inv}(\mathrm{id})=0$, while the bottom row of $(1\ 3)$ is $3,2,1$, which has 3 inversions. The error is the sentence "changes their relative order and nothing else". Think of an inversion as a pair of *entries* that appear in the wrong order, and interchange the entries a_k and a_l in positions $k<l$. An entry a_m with $m<k$ or $m>l$ stays on the same side of both a_k and a_l, so its pairs with them keep their status. But an entry a_m with $k<m<l$ used to stand to the right of a_k and to the left of a_l, and afterwards stands to the left of a_k and to the right of a_l. So the relative order of a_m with *both* interchanged entries changes, and the pairs $\{a_m,a_k\}$ and $\{a_m,a_l\}$ both change status. In the example, the entry 2 in the middle forms an inversion with 3 and with 1 afterwards, and with neither before. Each such m changes the count by -2, 0, or 2, which is why the total change is always odd (Lemma 5.2.22). When $l=k+1$ there is no entry in between, and then the claim is true.

Solution to Exercise 5.2.36

Uses. Definition 5.2.29 (Permutation matrix), Definition 2.4.1 (Matrix product), Proposition 5.2.30 (Algebra of permutation matrices)
Hint 1. Compute the $(\sigma(i),k)$ entry of $E_\sigma A$ from the definition of the product: $\sum_j (E_\sigma)_{\sigma(i)j}a_{jk}$. Which j gives a nonzero term?
Hint 2. For (b), $E_\sigma^{\mathsf T}=E_{\sigma^{-1}}$, so by Proposition 5.2.30 the columns of $AE_\sigma^{\mathsf T}$ are the columns of A rearranged. Combine this with (a).
Sketch. $(E_\sigma)_{\sigma(i)j}=1$ exactly when $j=i$, so $(E_\sigma A)_{\sigma(i)k}=a_{ik}$. Dually, column $\sigma(j)$ of $AE_{\sigma^{-1}}$ is column j of A. Together, $\left(E_\sigma AE_\sigma^{\mathsf T}\right)_{\sigma(i)\sigma(j)}=a_{ij}$: renaming vertex i as $\sigma(i)$ replaces the adjacency matrix A by $E_\sigma AE_\sigma^{\mathsf T}$.

(a) For every k, the $(\sigma(i),k)$ entry of $E_\sigma A$ is $\sum_{j=1}^{n}(E_\sigma)_{\sigma(i)j}a_{jk}$. By Definition 5.2.29, $(E_\sigma)_{\sigma(i)j}=1$ if $\sigma(i)=\sigma(j)$, that is, if $j=i$, and 0 otherwise. So the sum is a_{ik}, and row $\sigma(i)$ of $E_\sigma A$ is row i of A.

(b) Let $B = AE_\sigma^{\mathsf{T}} = AE_{\sigma^{-1}}$ (Proposition 5.2.30). By part (a) of that proposition, column p of B is column $\sigma^{-1}(p)$ of A; with $p = \sigma(j)$, column $\sigma(j)$ of B is column j of A, so $b_{i\sigma(j)} = a_{ij}$. By (a) applied to B, row $\sigma(i)$ of $E_\sigma B$ is row i of B. Hence the $(\sigma(i), \sigma(j))$ entry of $E_\sigma A E_\sigma^{\mathsf{T}} = E_\sigma B$ is $b_{i\sigma(j)} = a_{ij}$.

(c) Rename each vertex i as $\sigma(i)$. In the renamed graph, $\sigma(i)$ and $\sigma(j)$ are joined exactly when i and j were joined in the original, so its adjacency matrix A' satisfies $a'_{\sigma(i)\sigma(j)} = a_{ij}$ for all i, j. Since σ is a bijection, every entry of A' is of this form, so by (b), $A' = E_\sigma A E_\sigma^{\mathsf{T}} = E_\sigma A E_\sigma^{-1}$. Renaming the vertices replaces the adjacency matrix by a similar matrix, conjugated by a permutation matrix.

Solution to Exercise 5.2.37

> *Uses.* Proposition 5.2.25 (Multiplicativity of the sign), Definition 5.2.21 (Inversion)
> *Hint 1.* (a) and (d): try $\sigma = \tau = (1\ 2)$. (b): try two disjoint transpositions.
> *Hint 2.* (c): an inversion of σ is a pair $i < j$ with $\sigma(i) > \sigma(j)$. Rename $p = \sigma(j)$ and $q = \sigma(i)$, and see what σ^{-1} does to $p < q$.
> *Hint 3.* (e): use Proposition 5.2.25.
> *Sketch.* (a) False: $(1\ 2)^2 = \mathrm{id}$ is even. (b) False for $n \geq 4$: $(1\ 2)(3\ 4)$. (c) True: $(i, j) \mapsto (\sigma(j), \sigma(i))$ is a bijection from the inversions of σ to those of σ^{-1}. (d) False: $\mathrm{inv}\big((1\ 2)^2\big) = 0 \neq 2$. (e) True, by multiplicativity.

(a) False. $(1\ 2)$ is odd, but $(1\ 2)(1\ 2) = \mathrm{id}$ is even. In fact the product of two odd permutations is always even, by Proposition 5.2.25.

(b) False when $n \geq 4$. For $\sigma = (1\ 2)(3\ 4)$, the two disjoint transpositions commute, so $\sigma^2 = (1\ 2)^2(3\ 4)^2 = \mathrm{id}$, but σ moves four points, so it is neither the identity nor a transposition. (For $n \leq 3$ the statement is true, as Table 5.2.1 shows for $n = 3$.)

(c) True. Let (i, j) be an inversion of σ, so $i < j$ and $\sigma(i) > \sigma(j)$. Put $p = \sigma(j)$ and $q = \sigma(i)$. Then $p < q$ and $\sigma^{-1}(p) = j > i = \sigma^{-1}(q)$, so (p, q) is an inversion of σ^{-1}. The same argument applied to σ^{-1}, whose inverse is σ, sends an inversion (p, q) of σ^{-1} to the inversion $\big(\sigma^{-1}(q), \sigma^{-1}(p)\big)$ of σ, and the two maps are inverse to each other. So they are bijections, and $\mathrm{inv}\big(\sigma^{-1}\big) = \mathrm{inv}(\sigma)$.

(d) False. With $\sigma = \tau = (1\ 2)$, $\mathrm{inv}(\sigma\tau) = \mathrm{inv}(\mathrm{id}) = 0$ but $\mathrm{inv}(\sigma) + \mathrm{inv}(\tau) = 2$. Only the parities add (Corollary 5.2.24 and Proposition 5.2.25).

(e) True. By Proposition 5.2.25, $\mathrm{sgn}\big(\sigma\tau\sigma^{-1}\big) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)\,\mathrm{sgn}\big(\sigma^{-1}\big) = \mathrm{sgn}(\sigma)^2\,\mathrm{sgn}(\tau) = \mathrm{sgn}(\tau)$.

Solution to Exercise 5.2.38

> *Uses.* Theorem 5.3.4 (Existence and uniqueness of the determinant), Remark 5.2.32 (Two comments on the theorem), Theorem 5.2.31 (Alternating functions and permutations)
> *Hint 1.* How is the existence of a determinant function on $F^{n\times n}$ proved in Section 5.3, and what does that proof use?
> *Hint 2.* Suppose $F = \mathbb{F}_2$. What does $\det E_\sigma = (-1)^r$ tell you about r?
> *Sketch.* The argument is circular: the existence half of Theorem 5.3.4 is proved with a formula involving sgn, which is defined using Theorem 5.2.18. It also takes F arbitrary, but if $1 + 1 = 0$ then $(-1)^r = 1$ for every r and nothing about r follows. Over $\mathbb{Q}$, with a determinant function constructed without permutations, the argument would be valid.

There are two errors.

Circularity. The proof assumes that a determinant function exists on $F^{n\times n}$. For $n \geq 3$ the existence of a determinant function is the existence half of Theorem 5.3.4, which Section 5.3 proves by checking that a formula built from sgn satisfies the axioms. That formula only makes sense once sgn is well defined, that is, once Theorem 5.2.18 is known. So the argument assumes what it is trying to prove.

The field. Even granting a determinant function, the conclusion "$(-1)^r$ is determined, so the parity of r is determined" needs $1 \neq -1$ in F. Over $\mathbb{F}_2$, $(-1)^r = 1$ for every r, so the equation $\det E_\sigma = (-1)^r$ says nothing about r.

The core of the argument is sound, as Remark 5.2.32 explains: if one had a multilinear alternating D on $\mathbb{Q}^{n\times n}$ with $D(I_n) \neq 0$, constructed without permutations, then the induction in the proof of Theorem 5.2.31 gives $(-1)^r D(I_n) = D(E_\sigma) = (-1)^s D(I_n)$ for two factorizations, and $r \equiv s$ modulo 2 follows.

Solution to Exercise 5.2.39

Uses. Definition 5.2.15 (Cycle), Proposition 5.2.5 (Composition and inverses), Proposition 5.2.16 (Cycle decomposition)
Hint 1. Evaluate the left side at $\sigma(a_i)$, and at a point y not of this form.
Hint 2. For (b), find ρ with $\rho(1) = k$ and $\rho(2) = l$, and apply (a).
Hint 3. For (c), write the conjugate of a product as a product of conjugates: $\sigma\gamma\delta\sigma^{-1} = \big(\sigma\gamma\sigma^{-1}\big)\big(\sigma\delta\sigma^{-1}\big)$.
Sketch. Let $\gamma = (a_1 \ \dots \ a_k)$. At $\sigma(a_i)$, $\sigma\gamma\sigma^{-1}$ gives $\sigma(\gamma(a_i)) = \sigma(a_{i+1})$ (indices mod k), and at any other y, $\sigma^{-1}(y)$ is not an a_i, so y is fixed. For (b), take ρ with $\rho(1) = k$, $\rho(2) = l$. For (c), rename each entry by σ: (5 1 4)(3 2).

(a) Let $\gamma = (a_1 \ \dots \ a_k)$, and read indices modulo k, so $\gamma(a_i) = a_{i+1}$ and $a_{k+1} = a_1$. For each i,

$$\sigma\gamma\sigma^{-1}(\sigma(a_i)) = \sigma(\gamma(a_i)) = \sigma(a_{i+1}).$$

If $y \in [n]$ is not one of $\sigma(a_1), \dots, \sigma(a_k)$, then $\sigma^{-1}(y)$ is not one of $a_1, \dots, a_k$, so γ fixes it and $\sigma\gamma\sigma^{-1}(y) = \sigma\big(\sigma^{-1}(y)\big) = y$. Since $\sigma(a_1), \dots, \sigma(a_k)$ are distinct (as σ is injective), this is exactly the cycle $(\sigma(a_1) \ \dots \ \sigma(a_k))$.

(b) The sets $[n] \setminus \{1, 2\}$ and $[n] \setminus \{k, l\}$ both have $n - 2$ elements, so there is a bijection β between them. Define $\rho(1) = k$, $\rho(2) = l$, and $\rho(i) = \beta(i)$ for $i \geq 3$. Then ρ is a bijection of $[n]$, and by (a), $\rho(1\ 2)\rho^{-1} = (\rho(1)\ \rho(2)) = (k\ l)$.

(c) Since $\sigma^{-1}\sigma = \mathrm{id}$, we have $\sigma\gamma\delta\sigma^{-1} = \big(\sigma\gamma\sigma^{-1}\big)\big(\sigma\delta\sigma^{-1}\big)$. Here σ sends $1 \mapsto 5$, $2 \mapsto 1$, $3 \mapsto 4$, $4 \mapsto 3$, $5 \mapsto 2$, so by (a),

$$\sigma(1\ 2\ 3)(4\ 5)\sigma^{-1} = (5\ 1\ 4)(3\ 2) = (1\ 4\ 5)(2\ 3).$$

Solution to Exercise 5.2.40

Uses. Definition 5.2.23 (Sign of a permutation), Lemma 5.2.17 (Cycles as products of transpositions), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign)
Hint 1. An even permutation is a product of an even number of transpositions. Group them in pairs.
Hint 2. Show that a product of two transpositions is a product of at most two 3-cycles. Distinguish whether the two transpositions are equal, share one point, or are disjoint.
Hint 3. If they share one point, name them $(a\ b)$ and $(b\ c)$; by Lemma 5.2.17 their product is $(a\ b\ c)$. If they are disjoint, insert $(b\ c)(b\ c) = \mathrm{id}$.
Sketch. Pair up the transpositions of an even factorization. Equal transpositions multiply to id; $(a\ b)(b\ c) = (a\ b\ c)$; and for disjoint ones, $(a\ b)(c\ d) = (a\ b)(b\ c)(b\ c)(c\ d) = (a\ b\ c)(b\ c\ d)$. Conversely, 3-cycles are even.

Let σ be even, so $\sigma = \tau_1\tau_2\cdots\tau_{2q}$ for some transpositions τ_i, and $\sigma = (\tau_1\tau_2)(\tau_3\tau_4)\cdots\big(\tau_{2q-1}\tau_{2q}\big)$. It suffices to show that a product π of two transpositions is a product of 3-cycles (with the empty product allowed). There are three cases.

The transpositions are equal. Then $\pi = \mathrm{id}$, the empty product.

They share exactly one point. Since $(x\ y) = (y\ x)$, we can write them as $(a\ b)$ and $(b\ c)$ with a, b, c distinct, in the order they occur in π. By Lemma 5.2.17, $\pi = (a\ b)(b\ c) = (a\ b\ c)$.

They are disjoint, say $(a\ b)$ and $(c\ d)$ with a, b, c, d distinct. Inserting $(b\ c)^2 = \mathrm{id}$ and using the previous case twice,

$$\pi = (a\ b)(b\ c)(b\ c)(c\ d) = (a\ b\ c)(b\ c\ d).$$

Conversely, a 3-cycle is even by Proposition 5.2.26, so a product of 3-cycles is even by Proposition 5.2.25. Hence a permutation is even if and only if it is a product of 3-cycles. (We need $n \geq 3$ for 3-cycles to exist; for $n \leq 2$ the only even permutation is id, the empty product.)

Solution to Exercise 5.2.41

Uses. Proposition 5.2.16 (Cycle decomposition), Proposition 5.2.30 (Algebra of permutation matrices), Definition 5.2.15 (Cycle)
Hint 1. For a k-cycle $\gamma = (a_1 \ \dots \ a_k)$, compute $\gamma^m(a_i)$.
Hint 2. Disjoint cycles commute, so $\sigma^m = \gamma_1^m \cdots \gamma_t^m$, and γ_s^m moves only points of its own cycle. When is the product the identity?
Hint 3. For (b), $E_\sigma^m = E_{\sigma^m}$, and $E_\pi = I_n$ only when $\pi = \mathrm{id}$.

Sketch. $\gamma^m(a_i) = a_{i+m}$ (indices mod k), so $\gamma^m = \mathrm{id}$ iff $k \mid m$. Since the γ_s commute and move disjoint sets, $\sigma^m = \mathrm{id}$ iff each $\gamma_s^m = \mathrm{id}$ iff each $k_s \mid m$. In (b), $E_\sigma^m = E_{\sigma^m}$, so the answer is the order of σ, $\mathrm{lcm}(3,4) = 12$.

(a) Let $\gamma = (a_1 \ \ldots \ a_k)$ and read indices modulo k. By induction on m, $\gamma^m(a_i) = a_{i+m}$, and γ^m fixes every point outside $\{a_1, \ldots, a_k\}$. So $\gamma^m = \mathrm{id}$ if and only if $a_{i+m} = a_i$ for all i, that is, if and only if k divides m. The least such $m \geq 1$ is k.

Now let $\sigma = \gamma_1 \cdots \gamma_t$ with disjoint cycles γ_s of lengths k_s. Disjoint cycles commute, so $\sigma^m = \gamma_1^m \cdots \gamma_t^m$. Each γ_s^m maps the entries of γ_s among themselves and fixes all other points. So for x an entry of γ_s, $\sigma^m(x) = \gamma_s^m(x)$, and σ^m fixes the points that no γ_s moves. Hence $\sigma^m = \mathrm{id}$ if and only if $\gamma_s^m = \mathrm{id}$ for every s, that is, if and only if every k_s divides m. The least such $m \geq 1$ is $\mathrm{lcm}(k_1, \ldots, k_t)$.

(b) By Proposition 5.2.30, $E_\sigma^m = E_{\sigma^m}$. A permutation matrix E_π equals I_7 only if $\mathbf{e}_{\pi(j)} = \mathbf{e}_j$ for all j, that is, $\pi = \mathrm{id}$. So $E_\sigma^m = I_7$ if and only if $\sigma^m = \mathrm{id}$, and by (a) the least such m is $\mathrm{lcm}(3,4) = 12$.

Solution to Exercise 5.2.42

Uses. Exercise 5.2.28, Exercise 5.2.41, Proposition 5.2.26 (Sign of a cycle)
Hint 1. Square a 5-cycle $(b_1\ b_2\ b_3\ b_4\ b_5)$: where does b_1 go?
Hint 2. For (b), the square of a 4-cycle $(b_1\ b_2\ b_3\ b_4)$ is $(b_1\ b_3)(b_2\ b_4)$.
Hint 3. For (c), compute a sign (Exercise 5.2.28).
Sketch. (a) $(b_1 \ \ldots \ b_5)^2 = (b_1\ b_3\ b_5\ b_2\ b_4)$, so $\sigma = (1\ 4\ 2\ 5\ 3)$. (b) $\sigma = (1\ 3\ 2\ 4)$. (c) $(1\ 2)(3\ 4\ 5)$ is odd, and squares are even.

(a) For a 5-cycle $\gamma = (b_1\ b_2\ b_3\ b_4\ b_5)$, γ^2 sends $b_i \mapsto b_{i+2}$ (indices mod 5), so $\gamma^2 = (b_1\ b_3\ b_5\ b_2\ b_4)$. We want this to be $(1\ 2\ 3\ 4\ 5)$, so we take $b_1 = 1, b_3 = 2, b_5 = 3, b_2 = 4, b_4 = 5$, that is, $\sigma = (1\ 4\ 2\ 5\ 3)$. Check: σ sends $1 \mapsto 4 \mapsto 2 \mapsto 5 \mapsto 3 \mapsto 1$, so σ^2 sends $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 4$, $4 \mapsto 5$, $5 \mapsto 1$.

(b) For $\gamma = (b_1\ b_2\ b_3\ b_4)$, γ^2 sends $b_1 \mapsto b_3 \mapsto b_1$ and $b_2 \mapsto b_4 \mapsto b_2$, so $\gamma^2 = (b_1\ b_3)(b_2\ b_4)$. Taking $b_1 = 1, b_3 = 2, b_2 = 3, b_4 = 4$ gives $\sigma = (1\ 3\ 2\ 4)$ with $\sigma^2 = (1\ 2)(3\ 4)$.

(c) By Proposition 5.2.26, $\mathrm{sgn}((1\ 2)(3\ 4\ 5)) = (-1)(+1) = -1$. By Exercise 5.2.28, every square σ^2 is even. So $(1\ 2)(3\ 4\ 5)$ is not a square.

Solution to Exercise 5.2.43

Uses. Lemma 5.2.22 (A transposition changes inv by an odd number), Definition 5.2.21 (Inversion)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. For the lower bound, start from id and multiply on the right by adjacent transpositions one at a time. By how much can inv change at each step?
Hint 2. For the upper bound, induct on $\mathrm{inv}(\sigma)$. If $\sigma \neq \mathrm{id}$, find k with $\sigma(k) > \sigma(k+1)$, and show that $\mathrm{inv}(\sigma(k\ k+1)) = \mathrm{inv}(\sigma) - 1$.
Sketch. Each adjacent transposition changes inv by exactly 1 (Lemma 5.2.22), so a product of r of them has inv $\leq r$. Conversely, if $\mathrm{inv}(\sigma) > 0$, the bottom row has a descent $\sigma(k) > \sigma(k+1)$; interchanging those two entries removes exactly one inversion. Repeating (bubble sort) reaches id after $\mathrm{inv}(\sigma)$ steps, and reversing the steps writes σ as a product of $\mathrm{inv}(\sigma)$ adjacent transpositions.

Lower bound. Suppose $\sigma = s_1 \cdots s_r$ with adjacent transpositions s_i. By Lemma 5.2.22, $\mathrm{inv}(s_1 \cdots s_j) \leq \mathrm{inv}(s_1 \cdots s_{j-1}) + 1$ for each j, and $\mathrm{inv}(\mathrm{id}) = 0$, so by induction $\mathrm{inv}(s_1 \cdots s_j) \leq j$. In particular $\mathrm{inv}(\sigma) \leq r$.

Upper bound. We show by induction on $N = \mathrm{inv}(\sigma)$ that σ is a product of N adjacent transpositions. If $N = 0$, the bottom row of σ has no pair out of order, so it is increasing, and an increasing arrangement of $1, \ldots, n$ is $1, \ldots, n$. So $\sigma = \mathrm{id}$, the empty product. If $N > 0$, the bottom row is not increasing, so $\sigma(k) > \sigma(k+1)$ for some k. Let $s = (k\ k+1)$. By Remark 5.2.11, the bottom row of σs is that of σ with the entries in positions k and $k+1$ interchanged. The proof of Lemma 5.2.22 shows that, for an adjacent transposition, the pairs other than $(k, k+1)$ contribute the same number of inversions before and after, so $\mathrm{inv}(\sigma s) - \mathrm{inv}(\sigma)$ is the change in the status of $(k, k+1)$ alone. Here that pair changes from inverted to not inverted, so $\mathrm{inv}(\sigma s) = N - 1$, and by the induction hypothesis $\sigma s = s_1 \cdots s_{N-1}$ with adjacent transpositions s_i. Then $\sigma = \sigma s s = s_1 \cdots s_{N-1} s$, a product of N adjacent transpositions.

The upper bound is the analysis of *bubble sort*, which repeatedly interchanges an adjacent pair that is out of order. Each interchange lowers inv by exactly 1, so the algorithm stops after exactly $\operatorname{inv}(\sigma)$ interchanges, whichever out-of-order pairs it picks.

Solution to Exercise 5.2.44

Uses. Theorem 5.2.31 (Alternating functions and permutations), Definition 5.1.1 (Determinant function), Proposition 5.2.26 (Sign of a cycle)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. In (a), let $\sigma(j)$ be the row of the nonzero entry in column j. Why is σ a permutation?
Hint 2. Pull the scalars c_j out of the columns by multilinearity, and apply Theorem 5.2.31 to what is left.
Hint 3. In (b), read off σ column by column and find its cycle notation.
Sketch. (a) The rows $\sigma(j)$ of the nonzero entries are distinct because each row has one nonzero entry, so $\sigma \in S_n$. Multilinearity gives $D(A) = c_1 \cdots c_n D(E_\sigma) = c_1 \cdots c_n \operatorname{sgn}(\sigma) D(I_n)$. (b) $\sigma = (1\ 3\ 4\ 2)$, odd, and $c_1c_2c_3c_4 = i \cdot 2 \cdot (1-i)(-1) = -2-2i$, so $D(A) = (2+2i)D(I_4)$. (c) $\sigma = (1\ 3\ 2)$, even, so $D(B) = 8D(I_3) = 2D(I_3)$ in $\mathbb{F}_3$.

(a) For each column j, let $\sigma(j)$ be the row of its nonzero entry and c_j that entry, so column j is $c_j \mathbf{e}_{\sigma(j)}$. If $\sigma(j) = \sigma(j')$ with $j \neq j'$, row $\sigma(j)$ would have two nonzero entries; so σ is injective, hence a permutation. Pulling c_j out of column j for each j (multilinearity, Definition 5.1.1) and applying Theorem 5.2.31,

$$D(A) = c_1 \cdots c_n D\big(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}\big) = c_1 \cdots c_n D(E_\sigma) = \operatorname{sgn}(\sigma) c_1 \cdots c_n D(I_n).$$

(b) The columns of A are $i\mathbf{e}_3$, $2\mathbf{e}_1$, $(1-i)\mathbf{e}_4$, and $-\mathbf{e}_2$. So $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$, that is, $1 \mapsto 3 \mapsto 4 \mapsto 2 \mapsto 1$, the 4-cycle $(1\ 3\ 4\ 2)$, which is odd by Proposition 5.2.26. The product of the scalars is

$$i \cdot 2 \cdot (1-i) \cdot (-1) = -2i(1-i) = -2i + 2i^2 = -2 - 2i.$$

So $D(A) = -(-2-2i)D(I_4) = (2+2i)D(I_4)$.

(c) The columns of B are $2\mathbf{e}_3$, $2\mathbf{e}_1$, $2\mathbf{e}_2$, so $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1\ 3\ 2)$, a 3-cycle, which is even. Hence $D(B) = 2^3 D(I_3) = 8D(I_3) = 2D(I_3)$, since $8 = 2$ in $\mathbb{F}_3$.

Solution to Exercise 5.2.45

Uses. Definition 5.2.15 (Cycle), Proposition 5.2.16 (Cycle decomposition), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.27 (Half of all permutations are even)
Hint 1. Count lists $a_1, \ldots, a_k$ of distinct elements, then ask how many lists name the same cycle.
Hint 2. Two lists name the same k-cycle exactly when one is a rotation of the other. Why?
Hint 3. For (b), a product of two disjoint transpositions is determined by how it splits $[4]$ into two pairs.
Sketch. There are $n!/(n-k)!$ lists of k distinct elements, and each k-cycle has exactly k names (its rotations), giving $n!/((n-k)!k)$. In S_4: $1 + 6 + 8 + 6 + 3 = 24$. The even ones are id, the 3-cycles, and the double transpositions: $1 + 8 + 3 = 12 = 4!/2$.

(a) By the argument of Proposition 5.2.4, there are $n(n-1)\cdots(n-k+1) = n!/(n-k)!$ lists $a_1, \ldots, a_k$ of distinct elements of $[n]$, and each names the k-cycle $(a_1\ \ldots\ a_k)$. The k rotations $a_p, a_{p+1}, \ldots, a_k, a_1, \ldots, a_{p-1}$ name the same cycle. Conversely, if $(b_1\ \ldots\ b_k) = \gamma = (a_1\ \ldots\ a_k)$, then b_1 is moved by γ, so $b_1 = a_p$ for some p, and then $b_2 = \gamma(b_1) = a_{p+1}$, $b_3 = \gamma(b_2) = a_{p+2}$, and so on (indices mod k); so the list b is a rotation of the list a. The k rotations of a list of distinct elements are distinct lists, since their first entries differ. So each k-cycle has exactly k names, and there are $n!/((n-k)!\,k)$ of them.

(b) By Proposition 5.2.16, each $\sigma \in S_4$ is a product of disjoint cycles whose lengths add up to at most 4, so its cycle notation is one of: id; a transposition; a 3-cycle; a 4-cycle; two disjoint transpositions. By (a) there are $4!/(2! \cdot 2) = 6$ transpositions, $4!/(1! \cdot 3) = 8$ three-cycles, and $4!/(0! \cdot 4) = 6$ four-cycles. A product of two disjoint transpositions $(a\ b)(c\ d)$ is determined by the partition $\{\{a,b\},\{c,d\}\}$ of $[4]$, and there are 3 such partitions (the partner of 1 determines it). The total is $1 + 6 + 8 + 6 + 3 = 24 = 4!$, as it must be.

By Proposition 5.2.26, transpositions and 4-cycles are odd, while id, 3-cycles, and products of two disjoint transpositions are even. So there are $1 + 8 + 3 = 12$ even permutations and $6 + 6 = 12$ odd ones, as Proposition 5.2.27 predicts.

Solution to Exercise 5.2.46

Uses. Exercise 5.2.39, Proposition 5.2.13 (Products of transpositions), Proposition 1.1.17 (No zero divisors), Definition 5.2.23 (Sign of a permutation)
Hint 1. First find $f(\mathrm{id})$. Then show that f takes the same value c on every transposition, using Exercise 5.2.39.
Hint 2. Compute c^2 from $(1\ 2)^2 = \mathrm{id}$, and conclude that $c = \pm 1$.
Hint 3. Every σ is a product of transpositions, so $f(\sigma)$ is a power of c.
Sketch. $f(\mathrm{id})^2 = f(\mathrm{id}) \neq 0$ gives $f(\mathrm{id}) = 1$. Each transposition is $\rho(1\ 2)\rho^{-1}$, so f of it is $f(\rho)cf(\rho)^{-1} = c$ where $c = f((1\ 2))$. Then $c^2 = f(\mathrm{id}) = 1$, so $c = \pm 1$, and $f(\tau_1 \cdots \tau_r) = c^r$ is 1 or $(-1)^r = \mathrm{sgn}(\sigma)$.

$f(\mathrm{id}) = 1$. From $f(\mathrm{id}) = f(\mathrm{id}\ \mathrm{id}) = f(\mathrm{id})^2$ and $f(\mathrm{id}) \neq 0$, cancelling one factor $f(\mathrm{id})$ gives $f(\mathrm{id}) = 1$. Consequently $f(\rho)f\left(\rho^{-1}\right) = f(\mathrm{id}) = 1$, so $f\left(\rho^{-1}\right) = f(\rho)^{-1}$.

f is constant on transpositions. Let $c = f((1\ 2))$. By Exercise 5.2.39, every transposition has the form $\rho(1\ 2)\rho^{-1}$, and since F is commutative,

$$f\left(\rho(1\ 2)\rho^{-1}\right) = f(\rho)cf(\rho)^{-1} = c.$$

$c = \pm 1$. Since $(1\ 2)^2 = \mathrm{id}$, $c^2 = f(\mathrm{id}) = 1$, so $(c-1)(c+1) = 0$, and by Proposition 1.1.17, $c = 1$ or $c = -1$.

Conclusion. Let $\sigma \in S_n$, and write $\sigma = \tau_1 \cdots \tau_r$ with transpositions τ_i (Proposition 5.2.13). Then $f(\sigma) = f(\tau_1) \cdots f(\tau_r) = c^r$. If $c = 1$, then $f(\sigma) = 1$ for all σ. If $c = -1$, then $f(\sigma) = (-1)^r = \mathrm{sgn}(\sigma)$ by Definition 5.2.23. (If $1 + 1 = 0$ in F, the two cases coincide.)

Solution to Exercise 5.2.47

Uses. Proposition 5.2.25 (Multiplicativity of the sign), Remark 5.2.11 (Composing with a transposition)
Techniques. Parity via an invariant (Technique 5.2.19)
Hint 1. Number the cells $1, \ldots, 16$ row by row and call the empty space tile 16. A position is then a permutation $\pi \in S_{16}$, with $\pi(p)$ the tile in cell p. What does a move do to π?
Hint 2. A move interchanges the contents of two adjacent cells, so it multiplies π by a transposition, flipping $\mathrm{sgn}(\pi)$. What else flips with every move?
Hint 3. Color the cells like a checkerboard. Every move changes the color of the cell holding the empty space. Consider the product of $\mathrm{sgn}(\pi)$ and ± 1 for that color.
Sketch. Encode a position as $\pi \in S_{16}$ ($\pi(p)$ = tile in cell p, the empty space being 16). A move with the empty space in cell p and a tile in an adjacent cell q replaces π by $\pi(p\ q)$, which flips $\mathrm{sgn}(\pi)$, and also moves the empty space to a cell of the other checkerboard color. So $\mathrm{sgn}(\pi)$ times the color sign of the empty cell is invariant. It is $+1$ for the solved position and -1 for the target, which has $\pi = (14\ 15)$ and the empty space in the same cell.

Number the cells $1, \ldots, 16$ row by row from the top left, and call the empty space tile 16. A position is then described by the permutation $\pi \in S_{16}$ with $\pi(p)$ the tile in cell p. The solved position is $\pi = \mathrm{id}$, and the target position is $\pi = (14\ 15)$.

Suppose the empty space is in cell p, so $\pi(p) = 16$, and the tile in an adjacent cell q slides into it. Afterwards cell p holds $\pi(q)$, cell q holds 16, and all other cells are unchanged. By Remark 5.2.11 this is the position $\pi(p\ q)$, so by Proposition 5.2.25 every move multiplies $\mathrm{sgn}(\pi)$ by -1.

Give the cell in row r and column c the color $\varepsilon = (-1)^{r+c}$, and let $\varepsilon(\pi)$ be the color of the cell containing the empty space. Two cells that share an edge differ by 1 in exactly one of r, c, so they have opposite colors. Hence every move also multiplies $\varepsilon(\pi)$ by -1, and the product $Q(\pi) = \mathrm{sgn}(\pi)\varepsilon(\pi)$ is unchanged by every move.

In the solved position, $Q = \mathrm{sgn}(\mathrm{id}) \cdot (-1)^{4+4} = 1$. In the target position, the empty space is again in row 4, column 4, and $\mathrm{sgn}((14\ 15)) = -1$, so $Q = -1$. Since moves never change Q, the target cannot be reached from the solved position.

The same argument shows that every position reachable from the solved one with the empty space back in the bottom right corner is an even permutation. Neither $\mathrm{sgn}(\pi)$ nor $\varepsilon(\pi)$ is invariant on its own; the invariant is their product.

Solution to Exercise 5.2.48

Uses. Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign), Definition 5.2.23 (Sign of a permutation)
Techniques. Parity via an invariant (Technique 5.2.19)
Hint 1. Regard the row as the bottom row of a permutation σ. Which permutation γ satisfies "the row after the move is the bottom row of $\sigma\gamma$"?
Hint 2. γ is a 3-cycle, so moves preserve the sign. That rules out odd σ.
Hint 3. For even σ: a move at positions $p-2, p-1, p$ carries the entry at position p two places to the left. Use this to bring 1 to the front, then 2 to position 2, and so on. What happens with the last two entries?
Sketch. A move at positions $i, i+1, i+2$ replaces σ by $\sigma\gamma$ with $\gamma = (i\ i+2\ i+1)$, a 3-cycle, so sgn is invariant and odd orders cannot be sorted. For even orders, move 1 left two places at a time; if it gets stuck in position 2, two moves at positions $1, 2, 3$ bring it to the front. Then sort positions $2, \ldots, n$ the same way without touching position 1, and so on, until only positions $n-1, n$ remain. The result is id or $(n-1\ n)$, and since the sign is still 1, it is id. Answer: exactly the even orders.

Write the row as the bottom row of $\sigma \in S_n$, so that the entry in position i is $\sigma(i)$. A move at positions $i, i+1, i+2$ produces the row whose entries in these positions are $\sigma(i+2), \sigma(i), \sigma(i+1)$. This is the bottom row of $\sigma\gamma$, where $\gamma(i) = i+2$, $\gamma(i+1) = i$, $\gamma(i+2) = i+1$, and γ fixes everything else; that is, $\gamma = (i\ i+2\ i+1)$. By Proposition 5.2.26 and Proposition 5.2.25, $\operatorname{sgn}(\sigma\gamma) = \operatorname{sgn}(\sigma)$. So moves never change the sign, and since the sorted row is id, with sign 1, an odd starting order can never be sorted.

Now let σ be even. We sort the row in stages, placing the value v in position v at stage v, for $v = 1, \ldots, n-2$, using only moves within positions $v, \ldots, n$ at stage v. Suppose the values $1, \ldots, v-1$ are already in positions $1, \ldots, v-1$, and v is in position $p \geq v$. While $p \geq v+2$, a move at positions $p-2, p-1, p$ carries the entry v from position p to position $p-2$ and does not touch positions before $p-2 \geq v$. When this stops, $p = v$ or $p = v+1$. If $p = v+1$, then $v+2 \leq n$ (as $v \leq n-2$), and two moves at positions $v, v+1, v+2$ change x, v, z there into z, x, v and then into v, z, x. Either way v ends in position v, and positions $1, \ldots, v-1$ were not touched.

After stage $n-2$, the values $1, \ldots, n-2$ are in place, and positions $n-1$ and n hold $n-1$ and n in some order. So the current permutation is id or $(n-1\ n)$. Since moves preserve the sign and σ is even, it is even, so it is id: the row is sorted.

Answer: the row can be sorted exactly when the starting order is an even permutation.

Solution to Exercise 5.2.49

Uses. Definition 5.2.21 (Inversion), Exercise 5.2.34, Proposition 5.2.13 (Products of transpositions)
Hint 1. In (a), write $(\sigma\rho(i) - \sigma\rho(j))/(i-j)$ as a product of two fractions. Which two?
Hint 2. For (b), the reversal conditions are 0/1 quantities u, v, and "exactly one" is $u + v - 2uv$. Sum over all pairs, and note that $\{i, j\} \mapsto \{\rho(i), \rho(j)\}$ is a bijection on two-element sets.
Hint 3. For (c), by (b), $(-1)^{\operatorname{inv}(\sigma\rho)} = (-1)^{\operatorname{inv}(\sigma)}(-1)^{\operatorname{inv}(\rho)}$, and inv of a transposition is odd.
Sketch. (a) $(\sigma\rho(i) - \sigma\rho(j))/(i-j) = [(\sigma(\rho(i)) - \sigma(\rho(j)))/(\rho(i) - \rho(j))] \cdot [(\rho(i) - \rho(j))/(i-j)]$, and a product of two nonzero rationals is negative exactly when one factor is. (b) Summing $u + v - 2uv$ over pairs gives $\operatorname{inv}(\sigma\rho) = \operatorname{inv}(\rho) + \operatorname{inv}(\sigma) - 2M$, since the pairs $\{\rho(i), \rho(j)\}$ run over all pairs. (c) So $(-1)^{\operatorname{inv}}$ is multiplicative and is -1 on transpositions; hence $(-1)^r = (-1)^{\operatorname{inv}(\sigma)} = (-1)^s$ for factorizations of lengths r and s.

(a) Let $i \neq j$. Since ρ and σ are injective, $\rho(i) \neq \rho(j)$ and $\sigma(\rho(i)) \neq \sigma(\rho(j))$, so in $\mathbb{Q}$,

$$\frac{\sigma\rho(i) - \sigma\rho(j)}{i-j} = \frac{\sigma(\rho(i)) - \sigma(\rho(j))}{\rho(i) - \rho(j)} \cdot \frac{\rho(i) - \rho(j)}{i-j},$$

a product of two nonzero rationals. The left side is negative exactly when $\sigma\rho$ reverses $\{i, j\}$, the second factor is negative exactly when ρ reverses $\{i, j\}$, and the first factor is negative exactly when σ reverses $\{\rho(i), \rho(j)\}$. A product of two nonzero rationals is negative exactly when exactly one factor is negative.

(b) Let P be the set of two-element subsets of $[n]$. For $p = \{i, j\} \in P$, let $u_p = 1$ if ρ reverses p and $u_p = 0$ otherwise, and let $v_p = 1$ if σ reverses $\rho(p) = \{\rho(i), \rho(j)\}$ and $v_p = 0$ otherwise. By (a), $\sigma\rho$ reverses p exactly when $u_p + v_p = 1$, and the integer $u_p + v_p - 2u_pv_p$ is 1 in that case and 0 otherwise. So

$$\operatorname{inv}(\sigma\rho) = \sum_{p\in P}\left(u_p + v_p - 2u_pv_p\right) = \sum_{p\in P} u_p + \sum_{p\in P} v_p - 2\sum_{p\in P} u_pv_p.$$

The first sum is $\operatorname{inv}(\rho)$. Since ρ is a bijection, $p \mapsto \rho(p)$ is a bijection of P (with inverse $p \mapsto \rho^{-1}(p)$), so as p runs over P, $\rho(p)$ runs over P once, and the second sum is the number of elements of P reversed by σ, which is $\operatorname{inv}(\sigma)$. Hence $\operatorname{inv}(\sigma\rho) - \operatorname{inv}(\sigma) - \operatorname{inv}(\rho)$ is even.

(c) By (b), $(-1)^{\operatorname{inv}(\sigma\rho)} = (-1)^{\operatorname{inv}(\sigma)}(-1)^{\operatorname{inv}(\rho)}$ for all σ, ρ. By Exercise 5.2.34, $\operatorname{inv}((k\ l)) = 2(l-k)-1$ is odd, so $(-1)^{\operatorname{inv}(\tau)} = -1$ for every transposition τ. If $\sigma = \tau_1\cdots\tau_r$, then by induction on r, $(-1)^{\operatorname{inv}(\sigma)} = (-1)^{\operatorname{inv}(\tau_1)}\cdots(-1)^{\operatorname{inv}(\tau_r)} = (-1)^r$. So two factorizations of σ of lengths r and s give $(-1)^r = (-1)^{\operatorname{inv}(\sigma)} = (-1)^s$, and r and s have the same parity.

Compared with Approach 2 in the text, this proof establishes multiplicativity first and gets the parity theorem from it, while the text proves the parity theorem first and gets multiplicativity for free from the definition.

Solution to Exercise 5.2.50

> *Uses.* Proposition 5.2.16 (Cycle decomposition), Lemma 5.2.17 (Cycles as products of transpositions), Definition 5.2.23 (Sign of a permutation)
> *Hint 1.* $\sigma\tau$ agrees with σ except at a and b: $(\sigma\tau)(a) = \sigma(b)$ and $(\sigma\tau)(b) = \sigma(a)$. So only the orbits containing a or b can change.
> *Hint 2.* If a and b are in one orbit, list it as $a, x_1, \ldots, x_p, b, y_1, \ldots, y_q$ in the order σ visits it, and follow $\sigma\tau$ from a and from b. If they are in different orbits, follow $\sigma\tau$ from a.
> *Hint 3.* For (c), (a) shows that each transposition changes c by 1, and $c(\mathrm{id}) = n$.
> *Sketch.* (a) Same orbit $a, x_1, \ldots, x_p, b, y_1, \ldots, y_q$: under $\sigma\tau$, $a \mapsto y_1 \mapsto \cdots \mapsto y_q \mapsto a$ and $b \mapsto x_1 \mapsto \cdots \mapsto x_p \mapsto b$, so the orbit splits in two. Different orbits $a, x_1, \ldots, x_p$ and $b, y_1, \ldots, y_q$: $a \mapsto y_1 \mapsto \cdots \mapsto y_q \mapsto b \mapsto x_1 \mapsto \cdots \mapsto x_p \mapsto a$, so they merge. (b) Start from $c(\mathrm{id}) = n$; each factor changes c by ± 1. (c) r factors give $c(\sigma) \geq n - r$, and Lemma 5.2.17 gives $n - c(\sigma)$ factors.

(a) Since τ fixes every point other than a and b, $(\sigma\tau)(x) = \sigma(x)$ for $x \notin \{a, b\}$, while $(\sigma\tau)(a) = \sigma(b)$ and $(\sigma\tau)(b) = \sigma(a)$. An orbit of σ containing neither a nor b is mapped into itself by σ, hence by $\sigma\tau$, on which they agree; so it is also an orbit of $\sigma\tau$. It remains to see what happens to the orbits containing a or b.

Same orbit. By Proposition 5.2.16, the orbit of a is $a, \sigma(a), \sigma^2(a), \ldots$ until it returns to a, and b occurs in it. Write it as $a, x_1, \ldots, x_p, b, y_1, \ldots, y_q$ in this order, with $p, q \geq 0$, so σ sends $a \mapsto x_1 \mapsto \cdots \mapsto x_p \mapsto b \mapsto y_1 \mapsto \cdots \mapsto y_q \mapsto a$ (where $a \mapsto x_1$ is read as $a \mapsto b$ if $p = 0$, and so on). Then $\sigma\tau$ sends $a \mapsto \sigma(b) = y_1 \mapsto \cdots \mapsto y_q \mapsto \sigma\left(y_q\right) = a$ and $b \mapsto \sigma(a) = x_1 \mapsto \cdots \mapsto x_p \mapsto \sigma\left(x_p\right) = b$. (If $q = 0$, then $\sigma(b) = a$ and a is fixed by $\sigma\tau$; similarly if $p = 0$.) So the orbit splits into the two orbits $\left\{a, y_1, \ldots, y_q\right\}$ and $\left\{b, x_1, \ldots, x_p\right\}$ of $\sigma\tau$, and $c(\sigma\tau) = c(\sigma) + 1$.

Different orbits. Write them as $a, x_1, \ldots, x_p$ and $b, y_1, \ldots, y_q$ in the order σ visits them. Then $\sigma\tau$ sends $a \mapsto \sigma(b) = y_1 \mapsto \cdots \mapsto y_q \mapsto \sigma\left(y_q\right) = b \mapsto \sigma(a) = x_1 \mapsto \cdots \mapsto x_p \mapsto \sigma\left(x_p\right) = a$, with the same conventions when p or q is 0. So the two orbits merge into one orbit of $\sigma\tau$, and $c(\sigma\tau) = c(\sigma) - 1$.

(b) We show by induction on r that $c(\tau_1\cdots\tau_r) - (n - r)$ is even. For $r = 0$, $c(\mathrm{id}) = n$. For the step, (a) gives $c(\tau_1\cdots\tau_r) = c(\tau_1\cdots\tau_{r-1}) \pm 1$, which changes $c - (n - r)$ by $\pm 1 + 1$, an even number. So if $\sigma = \tau_1\cdots\tau_r = \tau'_1\cdots\tau'_s$, then r and s both have the parity of $n - c(\sigma)$, which proves Theorem 5.2.18, and $\operatorname{sgn}(\sigma) = (-1)^r = (-1)^{n-c(\sigma)}$.

(c) If $\sigma = \tau_1\cdots\tau_r$, then by (a) and induction, $c\left(\tau_1\cdots\tau_j\right) \geq n - j$ for each j, so $c(\sigma) \geq n - r$, that is, $r \geq n - c(\sigma)$. On the other hand, Proposition 5.2.16 and Lemma 5.2.17 write σ as a product of exactly $n - c(\sigma)$ transpositions. So the least number is $n - c(\sigma)$.

Solution to Exercise 5.2.51

> *Uses.* Corollary 5.2.24 (Formulas for the sign), Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.17 (No zero divisors), Proposition 5.2.26 (Sign of a cycle)
> *Hint 1.* For (a), multiply all the nonzero elements together, before and after applying μ_a.
> *Hint 2.* For (b), use Corollary 5.2.24 (b) over the field $\mathbb{F}_p$, with $\mathbf{x} = (1, 2, \ldots, p-1) \in \mathbb{F}_p^{p-1}$. What is $\mathbf{x}^{\mu_a}$?
> *Hint 3.* Each factor of $\Delta(\mathbf{x}^{\mu_a})$ is $aj - ai = a(j - i)$, and there are $(p-1)(p-2)/2$ factors. Then use (a) and that $p - 2$ is odd.

Sketch. (a) μ_a permutes the factors of $P = 1 \cdot 2 \cdots (p-1) \neq 0$, so $a^{p-1}P = P$. (b) With $\mathbf{x} = (1, \ldots, p-1)$, $\mathbf{x}^{\mu_a} = (a, 2a, \ldots, (p-1)a)$, so $\Delta(\mathbf{x}^{\mu_a}) = a^N \Delta(\mathbf{x})$ with $N = (p-1)(p-2)/2$, and $\Delta(\mathbf{x}) \neq 0$. By Corollary 5.2.24, $\operatorname{sgn}(\mu_a) = a^N$ in $\mathbb{F}_p$. With $h = (p-1)/2$, $a^h = \pm 1$ by (a), and $a^N = \left(a^h\right)^{p-2} = a^h$ since $p-2$ is odd. (c) $\mu_2 = (1\ 2\ 4)(3\ 6\ 5)$ is even and $2^3 = 1$; $\mu_3 = (1\ 3\ 2\ 6\ 4\ 5)$ is odd and $3^3 = -1$.

Since p is prime, $\mathbb{F}_p$ is a field (Proposition 1.1.24). If $ak = ak'$ then $k = k'$ (multiply by a^{-1}), and $ak \neq 0$ for $k \neq 0$, so μ_a is indeed a permutation of the $p-1$ nonzero elements.

(a) Let $P = 1 \cdot 2 \cdots (p-1) \in \mathbb{F}_p$, which is nonzero by Proposition 1.1.17. Since μ_a is a bijection of the nonzero elements, the products of $1, \ldots, p-1$ and of $a \cdot 1, a \cdot 2, \ldots, a(p-1)$ have the same factors in a different order, so they are equal: $a^{p-1}P = P$. Cancelling P gives $a^{p-1} = 1$.

(b) Let $\mathbf{x} = (1, 2, \ldots, p-1) \in \mathbb{F}_p^{p-1}$, so $x_k = k$. Its entries are distinct, so $\Delta(\mathbf{x}) \neq 0$. The k-th entry of $\mathbf{x}^{\mu_a}$ is $x_{\mu_a(k)} = ak$. So each factor of $\Delta(\mathbf{x}^{\mu_a})$ is $aj - ai = a(j-i)$, and since there are $N = \binom{p-1}{2} = (p-1)(p-2)/2$ factors,

$$\Delta(\mathbf{x}^{\mu_a}) = a^N \Delta(\mathbf{x}).$$

By Corollary 5.2.24 (b) over $F = \mathbb{F}_p$, the left side is $\operatorname{sgn}(\mu_a)\Delta(\mathbf{x})$. Cancelling $\Delta(\mathbf{x}) \neq 0$, $\operatorname{sgn}(\mu_a) = a^N$ in $\mathbb{F}_p$. Let $h = (p-1)/2$. By (a), $\left(a^h\right)^2 = 1$, so $a^h = \pm 1$ by Proposition 1.1.17. Since $N = h(p-2)$ and $p-2$ is odd, $a^N = \left(a^h\right)^{p-2} = a^h$. So $\operatorname{sgn}(\mu_a) = a^h$ in $\mathbb{F}_p$. Because p is odd, $1 \neq -1$ in $\mathbb{F}_p$, so this equation in $\mathbb{F}_p$ determines the integer $\operatorname{sgn}(\mu_a)$.

(c) For $p = 7$, μ_2 sends $1 \mapsto 2 \mapsto 4 \mapsto 1$ and $3 \mapsto 6 \mapsto 5 \mapsto 3$, so $\mu_2 = (1\ 2\ 4)(3\ 6\ 5)$, which is even by Proposition 5.2.26; and $2^3 = 8 = 1$ in $\mathbb{F}_7$. Also μ_3 sends $1 \mapsto 3 \mapsto 2 \mapsto 6 \mapsto 4 \mapsto 5 \mapsto 1$, a 6-cycle, which is odd; and $3^3 = 27 = -1$ in $\mathbb{F}_7$.

Section 5.3

Solution to Exercise 5.3.7

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Exercise 5.1.34, Definition 1.5.1 (Linear independence), Definition 4.1.5 (Dimension), Proposition 1.1.15 (Multiplying by zero)
Hint 1. The last sentence of Theorem 5.3.4 writes every $D \in \operatorname{Alt}_n$ as a scalar multiple of det.
Hint 2. For the dimension, the list consisting of det alone is independent because $\det \neq 0$. For the last claim, write $D = c \det$ and evaluate at A.
Sketch. Every $D \in \operatorname{Alt}_n$ equals $D(I_n) \det$, and $\det \in \operatorname{Alt}_n$, so $\operatorname{Alt}_n = \operatorname{span}(\det)$. Since $\det(I_n) = 1$, det is not the zero function, so it alone is a basis. If $D(A) \neq 0$, then $D = c \det$ with $c \det A \neq 0$, which forces $c \neq 0$ and $\det A \neq 0$.

By Theorem 5.3.4, every $D \in \operatorname{Alt}_n$ satisfies $D = D(I_n) \det \in \operatorname{span}(\det)$. Conversely, $\det \in \operatorname{Alt}_n$ and Alt_n is a subspace (Exercise 5.1.34), so $\operatorname{span}(\det) \subseteq \operatorname{Alt}_n$. Hence $\operatorname{Alt}_n = \operatorname{span}(\det)$. Since $\det(I_n) = 1$, det is not the zero function, so the one-element list det is linearly independent and is a basis of Alt_n. Thus $\dim \operatorname{Alt}_n = 1$.

Now let $D \in \operatorname{Alt}_n$ with $D(A) \neq 0$, and let $c = D(I_n)$, so $D = c \det$. If $c = 0$ or $\det A = 0$, then $D(A) = c \det A = 0$ by Proposition 1.1.15, which is false. So $c \neq 0$ and $\det A \neq 0$.

Solution to Exercise 5.3.8

Uses. Definition 5.3.6 (Determinant), Proposition 5.2.8 (Reindexing sums over S_n), Proposition 5.2.25 (Multiplicativity of the sign)
Hint 1. Fix σ and reorder the product $a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ by its row index. If $i = \sigma(j)$, what is j in terms of i?
Hint 2. You should get $\prod_j a_{\sigma(j)j} = \prod_i a_{i\sigma^{-1}(i)}$. Then use $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ and substitute σ^{-1} for σ in the sum.
Sketch. As j runs over $[n]$, $i = \sigma(j)$ runs over $[n]$ once each, with $j = \sigma^{-1}(i)$, so $\prod_j a_{\sigma(j)j} = \prod_i a_{i\sigma^{-1}(i)}$. With $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, $\det A = \sum_\sigma g(\sigma^{-1})$ for $g(\rho) = \operatorname{sgn}(\rho) \prod_i a_{i\rho(i)}$, and reindexing by $\sigma \mapsto \sigma^{-1}$ gives $\sum_\sigma g(\sigma)$.

Fix $\sigma \in S_n$. As j runs over $[n]$, the row index $i = \sigma(j)$ runs over $[n]$, taking each value once, and $j = \sigma^{-1}(i)$. Since multiplication in F is commutative, we may reorder the factors by their row index:

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}.$$

Let $g(\rho) = \operatorname{sgn}(\rho) a_{1\rho(1)} \cdots a_{n\rho(n)}$ for $\rho \in S_n$. Since $\operatorname{sgn}(\sigma) = \operatorname{sgn}\left(\sigma^{-1}\right)$ by Proposition 5.2.25, the term of σ in the Leibniz formula is $g\left(\sigma^{-1}\right)$. By Proposition 5.2.8,

$$\det A = \sum_{\sigma \in S_n} g\left(\sigma^{-1}\right) = \sum_{\sigma \in S_n} g(\sigma),$$

which is the claim. The right side is the Leibniz formula for the transpose A^{T}, whose (j,i) entry is a_{ij}. So this exercise shows $\det A^{\mathsf{T}} = \det A$; Theorem 5.4.14 records this and draws the consequences.

Solution to Exercise 5.3.16

Uses. Proposition 5.3.15 (Triangular matrices)
Hint 1. Both matrices in (a) are triangular. Multiply the diagonal entries.
Hint 2. In (b), N and $cI_n + N$ are upper triangular. What are their diagonal entries?
Sketch. (a) $2 \cdot (-1) \cdot 3 \cdot 1/2 = -3$ and $i(1+i)(1-i) = 2i$. (b) N is upper triangular with zero diagonal, and $cI_n + N$ is upper triangular with every diagonal entry c.

(a) The first matrix is upper triangular, so by Proposition 5.3.15 its determinant is $2 \cdot (-1) \cdot 3 \cdot 1/2 = -3$. The second is lower triangular, so its determinant is $i(1+i)(1-i) = i\left(1 - i^2\right) = 2i$.

(b) A strictly upper triangular N is upper triangular with all diagonal entries 0, so $\det N = 0 \cdot \cdots \cdot 0 = 0$ by Proposition 5.3.15 (here $n \geq 1$). The matrix $cI_n + N$ has (i,j) entry equal to that of N, which is 0, for $i > j$, and $c + 0 = c$ on the diagonal, so it is upper triangular with all diagonal entries c, and $\det(cI_n + N) = c^n$.

Solution to Exercise 5.3.23

Uses. Corollary 5.3.9 (Rule of Sarrus), Proposition 5.3.20 (Reduction modulo p), Proposition 5.1.17 (Dependent columns)
Techniques. Reduce modulo p (Technique 5.3.21)
Hint 1. The down-right diagonals give $1 \cdot 1 \cdot 1$, $0 \cdot 1 \cdot 1$, and $3 \cdot 3 \cdot 4$.
Hint 2. For (b), reduce the answer to (a) instead of recomputing (Proposition 5.3.20).
Hint 3. For (c), solve the system row by row over $\mathbb{F}_5$, starting with $x_1 + 3x_3 = 0$.
Sketch. (a) $(1 + 0 + 36) - (3 + 4 + 0) = 30$. (b) $30 = 2 \cdot 3 \cdot 5$, so $\det \overline{A}$ is 0 modulo 2, 3, 5, and 2 modulo 7. (c) $\mathbf{x} = (2, 3, 1)$; a dependence among the columns forces $\det \overline{A} = 0$, matching $5 \mid 30$.

(a) The diagonals running down to the right have products $1 \cdot 1 \cdot 1 = 1$, $0 \cdot 1 \cdot 1 = 0$, and $3 \cdot 3 \cdot 4 = 36$; the diagonals running up to the right have products $1 \cdot 1 \cdot 3 = 3$, $4 \cdot 1 \cdot 1 = 4$, and $1 \cdot 3 \cdot 0 = 0$. By Corollary 5.3.9, $\det A = 37 - 7 = 30$.

(b) By Proposition 5.3.20, $\det \overline{A} = 30 \bmod p$. Since $30 = 2 \cdot 3 \cdot 5$, we get $\det \overline{A} = 0$ for $p = 2, 3, 5$, and $\det \overline{A} = 30 \bmod 7 = 2$ for $p = 7$.

(c) Over $\mathbb{F}_5$, the rows of $\overline{A}\mathbf{x} = \mathbf{0}$ read $x_1 + 3x_3 = 0$, $3x_1 + x_2 + x_3 = 0$, and $x_1 + 4x_2 + x_3 = 0$. The first gives $x_1 = -3x_3 = 2x_3$, and then the second gives $x_2 = -7x_3 = 3x_3$. With $x_3 = 1$, $\mathbf{x} = (2, 3, 1)$, and the third row checks: $2 + 12 + 1 = 15 = 0$. So $2\mathbf{a}_1 + 3\mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$ is a nontrivial dependence among the columns of $\overline{A}$, which forces $\det \overline{A} = 0$ by Proposition 5.1.17. This agrees with (b), since $5 \mid 30$.

Solution to Exercise 5.3.24

Uses. Definition 5.3.6 (Determinant), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.25 (Multiplicativity of the sign)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. The term of σ is nonzero exactly when $\sigma(j)$ is one of the two rows where column j is nonzero, for every j.
Hint 2. Columns 1 and 3 are nonzero only in rows 2 and 4, and columns 2 and 4 only in rows 1 and 3. So σ must send $\{1,3\}$ onto $\{2,4\}$ and $\{2,4\}$ onto $\{1,3\}$.
Sketch. The surviving σ, written as bottom rows, are 2, 1, 4, 3 (even, product 120), 2, 3, 4, 1 (a 4-cycle, product 15), 4, 1, 2, 3 (a 4-cycle, product 8), and 4, 3, 2, 1 (even, product 1). So $\det B = 120 - 15 - 8 + 1 = 98$.

The term of σ is nonzero exactly when $b_{\sigma(j)j} \neq 0$ for every j. Column 1 is nonzero in rows 2 and 4, column 2 in rows 1 and 3, column 3 in rows 2 and 4, and column 4 in rows 1 and 3. Since σ is injective, $\sigma(1)$ and $\sigma(3)$ are the two rows 2 and 4 in some order, and $\sigma(2)$ and $\sigma(4)$ are 1 and 3 in some order. So exactly $2 \cdot 2 = 4$ permutations survive. Writing σ by its bottom row $\sigma(1), \sigma(2), \sigma(3), \sigma(4)$:

- 2, 1, 4, 3, that is, (1 2)(3 4), which is even, with product $b_{21}b_{12}b_{43}b_{34} = 3 \cdot 2 \cdot 5 \cdot 4 = 120$;

- $2, 3, 4, 1$, that is, $(1\ 2\ 3\ 4)$, which is odd by Proposition 5.2.26, with product $b_{21}b_{32}b_{43}b_{14} = 3 \cdot 1 \cdot 5 \cdot 1 = 15$;
- $4, 1, 2, 3$, that is, $(1\ 4\ 3\ 2)$, which is odd, with product $b_{41}b_{12}b_{23}b_{34} = 1 \cdot 2 \cdot 1 \cdot 4 = 8$;
- $4, 3, 2, 1$, that is, $(1\ 4)(2\ 3)$, which is even, with product $b_{41}b_{32}b_{23}b_{14} = 1$.

(The products of two disjoint transpositions are even by Proposition 5.2.25.) Hence $\det B = 120 - 15 - 8 + 1 = 98$.

Solution to Exercise 5.3.25

Uses. Definition 5.3.6 (Determinant), Proposition 5.2.26 (Sign of a cycle), Theorem 5.2.31 (Alternating functions and permutations), Remark 5.3.11 (Sarrus does not generalize)
Hint 1. How many terms does the Leibniz formula have for $n = 4$? How many products does the claim produce?
Hint 2. The diagonal from a_{12} is $a_{12}a_{23}a_{34}a_{41}$. Which permutation is this the term of, and what is its sign?
Hint 3. Try a permutation matrix whose nonzero entries are not all on one diagonal, such as I_4 with its first two columns interchanged.
Sketch. The argument shows that each diagonal is a term, not that each term is a diagonal: 8 diagonals cover only 8 of the 24 terms. And the signs are not those of the permutations: the downward diagonals from a_{12} and a_{14} are terms of 4-cycles, which are odd, and the upward diagonals from a_{41} and a_{43} are terms of even products of two transpositions. For I_4 with columns 1 and 2 interchanged, every diagonal contains a 0, so the claim gives 0, but $\det = -1$.

There are two errors. First, the argument shows that each diagonal is a term of the Leibniz formula, but not that each term is a diagonal. There are $4! = 24$ terms and only 8 diagonals, so 16 terms are missing. For example, the term $a_{21}a_{12}a_{33}a_{44}$ of $(1\ 2)$ lies on no diagonal. Second, the signs are wrong. The diagonal from a_{12} is $a_{12}a_{23}a_{34}a_{41}$, the term of the permutation σ with $\sigma(2) = 1$, $\sigma(3) = 2$, $\sigma(4) = 3$, $\sigma(1) = 4$, which is the 4-cycle $(1\ 4\ 3\ 2)$. By Proposition 5.2.26 it is odd, so this product enters $\det A$ with sign -1, while the claim adds it. The same happens for the downward diagonal from a_{14}, the term of the 4-cycle $(1\ 2\ 3\ 4)$, and in the other direction for the upward diagonals from a_{41} and a_{43}, the terms of the even permutations $(1\ 4)(2\ 3)$ and $(1\ 2)(3\ 4)$, which the claim subtracts. Only the four diagonals from a_{11}, a_{13}, a_{42}, and a_{44} get the right sign. For $n = 3$ the downward diagonals are id and two 3-cycles, which are even, and the upward ones are transpositions; that is why the rule of Sarrus works (Remark 5.3.11).

Each error gives a counterexample. Let A be I_4 with columns 1 and 2 interchanged, so its nonzero entries are $a_{21} = a_{12} = a_{33} = a_{44} = 1$. Then $\det A = -\det I_4 = -1$ by Theorem 5.2.31. But every diagonal contains a zero entry: the downward diagonals start at $a_{11} = 0$, a_{12} (followed by $a_{23} = 0$), $a_{13} = 0$, and $a_{14} = 0$, and the upward ones start at $a_{41} = a_{42} = a_{43} = 0$ and at a_{44} (followed by $a_{31} = 0$). So the claim gives 0. If $1 + 1 \neq 0$ in F, the sign error gives a second counterexample: let A be the permutation matrix whose nonzero entries are exactly $a_{12}, a_{23}, a_{34}, a_{41}$. Then $\det A = \operatorname{sgn}((1\ 4\ 3\ 2)) = -1$, while every diagonal except the one from a_{12} contains a 0, so the claim gives $+1 \neq -1$.

Solution to Exercise 5.3.26

Uses. Definition 5.3.6 (Determinant), Lemma 5.3.2 (Expanding every column), Proposition 5.3.15 (Triangular matrices)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Is $(x_1 + y_1)(x_2 + y_2) = x_1x_2 + y_1y_2$? Test the claim with two diagonal matrices that add up to I_n.
Hint 2. For (b), expand the product $\prod_j \left(a_{\sigma(j)j} + b_{\sigma(j)j}\right)$ by choosing one of the two summands in each factor, and record the set S of j where you chose a.
Sketch. The second equality replaces a product of sums by the sum of two products; a product of n sums expands into 2^n terms. With $A = \operatorname{diag}(1, 0, 1, \ldots, 1)$ and $B = \operatorname{diag}(0, 1, 0, \ldots, 0)$, $\det A = \det B = 0$ but $\det(A + B) = \det I_n = 1$. For (b), $\prod_j \left(a_{\sigma(j)j} + b_{\sigma(j)j}\right) = \sum_S \prod_{j \in S} a_{\sigma(j)j} \prod_{j \notin S} b_{\sigma(j)j}$, and exchanging the sums over σ and S gives $\sum_S \det C_S$.

(a) The second equality is false. A product of n sums is not the sum of the two "pure" products: by distributivity, $\prod_{j=1}^{n} \left(x_j + y_j\right)$ is a sum of 2^n terms, one for each way of choosing x_j or y_j in each factor, and the argument keeps only the two terms that choose x every time or y every time. For every $n \geq 2$ and every field the claim is false: let $A = \operatorname{diag}(1, 0, 1, \ldots, 1)$ and $B = \operatorname{diag}(0, 1, 0, \ldots, 0)$. Each has a zero diagonal entry, so $\det A = \det B = 0$ by Proposition 5.3.15, while $A + B = I_n$ has determinant 1. (For $n = 1$ the claim is true, since then each product has a single factor.)

(b) For each σ, distributivity gives

$$\prod_{j=1}^{n}\bigl(a_{\sigma(j)j}+b_{\sigma(j)j}\bigr)=\sum_{S\subseteq[n]}\prod_{j\in S}a_{\sigma(j)j}\prod_{j\notin S}b_{\sigma(j)j},$$

where S records the factors in which $a_{\sigma(j)j}$ was chosen. The (i,j) entry of C_S is a_{ij} if $j\in S$ and b_{ij} if $j\notin S$, so the term for S is $\prod_j (C_S)_{\sigma(j)j}$. Multiplying by $\operatorname{sgn}(\sigma)$, summing over σ, and exchanging the two finite sums,

$$\det(A+B)=\sum_{S\subseteq[n]}\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{j=1}^{n}(C_S)_{\sigma(j)j}=\sum_{S\subseteq[n]}\det C_S.$$

The same identity follows from multilinearity alone, by expanding each column $\mathbf{a}_j+\mathbf{b}_j$ as in Lemma 5.3.2. So it holds for every multilinear D in place of det; the terms $S=[n]$ and $S=\emptyset$ are $\det A$ and $\det B$, and the other 2^n-2 terms are what the argument in (a) dropped.

Solution to Exercise 5.3.27

Uses. Definition 5.3.6 (Determinant), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.27 (Half of all permutations are even), Definition 5.1.1 (Determinant function), Theorem 5.3.4 (Existence and uniqueness of the determinant)
Techniques. Cancel terms in pairs (Technique 5.3.5)
Hint 1. (a) In the term of σ, which entry comes from row r? Imitate the multilinearity part of the proof of Theorem 5.3.4.
Hint 2. (b) If rows r and s are equal, compare the terms of σ and $\tau\sigma$, where $\tau=(r\ s)$; note that the transposition now acts on the left.
Hint 3. (c) Row i of ΛA is d_i times row i of A, and column j of $A\Lambda$ is $d_j\mathbf{a}_j$.
Sketch. (a) The term of σ contains exactly one entry of row r, namely $a_{r\sigma^{-1}(r)}$, so det is a fixed linear combination of the entries of row r. (b) If rows r,s are equal and $\tau=(r\ s)$, then $a_{\tau(\sigma(j))j}=a_{\sigma(j)j}$ for all j and $\operatorname{sgn}(\tau\sigma)=-\operatorname{sgn}(\sigma)$; pair each even σ with $\tau\sigma$. (c) Pull d_i out of each row by (a), and out of each column by multilinearity.

(a) Fix $r\in[n]$ and all rows except row r, and let $\mathbf{x}=(x_1,\dots,x_n)$ be row r, so $a_{rj}=x_j$. In the term of σ, the entry from row r is $a_{\sigma(j)j}$ for the unique j with $\sigma(j)=r$, that is, $j=\sigma^{-1}(r)$. So

$$\det A=\sum_{\sigma\in S_n}b_\sigma x_{\sigma^{-1}(r)},\quad\text{where } b_\sigma=\operatorname{sgn}(\sigma)\prod_{j\neq\sigma^{-1}(r)}a_{\sigma(j)j}$$

does not depend on $\mathbf{x}$. As in the proof of Theorem 5.3.4, a function of this form is linear in $\mathbf{x}$.

(b) Suppose rows $r\neq s$ are equal, so $a_{rj}=a_{sj}$ for every j, and let $\tau=(r\ s)$. For every j, $\tau(\sigma(j))$ equals $\sigma(j)$ unless $\sigma(j)\in\{r,s\}$, in which case it is the other element of $\{r,s\}$; since rows r and s are equal, $a_{\tau(\sigma(j))j}=a_{\sigma(j)j}$ in all cases. Also $\operatorname{sgn}(\tau\sigma)=-\operatorname{sgn}(\sigma)$ by Proposition 5.2.25. So the term of $\tau\sigma$ is the negative of the term of σ. The map $\sigma\mapsto\tau\sigma$ maps even permutations to odd ones and odd ones to even ones, and it is its own inverse, so it maps the set S_n^+ of even permutations bijectively onto the set of odd ones (as in the proof of Proposition 5.2.27). Hence

$$\det A=\sum_{\sigma\in S_n^+}\bigl(t_\sigma(A)+t_{\tau\sigma}(A)\bigr)=0,$$

where $t_\sigma(A)$ denotes the term of σ.

(c) Row i of ΛA is d_i times row i of A. Pulling the factor d_i out of row i for each i, by (a), gives $\det(\Lambda A)=d_1\cdots d_n\det A$. Column j of $A\Lambda$ is $d_j\mathbf{a}_j$, and pulling the factors out of the columns by multilinearity gives $\det(A\Lambda)=d_1\cdots d_n\det A$.

Here is what (a) and (b) mean. Let $D(A)=\det A^{\mathsf T}$. The columns of A are the rows of $A^{\mathsf T}$, so by (a) and (b), D is multilinear and alternating, and $D(I_n)=1$. By Theorem 5.3.4, $D=\det$, that is, $\det A^{\mathsf T}=\det A$; Theorem 5.4.14 records this.

Solution to Exercise 5.3.28

Uses. Definition 5.3.6 (Determinant), Definition 1.1.29 (Complex numbers), Exercise 5.1.5
Techniques. det is a polynomial in the entries (Technique 5.3.18)

Hint 1. Check that $\overline{z+w} = \overline{z} + \overline{w}$ and $\overline{zw} = \overline{z} \cdot \overline{w}$, and apply this to the Leibniz formula.
Hint 2. For (b), compute $\det \overline{A}$ in two ways: by (a), and as $\det(-A)$.
Sketch. Conjugation respects sums and products, and fixes ± 1, so conjugating the Leibniz formula for A term by term gives the formula for $\overline{A}$. A real A has $\overline{A} = A$, so $\det A = \overline{\det A}$ is real. In (b), $\overline{\det A} = \det(-A) = (-1)^n \det A$.

(a) For $z = x + yi$ and $w = u + vi$, we have $\overline{z+w} = (x+u) - (y+v)i = \overline{z} + \overline{w}$, and

$$\overline{z} \cdot \overline{w} = (x - yi)(u - vi) = (xu - yv) - (xv + yu)i = \overline{zw},$$

since $zw = (xu - yv) + (xv + yu)i$. By induction, conjugation respects finite sums and products, and it fixes the real numbers ± 1. Conjugating the Leibniz formula term by term,

$$\overline{\det A} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \overline{a_{\sigma(1)1}} \cdots \overline{a_{\sigma(n)n}} = \det \overline{A}.$$

If every entry of A is real, then $\overline{A} = A$, so $\overline{\det A} = \det A$, which says that $\det A$ is real.

(b) By (a) and Exercise 5.1.5, $\overline{\det A} = \det \overline{A} = \det(-A) = (-1)^n \det A$. If n is even, $\overline{\det A} = \det A$, so $\det A$ is real. If n is odd, $\overline{\det A} = -\det A$, so the real part of $\det A$ is 0: $\det A$ is purely imaginary.

Solution to Exercise 5.3.29

Uses. Definition 5.3.6 (Determinant), Theorem 5.2.31 (Alternating functions and permutations), Definition 5.2.29 (Permutation matrix)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. Fix σ. The rows $\sigma(j)$ for $j \in C$ are $|C|$ different rows. Can they all avoid R?
Hint 2. For (b), if $a_{kk} = 0$, take $R = \{k, \ldots, n\}$ and $C = \{1, \ldots, k\}$.
Hint 3. For (c), look for a permutation matrix whose first $s = n - r$ columns have their 1s below row r.
Sketch. (a) The $|C|$ distinct rows $\sigma(j)$, $j \in C$, cannot all lie in $[n] \setminus R$, which has fewer than $|C|$ elements; so some $a_{\sigma(j)j}$ with $j \in C$ lies in the zero block, and every term vanishes. (b) A zero at (k,k) in an upper triangular matrix gives a zero block on rows $k, \ldots, n$ and columns $1, \ldots, k$, of total size $n+1$. (c) The permutation matrix with columns $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n, \mathbf{e}_1, \ldots, \mathbf{e}_r$ has a zero $r \times (n-r)$ block in its upper left corner and determinant ± 1.

(a) Let $\sigma \in S_n$. The rows $\sigma(j)$ for $j \in C$ are $|C|$ distinct elements of $[n]$. The set $[n] \setminus R$ has $n - |R| < |C|$ elements, so not all of these rows lie in it: there is $j \in C$ with $\sigma(j) \in R$. Then $a_{\sigma(j)j} = 0$, and the term of σ vanishes. Since every term vanishes, $\det A = 0$.

(b) Let A be upper triangular with $a_{kk} = 0$, and let $R = \{k, k+1, \ldots, n\}$ and $C = \{1, 2, \ldots, k\}$. For $i \in R$ and $j \in C$ we have $i \geq k \geq j$, so either $i > j$, and $a_{ij} = 0$ because A is upper triangular, or $i = j = k$, and $a_{ij} = a_{kk} = 0$. Since $|R| + |C| = (n - k + 1) + k = n + 1 > n$, part (a) gives $\det A = 0$, in agreement with Proposition 5.3.15.

(c) Let $1 \leq r \leq n-1$ and $s = n - r$. Let A be the matrix with columns $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n, \mathbf{e}_1, \ldots, \mathbf{e}_r$, that is, the permutation matrix E_σ with $\sigma(j) = r + j$ for $j \leq s$ and $\sigma(j) = j - s$ for $j > s$ (Definition 5.2.29). For $i \in R = \{1, \ldots, r\}$ and $j \in C = \{1, \ldots, s\}$, the column $\mathbf{e}_{r+j}$ has a 0 in row $i \leq r$, so $a_{ij} = 0$, and $|R| + |C| = n$. But $\det A = \operatorname{sgn}(\sigma) = \pm 1 \neq 0$ by Theorem 5.2.31.

Solution to Exercise 5.3.30

Uses. Definition 5.3.6 (Determinant), Proposition 5.3.15 (Triangular matrices), Proposition 5.3.17 (Block-diagonal matrices), Proposition 5.1.17 (Dependent columns), Definition 2.4.21 (Transpose)
Hint 1. (a) Try a matrix with two equal columns.
Hint 2. (b) If $\det A \neq 0$, some term $t_\sigma(A)$ is nonzero. Where are its n nonzero entries?
Hint 3. (c) What kind of matrix is A^{T}, and what are its diagonal entries? (d) Try $n = 2$ and $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.
Sketch. (a) False: the all-ones matrix with $n \geq 2$. (b) True: a nonzero term uses n nonzero entries, one in each row and each column, and these must be all of them. (c) True: A^{T} is lower triangular with the same diagonal. (d) False for $n \geq 2$: $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has determinant -1, but $\det \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 0$.

(a) False for every $n \geq 2$. The matrix with every entry 1 has two equal columns, so its determinant is 0. (For $n = 1$ the statement is true.)

(b) True. Since $\det A \neq 0$, some term of the Leibniz formula is nonzero, so the n entries $a_{\sigma(1)1}, \ldots, a_{\sigma(n)n}$ are nonzero for some σ. They lie in n different positions, and A has only n nonzero entries, so these are all of them. They lie one in each column, and, since σ is a bijection, one in each row.

(c) True. The (i,j) entry of A^{T} is a_{ji}, which is 0 when $j > i$, that is, when $i < j$. So A^{T} is lower triangular, with the same diagonal entries a_{jj} as A. By Proposition 5.3.15, both determinants equal $a_{11}\cdots a_{nn}$.

(d) False for every $n \geq 2$. For $n = 2$, take $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ and $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $\det A = 0$ (a zero column), while $\det(A+N) = \det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1$. For $n > 2$, take $A' = \begin{pmatrix} A & 0 \\ 0 & I_{n-2} \end{pmatrix}$ and $N' = \begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix}$, which is still strictly upper triangular. By Proposition 5.3.17, $\det A' = \det A = 0$ and $\det(A'+N') = \det(A+N) = -1$. (For $n = 1$, $N = 0$ and the statement is true.)

Solution to Exercise 5.3.31

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Proposition 5.2.8 (Reindexing sums over S_n), Proposition 5.2.30 (Algebra of permutation matrices), Definition 1.1.28 (Characteristic), Proposition 5.3.20 (Reduction modulo p)

Hint 1. For (a), the proof of Theorem 5.3.4 for normalization and multilinearity did not use the signs. For the last claim, the columns of AE_π are $\mathbf{a}_{\pi(1)}, \ldots, \mathbf{a}_{\pi(n)}$; reindex the sum.

Hint 2. For (b), if $1 + 1 = 0$, compare per with det term by term. If $1 + 1 \neq 0$, evaluate per at the matrix with columns $\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n$.

Hint 3. For (c), reduce both $\det A$ and $\operatorname{per} A$ modulo 2.

Sketch. (a) Part 2 of the proof of Theorem 5.3.4 works without signs. The term of σ in $\operatorname{per}(AE_\pi)$ is the term of $\sigma\pi^{-1}$ in $\operatorname{per} A$, and $\sigma \mapsto \sigma\pi^{-1}$ is a bijection. (b) If $1 + 1 = 0$, then $\operatorname{sgn}(\sigma) = 1$ in F, so per = det. Otherwise the matrix with two equal columns $\mathbf{e}_1 + \mathbf{e}_2$ has permanent $1 + 1 \neq 0$. (c) Each term is 0 or 1; modulo 2 the signs disappear, so $\det A \equiv \operatorname{per} A$.

(a) per is L with every sign replaced by 1, and the proof that L is normalized and multilinear (Part 2 of the proof of Theorem 5.3.4) never used the signs. So per is normalized and multilinear. By Proposition 5.2.30, the (i,j) entry of AE_π is $a_{i\pi(j)}$, so

$$\operatorname{per}(AE_\pi) = \sum_{\sigma \in S_n} \prod_{j=1}^{n} a_{\sigma(j)\pi(j)} = \sum_{\sigma \in S_n} \prod_{j=1}^{n} a_{\sigma(\pi^{-1}(j))j},$$

after reordering the product by $j' = \pi(j)$. The last sum is $\sum_\sigma g\left(\sigma\pi^{-1}\right)$ with $g(\rho) = \prod_j a_{\rho(j)j}$, which equals $\sum_\sigma g(\sigma) = \operatorname{per} A$ by Proposition 5.2.8.

(b) If $1 + 1 = 0$, then $-1 = 1$ in F, so $\operatorname{sgn}(\sigma) = 1$ in F for every σ, and $\operatorname{per} A = \det A$ for every A. In particular per is a determinant function. Conversely, suppose $1 + 1 \neq 0$, and let A have columns $\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n$. A term of $\operatorname{per} A$ is nonzero exactly when $\sigma(1), \sigma(2) \in \{1, 2\}$ and $\sigma(j) = j$ for $j \geq 3$, which leaves $\sigma = \mathrm{id}$ and $\sigma = (1\ 2)$, each with product 1. So $\operatorname{per} A = 1 + 1 \neq 0$, although A has two equal columns, and per is not alternating. (By (a) and Theorem 5.3.4, when $1 + 1 = 0$ the determinant function per must equal det, as we saw directly.)

(c) Each product $a_{\sigma(1)1}\cdots a_{\sigma(n)n}$ is 1 if all its factors are 1 and 0 otherwise, so $\operatorname{per} A$ counts the permutations with $a_{\sigma(j)j} = 1$ for every j. For $n = 1$, $\det A = \operatorname{per} A = a_{11}$. For $n \geq 2$, reducing modulo 2 with Proposition 5.3.20 gives $(\det A) \bmod 2 = \det \overline{A}$, and $\det \overline{A} = \operatorname{per} \overline{A}$ by (b), since $1 + 1 = 0$ in $\mathbb{F}_2$. Reducing the sum per A term by term in the same way gives $\operatorname{per} \overline{A} = (\operatorname{per} A) \bmod 2$. So $\det A$ and $\operatorname{per} A$ have the same remainder modulo 2.

Solution to Exercise 5.3.32

Uses. Lemma 5.1.8 (Column replacement), Definition 5.1.1 (Determinant function), Proposition 5.3.20 (Reduction modulo p)

Hint 1. What are the possible entries of $\mathbf{a}_j - \mathbf{a}_1$?

Hint 2. Subtract column 1 from every other column. This does not change $\det A$, and afterwards columns $2, \ldots, n$ have even entries.

Hint 3. Pull a factor 2 out of each of the columns $2, \ldots, n$. What is left is an integer matrix.

Sketch. For $j \geq 2$, $a_{ij} - a_{i1} \in \{-2, 0, 2\}$. Subtracting column 1 from each other column leaves $\det A$ unchanged (Lemma 5.1.8) and makes columns $2, \ldots, n$ equal to 2 times integer vectors. Pulling out the $n-1$ factors of 2 leaves the determinant of an integer matrix, which is an integer.

For $j \geq 2$, add -1 times column 1 to column j. By Lemma 5.1.8 none of these steps changes the determinant, and since column 1 is never changed, the result is the matrix A' with columns $\mathbf{a}_1, \mathbf{a}_2 - \mathbf{a}_1, \ldots, \mathbf{a}_n - \mathbf{a}_1$. Each entry $a_{ij} - a_{i1}$ with $j \geq 2$ is a difference of two elements of $\{1, -1\}$, so it lies in $\{-2, 0, 2\}$. Hence $\mathbf{a}_j - \mathbf{a}_1 = 2\mathbf{c}_j$ for a vector $\mathbf{c}_j$ with integer entries. Pulling the factor 2 out of each of the columns $2, \ldots, n$ by multilinearity,

$$\det A = \det A' = 2^{n-1} \det(\mathbf{a}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n).$$

The last matrix has integer entries, so its determinant is an integer (Proposition 5.3.20). Hence 2^{n-1} divides $\det A$.

Solution to Exercise 5.3.33

Uses. Definition 5.3.6 (Determinant), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Proposition 5.2.26 (Sign of a cycle), Proposition 1.1.17 (No zero divisors)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. $A + tI_3 = \begin{pmatrix} t & 0 & 1 \\ 1 & t & 0 \\ 0 & 1 & t \end{pmatrix}$. Which permutations avoid the zero entries?
Hint 2. Only id and (1 2 3) survive. For (b), factor $t^3 + 1 = (t+1)\left(t^2 - t + 1\right)$.
Sketch. The surviving terms are t^3 (identity) and $a_{21}a_{32}a_{13} = 1$ for the even permutation (1 2 3), so $\det(A + tI_3) = t^3 + 1$, with $c_2 = 0 = \operatorname{tr} A$ and $c_0 = 1 = \det A$. Roots: $-1, \left(1 \pm i\sqrt{3}\right)/2$ over $\mathbb{C}$; only -1 over $\mathbb{R}$; $t = 1$ over $\mathbb{F}_2$.

(a) $A + tI_3 = \begin{pmatrix} t & 0 & 1 \\ 1 & t & 0 \\ 0 & 1 & t \end{pmatrix}$, whose zero entries are in positions $(1,2)$, $(2,3)$, and $(3,1)$. A term is nonzero only if $\sigma(2) \neq 1$, $\sigma(3) \neq 2$, and $\sigma(1) \neq 3$. Checking the six permutations, only id (with product $t \cdot t \cdot t$) and (1 2 3) (with $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$ and product $a_{21}a_{32}a_{13} = 1$) survive; for instance the term of (1 3 2) contains $a_{31} = 0$. The 3-cycle is even (Proposition 5.2.26), so

$$\det(A + tI_3) = t^3 + 1.$$

This matches Proposition 5.3.19: $c_2 = 0 = \operatorname{tr} A$, and $c_0 = 1 = \det A$, since $A = E_{(1\,2\,3)}$ has determinant $\operatorname{sgn}((1\ 2\ 3)) = 1$.

(b) $t^3 + 1 = (t+1)\left(t^2 - t + 1\right)$, as multiplying out confirms. Over $\mathbb{C}$, let $z_\pm = \left(1 \pm i\sqrt{3}\right)/2$. Then $z_+ + z_- = 1$ and $z_+ z_- = (1+3)/4 = 1$, so $t^2 - t + 1 = (t - z_+)(t - z_-)$, and $t^3 + 1 = (t+1)(t - z_+)(t - z_-)$. A product in a field is 0 only if a factor is 0 (Proposition 1.1.17), so the roots are $t = -1$ and $t = z_\pm$. Over $\mathbb{R}$, $t^2 - t + 1 = (t - 1/2)^2 + 3/4 > 0$, so $t = -1$ is the only root. Over $\mathbb{F}_2$, $0^3 + 1 = 1$ and $1^3 + 1 = 0$, so $t = 1$ is the only root (consistent with the real root -1, since $-1 = 1$ in $\mathbb{F}_2$).

Solution to Exercise 5.3.34

Uses. Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Definition 5.2.10 (Transposition), Definition 5.1.14 (The 2×2 determinant)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. In the proof of Proposition 5.3.19, the term of σ involves powers of t only up to the number of fixed points of σ. Which σ fix at least $n-2$ points?
Hint 2. A permutation that moves exactly two points is a transposition. So only id and the transpositions contribute to t^{n-2}.
Sketch. Only σ with at least $n-2$ fixed points contribute to t^{n-2}: the identity and the transpositions. From $\prod_k (a_{kk} + t)$ the coefficient is $\sum_{i<j} a_{ii}a_{jj}$. The term of $(i\ j)$ is $-a_{ji}a_{ij} \prod_{k \neq i,j} (a_{kk} + t)$, contributing $-a_{ij}a_{ji}$.

We use the notation of the proof of Proposition 5.3.19. The term $T_\sigma(t)$ involves only the powers t^p with $p \leq f_\sigma$, the number of fixed points of σ. So only permutations with at least $n-2$ fixed points contribute to c_{n-2}. Such a σ is either id or moves exactly two points i, j; in the second case $\sigma(i)$ is a moved point different from i, so $\sigma(i) = j$, and likewise $\sigma(j) = i$, so $\sigma = (i\ j)$.

For $\sigma = \mathrm{id}$, $T_{\mathrm{id}}(t) = \prod_{k=1}^{n}(a_{kk}+t)$. A product of t^{n-2} arises by choosing a_{kk} in exactly two factors, $k=i$ and $k=j$, so the coefficient of t^{n-2} is $\sum_{i<j} a_{ii}a_{jj}$.

For $\sigma = (i\ j)$ with $i<j$, which is odd,

$$T_\sigma(t) = -a_{ji}a_{ij}\prod_{k\notin\{i,j\}}(a_{kk}+t),$$

and the coefficient of t^{n-2} is $-a_{ij}a_{ji}$, from choosing t in all $n-2$ factors. Adding up,

$$c_{n-2} = \sum_{i<j} a_{ii}a_{jj} - \sum_{i<j} a_{ij}a_{ji} = \sum_{i<j}\bigl(a_{ii}a_{jj} - a_{ij}a_{ji}\bigr),$$

and $a_{ii}a_{jj} - a_{ij}a_{ji} = \det\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj}\end{pmatrix}$. For $n=2$ this gives $c_0 = \det A$, as it should.

Solution to Exercise 5.3.35

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Definition 1.6.14 (Coordinates), Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 2.1.17 (Compositions of linear maps are linear), Remark 5.1.6
Hint 1. For existence, the coordinate map $\mathbf{w} \mapsto [\mathbf{w}]_{\mathcal{B}}$ is linear and $\left[\mathbf{v}_j\right]_{\mathcal{B}} = \mathbf{e}_j$.
Hint 2. For uniqueness, given g, define $D(\mathbf{x}_1,\dots,\mathbf{x}_n) = g(\Phi(\mathbf{x}_1),\dots,\Phi(\mathbf{x}_n))$ for $\mathbf{x}_j \in F^n$, where $\Phi(\mathbf{x}) = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$. Show $D = \det$.
Sketch. The given f is det composed with the linear coordinate map in each slot, so it is multilinear; equal vectors have equal coordinates, so it is alternating; and $f(\mathbf{v}_1,\dots,\mathbf{v}_n) = \det I_n = 1$. If g is another, then $D = g\circ(\Phi,\dots,\Phi)$ is a determinant function on $F^{n\times n}$, so $D = \det$, and $g(\mathbf{w}_1,\dots) = D\bigl([\mathbf{w}_1]_{\mathcal{B}},\dots\bigr) = f(\mathbf{w}_1,\dots)$.

Existence. Fix j and the vectors $\mathbf{w}_i$ for $i\neq j$. The map $\mathbf{w}\mapsto f(\mathbf{w}_1,\dots,\mathbf{w},\dots,\mathbf{w}_n)$ is the composite of the coordinate map $\mathbf{w}\mapsto[\mathbf{w}]_{\mathcal{B}}$, which is linear (Proposition 1.6.19), with the map $\mathbf{x}\mapsto\det(\dots,\mathbf{x},\dots)$, which is linear because det is multilinear. So it is linear (Proposition 2.1.17), and f is multilinear. If $\mathbf{w}_i = \mathbf{w}_j$ with $i\neq j$, then $[\mathbf{w}_i]_{\mathcal{B}} = \left[\mathbf{w}_j\right]_{\mathcal{B}}$, so $f(\mathbf{w}_1,\dots,\mathbf{w}_n) = 0$ because det is alternating. Finally, $\left[\mathbf{v}_j\right]_{\mathcal{B}} = \mathbf{e}_j$, so $f(\mathbf{v}_1,\dots,\mathbf{v}_n) = \det I_n = 1$.

Uniqueness. Let g be multilinear and alternating with $g(\mathbf{v}_1,\dots,\mathbf{v}_n) = 1$. Let $\Phi: F^n\to V$, $\Phi(\mathbf{x}) = x_1\mathbf{v}_1+\cdots+x_n\mathbf{v}_n$, which is linear, and define $D: F^{n\times n}\to F$ by $D(\mathbf{x}_1,\dots,\mathbf{x}_n) = g(\Phi(\mathbf{x}_1),\dots,\Phi(\mathbf{x}_n))$. As in the existence part, D is multilinear and alternating, and $D(I_n) = g(\Phi(\mathbf{e}_1),\dots,\Phi(\mathbf{e}_n)) = g(\mathbf{v}_1,\dots,\mathbf{v}_n) = 1$. So D is a determinant function, and $D=\det$ by Theorem 5.3.4. Now let $\mathbf{w}_1,\dots,\mathbf{w}_n\in V$. By the definition of coordinates (Definition 1.6.14), $\Phi\bigl(\left[\mathbf{w}_j\right]_{\mathcal{B}}\bigr) = \mathbf{w}_j$, so

$$\begin{aligned} g(\mathbf{w}_1,\dots,\mathbf{w}_n) &= D\bigl([\mathbf{w}_1]_{\mathcal{B}},\dots,[\mathbf{w}_n]_{\mathcal{B}}\bigr) \\ &= \det\bigl([\mathbf{w}_1]_{\mathcal{B}},\dots,[\mathbf{w}_n]_{\mathcal{B}}\bigr) = f(\mathbf{w}_1,\dots,\mathbf{w}_n).\end{aligned}$$

Hence $g=f$. Section 13.4 uses this to define the determinant of an operator without choosing a basis.

Solution to Exercise 5.3.36

Uses. Proposition 5.3.20 (Reduction modulo p), Theorem 5.2.31 (Alternating functions and permutations), Proposition 5.3.15 (Triangular matrices), Corollary 5.1.18 (Non-invertible matrices)
Techniques. Reduce modulo p (Technique 5.3.21)
Hint 1. Reduce A modulo 2. What kind of matrix do you get?
Hint 2. Modulo 2, A becomes a permutation matrix, whose determinant is $\pm1 = 1$ in $\mathbb{F}_2$.
Hint 3. Modulo 2, B has a zero row. Modulo 3, it is upper triangular.
Sketch. (a) Modulo 2, A is the permutation matrix with columns $\mathbf{e}_2,\mathbf{e}_1,\mathbf{e}_4,\mathbf{e}_3$, of determinant 1 in $\mathbb{F}_2$, so $\det A$ is odd. (b) Modulo 2, B has a zero row, so $\det\overline{B} = 0$, which only says $\det B$ is even. Modulo 3, B is upper triangular with diagonal $1,2,1$, so $\det B\equiv 2 \bmod 3$ and $\det B\neq 0$.

(a) Modulo 2, the entries of A become

$$\overline{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

the permutation matrix with columns $\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_3$, that is, E_σ for $\sigma = (1\ 2)(3\ 4)$. By Theorem 5.2.31, $\det \overline{A} = \operatorname{sgn}(\sigma) = 1$ in $\mathbb{F}_2$. By Proposition 5.3.20, $\det A$ is odd, so $\det A \neq 0$, and A is invertible by Corollary 5.1.18. (In fact $\det A = 39$.)

(b) Modulo 2, the second row of B becomes $(0, 0, 0)$. Every term of the Leibniz formula contains an entry of the second row, so $\det \overline{B} = 0$ in $\mathbb{F}_2$. By Proposition 5.3.20 this only says that $\det B$ is even, which is compatible with $\det B = 0$ and with $\det B \neq 0$. Modulo 3,

$$\overline{B} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is upper triangular, so $\det \overline{B} = 1 \cdot 2 \cdot 1 = 2 \neq 0$ in $\mathbb{F}_3$ by Proposition 5.3.15. By Proposition 5.3.20, $\det B$ is not divisible by 3, so $\det B \neq 0$ and B is invertible by Corollary 5.1.18. (In fact $\det B = -64$.)

Solution to Exercise 5.3.37

Uses. Proposition 5.2.30 (Algebra of permutation matrices), Theorem 5.2.31 (Alternating functions and permutations), Proposition 5.3.17 (Block-diagonal matrices), Corollary 5.2.24 (Formulas for the sign), Definition 5.2.21 (Inversion)
Hint 1. Rearrange the columns: moving the last k columns to the front turns the matrix into $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$.
Hint 2. This rearrangement is $M \mapsto ME_\pi$ for a permutation π. Count the inversions of π.
Sketch. Let M be the given matrix, $n = k + m$, and $\pi(j) = m + j$ for $j \leq k$, $\pi(k + i) = i$ for $i \leq m$. Then $ME_\pi = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. The inversions of π are exactly the km pairs $(j, k + i)$, so $\operatorname{sgn}(\pi) = (-1)^{km}$, and $\det M = \operatorname{sgn}(\pi) \det(ME_\pi) = (-1)^{km} \det A \det B$.

Let M be the given matrix and $n = k + m$. Its first m columns are $(\mathbf{0}, \mathbf{b}_i)$ for $i \leq m$, and its last k columns are $(\mathbf{a}_j, \mathbf{0})$ for $j \leq k$. Define $\pi \in S_n$ by $\pi(j) = m + j$ for $j \leq k$ and $\pi(k + i) = i$ for $i \leq m$. By Proposition 5.2.30, the matrix ME_π has columns $\mathbf{c}_{\pi(1)}, \ldots, \mathbf{c}_{\pi(n)}$, where $\mathbf{c}_p$ denotes column p of M; that is, $(\mathbf{a}_1, \mathbf{0}), \ldots, (\mathbf{a}_k, \mathbf{0}), (\mathbf{0}, \mathbf{b}_1), \ldots, (\mathbf{0}, \mathbf{b}_m)$. So $ME_\pi = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, and $\det(ME_\pi) = \operatorname{sgn}(\pi) \det M$ by Theorem 5.2.31.

The bottom row of π is $m + 1, \ldots, m + k, 1, \ldots, m$. Two positions that both lie among the first k, or both among the last m, carry increasing entries, so they are not inversions. A position $j \leq k$ and a position $k + i$ form an inversion, since $\pi(j) = m + j > i = \pi(k + i)$. So π has exactly km inversions, and $\operatorname{sgn}(\pi) = (-1)^{km}$ by Corollary 5.2.24. Using Proposition 5.3.17 and $\operatorname{sgn}(\pi)^2 = 1$,

$$\det M = \operatorname{sgn}(\pi) \det(ME_\pi) = (-1)^{km} \det A \det B.$$

Solution to Exercise 5.3.38

Uses. Definition 5.3.6 (Determinant), Figure 5.3.1, Proposition 5.1.17 (Dependent columns)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. Write $\det A = P - N$, where P is the sum of the three products for even permutations and N the sum for odd ones. Each product is 0 or 1.
Hint 2. So $\det A \leq 3$, with equality only if all three even products are 1. Look at the top row of Figure 5.3.1: which positions do the three even permutations use together?
Sketch. (a) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ has determinant $1 + 1 - 0 = 2$. (b) Each of the six products in (5.3.2) is 0 or 1, so $\det A \leq 3$, and $\det A = 3$ forces all three even products to be 1. The three even permutations together use each of the nine positions exactly once, so then every entry is 1, and $\det A = 0$, a contradiction.

(a) Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. In (5.3.2), the even terms are $a_{11}a_{22}a_{33} = 1$, $a_{21}a_{32}a_{13} = 0$, and $a_{31}a_{12}a_{23} = 1$, and the odd terms are $a_{21}a_{12}a_{33} = 0$, $a_{11}a_{32}a_{23} = 0$, and $a_{31}a_{22}a_{13} = 0$. So $\det A = 2$.

(b) Let A have entries in $\{0,1\}$. In (5.3.2), $\det A = P - N$, where P is the sum of the three products for the even permutations and N is the sum of the three for the odd ones. Each product is 0 or 1, so $P \leq 3$ and $N \geq 0$, and $\det A \leq 3$. Suppose $\det A = 3$. Then $P = 3$, so all three even products equal 1, and all the entries they use equal 1. The three even permutations id, (1 2 3), (1 3 2) use the positions (j,j), $(j+1,j)$, and $(j+2,j)$ in column j, with row indices read modulo 3; in each column these are the three different rows. So together they use all nine positions (the top row of Figure 5.3.1), and every entry of A is 1. But then A has equal columns and $\det A = 0 \neq 3$, a contradiction. Since $\det A$ is an integer, $\det A \leq 2$.

Solution to Exercise 5.3.39

Uses. Definition 5.1.1 (Determinant function), Exercise 5.1.35, Theorem 5.3.4 (Existence and uniqueness of the determinant), Remark 5.2.32 (Two comments on the theorem), Proposition 2.1.17 (Compositions of linear maps are linear)
Techniques. Cancel terms in pairs (Technique 5.3.5)
Hint 1. (a) For column 1, the minors $A^{(i)}$ do not involve column 1. For column $k \geq 2$, the entries a_{i1} are fixed, and column $k-1$ of $A^{(i)}$ is column k of A with its i-th entry deleted.
Hint 2. (b) If columns $k, k+1$ are equal with $k \geq 2$, every minor $A^{(i)}$ has two equal adjacent columns. The case $k = 1$ is the real work: expand each $D_{n-1}\big(A^{(i)}\big)$ once more along its first column.
Hint 3. For $k = 1$ with $x_i = a_{i1} = a_{i2}$, you get a double sum over $i \neq l$ of $\pm x_i x_l D_{n-2}\big(A^{\{i,l\}}\big)$, where $A^{\{i,l\}}$ deletes rows i, l and columns $1, 2$. Compare the signs of the (i,l) and (l,i) terms. Then use Exercise 5.1.35.
Sketch. (a) $D_{n(I_n)} = D_{n-1}(I_{n-1}) = 1$; in column 1, D_n is $\sum (-1)^{i+1} a_{i1} \cdot (\text{const})$; in column $k \geq 2$ each summand is a constant times D_{n-1} of a matrix depending linearly on column k. (b) Equal columns $k, k+1$ with $k \geq 2$ make every minor vanish. For $k = 1$, expanding twice gives $\sum_{i \neq l} \varepsilon_{il} x_i x_l D_{n-2}\big(A^{\{i,l\}}\big)$, with $\varepsilon_{il} = (-1)^{i+l}$ for $l < i$ and $-(-1)^{i+l}$ for $l > i$; the terms (i,l) and (l,i) cancel. With Exercise 5.1.35, D_n is alternating. (c) D_n is a nonzero multilinear alternating function on $\mathbb{Q}^{n \times n}$ built without sgn.

We write $\mathbf{a}_k^{(i)}$ for column k of A with its i-th entry deleted, so $A^{(i)}$ has columns $\mathbf{a}_2^{(i)}, \ldots, \mathbf{a}_n^{(i)}$.

(a) We show by induction on n that D_n is normalized and multilinear. For $n = 1$, $D_1(1) = 1$ and D_1 is linear. Let $n \geq 2$ and assume the claim for $n-1$.

Normalized. The first column of I_n is $\mathbf{e}_1$, so only $i = 1$ contributes, and $I_n^{(1)} = I_{n-1}$. Hence $D_n(I_n) = D_{n-1}(I_{n-1}) = 1$.

Linear in column 1. The minors $A^{(i)}$ do not involve column 1, so for fixed columns $2, \ldots, n$, $D_n(A) = \sum_i (-1)^{i+1} a_{i1} \gamma_i$ with constants γ_i. This is linear in $(a_{11}, \ldots, a_{n1})$.

Linear in column $k \geq 2$. Fix all columns except column k. Each a_{i1} is then constant, and $A^{(i)}$ has column $k-1$ equal to $\mathbf{a}_k^{(i)}$, which depends linearly on $\mathbf{a}_k$ (deleting an entry is a linear map $F^n \to F^{n-1}$), while its other columns are fixed. By the induction hypothesis D_{n-1} is linear in column $k-1$, so $\mathbf{a}_k \mapsto D_{n-1}\big(A^{(i)}\big)$ is a composite of linear maps, hence linear (Proposition 2.1.17). A linear combination of linear maps is linear, so D_n is linear in column k.

(b) We show by induction on $n \geq 2$ that $D_n(A) = 0$ whenever two adjacent columns of A are equal. For $n = 2$, $D_2(A) = a_{11}a_{22} - a_{21}a_{12}$, which vanishes when the two columns are equal. Let $n \geq 3$, assume the claim for $n-1$, and suppose $\mathbf{a}_k = \mathbf{a}_{k+1}$.

Case $k \geq 2$. Each $A^{(i)}$ has equal adjacent columns $\mathbf{a}_k^{(i)} = \mathbf{a}_{k+1}^{(i)}$, in positions $k-1$ and k. By the induction hypothesis, $D_{n-1}\big(A^{(i)}\big) = 0$ for every i, so $D_n(A) = 0$.

Case $k = 1$. Write $x_i = a_{i1} = a_{i2}$. For $i \neq l$, let $A^{\{i,l\}}$ be A with rows i, l and columns $1, 2$ deleted; it depends only on the set $\{i,l\}$. The first column of $A^{(i)}$ is $\mathbf{a}_2^{(i)}$, whose entries are x_l for $l \neq i$, with x_l in position l if $l < i$ and in position $l-1$ if $l > i$. Deleting that position and the first column of $A^{(i)}$ leaves $A^{\{i,l\}}$. So the definition of D_{n-1} gives

$$D_{n-1}\big(A^{(i)}\big) = \sum_{l<i} (-1)^{l+1} x_l D_{n-2}\big(A^{\{i,l\}}\big)$$

$$+\sum_{l>i}(-1)^l x_l D_{n-2}\left(A^{\{i,l\}}\right).$$

Substituting into the definition of D_n,

$$D_n(A)=\sum_{i\neq l}\varepsilon_{il}x_i x_l D_{n-2}\left(A^{\{i,l\}}\right),$$

where the sum runs over ordered pairs (i,l) of distinct indices and

$$\varepsilon_{il}=\begin{cases}(-1)^{i+1}(-1)^{l+1}=(-1)^{i+l} & \text{if } l<i\\ (-1)^{i+1}(-1)^{l}=-(-1)^{i+l} & \text{if } l>i.\end{cases}$$

For $i<l$, the ordered pairs (i,l) and (l,i) give the same product $x_i x_l D_{n-2}\left(A^{\{i,l\}}\right)$ with the signs $\varepsilon_{il}=-(-1)^{i+l}$ and $\varepsilon_{li}=(-1)^{i+l}$, so they cancel. Grouping the terms in these pairs (Technique 5.3.5), $D_n(A)=0$.

By (a) and the above, D_n is multilinear and vanishes whenever two adjacent columns are equal, so it is alternating by Exercise 5.1.35 (for $n=1$ there is nothing to check). Together with $D_n(I_n)=1$, D_n is a determinant function, and $D_n=\det$ by Theorem 5.3.4.

(c) The construction of D_n and parts (a) and (b) never mention permutations or signs. Over $F=\mathbb{Q}$, D_n is a multilinear alternating function with $D_n(I_n)=1\neq 0$. By part (2) of Remark 5.2.32, if σ were a product of r transpositions and also of s transpositions, then $(-1)^r=D_n(E_\sigma)=(-1)^s$, so r and s would have the same parity. This proves Theorem 5.2.18 without Approach 1: a product of differences or Approach 2: counting inversions. Expanding along the first column is the subject of Section 5.5.

Solution to Exercise 5.3.40

Uses. Proposition 5.3.20 (Reduction modulo p), Exercise 5.3.31, Proposition 5.2.16 (Cycle decomposition), Corollary 5.1.18 (Non-invertible matrices), Proposition 5.1.17 (Dependent columns)
Techniques. Reduce modulo p (Technique 5.3.21), Cancel terms in pairs (Technique 5.3.5)
Hint 1. Work modulo 2. What is $\overline{A}$, and what does the Leibniz formula over $\mathbb{F}_2$ count?
Hint 2. Over $\mathbb{F}_2$, $\det\overline{A}$ is the number of permutations with no fixed points (derangements), modulo 2. To find its parity, pair each derangement σ with σ^{-1}.
Hint 3. The unpaired derangements satisfy $\sigma^2=\mathrm{id}$, so they are products of $n/2$ disjoint transpositions. Count them: how many choices are there for the partner of 1?
Sketch. Modulo 2, $\overline{A}$ has 0 on the diagonal and 1 elsewhere, and $\det\overline{A}$ is the number of derangements of $[n]$ modulo 2. The involution $\sigma\mapsto\sigma^{-1}$ on derangements shows that this number has the parity of the number of fixed-point-free involutions, which is $(n-1)(n-3)\cdots 3\cdot 1$, an odd number. So $\det A$ is odd. For $n=3$, the matrix $\begin{pmatrix}0&1&-1\\-1&0&1\\1&-1&0\end{pmatrix}$ has determinant 0.

Reduce modulo 2: $\overline{A}$ has 0 on the diagonal and 1 off the diagonal, since $-1=1$ in $\mathbb{F}_2$. In $\mathbb{F}_2$, $\operatorname{sgn}(\sigma)=1$ for every σ, and the term of σ is 1 if $\sigma(j)\neq j$ for every j and 0 otherwise. So $\det\overline{A}$ is the number d of *derangements* of $[n]$, the permutations without fixed points, reduced modulo 2 (compare part (c) of Exercise 5.3.31).

Pairing. If σ is a derangement, so is σ^{-1}, and $\left(\sigma^{-1}\right)^{-1}=\sigma$. So $\sigma\mapsto\sigma^{-1}$ is an involution of the set of derangements, and the derangements with $\sigma^{-1}\neq\sigma$ fall into pairs $\{\sigma,\sigma^{-1}\}$. Hence d has the same parity as the number m_n of derangements with $\sigma^{-1}=\sigma$, that is, with $\sigma^2=\mathrm{id}$ (Technique 5.3.5).

Counting involutions. Let σ be a derangement with $\sigma^2=\mathrm{id}$. Every orbit of σ has size 2: for each i, $\sigma(i)\neq i$ but $\sigma^2(i)=i$. By Proposition 5.2.16, σ is a product of $n/2$ disjoint transpositions, so it pairs up the elements of $[n]$. Conversely, every way of splitting $[n]$ into $n/2$ pairs gives such a σ. We count these by induction on the even number n: there are $n-1$ choices for the partner p of 1, and then the remaining $n-2$ elements must be split into pairs, in m_{n-2} ways. With $m_2=1$, we get $m_n=(n-1)(n-3)\cdots 3\cdot 1$, a product of odd numbers, which is odd.

So d is odd, $\det \overline{A} = 1$ in $\mathbb{F}_2$, and $\det A$ is odd by Proposition 5.3.20. In particular $\det A \neq 0$, and A is invertible by Corollary 5.1.18.

For odd n the conclusion can fail. With $n = 3$ and $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$, every row sums to 0, so $\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$ and $\det A = 0$ by Proposition 5.1.17. Here the argument breaks down because $m_3 = 0$: a permutation of an odd number of points cannot pair them all up.

Section 5.4

Solution to Exercise 5.4.5

Uses. Proposition 5.4.1 (Determinants of elementary matrices), Corollary 5.4.4 (Products of elementary matrices), Proposition 3.2.28 (Column operations are right multiplications), Remark 3.2.29 (One matrix, two operations)
Hint 1. Read off the type of each factor from its subscripts: two subscripts without a comma is a swap, two with a comma a scaling, three a replacement.
Hint 2. $AG = \big((AE_{13})E_{2,4}\big)E_{3,-7,1}$, so the leftmost factor acts first. As a right factor, E_{icj} adds c times column j to column i.
Sketch. $\det G = (-1) \cdot 4 \cdot 1 = -4$ and $\det(AG) = 5 \cdot (-4) = -20$. The operations are: swap columns 1 and 3, then scale column 2 by 4, then add -7 times column 1 to column 3.

E_{13} swaps, $E_{2,4}$ scales by 4, and $E_{3,-7,1}$ is a replacement, so by Corollary 5.4.4 and Proposition 5.4.1,

$$\begin{aligned} \det G &= \det E_{13} \det E_{2,4} \det E_{3,-7,1} \\ &= (-1) \cdot 4 \cdot 1 = -4, \end{aligned}$$

and $\det(AG) = \det A \det G = -20$. Since $AG = \big((AE_{13})E_{2,4}\big)E_{3,-7,1}$, the column operations are performed from the leftmost factor to the rightmost (Remark 3.2.29). By Proposition 3.2.28 they are: swap columns 1 and 3; scale column 2 by 4; add -7 times column 1 to column 3. Their effects on the determinant, -1, 4, and 1, multiply to $\det G$, as they should.

Solution to Exercise 5.4.9

Uses. Corollary 5.3.9 (Rule of Sarrus), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Compute $\det A(t)$ with the rule of Sarrus. It does not depend on the field.
Hint 2. You should get $1 - 2t^2$. In each field, decide when $2t^2 = 1$ has a solution.
Sketch. $\det A(t) = 1 - 2t^2$. Over $\mathbb{R}$: invertible unless $t = \pm 1/\sqrt{2}$. Over $\mathbb{Q}$: always, since $\sqrt{2}$ is irrational. Over $\mathbb{F}_3$: $1 - 2t^2 = 1 + t^2 \in \{1, 2\}$, so always. Over $\mathbb{F}_7$: $2t^2 = 1$ means $t^2 = 4$, so invertible unless $t \in \{2, 5\}$.

The diagonals running down to the right have products 1, $t \cdot t \cdot 0 = 0$, and 0, and those running up to the right have products 0, $t \cdot t \cdot 1 = t^2$, and $1 \cdot t \cdot t = t^2$. So $\det A(t) = 1 - 2t^2$ by Corollary 5.3.9, and by Theorem 5.4.7, $A(t)$ is invertible exactly when $2t^2 \neq 1$.

Over $\mathbb{R}$, $2t^2 = 1$ exactly for $t = \pm 1/\sqrt{2}$, so $A(t)$ is invertible for all other t. Over $\mathbb{Q}$ there is no solution, since $2t^2 = 1$ would give $(2t)^2 = 2$ with $2t$ rational; so $A(t)$ is always invertible. Over $\mathbb{F}_3$, $-2 = 1$, so $\det A(t) = 1 + t^2$, and $t^2 \in \{0, 1\}$ for $t \in \{0, 1, 2\}$. So $\det A(t) \in \{1, 2\}$ is never 0, and $A(t)$ is always invertible. Over $\mathbb{F}_7$, $2^{-1} = 4$, so $2t^2 = 1$ means $t^2 = 4$, that is, $t = 2$ or $t = -2 = 5$. So $A(t)$ is invertible exactly when $t \notin \{2, 5\}$.

Solution to Exercise 5.4.12

Uses. Theorem 5.4.10 (Multiplicativity), induction on k
Hint 1. Determinants are elements of F, so they commute even when the matrices do not.
Hint 2. For the inverse, take determinants of $AA^{-1} = I_n$.
Sketch. $\det(AB) = \det A \det B = \det B \det A = \det(BA)$. Induction gives $\det\big(A^{k+1}\big) = \det\big(A^k\big) \det A = (\det A)^{k+1}$. From $AA^{-1} = I_n$, $\det A \det A^{-1} = 1$. For $k > 0$, $\det\big(A^{-k}\big) = \det\Big(\big(A^{-1}\big)^k\Big) = \big(\det A^{-1}\big)^k = (\det A)^{-k}$.

By Theorem 5.4.10 and commutativity of multiplication in F, $\det(AB) = \det A \det B = \det B \det A = \det(BA)$.

For powers, $\det\big(A^0\big) = \det I_n = 1 = (\det A)^0$, and if $\det\big(A^k\big) = (\det A)^k$, then $\det\big(A^{k+1}\big) = \det\big(A^k A\big) = \det\big(A^k\big) \det A = (\det A)^{k+1}$. By induction, $\det\big(A^k\big) = (\det A)^k$ for all $k \geq 0$.

If A is invertible, then $\det A \det\left(A^{-1}\right) = \det\left(AA^{-1}\right) = \det I_n = 1$, so $\det\left(A^{-1}\right) = (\det A)^{-1}$. For $k > 0$, the case already proved, applied to A^{-1}, gives $\det\left(A^{-k}\right) = \det\left(\left(A^{-1}\right)^k\right) = \left(\det A^{-1}\right)^k = (\det A)^{-k}$.

Solution to Exercise 5.4.16

Uses. Corollary 5.4.15 (Rows), Theorem 5.4.14 (The transpose), Exercise 5.1.5
Hint 1. Describe each matrix as the result of row operations on the original matrix, or on its transpose.
Hint 2. In (b), the new second row is 2 times the old one plus -3 times the first row. In (c), take the transpose first.
Sketch. (a) One swap: -7. (b) Scale row 2 by 2, then add -3 times row 1: 14. (c) The transpose is the original matrix with 5 times row 1 added to row 3: 7. (d) $3^3 \cdot 7 = 189$.

Call the given matrix M.

(a) This is M with rows 1 and 2 swapped, so its determinant is -7 by Corollary 5.4.15(b).

(b) Scaling row 2 of M by 2 gives determinant 14, and adding -3 times row 1 to the new row 2 does not change it. The result is the given matrix, so its determinant is 14.

(c) The transpose of this matrix is M with 5 times row 1 added to row 3. By Theorem 5.4.14 and Corollary 5.4.15(b), its determinant is $\det M = 7$.

(d) This is $3M$, so by Exercise 5.1.5 its determinant is $3^3 \det M = 189$.

Solution to Exercise 5.4.30

Uses. Algorithm 5.4.22 (Determinant by elimination), Proposition 5.3.20 (Reduction modulo p), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Stage 1 clears column 1 using row 1. Check the $(2,2)$ entry afterwards.
Hint 2. For (b), reduce your answer to (a) modulo p (Proposition 5.3.20).
Sketch. After stage 1 the rows below row 1 are $(0,0,2,1)$, $(0,1,-1,3)$, $(0,-2,4,10)$. Swap rows 2 and 3, add 2 times row 2 to row 4, then subtract row 3 from row 4. The diagonal is $1,1,2,15$ with one swap, so $\det A = -30$. The reduction is invertible exactly for $p \notin \{2,3,5\}$.

(a) Stage 1 adds 2, -3, and -1 times row 1 to rows 2, 3, and 4. Stage 2 finds a 0 in position $(2,2)$, swaps rows 2 and 3, and adds 2 times the new row 2 to row 4. Stage 3 adds -1 times row 3 to row 4:

$$A \xrightarrow{\text{stage 1}} \begin{pmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & -1 & 3 \\ 0 & -2 & 4 & 10 \end{pmatrix},$$

$$\xrightarrow{\text{stage 2}} \begin{pmatrix} 1 & 3 & 0 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 16 \end{pmatrix} \xrightarrow{\text{stage 3}} \begin{pmatrix} 1 & 3 & 0 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 15 \end{pmatrix}.$$

There was one swap, so $\det A = (-1) \cdot 1 \cdot 1 \cdot 2 \cdot 15 = -30$.

(b) By Proposition 5.3.20, the determinant of the reduction is $-30 \bmod p$, which is 0 exactly when p divides $30 = 2 \cdot 3 \cdot 5$. By Theorem 5.4.7 over $\mathbb{F}_p$, the reduction is invertible exactly when $p \notin \{2,3,5\}$.

Solution to Exercise 5.4.31

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Exercise 5.1.5, Exercise 5.4.12, Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Both matrices are invertible by Theorem 5.4.7, so the inverses make sense.
Hint 2. Use $\det(cA) = c^3 \det A$ for 3×3 matrices, and $\det\left(A^{-1}\right) = (\det A)^{-1}$.
Sketch. -6, $9/2$, 16, 2, 4, -8.

Since $\det A$ and $\det B$ are nonzero, A and B are invertible (Theorem 5.4.7). By Theorem 5.4.10, Theorem 5.4.14, Exercise 5.4.12, and Exercise 5.1.5 (with $n = 3$):

- $\det(AB) = 2 \cdot (-3) = -6$;
- $\det\left(A^{-1}B^2\right) = (1/2) \cdot 9 = 9/2$;

- $\det\left(2A^{\mathsf{T}}\right) = 2^3 \det A = 16$;
- $\det\left(BAB^{-1}\right) = \det B \det A (\det B)^{-1} = 2$;
- $\det\left(A^{\mathsf{T}}A\right) = (\det A)^2 = 4$;
- $\det\left(-A^3\right) = (-1)^3 (\det A)^3 = -8$.

Solution to Exercise 5.4.32

Uses. Theorem 5.4.10 (Multiplicativity), Exercise 5.1.5, Theorem 2.4.34 (Block multiplication)
Hint 1. Take determinants: what does $A^2 = -I_n$ say about $(\det A)^2$?
Hint 2. For odd n you need -1 to be a square in F. For even n, find a 2×2 example and place copies of it along the diagonal.
Sketch. $(\det A)^2 = \det(-I_n) = (-1)^n$. If n is odd this needs a square root of -1 in F, which $\mathbb{R}$ and $\mathbb{F}_3$ lack. For even n, the block diagonal matrix with blocks $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ works over every field, as $J^2 = -I_2$. Over $\mathbb{C}$ and $\mathbb{F}_5$, iI_n and $2I_n$ work for every n. So: (a), (c) exactly the even n; (b), (d) every n.

If $A^2 = -I_n$, then by Theorem 5.4.10 and Exercise 5.1.5,

$$(\det A)^2 = \det\left(A^2\right) = \det(-I_n) = (-1)^n.$$

So if n is odd, $\det A$ is a square root of -1 in F.

Conversely, let $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Then $J^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = -I_2$ over every field. If $n = 2l$ is even, let A be block diagonal with l diagonal blocks equal to J. Block diagonal matrices multiply block by block (Theorem 2.4.34), so A^2 is block diagonal with blocks $J^2 = -I_2$, that is, $A^2 = -I_n$.

(a) Squares of real numbers are nonnegative, so $(\det A)^2 = -1$ is impossible: there is no such A for odd n. For even n there is. Answer: exactly the even n.

(b) $A = iI_n$ has $A^2 = i^2 I_n = -I_n$, for every n.

(c) In $\mathbb{F}_3$, $0^2 = 0$, $1^2 = 1$, and $2^2 = 4 = 1$, so $-1 = 2$ is not a square. As in (a), the answer is exactly the even n.

(d) In $\mathbb{F}_5$, $2^2 = 4 = -1$, so $A = 2I_n$ has $A^2 = -I_n$ for every n.

Solution to Exercise 5.4.33

Uses. Theorem 5.4.14 (The transpose), Exercise 5.1.5, Proposition 1.1.17 (No zero divisors), Definition 2.4.21 (Transpose), Corollary 5.4.8 (Vanishing determinants)
Hint 1. Compute $\det A^{\mathsf{T}}$ in two ways: by Theorem 5.4.14 and from $A^{\mathsf{T}} = -A$.
Hint 2. For (b), read off what $a_{ji} = -a_{ij}$ says on the diagonal and off it. For the kernel vector, try a vector built from a, b, c.
Hint 3. Over $\mathbb{F}_2$, $-1 = 1$. What does skew-symmetric mean there?
Sketch. (a) $\det A = \det A^{\mathsf{T}} = \det(-A) = (-1)^n \det A = -\det A$, so $(1+1)\det A = 0$. (b) $a_{ii} = -a_{ii}$ forces $a_{ii} = 0$, and the entries below the diagonal are the negatives of those above; $\mathbf{x} = (c, -b, a)$. (c) Over $\mathbb{F}_2$ skew-symmetric means symmetric, and I_3 is symmetric with determinant 1.

(a) By Theorem 5.4.14 and Exercise 5.1.5,

$$\begin{aligned} \det A &= \det A^{\mathsf{T}} = \det(-A) \\ &= (-1)^n \det A = -\det A, \end{aligned}$$

since n is odd. So $(1+1)\det A = 0$, and as $1 + 1 \neq 0$, $\det A = 0$ (Proposition 1.1.17).

(b) $A^{\mathsf{T}} = -A$ says $a_{ji} = -a_{ij}$ for all i, j (Definition 2.4.21). For $i = j$ this is $(1+1)a_{ii} = 0$, so $a_{ii} = 0$. For $i \neq j$ it says that the entries below the diagonal are the negatives of those above it. With $a = a_{12}$, $b = a_{13}$, and $c = a_{23}$, A has the stated form, and conversely every matrix of that form satisfies $A^{\mathsf{T}} = -A$. Let $\mathbf{x} = (c, -b, a)$. Then

$$A\mathbf{x} = (-ab + ba, -ac + ca, -bc + cb) = \mathbf{0},$$

and $\mathbf{x} \neq \mathbf{0}$ because $A \neq 0$ means that a, b, c are not all 0. So the columns of A are dependent, and $\det A = 0$ by Corollary 5.4.8, as (a) predicts.

(c) In $\mathbb{F}_2$, $-1 = 1$, so $A^{\mathsf{T}} = -A$ means $A^{\mathsf{T}} = A$. The identity matrix I_3 is symmetric, hence skew-symmetric over $\mathbb{F}_2$, and $\det I_3 = 1 \neq 0$. So the hypothesis $1 + 1 \neq 0$ in (a) cannot be dropped.

Solution to Exercise 5.4.34

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Exercise 5.4.12, Proposition 1.1.17 (No zero divisors), Proposition 5.4.1 (Determinants of elementary matrices)
Hint 1. Apply det to the defining equation. You get a polynomial equation for $d = \det A$.
Hint 2. Factor: $d^2 - d = d(d-1)$ and $d^2 - 1 = (d-1)(d+1)$. For the examples, try 0, I_n, and a scaling matrix.
Sketch. (a) $d^2 = d$, so $d \in \{0, 1\}$: the zero matrix and I_n. (b) $d^k = 0$, so $d = 0$: the zero matrix. (c) $d^2 = 1$, so $d = \pm 1$: I_n and $E_{1,-1}$. (d) $d^2 = 1$: again I_n and $E_{1,-1}$.

Write $d = \det A$, and use Exercise 5.4.12 for $\det\left(A^k\right) = d^k$. The zero matrix has determinant 0 (it has a zero column), and I_n has determinant 1.

(a) $d^2 = d$, so $d(d-1) = 0$ and $d \in \{0, 1\}$ (Proposition 1.1.17). Both occur: $0^2 = 0$ and $I_n^2 = I_n$.

(b) $d^k = \det 0 = 0$, and a product of nonzero scalars is nonzero, so $d = 0$. It occurs for the zero matrix.

(c) $d^2 = \det I_n = 1$, so $(d-1)(d+1) = 0$ and $d = \pm 1$. Both occur: $I_n^2 = I_n$, and the scaling $E_{1,-1}$ satisfies $E_{1,-1}^2 = E_{1,1} = I_n$ and has determinant -1 (Proposition 5.4.1). (If $1 + 1 = 0$, the two values coincide.)

(d) $d^2 = \det A^{\mathsf{T}} \det A = \det\left(A^{\mathsf{T}} A\right) = 1$ by Theorem 5.4.14, so $d = \pm 1$. Both occur: I_n, and $E_{1,-1}$, which is diagonal, so $E_{1,-1}^{\mathsf{T}} E_{1,-1} = E_{1,-1}^2 = I_n$.

Solution to Exercise 5.4.35

Uses. Theorem 3.5.19 (Invertible matrices are products of elementary matrices), Corollary 3.2.11 (Products of elementary matrices), Theorem 3.5.23 (One-sided inverses of square matrices), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Reduce to elementary matrices (Technique 5.4.6)
Hint 1. What is the hypothesis of Theorem 3.5.19? Is the zero matrix a product of elementary matrices?
Hint 2. To repair the proof, show that if A or B is not invertible, then neither is AB.
Sketch. The factorization exists only for invertible matrices; a product of elementary matrices is always invertible, so for example 0 is not one. The argument proves the claim when A and B are both invertible. If A or B is not invertible, then AB is not invertible (an inverse C of AB would give A the right inverse BC and B the left inverse CA), so both sides are 0.

The claim is true, but the proof covers only invertible A and B. Theorem 3.5.19 applies to invertible matrices, and it cannot apply to any other matrix, since a product of elementary matrices is invertible (Corollary 3.2.11). For instance, the zero matrix is not a product of elementary matrices.

To repair the argument, it remains to treat the case where A or B is not invertible; then $\det A \det B = 0$ by Theorem 5.4.7. Suppose AB were invertible, with inverse C. Then $A(BC) = I_n$ and $(CA)B = I_n$, so A has a right inverse and B has a left inverse, and both are invertible by Theorem 3.5.23. So if A or B is not invertible, then AB is not invertible, and $\det(AB) = 0 = \det A \det B$. (The first proof of Theorem 5.4.10 avoids factoring A at all.)

Solution to Exercise 5.4.36

Uses. Theorem 5.1.13 (The 2×2 determinant), Theorem 5.2.31 (Alternating functions and permutations), Theorem 5.4.26 (Block-triangular matrices), Definition 5.2.29 (Permutation matrix)
Hint 1. What kind of object is Theorem 5.1.13 about? Are the blocks elements of F?
Hint 2. For (b), look for blocks that each have a zero column, while M is a permutation matrix.
Sketch. (a) Theorem 5.1.13 is about 2×2 matrices with entries in the field F. Blocks are not elements of F, and M is a $2m \times 2m$ matrix, whose determinant is defined through its $2m$ columns; that blocks multiply like entries says nothing about determinants. (b) $A = D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ make M the permutation matrix of $(2\ 4)$, with determinant -1, while every block has determinant 0. (c) This is Theorem 5.4.26, since $\det 0 = 0$.

(a) Theorem 5.1.13 is about functions on $F^{2\times 2}$, matrices whose four entries are elements of the field F. The blocks are matrices, not elements of F (their products do not even commute), and M is a $2m \times 2m$ matrix whose determinant is defined through its $2m$ columns. Theorem 2.4.34 says only how products of block matrices are computed. It says nothing about determinants, so the formula of Theorem 5.1.13 cannot simply be transplanted.

(b) Let $A = D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Then

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

whose columns are $\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_3, \mathbf{e}_2$. So $M = E_\sigma$ for the transposition $\sigma = (2\ 4)$ (Definition 5.2.29), and $\det M = \operatorname{sgn}(\sigma) = -1$ by Theorem 5.2.31. Each of A, B, C, D has a zero column, hence determinant 0, so the claimed formula gives 0. Since $-1 \neq 0$ in every field, the claim is false.

(c) If $C = 0$, then $\det C = 0$ (as $m \geq 1$, C has a zero column), so the claim reads $\det M = \det A \det D$, which is Theorem 5.4.26.

Solution to Exercise 5.4.37

Uses. Corollary 5.4.15 (Rows), Proposition 5.3.15 (Triangular matrices), Theorem 5.2.31 (Alternating functions and permutations), Definition 5.2.29 (Permutation matrix), Proposition 5.2.26 (Sign of a cycle)
Hint 1. Subtract each row from the row below it, starting at the bottom, so that every subtraction uses an original row.
Hint 2. For P_n you get an upper triangular matrix. For Q_n you get rows of the form $(1, \ldots, 1, 0, \ldots, 0)$; subtract again, starting at the bottom.
Sketch. For $i = n, \ldots, 2$, replace row i by row i minus row $i-1$. In P_n, row i becomes $(0, \ldots, 0, 1, \ldots, 1)$ with its first 1 in column i: upper triangular with 1s on the diagonal. In Q_n, row i becomes $(1, \ldots, 1, 0, \ldots, 0)$ with $i-1$ ones; repeating the subtraction turns row i into $\mathbf{e}_{i-1}^{\mathsf{T}}$, and clearing row 1 with these leaves $n\mathbf{e}_n^{\mathsf{T}}$. The rows $n\mathbf{e}_n^{\mathsf{T}}, \mathbf{e}_1^{\mathsf{T}}, \ldots, \mathbf{e}_{n-1}^{\mathsf{T}}$ give n times a cyclic permutation matrix of sign $(-1)^{n-1}$.

Row replacements do not change the determinant (Corollary 5.4.15). In both cases, first replace row i by row i minus row $i-1$ for $i = n, n-1, \ldots, 2$, in this order, so that row $i-1$ is still the original row when it is used.

P_n. The new row i (for $i \geq 2$) has entries $\min(i, j) - \min(i-1, j)$, which is 0 for $j \leq i-1$ and 1 for $j \geq i$. Row 1 is $(1, \ldots, 1)$. The result is upper triangular with every diagonal entry 1, so $\det P_n = 1$ by Proposition 5.3.15.

Q_n. The new row i (for $i \geq 2$) has entries $\max(i, j) - \max(i-1, j)$, which is 1 for $j \leq i-1$ and 0 for $j \geq i$. Row 1 is $(1, 2, \ldots, n)$. Repeat the procedure on rows $2, \ldots, n$: for $i = n, \ldots, 3$, subtract the current row $i-1$ from row i. Since row i has $i-1$ leading ones and row $i-1$ has $i-2$, row i becomes $\mathbf{e}_{i-1}^{\mathsf{T}}$, and row 2 is $\mathbf{e}_1^{\mathsf{T}}$ already. Finally subtract j times row $j+1$ (which is $\mathbf{e}_j^{\mathsf{T}}$) from row 1, for $j = 1, \ldots, n-1$; this leaves row 1 equal to $n\mathbf{e}_n^{\mathsf{T}}$. None of these steps changes the determinant. Pulling the factor n out of row 1 (Corollary 5.4.15), $\det Q_n = n \det R$, where R has rows $\mathbf{e}_n^{\mathsf{T}}, \mathbf{e}_1^{\mathsf{T}}, \ldots, \mathbf{e}_{n-1}^{\mathsf{T}}$. The columns of R are $\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n, \mathbf{e}_1$, so $R = E_\sigma$ for the n-cycle σ with $\sigma(j) = j+1$ for $j < n$ and $\sigma(n) = 1$ (Definition 5.2.29). By Theorem 5.2.31 and Proposition 5.2.26, $\det R = \operatorname{sgn}(\sigma) = (-1)^{n-1}$. Hence $\det Q_n = (-1)^{n-1} n$. (For $n = 1$ there is nothing to do: $Q_1 = (1)$.)

Solution to Exercise 5.4.38

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.10 (Multiplicativity), Exercise 5.1.5
Hint 1. A and B are not square, so Theorem 5.4.10 does not apply directly. Make them square without changing AB.
Hint 2. Add $n-m$ zero columns to A and $n-m$ zero rows to B.
Sketch. With $A' = (A\ \ 0) \in F^{n\times n}$ and $B' = \begin{pmatrix} B \\ 0 \end{pmatrix} \in F^{n\times n}$, block multiplication gives $A'B' = AB + 0 = AB$. Since A' has a zero column, $\det(AB) = \det A' \det B' = 0$. For $m \geq n$ the conclusion fails: $A = (I_n\ \ 0) \in F^{n\times m}$ and $B = A^{\mathsf{T}}$ give $AB = I_n$.

Let $A' = (A\ \ 0) \in F^{n\times n}$ be A followed by $n-m$ zero columns, and let $B' = \begin{pmatrix} B \\ 0 \end{pmatrix} \in F^{n\times n}$ be B followed by $n-m$ zero rows. By Theorem 2.4.34, $A'B' = AB + 0\cdot 0 = AB$. Since $m < n$, A' has a zero column, so $\det A' = 0$ (Exercise 5.1.5), and by Theorem 5.4.10,

$$\det(AB) = \det(A'B') = \det A' \det B' = 0.$$

For $m \geq n$ the conclusion fails. If $m = n$, take $A = B = I_n$. If $m > n$, let $A \in F^{n\times m}$ be I_n followed by $m-n$ zero columns, and $B = A^{\mathsf{T}} \in F^{m\times n}$. Then $AB = I_n$ by Theorem 2.4.34, so $\det(AB) = 1$.

Solution to Exercise 5.4.39

Uses. Definition 5.4.18 (Determinant of an operator), Definition 2.3.1 (Matrix of a linear map), Definition 2.4.1 (Matrix product), Corollary 5.4.27 (Several diagonal blocks), Theorem 5.4.14 (The transpose)
Hint 1. Let U_{ij} be the matrix with 1 in position (i,j) and 0 elsewhere. These n^2 matrices form a basis of $F^{n\times n}$; the order in which you list them matters.
Hint 2. AU_{ij} has column j equal to column i of A, and zeros elsewhere. List the U_{ij} column by column, that is, $U_{11}, U_{21}, \ldots, U_{n1}, U_{12}, \ldots$.
Hint 3. For R_A, list them row by row instead. You should get a block diagonal matrix with blocks A^{T}.
Sketch. With the basis $U_{11}, \ldots, U_{n1}, U_{12}, \ldots, U_{nn}$, the coordinates of X are its columns stacked, and L_A multiplies each column by A: its matrix is block diagonal with n blocks A. With the basis listed row by row, R_A replaces each row $\mathbf{r}^{\mathsf{T}}$ by $\mathbf{r}^{\mathsf{T}}A = \left(A^{\mathsf{T}}\mathbf{r}\right)^{\mathsf{T}}$, so its matrix is block diagonal with n blocks A^{T}. Then use Corollary 5.4.27 and $\det A^{\mathsf{T}} = \det A$.

For $i, j \in [n]$, let U_{ij} have 1 in position (i,j) and 0 elsewhere. Every $X \in F^{n\times n}$ is $\sum_{i,j} x_{ij}U_{ij}$, and the coefficients are the entries of X, so they are unique. Hence the U_{ij}, in any order, form a basis of $F^{n\times n}$, and the coordinates of X are its entries in the corresponding order.

L_A. List the basis column by column: $\mathcal{B} = U_{11}, U_{21}, \ldots, U_{n1}, U_{12}, \ldots, U_{nn}$, so that U_{kj} is the basis vector number $(j-1)n + k$. By Definition 2.4.1, column l of AU_{ij} is A times column l of U_{ij}, which is $A\mathbf{e}_i = \mathbf{a}_i$ for $l = j$ and $\mathbf{0}$ otherwise. So

$$L_A(U_{ij}) = AU_{ij} = \sum_{k=1}^{n} a_{ki}U_{kj}.$$

Hence the column of $[L_A]_{\mathcal{B}\mathcal{B}}$ belonging to U_{ij} has the entries $a_{1i}, \ldots, a_{ni}$ in the positions $(j-1)n+1, \ldots, jn$, and zeros elsewhere (Definition 2.3.1). That is, $[L_A]_{\mathcal{B}\mathcal{B}}$ is block diagonal with n diagonal blocks, each equal to A. By Corollary 5.4.27, $\det L_A = (\det A)^n$.

R_A. List the basis row by row: $\mathcal{B}' = U_{11}, U_{12}, \ldots, U_{1n}, U_{21}, \ldots, U_{nn}$, so that U_{ik} is number $(i-1)n + k$. The (r,k) entry of $U_{ij}A$ is $\sum_l (U_{ij})_{rl}a_{lk}$, which is a_{jk} if $r = i$ and 0 otherwise. So

$$R_A(U_{ij}) = U_{ij}A = \sum_{k=1}^{n} a_{jk}U_{ik},$$

and the column of $[R_A]_{\mathcal{B}'\mathcal{B}'}$ belonging to U_{ij} has the entries $a_{j1}, \ldots, a_{jn}$ in the positions $(i-1)n+1, \ldots, in$, and zeros elsewhere. So $[R_A]_{\mathcal{B}'\mathcal{B}'}$ is block diagonal with n diagonal blocks, each the matrix whose column j is $(a_{j1}, \ldots, a_{jn})$, that is, A^{T}. By Corollary 5.4.27 and Theorem 5.4.14, $\det R_A = \left(\det A^{\mathsf{T}}\right)^n = (\det A)^n$.

Solution to Exercise 5.4.40

Uses. Definition 5.4.18 (Determinant of an operator), Definition 5.2.29 (Permutation matrix), Theorem 5.2.31 (Alternating functions and permutations), Definition 5.2.23 (Sign of a permutation), Example 5.4.19 (Two operators on matrices and polynomials)
Hint 1. Use the basis of matrices U_{ij} with a single 1, as in Example 5.4.19. What does T do to it?
Hint 2. T permutes the basis, so its matrix is a permutation matrix. Count the transpositions in that permutation.
Sketch. $T(U_{ij}) = U_{ji}$, so T permutes the basis by the permutation π of $[n^2]$ exchanging the positions of U_{ij} and U_{ji} for $i < j$. Its matrix is E_π, and π is a product of $n(n-1)/2$ disjoint transpositions, so $\det T = \operatorname{sgn}(\pi) = (-1)^{n(n-1)/2}$. This is 1 exactly when $n \equiv 0$ or 1 modulo 4.

Let U_{ij} be the matrix with 1 in position (i,j) and 0 elsewhere, and list these n^2 matrices in some order $\mathbf{b}_1, \ldots, \mathbf{b}_{n^2}$; they form a basis of $F^{n\times n}$, as in Exercise 5.4.39. Since $U_{ij}^{\mathsf{T}} = U_{ji}$, there is a permutation π of $[n^2]$ with $T(\mathbf{b}_l) = \mathbf{b}_{\pi(l)}$ for every l: π exchanges the positions of U_{ij} and U_{ji} for $i \neq j$ and fixes the positions of the U_{ii}. So column l of the matrix of T is $\mathbf{e}_{\pi(l)}$, that is, the matrix is the permutation matrix E_π (Definition 5.2.29), and $\det T = \operatorname{sgn}(\pi)$ by Theorem 5.2.31.

The permutation π is the product of the transpositions exchanging the positions of U_{ij} and U_{ji}, one for each pair $i < j$. There are $n(n-1)/2$ such pairs, so $\operatorname{sgn}(\pi) = (-1)^{n(n-1)/2}$ by Definition 5.2.23. Hence $\det T = (-1)^{n(n-1)/2}$, in agreement with Example 5.4.19 for $n = 2$.

If $1 + 1 \neq 0$, then $-1 \neq 1$, and $\det T = 1$ exactly when $n(n-1)/2$ is even, that is, when $n(n-1)$ is divisible by 4. One of n and $n-1$ is odd, so this happens exactly when n or $n-1$ is divisible by 4: $n \equiv 0$ or 1 modulo 4. For instance, $\det T = -1$ for $n = 2, 3$ and $\det T = 1$ for $n = 4, 5$.

Solution to Exercise 5.4.41

Uses. Theorem 5.4.10 (Multiplicativity), Exercise 5.4.12, Exercise 5.1.5, Theorem 5.4.7 (Invertibility and the determinant), Definition 4.5.23 (Similar matrices), Definition 2.4.26 (Trace)
Hint 1. (a) I_2 is similar only to itself. (d) Try matrices with a zero column.
Hint 2. (c) Take determinants of $AB = -BA$, and remember that n is odd.
Sketch. (a) False: $\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ and I_2. (b) True: $\det A \det B (\det A)^{-1}(\det B)^{-1} = 1$. (c) True: $\det A \det B = (-1)^n \det B \det A = -\det A \det B$, so $\det A \det B = 0$. (d) False: $A = 0$, $B = \left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right)$, $C = \left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)$.

(a) False. $A = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ and $B = I_2$ have determinant 1 and trace 2. But $S^{-1} I_2 S = I_2 \neq A$ for every invertible S, so A is not similar to I_2.

(b) True. By Theorem 5.4.10 and Exercise 5.4.12,

$$\begin{aligned}&\det\left(ABA^{-1}B^{-1}\right)\\ &= \det A \det B (\det A)^{-1}(\det B)^{-1} = 1,\end{aligned}$$

because scalars commute.

(c) True. Taking determinants, with Exercise 5.1.5 for the factor -1,

$$\begin{aligned}\det A \det B = \det(AB) &= \det(-BA)\\ &= (-1)^n \det B \det A = -\det A \det B.\end{aligned}$$

So $2 \det A \det B = 0$, hence $\det A = 0$ or $\det B = 0$, and the corresponding matrix is not invertible (Theorem 5.4.7). (For even n the statement fails: $A = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$ are invertible, and $AB = \left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right) = -BA$.)

(d) False. Let $n = 2$, $A = 0$, $B = \left(\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right)$, and $C = \left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)$. Then $\det A = \det B = 0$, but $\det(A + C) = \det C = 0$ while $\det(B + C) = \det I_2 = 1$.

Solution to Exercise 5.4.42

Uses. Proposition 5.3.20 (Reduction modulo p), Theorem 5.4.10 (Multiplicativity)
Hint 1. The determinant of a matrix with integer entries is an integer.
Hint 2. Which integers have an integer reciprocal?
Sketch. $\det A$ and $\det A^{-1}$ are integers whose product is $\det I_n = 1$, so both are ± 1.

By Proposition 5.3.20, the determinant of a matrix with integer entries is an integer, so $\det A$ and $\det\left(A^{-1}\right)$ are integers. By Theorem 5.4.10, their product is $\det\left(AA^{-1}\right) = \det I_n = 1$. The only integers u with $uv = 1$ for some integer v are ± 1, since $|u| \geq 2$ would give $|uv| \geq 2$ for $v \neq 0$. Hence $\det A = \pm 1$.

Solution to Exercise 5.4.43

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.10 (Multiplicativity)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Imitate what you would do for a 2×2 matrix $\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)$: add the second column to the first, then subtract the first row from the second.
Hint 2. With blocks, these operations are multiplications by $Q = \left(\begin{smallmatrix} I & 0 \\ I & I \end{smallmatrix}\right)$ on the right and $P = \left(\begin{smallmatrix} I & 0 \\ -I & I \end{smallmatrix}\right)$ on the left.
Sketch. With $P = \left(\begin{smallmatrix} I & 0 \\ -I & I \end{smallmatrix}\right)$ and $Q = \left(\begin{smallmatrix} I & 0 \\ I & I \end{smallmatrix}\right)$, block multiplication gives $PMQ = \left(\begin{smallmatrix} A+B & B \\ 0 & A-B \end{smallmatrix}\right)$. Both P and Q have determinant 1, so $\det M = \det(A+B)\det(A-B)$ by Theorem 5.4.26.

Let $M = \left(\begin{smallmatrix} A & B \\ B & A \end{smallmatrix}\right)$, $P = \left(\begin{smallmatrix} I_m & 0 \\ -I_m & I_m \end{smallmatrix}\right)$, and $Q = \left(\begin{smallmatrix} I_m & 0 \\ I_m & I_m \end{smallmatrix}\right)$. By Theorem 2.4.34,

$$MQ = \begin{pmatrix} A+B & B \\ B+A & A \end{pmatrix},$$

$$\begin{aligned} PMQ &= \begin{pmatrix} A+B & B \\ -(A+B)+(B+A) & -B+A \end{pmatrix} \\ &= \begin{pmatrix} A+B & B \\ 0 & A-B \end{pmatrix}. \end{aligned}$$

P and Q are block lower triangular with diagonal blocks I_m, so $\det P = \det Q = 1$ by Theorem 5.4.26. By Theorem 5.4.10 and Theorem 5.4.26,

$$\begin{aligned} \det M &= \det P \det M \det Q = \det(PMQ) \\ &= \det(A+B)\det(A-B). \end{aligned}$$

Solution to Exercise 5.4.44

Uses. Corollary 5.4.15 (Rows), Lemma 5.1.8 (Column replacement), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Every row sums to $s = 1+z+z^2$. Add the second and third columns to the first, and pull out s.
Hint 2. Then subtract row 1 from rows 2 and 3, and look for the factor $1-z$ in the new rows. Recall $(1-z)s = 1-z^3$.
Sketch. Adding columns 2 and 3 to column 1 and pulling out $s = 1+z+z^2$ leaves first column $(1,1,1)$. Subtracting row 1 from rows 2, 3 gives rows $(1-z)(0,1,z)$ and $(1-z)(0,-z,1+z)$. So $\det C(z) = s(1-z)^2\left(1+z+z^2\right) = \left(1-z^3\right)^2$. Non-invertible exactly when $z^3 = 1$: $z = 1, \omega, \omega^2$ in $\mathbb{C}$; $z = 1$ in $\mathbb{R}$; $z \in \{1,2,4\}$ in $\mathbb{F}_7$.

(a) Let $s = 1+z+z^2$. Adding columns 2 and 3 to column 1 does not change the determinant (Lemma 5.1.8), and makes every entry of column 1 equal to s. Pulling out s, and then subtracting row 1 from rows 2 and 3 (Corollary 5.4.15),

$$\begin{aligned} \det C(z) &= s\det\begin{pmatrix} 1 & z & z^2 \\ 1 & 1 & z \\ 1 & z^2 & 1 \end{pmatrix} \\ &= s\det\begin{pmatrix} 1 & z & z^2 \\ 0 & 1-z & z-z^2 \\ 0 & z^2-z & 1-z^2 \end{pmatrix}. \end{aligned}$$

Row 2 of the last matrix is $(1-z)(0,1,z)$, and row 3 is $(1-z)(0,-z,1+z)$. Pulling out both factors $1-z$ and using Theorem 5.4.26 with a 1×1 upper left block,

$$\begin{aligned} \det C(z) &= s(1-z)^2\det\begin{pmatrix} 1 & z \\ -z & 1+z \end{pmatrix} \\ &= s(1-z)^2\left(1+z+z^2\right) = ((1-z)s)^2 = \left(1-z^3\right)^2, \end{aligned}$$

since $(1-z)\left(1+z+z^2\right) = 1-z^3$. Every step is valid over any field.

(b) By Theorem 5.4.7, $C(z)$ fails to be invertible exactly when $\left(1-z^3\right)^2 = 0$, that is, when $z^3 = 1$. Over $\mathbb{C}$ the solutions are the cube roots of unity 1, $\omega = \left(-1+i\sqrt{3}\right)/2$, and $\omega^2 = \overline{\omega}$. Over $\mathbb{R}$, only $z = 1$. Over $\mathbb{F}_7$, cubing the nonzero elements gives $1^3 = 1$, $2^3 = 8 = 1$, $3^3 = 27 = 6$, $4^3 = 64 = 1$, $5^3 = 125 = 6$, and $6^3 = 216 = 6$, so $z \in \{1, 2, 4\}$.

Solution to Exercise 5.4.45

Uses. Theorem 2.4.34 (Block multiplication), Proposition 5.4.17 (Similar matrices have equal determinants), Theorem 5.4.26 (Block-triangular matrices), Exercise 5.3.28, Definition 5.3.6 (Determinant)
Hint 1. Regard $M = \left(\begin{smallmatrix} B & -C \\ C & B \end{smallmatrix}\right)$ as a complex matrix. Its determinant is the same whether computed over $\mathbb{R}$ or over $\mathbb{C}$.
Hint 2. Conjugate by $P = \left(\begin{smallmatrix} I & 0 \\ iI & I \end{smallmatrix}\right)$, whose inverse is $\left(\begin{smallmatrix} I & 0 \\ -iI & I \end{smallmatrix}\right)$, and simplify $P^{-1}MP$ with block multiplication.
Hint 3. $B - iC$ is the entrywise complex conjugate of A. Use Exercise 5.3.28.
Sketch. Over $\mathbb{C}$, $P^{-1}MP = \left(\begin{smallmatrix} B-iC & -C \\ 0 & B+iC \end{smallmatrix}\right)$ with $P = \left(\begin{smallmatrix} I & 0 \\ iI & I \end{smallmatrix}\right)$. So $\det M = \det(B-iC)\det(B+iC) = \overline{\det A}\det A = |\det A|^2$.

Let $M = \left(\begin{smallmatrix} B & -C \\ C & B \end{smallmatrix}\right)$. Its determinant is given by the Leibniz formula (Definition 5.3.6), which involves only sums and products of the (real) entries, so it is the same whether we regard M as a real or a complex matrix. We work over $\mathbb{C}$. Let $P = \left(\begin{smallmatrix} I_n & 0 \\ iI_n & I_n \end{smallmatrix}\right)$ and $P' = \left(\begin{smallmatrix} I_n & 0 \\ -iI_n & I_n \end{smallmatrix}\right)$. By Theorem 2.4.34, $PP' = P'P = I_{2n}$, so $P' = P^{-1}$, and

$$MP = \begin{pmatrix} B - iC & -C \\ C + iB & B \end{pmatrix},$$

$$\begin{aligned} P^{-1}MP &= \begin{pmatrix} B - iC & -C \\ -i(B - iC) + C + iB & iC + B \end{pmatrix} \\ &= \begin{pmatrix} B - iC & -C \\ 0 & B + iC \end{pmatrix}, \end{aligned}$$

using $-i(-iC) = i^2C = -C$. By Proposition 5.4.17 and Theorem 5.4.26,

$$\begin{aligned} \det M &= \det\left(P^{-1}MP\right) \\ &= \det(B - iC)\det(B + iC). \end{aligned}$$

Since B and C are real, $B - iC = \overline{A}$, the entrywise complex conjugate of A, and $\det\overline{A} = \overline{\det A}$ by Exercise 5.3.28. Hence $\det M = \overline{\det A}\det A = |\det A|^2 \geq 0$.

Solution to Exercise 5.4.46

Uses. Theorem 5.4.10 (Multiplicativity), Exercise 5.3.28
Hint 1. Over $\mathbb{C}$, the scalar $a^2 + b^2$ factors as $(a + ib)(a - ib)$. Does the same factorization work for commuting matrices?
Hint 2. $\det(A - iB)$ is the complex conjugate of $\det(A + iB)$, by Exercise 5.3.28.
Hint 3. For the example, take $n = 2$ and look for A and B with small entries such that $A^2 + B^2$ has determinant -1.
Sketch. $(A + iB)(A - iB) = A^2 - iAB + iBA + B^2 = A^2 + B^2$ when $AB = BA$. So $\det\left(A^2 + B^2\right) = z\overline{z} = |z|^2 \geq 0$ with $z = \det(A + iB)$. Without commuting: $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ give $A^2 + B^2 = \left(\begin{smallmatrix} 2 & 3 \\ 1 & 1 \end{smallmatrix}\right)$, with determinant -1.

Regard A and B as complex matrices; this does not change $\det\left(A^2 + B^2\right)$, which is computed from the entries by the Leibniz formula. Since $AB = BA$,

$$\begin{aligned} (A + iB)(A - iB) &= A^2 - iAB + iBA - i^2B^2 \\ &= A^2 + B^2. \end{aligned}$$

Let $z = \det(A + iB)$. Since A and B are real, $A - iB$ is the entrywise conjugate of $A + iB$, so $\det(A - iB) = \overline{z}$ by Exercise 5.3.28. By Theorem 5.4.10,

$$\det\left(A^2 + B^2\right) = z\overline{z} = |z|^2 \geq 0.$$

Without the hypothesis, let $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Then $A^2 = \left(\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right)$ and $B^2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$, so $A^2 + B^2 = \left(\begin{smallmatrix} 2 & 3 \\ 1 & 1 \end{smallmatrix}\right)$, whose determinant is $2 - 3 = -1 < 0$. Indeed $AB = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $BA = \left(\begin{smallmatrix} 2 & 2 \\ -1 & -1 \end{smallmatrix}\right)$ are different.

Solution to Exercise 5.4.47

Uses. Lemma 5.4.2 (Multiplying by an elementary matrix), Proposition 5.4.1 (Determinants of elementary matrices), Theorem 5.4.7 (Invertibility and the determinant), Corollary 5.4.8 (Vanishing determinants), Proposition 1.2.35 (Zero products)
Hint 1. (a) Compute $\det(XE_{1c})$, and write down the inverse map. Which values can the determinant of an invertible matrix take?
Hint 2. (b) A 2×2 matrix is invertible exactly when its columns are independent. Count the choices of the first column, then of the second.
Sketch. (a) $\det(XE_{1c}) = c\det X$, and $Y \mapsto YE_{1,c^{-1}}$ is the inverse map. G is the disjoint union of the $p-1$ sets $\{\det = c\}$, $c \neq 0$, each in bijection with H. (b) The first column is any of the p^2-1 nonzero vectors, and the second any of the p^2-p vectors that are not multiples of the first. (c) For $p=2$, $H=G$ has the 6 matrices I_2, E_{12}, and the four with exactly three entries 1.

(a) For $X \in H$, $\det(XE_{1c}) = \det X \det E_{1c} = c$ by Lemma 5.4.2 and Proposition 5.4.1, and XE_{1c} is invertible. So the map sends H into $H_c = \{A \in G \mid \det A = c\}$. Similarly $Y \mapsto YE_{1,c^{-1}}$ sends H_c into H, and the two maps are inverse to each other, since $E_{1c}E_{1,c^{-1}} = E_{1,c^{-1}}E_{1c} = I_n$. So $|H_c| = |H|$. By Theorem 5.4.7, every $A \in G$ has $\det A \neq 0$, so G is the disjoint union of the $p-1$ sets H_c with $c \neq 0$. Hence $|G| = (p-1)|H|$.

(b) By Corollary 5.4.8, $A \in \mathbb{F}_p^{2\times 2}$ is invertible exactly when its columns $\mathbf{a}_1, \mathbf{a}_2$ are linearly independent. This holds exactly when $\mathbf{a}_1 \neq \mathbf{0}$ and $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$. Indeed, if $x\mathbf{a}_1 + y\mathbf{a}_2 = \mathbf{0}$ with $(x,y) \neq (0,0)$, then either $y \neq 0$ and $\mathbf{a}_2 = -y^{-1}x\mathbf{a}_1$, or $y = 0$, $x \neq 0$, and $\mathbf{a}_1 = \mathbf{0}$; conversely, $\mathbf{a}_1 = \mathbf{0}$ or $\mathbf{a}_2 = t\mathbf{a}_1$ gives a nontrivial relation. There are p^2-1 nonzero vectors in $\mathbb{F}_p^2$. For each of them, the multiples $t\mathbf{a}_1$ with $t \in \mathbb{F}_p$ are p distinct vectors, since $t\mathbf{a}_1 = t'\mathbf{a}_1$ forces $t = t'$ (Proposition 1.2.35), so there are p^2-p choices of $\mathbf{a}_2$. Hence $|G| = (p^2-1)(p^2-p)$, and by (a),

$$|H| = \frac{(p^2-1)(p^2-p)}{p-1} = p(p^2-1).$$

(c) For $p=2$ the formula gives $2\cdot 3 = 6$. Over $\mathbb{F}_2$ the only nonzero scalar is 1, so $H=G$. The invertible 2×2 matrices over $\mathbb{F}_2$ are those with $ad-bc = ad+bc = 1$, that is, with exactly one of the products ad and bc equal to 1:

$$\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \begin{pmatrix}1&1\\0&1\end{pmatrix},$$
$$\begin{pmatrix}1&0\\1&1\end{pmatrix},\quad \begin{pmatrix}0&1\\1&1\end{pmatrix},\quad \begin{pmatrix}1&1\\1&0\end{pmatrix}.$$

Solution to Exercise 5.4.48

Uses. Exercise 3.2.53, Proposition 3.2.7 (Row operations are left multiplications), Proposition 5.4.1 (Determinants of elementary matrices), Theorem 5.4.10 (Multiplicativity), Proposition 5.1.17 (Dependent columns), Proposition 5.3.15 (Triangular matrices), Proposition 3.2.10 (Inverses of elementary matrices)
Techniques. Algorithms as proofs: invariants and termination (Technique 3.3.5)
Hint 1. Imitate Exercise 3.2.53, column by column: for $k = 1, \ldots, n-1$, use replacements to make column k equal to $\mathbf{e}_k$ without disturbing columns $1, \ldots, k-1$.
Hint 2. To make the (k,k) entry equal to 1, you need a nonzero entry in column k below row k. If there is none, first add row k to row $k+1$. Why is the (k,k) entry then nonzero?
Hint 3. Replacements do not change the determinant. After column $n-1$, read off the last diagonal entry, and clear the last column above it.
Sketch. Keep the invariant that columns $1, \ldots, k-1$ are $\mathbf{e}_1, \ldots, \mathbf{e}_{k-1}$; then rows $k, \ldots, n$ vanish in those columns, so adding them to other rows does not break the invariant. Column k has a nonzero entry in rows $k, \ldots, n$ (else the columns are dependent); arrange one in a row $j > k$, add $(1-m_{kk})/m_{jk}$ times row j to row k to get a 1, and clear the rest of column k. At the end the matrix is upper triangular with diagonal $1, \ldots, 1, d$, where $d = \det A$ because replacements preserve det; clearing column n above d gives $\operatorname{diag}(1, \ldots, 1, d)$. Then $A = E_1^{-1}\cdots E_t^{-1}E_{n,d}$.

(a) Every matrix below is obtained from A by replacements, so it is invertible with determinant $\det A$ (Proposition 3.2.7, Proposition 5.4.1, Theorem 5.4.10). We process $k = 1, \ldots, n-1$ in turn, keeping the

following invariant: before step k, columns $1, \dots, k-1$ of the current matrix M are $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$. Then for $l < k$, every row other than row l has entry 0 in column l. In particular rows $k, \dots, n$ vanish in columns $1, \dots, k-1$, so adding a multiple of one of these rows to another row does not change columns $1, \dots, k-1$. All replacements in step k add multiples of rows numbered at least k.

Step k. Column k of M has a nonzero entry in one of the rows $k, \dots, n$: otherwise column k would be a combination of $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$, that is, of columns $1, \dots, k-1$, and $\det M = 0$ by Proposition 5.1.17. If $m_{jk} = 0$ for every $j > k$, then $m_{kk} \neq 0$, and we add row k to row $k+1$ (possible as $k+1 \leq n$); now $m_{k+1,k} = m_{kk} \neq 0$. So we may choose $j > k$ with $c = m_{jk} \neq 0$. Adding $(1 - m_{kk})c^{-1}$ times row j to row k changes the (k,k) entry to $m_{kk} + (1 - m_{kk}) = 1$. Finally, for each $i \neq k$, add $-m_{ik}$ times row k to row i; this makes column k equal to $\mathbf{e}_k$. The invariant now holds before step $k+1$.

After step $n-1$, columns $1, \dots, n-1$ are $\mathbf{e}_1, \dots, \mathbf{e}_{n-1}$, so the matrix is upper triangular with diagonal entries $1, \dots, 1, m_{nn}$, and $m_{nn} = \det A$ by Proposition 5.3.15. In particular $m_{nn} \neq 0$. Row n is $(0, \dots, 0, m_{nn})$, so adding $-m_{in}m_{nn}^{-1}$ times row n to row i, for each $i < n$, clears column n above the diagonal and changes nothing else. The result is $\operatorname{diag}(1, \dots, 1, \det A)$, reached by replacements $E_1, \dots, E_t$.

(b) By (a), $A = E_1^{-1} \cdots E_t^{-1} \operatorname{diag}(1, \dots, 1, \det A)$, and each E_l^{-1} is a replacement by Proposition 3.2.10. If $\det A = 1$, the last factor is I_n, and A is a product of replacements. In general the last factor is the scaling $E_{n,\det A}$.

Solution to Exercise 5.4.49

Uses. Exercise 5.4.48, Proposition 5.2.30 (Algebra of permutation matrices), Theorem 3.3.10 ($GA = \operatorname{rref}(A)$), Theorem 3.3.23 (Invertibility and RREF), Proposition 3.3.8 (Properties of the output), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Reduce to elementary matrices (Technique 5.4.6)
Hint 1. First show $f(I_n) = 1$, that $f(G) \neq 0$ for invertible G, and that $f\left(GXG^{-1}\right) = f(X)$.
Hint 2. Show that all replacements are conjugate to $E_{1,1,2}$, using a scaling and a permutation matrix. Then compare $E_{1,1,2}^2$ with $E_{1,2,2}$; treat $1 + 1 = 0$ separately.
Hint 3. For invertible A, use Exercise 5.4.48. For A not invertible, write $A = G^{-1} \operatorname{rref}(A)$, and note that $\operatorname{rref}(A)$ is unchanged by left multiplication with $\operatorname{diag}(1, \dots, 1, 0)$.
Sketch. f is invariant under conjugation by invertible matrices. Conjugating $E_{i,1,j}$ by E_{jc} gives E_{icj}, and conjugating by a permutation matrix moves it to $E_{1,1,2}$; so all replacements have the same value $s \neq 0$. From $E_{1,1,2}^2 = E_{1,2,2}$ we get $s^2 = s$ (or $s^2 = 1$ if $1 + 1 = 0$), so $s = 1$. Conjugating $\operatorname{diag}(t, 1, \dots, 1)$ by E_{1n} gives $f(\operatorname{diag}(1, \dots, 1, t)) = t$, so $f(A) = \det A$ for invertible A by Exercise 5.4.48. If A is not invertible, $\operatorname{rref}(A) = Z \operatorname{rref}(A)$ with $Z = \operatorname{diag}(1, \dots, 1, 0)$ and $f(Z) = 0$.

Throughout, let U_{ji} be the matrix with 1 in position (j,i) and 0 elsewhere, so that $E_{icj} = I_n + cU_{ji}$ for $i \neq j$.

Step 0. $f(I_n) = f(\operatorname{diag}(1, \dots, 1)) = 1$. If G is invertible, then $f(G)f\left(G^{-1}\right) = f(I_n) = 1$, so $f(G) \neq 0$ and $f\left(G^{-1}\right) = f(G)^{-1}$. Hence $f\left(GXG^{-1}\right) = f(G)f(X)f(G)^{-1} = f(X)$ for every X, because values of f commute.

Step 1: replacements. Let $i \neq j$ and $c \neq 0$, and let $D = E_{jc}$, so $D^{-1} = E_{j,c^{-1}}$. Multiplying U_{ji} on the left by D scales its row j by c. Multiplying on the right by D^{-1} scales its column j by c^{-1}, and that column is zero since $i \neq j$. So $DU_{ji}D^{-1} = cU_{ji}$, hence $DE_{i,1,j}D^{-1} = E_{icj}$, and $f\left(E_{icj}\right) = f\left(E_{i,1,j}\right)$ by Step 0. Next choose $\sigma \in S_n$ with $\sigma(i) = 1$ and $\sigma(j) = 2$, possible as $i \neq j$ and $n \geq 2$. By Proposition 5.2.30, $E_\sigma^{-1} = E_{\sigma^{-1}}$, and $XE_{\sigma^{-1}}$ has column l equal to column $\sigma^{-1}(l)$ of X. So $U_{ji}E_{\sigma^{-1}}$ has its only nonzero column in position $\sigma(i)$, equal to $\mathbf{e}_j$; multiplying by E_σ on the left turns this column into $E_\sigma \mathbf{e}_j = \mathbf{e}_{\sigma(j)}$. Hence $E_\sigma U_{ji} E_\sigma^{-1} = U_{\sigma(j)\sigma(i)} = U_{21}$, and $E_\sigma E_{i,1,j} E_\sigma^{-1} = I_n + U_{21} = E_{1,1,2}$. So every replacement E_{icj} with $c \neq 0$ has the value $s = f\left(E_{1,1,2}\right)$, and $s \neq 0$ by Step 0. (A replacement with $c = 0$ is I_n, with $f(I_n) = 1$.)

Since $U_{21}^2 = 0$, we have $E_{1,1,2}^2 = I_n + 2U_{21}$, where $2 = 1 + 1$. If $1 + 1 \neq 0$, this is the replacement $E_{1,2,2}$, so $s^2 = s$, and $s = 1$ because $s \neq 0$. If $1 + 1 = 0$, then $E_{1,1,2}^2 = I_n$, so $s^2 = 1$, and $(s-1)^2 = s^2 - (1+1)s + 1 = s^2 + 1 = 1 + 1 = 0$, so again $s = 1$. Thus $f(E) = 1 = \det E$ for every replacement E.

Step 2: invertible matrices. E_{1n} is its own inverse, and $E_{1n} \operatorname{diag}(t, 1, \dots, 1) E_{1n} = \operatorname{diag}(1, \dots, 1, t) = E_{n,t}$, as conjugating by E_{1n} swaps rows 1 and n and then columns 1 and n. So $f(\operatorname{diag}(1, \dots, 1, t)) = t$ for every $t \in F$, by

Step 0; for $t \neq 0$ this matrix is the scaling E_{nt}. If A is invertible, then by Exercise 5.4.48(b), $A = R_1 \cdots R_q E_{n,\det A}$ with replacements R_l, and

$$f(A) = f(R_1)\cdots f\big(R_q\big) f\big(E_{n,\det A}\big) = \det A.$$

Step 3: matrices that are not invertible. Let A be not invertible, so $\det A = 0$ (Theorem 5.4.7). By Theorem 3.3.10, $A = G^{-1}R$ with G invertible and $R = \operatorname{rref}(A)$, and $R \neq I_n$ by Theorem 3.3.23. So R has fewer than n pivots, and its last row is zero by Proposition 3.3.8. Let $Z = \operatorname{diag}(1, \ldots, 1, 0)$. Multiplying by Z on the left replaces the last row by zeros, so $ZR = R$. By the computation in Step 2 with $t = 0$, $f(Z) = 0$. Hence $f(R) = f(Z)f(R) = 0$, and $f(A) = f\big(G^{-1}\big)f(R) = 0 = \det A$.

So $f(A) = \det A$ for every A.

Solution to Exercise 5.4.50

Uses. Proposition 5.4.29 (Schur complement), Theorem 5.4.10 (Multiplicativity)
Techniques. Block elimination (Technique 5.4.28)
Hint 1. Start from Proposition 5.4.29, and combine the two determinants on the right into one.
Hint 2. $ACA^{-1} = C$ when A and C commute. For the example, a zero row in M is an easy way to make $\det M = 0$.
Sketch. $\det M = \det A \det\big(D - CA^{-1}B\big) = \det\big(AD - ACA^{-1}B\big) = \det(AD - CB)$. Example: $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B = C = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $D = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$; then M has a zero row, while $AD - CB = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

By Proposition 5.4.29 and Theorem 5.4.10,

$$\begin{aligned} \det\begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \det A \det\big(D - CA^{-1}B\big) \\ &= \det\big(A\big(D - CA^{-1}B\big)\big) \\ &= \det\big(AD - ACA^{-1}B\big). \end{aligned}$$

Since $AC = CA$, we have $ACA^{-1} = CAA^{-1} = C$, so the right side is $\det(AD - CB)$.

For the example, let $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B = C = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, and $D = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Then A is invertible, and $AC = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \neq \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) = CA$. The third row of $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is the first row of C followed by the first row of D, which is zero, so $\det\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) = 0$. But $AD = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $CB = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, so $\det(AD - CB) = \det\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = -1$.

Section 5.5

Solution to Exercise 5.5.5

Uses. Theorem 5.5.3 (Cofactor expansion), Lemma 5.5.2 (A row or column with one nonzero entry)
Hint 1. Look for a row or a column with a single nonzero entry.
Hint 2. Row 3 is $(0, 0, 1, 0)$. After deleting row 3 and column 3, look at the first row of what is left.
Sketch. Expanding along row 3 leaves $M_{\widehat{33}} = \left(\begin{smallmatrix} i & 0 & 0 \\ 1 & 3 & 1 \\ 2 & 1 & 1 \end{smallmatrix}\right)$ with sign $+$. Expanding that along its first row gives $i(3-1) = 2i$.
So $\det M = 2i$.

Row 3 of M is $(0, 0, 1, 0)$, so by Lemma 5.5.2,

$$\det M = (-1)^{3+3} \cdot 1 \cdot \det M_{\widehat{33}} = \det\begin{pmatrix} i & 0 & 0 \\ 1 & 3 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

The first row of this matrix is $(i, 0, 0)$, so expanding along it,

$$\det M = (-1)^{1+1} i \det\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} = i(3-1) = 2i.$$

Rows with a single nonzero entry are the best choice, since each expansion then has a single term.

Solution to Exercise 5.5.13

Uses. Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Theorem 5.5.9 (The adjugate and the inverse), Proposition 5.3.20 (Reduction modulo p)
Hint 1. Compute the cofactors over $\mathbb{Z}$ and reduce modulo 7 at the end. Remember to transpose.
Hint 2. $\det A = -19$ over $\mathbb{Z}$, which is 2 in $\mathbb{F}_7$, and $2^{-1} = 4$ in $\mathbb{F}_7$.
Sketch. Over $\mathbb{Z}$, $\operatorname{adj} A = \begin{pmatrix} -1 & -6 & 2 \\ -9 & 3 & -1 \\ 12 & -4 & -5 \end{pmatrix}$ and $\det A = -19$. In $\mathbb{F}_7$, $\det A = 2$ and $\operatorname{adj} A = \begin{pmatrix} 6 & 1 & 2 \\ 5 & 3 & 6 \\ 5 & 3 & 2 \end{pmatrix}$, so $A^{-1} = 4 \operatorname{adj} A = \begin{pmatrix} 3 & 4 & 1 \\ 6 & 5 & 3 \\ 6 & 5 & 1 \end{pmatrix}$.

We compute over $\mathbb{Z}$ and reduce at the end (Proposition 5.3.20). The cofactors are

$$\begin{aligned} c_{11} &= 1 \cdot 3 - 1 \cdot 4 = -1, & c_{12} &= -(3 \cdot 3 - 1 \cdot 0) = -9, & c_{13} &= 3 \cdot 4 - 1 \cdot 0 = 12, \\ c_{21} &= -(2 \cdot 3 - 0 \cdot 4) = -6, & c_{22} &= 1 \cdot 3 - 0 \cdot 0 = 3, & c_{23} &= -(1 \cdot 4 - 2 \cdot 0) = -4, \\ c_{31} &= 2 \cdot 1 - 0 \cdot 1 = 2, & c_{32} &= -(1 \cdot 1 - 0 \cdot 3) = -1, & c_{33} &= 1 \cdot 1 - 2 \cdot 3 = -5. \end{aligned}$$

Expanding along row 1, $\det A = 1 \cdot (-1) + 2 \cdot (-9) + 0 \cdot 12 = -19$, which is 2 in $\mathbb{F}_7$. Transposing C_A and reducing modulo 7,

$$\operatorname{adj} A = \begin{pmatrix} -1 & -6 & 2 \\ -9 & 3 & -1 \\ 12 & -4 & -5 \end{pmatrix} = \begin{pmatrix} 6 & 1 & 2 \\ 5 & 3 & 6 \\ 5 & 3 & 2 \end{pmatrix} \quad \text{in } \mathbb{F}_7^{3\times 3}.$$

Since $2 \cdot 4 = 8 = 1$ in $\mathbb{F}_7$, $(\det A)^{-1} = 4$, and by Theorem 5.5.9

$$A^{-1} = 4 \operatorname{adj} A = \begin{pmatrix} 24 & 4 & 8 \\ 20 & 12 & 24 \\ 20 & 12 & 8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 1 \\ 6 & 5 & 3 \\ 6 & 5 & 1 \end{pmatrix}.$$

Check: the first row of AA^{-1} is $(3 + 12, 4 + 10, 1 + 6) = (15, 14, 7) = (1, 0, 0)$, the second is $(9 + 6 + 6, 12 + 5 + 5, 3 + 3 + 1) = (21, 22, 7) = (0, 1, 0)$, and the third is $(24 + 18, 20 + 15, 12 + 3) = (42, 35, 15) = (0, 0, 1)$, all modulo 7.

Solution to Exercise 5.5.16

Uses. Theorem 5.5.14 (Cramer's rule), Lemma 5.5.2 (A row or column with one nonzero entry), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate)
Hint 1. Column j of B^{-1} is the solution of $B\mathbf{x} = \mathbf{e}_j$.
Hint 2. In $B_i(\mathbf{e}_j)$, column i has a single nonzero entry. Where is it?
Sketch. Column j of B^{-1} is $B^{-1}\mathbf{e}_j$, the solution of $B\mathbf{x} = \mathbf{e}_j$. Its i-th entry is $\det B_i(\mathbf{e}_j)/\det B$. Column i of $B_i(\mathbf{e}_j)$ has a single 1 in row j, and deleting row j and column i leaves $B_{\hat{j}\hat{i}}$. So $\det B_i(\mathbf{e}_j) = (-1)^{i+j} \det B_{\hat{j}\hat{i}} = c_{ji}$.

The (i, j) entry of B^{-1} is the i-th entry of its j-th column $B^{-1}\mathbf{e}_j$, which is the solution of $B\mathbf{x} = \mathbf{e}_j$. By Theorem 5.5.14 it equals $\det B_i(\mathbf{e}_j)/\det B$. Column i of $B_i(\mathbf{e}_j)$ is $\mathbf{e}_j$, whose only nonzero entry is the 1 in row j. Deleting row j and column i from $B_i(\mathbf{e}_j)$ removes the replaced column, so it leaves $B_{\hat{j}\hat{i}}$. By Lemma 5.5.2,

$$\det B_i(\mathbf{e}_j) = (-1)^{j+i} \det B_{\hat{j}\hat{i}} = c_{ji}.$$

Hence the (i, j) entry of B^{-1} is $c_{ji}/\det B$, the (i, j) entry of $(\det B)^{-1} C_B^{\mathsf{T}}$.

Solution to Exercise 5.5.26

Uses. Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.7 (Invertibility and the determinant), Theorem 3.4.24 (Polynomial interpolation)
Hint 1. Group the factors $j - i$ of the product by their larger index j.
Hint 2. For (b), the nodes $0, 1, 2, 3, 4$ are all the elements of $\mathbb{F}_5$. Reduce the integer from (a) modulo 5, or compute directly. Then think of g as a list of 5 values.
Sketch. (a) For fixed j, the factors $j - i$ with $1 \le i < j$ are $j - 1, j - 2, \ldots, 1$, with product $(j - 1)!$. So the determinant is $\prod_{j=2}^{n} (j - 1)! = 1! \cdots (n - 1)!$, which is 12 for $n = 4$. (b) The same grouping gives $1!2!3!4! = 288 = 3$ in $\mathbb{F}_5$, which is nonzero. So V is invertible, and for any values $g(0), \ldots, g(4)$ there is exactly one P of degree at most 4 taking them.

(a) With $c_k = k+1$, Theorem 5.5.21 gives $\det V(1,\dots,n) = \prod_{1\le i<j\le n}(j-i)$. For a fixed j, the factors with $i = 1,\dots,j-1$ are $j-1, j-2,\dots,1$, whose product is $(j-1)!$. So

$$\det V(1,\dots,n) = \prod_{j=2}^{n}(j-1)! = 1!\cdot 2!\cdots(n-1)!.$$

For $n=4$ this is $1\cdot 2\cdot 6 = 12$.

(b) By Theorem 5.5.21, $\det V(0,1,2,3,4) = \prod_{0\le i<j\le 4}(j-i)$, computed in $\mathbb{F}_5$. Grouping by j as in (a), this is $1!\cdot 2!\cdot 3!\cdot 4! = 1\cdot 2\cdot 6\cdot 24 = 288$, and $288 = 57\cdot 5 + 3$, so the determinant is $3 \neq 0$ in $\mathbb{F}_5$. (Each factor $j-i$ lies in $\{1,2,3,4\}$, so the product is nonzero by Proposition 1.1.17, without computing it.) So $V = V(0,1,2,3,4)$ is invertible by Theorem 5.4.7.

Given $g : \mathbb{F}_5 \to \mathbb{F}_5$, a polynomial $P = \sum_{k=0}^{4} a_k x^k$ satisfies $P(t) = g(t)$ for all $t \in \mathbb{F}_5$ if and only if $V\mathbf{a} = (g(0), g(1), g(2), g(3), g(4))$, where $\mathbf{a} = (a_0,\dots,a_4)$. Since V is invertible, this system has exactly one solution. (This is Theorem 3.4.24 with all five elements of $\mathbb{F}_5$ as nodes.)

Solution to Exercise 5.5.33

Uses. Lemma 5.5.2 (A row or column with one nonzero entry), Theorem 5.5.3 (Cofactor expansion), Proposition 5.3.20 (Reduction modulo p), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Row 2 has a single nonzero entry. After deleting its row and column, look for another row or column with a single nonzero entry.
Hint 2. For (b), reduce the answer to (a) modulo p.
Sketch. Row 2 gives $\det A = 2\det\begin{pmatrix}3&0&2\\1&5&1\\2&0&1\end{pmatrix}$, and column 2 of that minor gives $5\det\begin{pmatrix}3&2\\2&1\end{pmatrix} = -5$. So $\det A = -10$, and the reduction is invertible exactly for $p \notin \{2,5\}$.

(a) Row 2 is $(0,2,0,0)$, so by Lemma 5.5.2,

$$\det A = (-1)^{2+2}\cdot 2\det A_{\widehat{22}} = 2\det\begin{pmatrix}3&0&2\\1&5&1\\2&0&1\end{pmatrix}.$$

Column 2 of this minor is $(0,5,0)$, so, again by Lemma 5.5.2,

$$\det\begin{pmatrix}3&0&2\\1&5&1\\2&0&1\end{pmatrix} = (-1)^{2+2}\cdot 5\det\begin{pmatrix}3&2\\2&1\end{pmatrix} = 5(3-4) = -5.$$

Hence $\det A = -10$.

(b) By Proposition 5.3.20, the determinant of the reduction is $-10 \bmod p$, which is 0 exactly when p divides 10. By Theorem 5.4.7 over $\mathbb{F}_p$, the reduction is invertible exactly when $p \neq 2$ and $p \neq 5$.

Solution to Exercise 5.5.34

Uses. Lemma 5.5.2 (A row or column with one nonzero entry), Definition 3.3.31 (Triangular matrices), Theorem 5.4.14 (The transpose), induction on n
Techniques. Induction on matrix size (Technique 5.5.17)
Hint 1. For an upper triangular matrix, look at the first column.
Hint 2. Deleting row 1 and column 1 of an upper triangular matrix leaves an upper triangular matrix. Which diagonal entries does it have?
Sketch. If A is upper triangular, its first column is $(a_{11}, 0, \dots, 0)$, so $\det A = a_{11}\det A_{\widehat{11}}$, and $A_{\widehat{11}}$ is upper triangular with diagonal $a_{22},\dots,a_{nn}$. Induction finishes. A lower triangular matrix is the transpose of an upper triangular one.

We induct on n; for $n=1$ there is nothing to prove. Let $n \ge 2$ and let A be upper triangular, so $a_{ij} = 0$ whenever $i > j$. The first column of A is $(a_{11}, 0, \dots, 0)$, so by Lemma 5.5.2,

$$\det A = (-1)^{1+1}a_{11}\det A_{\widehat{11}}.$$

The (p,q) entry of $A_{\widehat{1}\widehat{1}}$ is $a_{p+1,q+1}$, which is 0 when $p > q$. So $A_{\widehat{1}\widehat{1}}$ is upper triangular with diagonal entries $a_{22}, \ldots, a_{nn}$, and by the induction hypothesis $\det A_{\widehat{1}\widehat{1}} = a_{22}\cdots a_{nn}$. Hence $\det A = a_{11}a_{22}\cdots a_{nn}$.

If A is lower triangular, then A^{T} is upper triangular with the same diagonal, and $\det A = \det A^{\mathsf{T}}$ by Theorem 5.4.14. (Alternatively, expand along the first row.)

Solution to Exercise 5.5.35

Uses. Theorem 5.5.3 (Cofactor expansion), Theorem 5.5.14 (Cramer's rule), Theorem 5.4.7 (Invertibility and the determinant), (5.5.3)
Hint 1. Expand $\det B(t)$ along the first row and factor. Check your factorization at $t = 1$.
Hint 2. For (b), $\det B_2(\mathbf{y}) = \det\begin{pmatrix} 1 & 1 & t \\ 1 & 2 & 1 \\ t & 3 & 1 \end{pmatrix}$; it factors as $-2(t-1)^2$.
Hint 3. For $t = 1$, every column of $B(1)$ is $(1,1,1)$. For $t = -2$, add the three equations.
Sketch. (a) Expanding along row 1, $\det B(t) = (t-1) - (1-t) + t\big(1-t^2\big) = -t^3 + 3t - 2 = -(t-1)^2(t+2)$, so the solution is unique exactly for $t \notin \{1, -2\}$. (b) $\det B_2(\mathbf{y}) = -2(t-1)^2$, so $x_2 = 2/(t+2)$. (c) At $t = 1$, each $B_i(\mathbf{y})$ keeps two columns $(1,1,1)$, so all three determinants vanish, but the equations say $x_1 + x_2 + x_3 = 1, 2, 3$ at once. At $t = -2$, $\det B_2(\mathbf{y}) = -18 \neq 0$, so there is no solution by (5.5.3).

(a) Expanding along the first row,

$$\begin{aligned} \det B(t) &= 1\cdot(t-1) - 1\cdot(1-t) + t\big(1-t^2\big) \\ &= -t^3 + 3t - 2 = -(t-1)^2(t+2), \end{aligned}$$

since $(t-1)^2(t+2) = \big(t^2 - 2t + 1\big)(t+2) = t^3 - 3t + 2$. By Theorem 5.4.7 and Theorem 5.5.14, the system has exactly one solution when $\det B(t) \neq 0$, that is, when $t \neq 1$ and $t \neq -2$. (If $\det B(t) = 0$, then $B(t)$ is not invertible, so the system has either no solution or more than one.)

(b) Replacing column 2 by $\mathbf{y}$ and expanding along the first row,

$$\begin{aligned} \det B_2(\mathbf{y}) = \det\begin{pmatrix} 1 & 1 & t \\ 1 & 2 & 1 \\ t & 3 & 1 \end{pmatrix} &= (2-3) - (1-t) + t(3-2t) \\ &= -2t^2 + 4t - 2 = -2(t-1)^2. \end{aligned}$$

By Theorem 5.5.14, for $t \notin \{1, -2\}$,

$$x_2 = \frac{-2(t-1)^2}{-(t-1)^2(t+2)} = \frac{2}{t+2}.$$

(c) For $t = 1$, every column of $B(1)$ is $(1,1,1)$. Replacing one column by $\mathbf{y}$ leaves two columns equal to $(1,1,1)$, so $\det B_i(\mathbf{y}) = 0$ for $i = 1, 2, 3$. But the three equations read $x_1 + x_2 + x_3 = 1$, $x_1 + x_2 + x_3 = 2$, and $x_1 + x_2 + x_3 = 3$, which contradict each other. So the vanishing of all the determinants in (5.5.3) does not imply that a solution exists.

For $t = -2$, part (b) gives $\det B_2(\mathbf{y}) = -2 \cdot 9 = -18 \neq 0$, while $\det B(-2) = 0$. By (5.5.3), a solution would satisfy $x_2 \cdot 0 = -18$, so there is none. Directly: the rows of $B(-2)$ add up to $\mathbf{0}^{\mathsf{T}}$, so adding the three equations gives $0 = 1 + 2 + 3$.

Solution to Exercise 5.5.36

Uses. Corollary 5.5.12 (Integer inverses), Theorem 5.5.9 (The adjugate and the inverse), Corollary 5.4.15 (Rows), Proposition 5.3.15 (Triangular matrices)
Hint 1. By Corollary 5.5.12 it is enough to have $\det A = \pm 1$.
Hint 2. Start from an upper triangular matrix with 1s on the diagonal, and add rows to other rows until no entry is 0.

Sketch. From $\begin{pmatrix}1&2&1\\0&1&1\\0&0&1\end{pmatrix}$, add row 1 to row 2 and then the new row 2 to row 3: $A = \begin{pmatrix}1&2&1\\1&3&2\\1&3&3\end{pmatrix}$, with $\det A = 1$. Its adjugate is $\begin{pmatrix}3&-3&1\\-1&2&-1\\0&-1&1\end{pmatrix}$, which is A^{-1}.

Start with $U = \begin{pmatrix}1&2&1\\0&1&1\\0&0&1\end{pmatrix}$, which has $\det U = 1$ by Proposition 5.3.15. Adding row 1 to row 2, and then the new row 2 to row 3, gives

$$A = \begin{pmatrix}1&2&1\\1&3&2\\1&3&3\end{pmatrix},$$

and these replacements do not change the determinant (Corollary 5.4.15), so $\det A = 1$. No entry of A is 0. The cofactors are

$$\begin{aligned} c_{11} &= 9-6 = 3, & c_{12} &= -(3-2) = -1, & c_{13} &= 3-3 = 0,\\ c_{21} &= -(6-3) = -3, & c_{22} &= 3-1 = 2, & c_{23} &= -(3-2) = -1,\\ c_{31} &= 4-3 = 1, & c_{32} &= -(2-1) = -1, & c_{33} &= 3-2 = 1. \end{aligned}$$

(As a check, $1\cdot 3 + 2\cdot(-1) + 1\cdot 0 = 1 = \det A$.) By Theorem 5.5.9,

$$A^{-1} = \operatorname{adj} A = C_A^{\mathsf{T}} = \begin{pmatrix}3&-3&1\\-1&2&-1\\0&-1&1\end{pmatrix}.$$

Indeed, the first row of AA^{-1} is $(3-2+0, -3+4-1, 1-2+1) = (1,0,0)$, and the other rows check the same way.

Solution to Exercise 5.5.37

Uses. Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Proposition 5.3.15 (Triangular matrices), Theorem 5.5.9 (The adjugate and the inverse), Definition 3.3.31 (Triangular matrices)
Hint 1. The (i,j) entry of $\operatorname{adj} B$ is $c_{ji} = (-1)^{i+j}\det B_{\hat{j}\hat{i}}$. For $i > j$, show that $B_{\hat{j}\hat{i}}$ is upper triangular with a zero on its diagonal.
Hint 2. Row p of $B_{\hat{j}\hat{i}}$ is row p or $p+1$ of B, and column q is column q or $q+1$. Look at the diagonal position $p = j$.
Sketch. For $j < i$, row p of $B_{\hat{j}\hat{i}}$ is row r_p of B and column q is column s_q, with $r_p = p$ for $p < j$, $r_p = p+1$ otherwise, $s_q = q$ for $q < i$, $s_q = q+1$ otherwise. For $p > q$ one checks $r_p > s_q$, so the minor matrix is upper triangular, and its (j,j) entry is $b_{j+1,j} = 0$. So $c_{ji} = 0$. The diagonal entry $c_{ii} = \det B_{\hat{i}\hat{i}}$ is the product of the other diagonal entries. Then $B^{-1} = (\det B)^{-1}\operatorname{adj} B$.

(a) The (i,j) entry of $\operatorname{adj} B$ is $c_{ji} = (-1)^{i+j}\det B_{\hat{j}\hat{i}}$. Let $i > j$, and let $N = B_{\hat{j}\hat{i}}$. For $p, q \in [n-1]$, the (p,q) entry of N is $b_{r_p s_q}$, where

$$r_p = \begin{cases} p & \text{if } p < j \\ p+1 & \text{if } p \geq j \end{cases}, \quad s_q = \begin{cases} q & \text{if } q < i \\ q+1 & \text{if } q \geq i \end{cases}.$$

Let $p > q$. Then $r_p \geq p \geq q+1 \geq s_q$. If $r_p = s_q$, all these are equalities, so $r_p = p$ and $s_q = q+1$, which means $p < j$ and $q \geq i$; but then $p = q+1 > i > j > p$, which is impossible. So $r_p > s_q$ and $b_{r_p s_q} = 0$, since B is upper triangular. Hence N is upper triangular. Its diagonal entry in position (j,j), which exists since $j < i \leq n$, is $b_{r_j s_j} = b_{j+1,j} = 0$ (as $j \geq j$ and $j < i$). By Proposition 5.3.15, $\det N = 0$, so the (i,j) entry of $\operatorname{adj} B$ is 0. Thus $\operatorname{adj} B$ is upper triangular.

The (i,i) entry of $\operatorname{adj} B$ is $c_{ii} = \det B_{\hat{i}\hat{i}}$. Deleting row i and column i of an upper triangular matrix leaves an upper triangular matrix whose diagonal entries are the b_{kk} with $k \neq i$ (the same index computation with $i = j$ gives $r_p = s_p$). So $c_{ii} = \prod_{k\neq i} b_{kk}$ by Proposition 5.3.15.

(b) If B is invertible, then $\det B = b_{11}\cdots b_{nn} \neq 0$, and $B^{-1} = (\det B)^{-1} \operatorname{adj} B$ by Theorem 5.5.9. This is upper triangular by (a), and its (i,i) entry is $\prod_{k\neq i} b_{kk} / \prod_k b_{kk} = b_{ii}^{-1}$.

Solution to Exercise 5.5.38

Uses. Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Theorem 5.4.14 (The transpose), Exercise 5.1.5, Example 5.5.7 (The 2×2 case)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a) and (b), compare the minors of A^{T} and of cA with those of A.
Hint 2. For $n = 2$, Example 5.5.7 gives adj explicitly; is it linear? For $n \geq 3$, compare $\operatorname{adj}(I_n + U_{11})$ with $\operatorname{adj} I_n + \operatorname{adj} U_{11}$ in (c), where U_{11} has a single entry 1, and use U_{11} again in (d).
Sketch. (a) True: $\left(A^{\mathsf{T}}\right)_{\hat{i}\hat{j}} = \left(A_{\hat{j}\hat{i}}\right)^{\mathsf{T}}$, so the cofactor matrix of A^{T} is C_A^{T}. (b) True: $(cA)_{\hat{i}\hat{j}} = cA_{\hat{i}\hat{j}}$ has determinant $c^{n-1}\det A_{\hat{i}\hat{j}}$. (c) True for $n = 2$, where $\operatorname{adj}\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = \left(\begin{smallmatrix} s & -q \\ -r & p \end{smallmatrix}\right)$ is linear; false for $n \geq 3$: the $(2,2)$ entry of $\operatorname{adj}(I_n + U_{11})$ is 2, but that of $\operatorname{adj} I_n + \operatorname{adj} U_{11} = I_n + 0$ is 1. (d) True for $n = 2$ (applying adj twice gives the identity); false for $n \geq 3$, since $\operatorname{adj} U_{11} = 0 = \operatorname{adj} 0$.

(a) True. Deleting row i and column j of A^{T} gives $\left(A_{\hat{j}\hat{i}}\right)^{\mathsf{T}}$, so by Theorem 5.4.14 the (i,j) cofactor of A^{T} is $(-1)^{i+j}\det A_{\hat{j}\hat{i}} = c_{ji}$, where c_{ji} is a cofactor of A. So $C_{A^{\mathsf{T}}} = C_A^{\mathsf{T}}$, and $\operatorname{adj}\left(A^{\mathsf{T}}\right) = C_{A^{\mathsf{T}}}^{\mathsf{T}} = C_A = (\operatorname{adj} A)^{\mathsf{T}}$.

(b) True. Deleting a row and a column of cA gives $cA_{\hat{i}\hat{j}}$, an $(n-1)\times(n-1)$ matrix, so its determinant is $c^{n-1}\det A_{\hat{i}\hat{j}}$ (Exercise 5.1.5). So every cofactor, and hence adj, is multiplied by c^{n-1}.

(c) True for $n = 2$ and false for $n \geq 3$. For $n = 2$, Example 5.5.7 shows $\operatorname{adj}\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = \left(\begin{smallmatrix} s & -q \\ -r & p \end{smallmatrix}\right)$, whose entries are linear in p, q, r, s, so $\operatorname{adj}(A+B) = \operatorname{adj} A + \operatorname{adj} B$. For $n \geq 3$, first note that for a diagonal matrix D with diagonal entries $d_1, \ldots, d_n$ and $i \neq j$, the matrix $D_{\hat{i}\hat{j}}$ has a zero row, namely what is left of row j of D after its only nonzero entry, in column j, is deleted. So $\operatorname{adj} D$ is diagonal, with (i,i) entry $\det D_{\hat{i}\hat{i}} = \prod_{k\neq i} d_k$. Now take $A = I_n$ and $B = U_{11}$, the matrix with a single entry 1 in position $(1,1)$. Then $A + B$ is diagonal with entries $2, 1, \ldots, 1$, so the $(2,2)$ entry of $\operatorname{adj}(A+B)$ is $2\cdot 1\cdots 1 = 2$. On the other hand, $\operatorname{adj} I_n = I_n$ and $\operatorname{adj} U_{11} = 0$ (see (d)), so the $(2,2)$ entry of $\operatorname{adj} A + \operatorname{adj} B$ is 1. Since $2 \neq 1$ in every field, $\operatorname{adj}(A+B) \neq \operatorname{adj} A + \operatorname{adj} B$.

(d) True for $n = 2$ and false for $n \geq 3$. For $n = 2$, applying the formula of Example 5.5.7 twice returns $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$, so adj has a left inverse (itself) and is injective. For $n \geq 3$, let U_{11} have a single entry 1 in position $(1,1)$. Every $(n-1)\times(n-1)$ submatrix $U_{11\hat{i}\hat{j}}$ has at most one nonzero entry, so it has a zero row (as $n - 1 \geq 2$) and determinant 0. So $\operatorname{adj} U_{11} = 0 = \operatorname{adj} 0$, with $U_{11} \neq 0$.

Solution to Exercise 5.5.39

Uses. Theorem 5.5.9 (The adjugate and the inverse), Theorem 5.4.10 (Multiplicativity), Exercise 5.1.5, Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Take determinants of $A \operatorname{adj} A = (\det A) I_n$. This settles the invertible case.
Hint 2. If $\det A = 0$, suppose $\operatorname{adj} A$ were invertible and multiply $A \operatorname{adj} A = 0$ on the right by $(\operatorname{adj} A)^{-1}$. What is the adjugate of the zero matrix?
Sketch. From $A \operatorname{adj} A = (\det A) I_n$, $\det A \det(\operatorname{adj} A) = (\det A)^n$, so we can divide when $\det A \neq 0$. If $\det A = 0$ and $\operatorname{adj} A$ were invertible, then $A = A \operatorname{adj} A (\operatorname{adj} A)^{-1} = 0$, so $\operatorname{adj} A = \operatorname{adj} 0 = 0$ (as $n \geq 2$), which is not invertible. So $\det(\operatorname{adj} A) = 0 = (\det A)^{n-1}$.

By Theorem 5.5.9, Theorem 5.4.10, and Exercise 5.1.5,

$$\det A \det(\operatorname{adj} A) = \det(A \operatorname{adj} A) = \det((\det A) I_n) = (\det A)^n.$$

If $\det A \neq 0$, dividing by $\det A$ gives $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

Suppose $\det A = 0$; we show $\det(\operatorname{adj} A) = 0$, which equals $(\det A)^{n-1}$ since $n - 1 \geq 1$. Suppose instead that $\operatorname{adj} A$ is invertible (Theorem 5.4.7). Since $A \operatorname{adj} A = 0 \cdot I_n = 0$,

$$A = A(\operatorname{adj} A)(\operatorname{adj} A)^{-1} = 0(\operatorname{adj} A)^{-1} = 0.$$

But every minor of the zero matrix is the determinant of a zero matrix of size $n-1 \geq 1$, hence 0, so $\operatorname{adj} A = \operatorname{adj} 0 = 0$, which is not invertible. This contradiction shows $\det(\operatorname{adj} A) = 0$.

Solution to Exercise 5.5.40

Uses. Theorem 5.5.9 (The adjugate and the inverse), Exercise 5.5.39, Proposition 5.3.19 ($\det(A+tI_n)$ as a function of t), Lemma 3.4.23 (Too many roots), Technique 5.3.18 (det is a polynomial in the entries)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. For invertible A, $\operatorname{adj} A = (\det A)A^{-1}$, and $\operatorname{adj} A$ is invertible too. Apply the same formula to $\operatorname{adj} A$, using Exercise 5.5.39.
Hint 2. For general A, replace A by $A+tI_n$. Each entry of both sides is a polynomial in t, and the two sides agree whenever $\det(A+tI_n) \neq 0$.
Hint 3. How many t can make $\det(A+tI_n) = 0$? If two polynomials of degree at most d agree at more than d points, they are equal. Then put $t = 0$.
Sketch. If A is invertible, $\operatorname{adj} A = (\det A)A^{-1}$ is invertible with determinant $(\det A)^{n-1}$, so $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-1}\big((\det A)A^{-1}\big)^{-1} = (\det A)^{n-2}A$. In general, every entry of $\operatorname{adj}(\operatorname{adj}(A+tI_n))$ and of $\det(A+tI_n)^{n-2}(A+tI_n)$ is a polynomial in t of bounded degree, and they agree for all t outside the at most n roots of $\det(A+tI_n)$. Over an infinite field that is more points than the degree, so they agree as polynomials, in particular at $t = 0$.

Invertible A. By Theorem 5.5.9, $\operatorname{adj} A = (\det A)A^{-1}$, which is invertible, and $\det(\operatorname{adj} A) = (\det A)^{n-1}$ by Exercise 5.5.39. Applying Theorem 5.5.9 to $\operatorname{adj} A$,

$$\begin{aligned}\operatorname{adj}(\operatorname{adj} A) &= \det(\operatorname{adj} A)(\operatorname{adj} A)^{-1} = (\det A)^{n-1}\big((\det A)A^{-1}\big)^{-1} \\ &= (\det A)^{n-1}(\det A)^{-1}A = (\det A)^{n-2}A.\end{aligned}$$

All A, F infinite. For $t \in F$, let $A_t = A + tI_n$. Every entry of A_t is a polynomial in t of degree at most 1. A determinant of size m whose entries are polynomials in t of degree at most d is a polynomial in t of degree at most md, by the Leibniz formula (Technique 5.3.18). So each entry of $\operatorname{adj} A_t$ is a polynomial in t of degree at most $n-1$, and each entry of $\operatorname{adj}(\operatorname{adj} A_t)$ is one of degree at most $(n-1)^2$. Also $\det A_t$ is a polynomial of degree at most n, and a product of polynomials in t of degrees at most d_1 and d_2 is a polynomial of degree at most $d_1 + d_2$ (expand by distributivity, as in the proof of Lemma 5.5.23). So each entry of $\det(A_t)^{n-2}A_t$ is a polynomial of degree at most $n(n-2)+1 = (n-1)^2$. Let $d = (n-1)^2$, which bounds both degrees.

Fix a position (i,j), and let P and Q be the polynomials of degree at most d that give the (i,j) entries of the two sides. By Proposition 5.3.19, $\det A_t$ is the value at t of a polynomial of degree n with leading coefficient 1, which is nonzero, so it has at most n roots (Lemma 3.4.23). For every other t, A_t is invertible, and the first part gives $P(t) = Q(t)$. Since F is infinite, there are more than $d+1$ such t, so $P - Q$ is a polynomial of degree at most d with more than $d+1$ roots, and $P = Q$ by Lemma 3.4.23. In particular $P(0) = Q(0)$. As this holds for every (i,j), $\operatorname{adj}(\operatorname{adj} A) = (\det A)^{n-2}A$.

Solution to Exercise 5.5.40 BEYOND 21-242

Uses. Exercise 5.5.51, Theorem 5.5.30 (Rank via minors), Example 5.5.7 (The 2×2 case)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. For $n = 2$, compute $\operatorname{adj}(\operatorname{adj} A)$ from Example 5.5.7.
Hint 2. For $n \geq 3$ and $\det A = 0$, the right side is 0. What is the rank of $\operatorname{adj} A$ (Exercise 5.5.51), and what does that say about its minors of order $n-1$?
Sketch. For $n = 2$, adj swaps the diagonal entries and negates the others, so applying it twice gives $A = (\det A)^0 A$. For $n \geq 3$ and A not invertible, $\operatorname{rank}(\operatorname{adj} A) \leq 1 \leq n-2$, so all minors of order $n-1$ of $\operatorname{adj} A$ vanish (Theorem 5.5.30), and $\operatorname{adj}(\operatorname{adj} A) = 0 = (\det A)^{n-2}A$. Invertible A was handled in the first approach.

For $n = 2$, Example 5.5.7 gives $\operatorname{adj}\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$, and applying it again gives $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$. So $\operatorname{adj}(\operatorname{adj} A) = A = (\det A)^0 A$. Let $n \geq 3$. Invertible A was handled in the first approach. If A is not invertible, then $(\det A)^{n-2}A = 0$, and by Exercise 5.5.51, $\operatorname{rank}(\operatorname{adj} A) \leq 1 \leq n-2$. By Theorem 5.5.30, every minor of order $n-1$ of $\operatorname{adj} A$ is 0 (and if $\operatorname{adj} A = 0$, trivially so). These minors are, up to sign, the entries of $\operatorname{adj}(\operatorname{adj} A)$, so $\operatorname{adj}(\operatorname{adj} A) = 0$.

Solution to Exercise 5.5.41

Uses. Theorem 5.5.9 (The adjugate and the inverse), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.19 ($\det(A+tI_n)$ as a function of t), Lemma 3.4.23 (Too many roots), Technique 5.3.18 (det is a polynomial in the entries)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. Prove it first for invertible A and B, using $\operatorname{adj} X = (\det X)X^{-1}$.
Hint 2. Then replace A and B by $A+tI_n$ and $B+tI_n$. For how many t can either fail to be invertible? Compare the entries of both sides as polynomials in t.
Sketch. For invertible A, B: $\operatorname{adj}(AB) = \det(AB)(AB)^{-1} = \big(\det BB^{-1}\big)\big(\det AA^{-1}\big) = (\operatorname{adj} B)(\operatorname{adj} A)$. In general, the entries of $\operatorname{adj}((A+tI)(B+tI))$ and of $\operatorname{adj}(B+tI)\operatorname{adj}(A+tI)$ are polynomials in t of degree at most $2(n-1)$, and they agree except at the at most $2n$ values of t where $A+tI$ or $B+tI$ is singular. An infinite field has more points than that, so the polynomials are equal; set $t=0$.

For $n=1$ both sides are (1), so let $n \geq 2$.

Invertible A and B. By Theorem 5.5.9 and Theorem 5.4.10, AB is invertible and

$$\operatorname{adj}(AB) = \det(AB)(AB)^{-1} = \det A \det B\, B^{-1}A^{-1} = \big(\det B\, B^{-1}\big)\big(\det A\, A^{-1}\big) = (\operatorname{adj} B)(\operatorname{adj} A).$$

General A and B. For $t \in F$, let $A_t = A+tI_n$ and $B_t = B+tI_n$. The entries of $A_tB_t = AB + t(A+B) + t^2I_n$ are polynomials in t of degree at most 2, so by the Leibniz formula (Technique 5.3.18) the entries of $\operatorname{adj}(A_tB_t)$, which are $\pm$ determinants of size $n-1$, are polynomials in t of degree at most $2(n-1)$. The entries of $\operatorname{adj} A_t$ and $\operatorname{adj} B_t$ have degree at most $n-1$, so the entries of $(\operatorname{adj} B_t)(\operatorname{adj} A_t)$, which are sums of products of two such, have degree at most $2(n-1)$ (a product of polynomials in t of degrees at most d_1 and d_2 is one of degree at most d_1+d_2, by distributivity).

Fix a position, and let $P, Q \in F[x]_{\leq 2(n-1)}$ give the corresponding entries of the two sides. By Proposition 5.3.19, $\det A_t$ and $\det B_t$ are values of polynomials of degree n with leading coefficient 1, so each vanishes for at most n values of t (Lemma 3.4.23). For all other t, A_t and B_t are invertible, and $P(t) = Q(t)$ by the first part. Since F is infinite, this happens for more than $2(n-1)+1$ values of t, so $P = Q$ by Lemma 3.4.23. Setting $t=0$ at every position gives $\operatorname{adj}(AB) = (\operatorname{adj} B)(\operatorname{adj} A)$.

Solution to Exercise 5.5.42

Uses. Lemma 5.5.23 (A polynomial with prescribed roots), Theorem 5.5.3 (Cofactor expansion), Corollary 5.4.15 (Rows), Lemma 5.5.2 (A row or column with one nonzero entry)
Techniques. Expose a factor with row and column operations (Technique 5.5.22), Polynomial identity arguments (Technique 5.5.24)
Hint 1. What is the coefficient of c^3 in $D(a,b,c)$? Compare with the right side. Then test $a=0, b=1, c=2$.
Hint 2. For (b), subtract row 1 from rows 2 and 3, and use $b^3-a^3 = (b-a)\big(b^2+ab+a^2\big)$.
Sketch. (a) Two polynomials with the same roots need not be equal: D has degree 3 in c with leading coefficient $b-a$, two known roots do not determine it, and the right side has degree 2. At $(0,1,2)$, $D=6$ but the right side is 2. (b) Row operations give $D = (b-a)(c-a)\det\begin{pmatrix} 1 & b^2+ab+a^2 \\ 1 & c^2+ca+a^2 \end{pmatrix} = (b-a)(c-a)(c-b)(a+b+c)$.

(a) The step "two polynomials with the same roots are equal" is false: x and $2x$ have the same roots. Lemma 5.5.23 says that a polynomial of degree at most n is determined by n distinct roots *together with* its coefficient of x^n. Here $n=3$, but only two roots are known, and the coefficients are not compared. In fact, expanding along the last row, the coefficient of c^3 in $D(a,b,c)$ is the cofactor $(-1)^{3+3}\det\begin{pmatrix} 1 & a \\ 1 & b \end{pmatrix} = b-a$, which is nonzero, while the right side has degree 2 in c. (The claim "the right side vanishes exactly at $c=a$ and $c=b$" does not repair this, since D may have a third root.) Concretely, for $a=0, b=1, c=2$,

$$D(0,1,2) = \det\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 8 \end{pmatrix} = 8-2 = 6,$$

while $(1-0)(2-0)(2-1) = 2$, in $\mathbb{Q}$. (Over some fields the claim happens to hold, for instance over $\mathbb{F}_2$, where three distinct elements do not exist; by (b), it fails exactly when there are distinct a,b,c with $a+b+c \neq 1$.)

(b) We show $D(a,b,c) = (a+b+c)(b-a)(c-a)(c-b)$ for all $a,b,c \in F$. Subtract row 1 from rows 2 and 3; this does not change the determinant (Corollary 5.4.15). Since $b^3 - a^3 = (b-a)\big(b^2+ab+a^2\big)$, the new rows 2 and 3 are

$$(b-a)\big(0,\,1,\,b^2+ab+a^2\big) \quad \text{and} \quad (c-a)\big(0,\,1,\,c^2+ca+a^2\big).$$

Pulling out the factors $b-a$ and $c-a$ (Corollary 5.4.15) and expanding along the first column, whose only nonzero entry is the 1 in row 1 (Lemma 5.5.2),

$$\begin{aligned} D(a,b,c) &= (b-a)(c-a)\det\begin{pmatrix} 1 & b^2+ab+a^2 \\ 1 & c^2+ca+a^2 \end{pmatrix} \\ &= (b-a)(c-a)\big(c^2-b^2+a(c-b)\big) = (b-a)(c-a)(c-b)(a+b+c). \end{aligned}$$

No division was used, so this holds in every field, including when a,b,c are not distinct. At $(0,1,2)$ it gives $3\cdot 2 = 6$, as computed in (a).

Solution to Exercise 5.5.43

Uses. Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.14 (The transpose), Corollary 5.4.8 (Vanishing determinants), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. A linear relation among infinite sequences gives in particular a relation among their first k entries. Which matrix has columns $\big(1,\lambda_i,\dots,\lambda_i^{k-1}\big)$?
Hint 2. For (b), differentiate a relation $\sum_i a_i e^{\lambda_i t} = 0$ repeatedly and set $t=0$.
Sketch. (a) If $\sum_i a_i \mathbf{s}_{\lambda_i} = \mathbf{0}$, the entries in positions $0,\dots,k-1$ give $V^{\mathsf T}\mathbf{a} = \mathbf{0}$ with $V = V(\lambda_1,\dots,\lambda_k)$. Since $\det V^{\mathsf T} = \det V = \prod_{i<j}(\lambda_j-\lambda_i) \neq 0$, $\mathbf{a} = \mathbf{0}$. (b) Differentiating m times and setting $t=0$ gives $\sum_i a_i \lambda_i^m = 0$ for $m = 0,\dots,k-1$, the same system.

Let $V = V(\lambda_1,\dots,\lambda_k) \in F^{k\times k}$, whose row i is $\big(1,\lambda_i,\dots,\lambda_i^{k-1}\big)$. The $(m+1)$-th row of $V^{\mathsf T}$ is $\big(\lambda_1^m,\dots,\lambda_k^m\big)$, so for $\mathbf{a} = (a_1,\dots,a_k)$, the system $V^{\mathsf T}\mathbf{a} = \mathbf{0}$ says that $\sum_i a_i\lambda_i^m = 0$ for $m = 0,\dots,k-1$. By Theorem 5.5.21 and Theorem 5.4.14, $\det V^{\mathsf T} = \prod_{i<j}(\lambda_j-\lambda_i) \neq 0$, so $V^{\mathsf T}$ is invertible and the only solution is $\mathbf{a} = \mathbf{0}$ (Corollary 5.4.8).

(a) If $\sum_i a_i \mathbf{s}_{\lambda_i} = \mathbf{0}$, then the entries in positions $m = 0,\dots,k-1$ give $\sum_i a_i\lambda_i^m = 0$. So $\mathbf{a} = \mathbf{0}$, and the sequences are independent by Theorem 1.5.5.

(b) If $\sum_i a_i e^{\lambda_i t} = 0$ for all $t \in \mathbb{R}$, then differentiating m times gives $\sum_i a_i\lambda_i^m e^{\lambda_i t} = 0$ for all t, and setting $t=0$ gives $\sum_i a_i\lambda_i^m = 0$, for $m = 0,\dots,k-1$. So $\mathbf{a} = \mathbf{0}$.

Solution to Exercise 5.5.44

Uses. Theorem 5.5.14 (Cramer's rule), Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.5.3 (Cofactor expansion), Definition 3.4.22 (Vandermonde matrix), Theorem 3.4.24 (Polynomial interpolation)
Hint 1. The coefficient vector $\mathbf{a}$ of p solves $V\mathbf{a} = \mathbf{d}$, and the coefficient of x^n is its last entry. Replace the last column of V by $\mathbf{d}$ and expand along that column.
Hint 2. Deleting row $i+1$ and the last column of V leaves $V(c_0,\dots,c_{i-1},c_{i+1},\dots,c_n)$. Divide its Vandermonde product into that of V: which factors are left over?
Hint 3. For (b), which polynomial of degree at most n takes the values c_i^m?
Sketch. By Cramer, $a_n = \det V_{n+1}(\mathbf{d})/\det V$. Expanding along the last column, $\det V_{n+1}(\mathbf{d}) = \sum_i (-1)^{i+n} d_i \det V_i$, where V_i omits c_i. The quotient $\det V/\det V_i$ is the product of the factors of $\det V$ involving c_i, namely $\prod_{j<i}(c_i-c_j)\prod_{k>i}(c_k-c_i) = (-1)^{n-i}\prod_{j\neq i}(c_i-c_j)$. The signs cancel. For (b), the interpolating polynomial of the values c_i^m is x^m.

(a) Write $p = a_0 + a_1x + \cdots + a_nx^n$. As in Section 3.4, $\mathbf{a} = (a_0,\dots,a_n)$ solves $V\mathbf{a} = \mathbf{d}$, where $V = V(c_0,\dots,c_n)$ and $\mathbf{d} = (d_0,\dots,d_n)$, and $\det V \neq 0$ by Theorem 5.5.21. The coefficient a_n is the entry of $\mathbf{a}$ in position $n+1$, so by Theorem 5.5.14, $a_n = \det V_{n+1}(\mathbf{d})/\det V$, where $V_{n+1}(\mathbf{d})$ is V with its last column replaced by $\mathbf{d}$.

Expanding along the last column, whose entry in row $i+1$ is d_i,

$$\det V_{n+1}(\mathbf{d}) = \sum_{i=0}^{n} (-1)^{(i+1)+(n+1)} d_i \det W_i,$$

where W_i is V with row $i+1$ and the last column deleted. Its rows are $\left(1, c_k, \ldots, c_k^{n-1}\right)$ for $k \neq i$, so $W_i = V(c_0, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n)$, and by Theorem 5.5.21 $\det W_i$ is the product of $c_l - c_k$ over all $k < l$ with $k, l \neq i$. The product for $\det V$ has, in addition, exactly the factors that involve i:

$$\det V = \det W_i \cdot \prod_{j<i}(c_i - c_j) \cdot \prod_{k>i}(c_k - c_i) = (-1)^{n-i} \det W_i \prod_{j \neq i}(c_i - c_j),$$

since there are $n - i$ indices $k > i$, and $c_k - c_i = -(c_i - c_k)$. Hence

$$a_n = \sum_{i=0}^{n} (-1)^{i+n} d_i \frac{\det W_i}{\det V} = \sum_{i=0}^{n} (-1)^{i+n}(-1)^{n-i} \frac{d_i}{\prod_{j \neq i}(c_i - c_j)} = \sum_{i=0}^{n} \frac{d_i}{\prod_{j \neq i}(c_i - c_j)},$$

since $(-1)^{i+n}(-1)^{n-i} = (-1)^{2n} = 1$.

(b) Let $0 \leq m \leq n$ and $d_i = c_i^m$. The polynomial x^m has degree at most n and takes the values c_i^m at the nodes, so by the uniqueness in Theorem 3.4.24 it is p. Its coefficient of x^n is 0 if $m < n$ and 1 if $m = n$, and by (a) this coefficient is $\sum_i c_i^m / \prod_{j \neq i}(c_i - c_j)$.

Solution to Exercise 5.5.45

Uses. Proposition 5.5.18 (Tridiagonal determinants), Lemma 5.5.23 (A polynomial with prescribed roots), Theorem 5.4.7 (Invertibility and the determinant), $\sin(a+b) + \sin(a-b) = 2\sin a \cos b$
Techniques. Induction on matrix size (Technique 5.5.17), Polynomial identity arguments (Technique 5.5.24)
Hint 1. (a) is Proposition 5.5.18 with $a_k = x$ and $b_k c_k = 1$. For the polynomial, induct with the recurrence.
Hint 2. (b) Induct on n. The identity $\sin((n+2)\theta) + \sin(n\theta) = 2\sin((n+1)\theta)\cos\theta$ is the recurrence in disguise.
Hint 3. (c) The numbers $2\cos(k\pi/(n+1))$, $k = 1, \ldots, n$, are distinct roots of P_n. How many roots can a polynomial of degree n with leading coefficient 1 have, given Lemma 5.5.23?
Sketch. (a) By Proposition 5.5.18, $d_n = x d_{n-1} - d_{n-2}$, and by induction $P_n = xP_{n-1} - P_{n-2}$ has degree n and leading coefficient 1. (b) $d_0 = 1$ and $d_1 = 2\cos\theta = \sin 2\theta / \sin\theta$; the sine identity turns the recurrence into the claim for $n+1$. (c) $\sin((n+1)\theta) = 0$ for $\theta = k\pi/(n+1)$, giving n distinct roots $2\cos(k\pi/(n+1))$ in $(-2, 2)$. By Lemma 5.5.23, $P_n(x) = \prod_k (x - 2\cos(k\pi/(n+1)))$ for all x, which vanishes only at these points.

(a) By Proposition 5.5.18 with $a_k = x$ and $b_k = c_k = 1$, $d_n(x) = x d_{n-1}(x) - d_{n-2}(x)$ for $n \geq 2$, with $d_0(x) = 1$ and $d_1(x) = x$. Define $P_0 = 1$, $P_1 = x$, and $P_n = xP_{n-1} - P_{n-2}$, where multiplying by x shifts the coefficients up by one place. By induction, $P_n \in \mathbb{R}[x]_{\leq n}$ has coefficient 1 of x^n (it comes from xP_{n-1} alone, as P_{n-2} has degree at most $n-2$), and $P_n(x) = d_n(x)$ for every x, since both sides satisfy the same recurrence with the same initial values.

(b) Let $x = 2\cos\theta$ with $0 < \theta < \pi$, so $\sin\theta \neq 0$, and let $e_n = \sin((n+1)\theta)/\sin\theta$. Then $e_0 = 1 = d_0$ and $e_1 = \sin 2\theta / \sin\theta = 2\cos\theta = d_1$. With $a = (n+1)\theta$ and $b = \theta$, the identity $\sin(a+b) + \sin(a-b) = 2\sin a \cos b$ gives

$$e_{n+1} + e_{n-1} = 2\cos\theta\, e_n = x e_n,$$

so (e_n) satisfies the same recurrence as $(d_n(x))$. By induction, $d_n(x) = e_n$ for all n.

(c) For $k = 1, \ldots, n$, let $\theta_k = k\pi/(n+1) \in (0, \pi)$ and $x_k = 2\cos\theta_k$. Since cos is strictly decreasing on $(0, \pi)$, the x_k are distinct. By (b), $d_n(x_k) = \sin(k\pi)/\sin\theta_k = 0$. So P_n is a polynomial of degree at most n with coefficient 1 of x^n and the n distinct roots $x_1, \ldots, x_n$, and Lemma 5.5.23 gives

$$d_n(x) = P_n(x) = (x - x_1)(x - x_2)\cdots(x - x_n) \quad \text{for every } x \in \mathbb{R}.$$

A product of real numbers is 0 only if a factor is, so $d_n(x) = 0$ exactly when x is one of the x_k, and by Theorem 5.4.7 these are exactly the x for which $M_n(x)$ is singular.

Solution to Exercise 5.5.46

Uses. Proposition 5.5.20 (Companion matrices), Theorem 5.4.7 (Invertibility and the determinant)
Hint 1. Use Proposition 5.5.20 with $a_0 = 2$, $a_1 = 1$, $a_2 = a_3 = 0$, and reduce $-a_k$ modulo 3.
Hint 2. Evaluate $t^4 + t + 2$ at $t = 0, 1, 2$.

Sketch. The matrix A of Proposition 5.5.20 has last column $-(2,1,0,0) = (1,2,0,0)$ in $\mathbb{F}_3$ and 1s just below the diagonal. The values $p(0) = 2$, $p(1) = 1$, $p(2) = 2$ are nonzero.

Let $p = 2 + x + x^4 \in \mathbb{F}_3[x]$, so $a_0 = 2$, $a_1 = 1$, $a_2 = a_3 = 0$. In $\mathbb{F}_3$, $-2 = 1$ and $-1 = 2$, so the matrix of Proposition 5.5.20 is

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

and $\det(tI_4 - A) = p(t)$ for every $t \in \mathbb{F}_3$. Now $p(0) = 2$, $p(1) = 1 + 1 + 2 = 4 = 1$, and $p(2) = 16 + 2 + 2 = 20 = 2$ in $\mathbb{F}_3$. None of these is 0, so $tI_4 - A$ is invertible for every $t \in \mathbb{F}_3$ by Theorem 5.4.7. (In Section 7.2 this says that A has no eigenvalue in $\mathbb{F}_3$.)

Solution to Exercise 5.5.47

Uses. Theorem 2.4.34 (Block multiplication), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.10 (Multiplicativity)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Multiply M on the left by $\left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$, and on the right by the same matrix.
Hint 2. For (b), apply (a) to $\mathbf{u} \in F^{n\times 1}$ and $\mathbf{v}^{\mathsf{T}} \in F^{1\times n}$, so that one of the two products is 1×1.
Sketch. With $P = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$, block multiplication gives $PM = \left(\begin{smallmatrix} I_m+AB & 0 \\ B & I_n \end{smallmatrix}\right)$ and $MP = \left(\begin{smallmatrix} I_m & 0 \\ B & I_n+BA \end{smallmatrix}\right)$. Since $\det P = 1$, both $\det(I_m + AB)$ and $\det(I_n + BA)$ equal $\det M$. For (b), $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is $n \times n$ and $\mathbf{v}^{\mathsf{T}}\mathbf{u}$ is 1×1; with $\mathbf{u} = \mathbf{v} = (1, \ldots, 1)$, $\det(I_n + J) = 1 + n$.

(a) Let $P = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$. By Theorem 2.4.34,

$$PM = \begin{pmatrix} I_m + AB & -A + A \\ B & I_n \end{pmatrix} = \begin{pmatrix} I_m + AB & 0 \\ B & I_n \end{pmatrix}, \quad MP = \begin{pmatrix} I_m & A - A \\ B & BA + I_n \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ B & I_n + BA \end{pmatrix}.$$

By Theorem 5.4.26, $\det P = \det I_m \det I_n = 1$, $\det(PM) = \det(I_m + AB)$, and $\det(MP) = \det(I_n + BA)$. By Theorem 5.4.10, both equal $\det P \det M = \det M$.

(b) Apply (a) with $A = \mathbf{u} \in F^{n\times 1}$ and $B = \mathbf{v}^{\mathsf{T}} \in F^{1\times n}$, so that $AB = \mathbf{u}\mathbf{v}^{\mathsf{T}} \in F^{n\times n}$ and $BA = \mathbf{v}^{\mathsf{T}}\mathbf{u} \in F^{1\times 1}$:

$$\det\left(I_n + \mathbf{u}\mathbf{v}^{\mathsf{T}}\right) = \det\left(1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}\right) = 1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}.$$

With $\mathbf{u} = \mathbf{v} = (1, \ldots, 1)$, we have $\mathbf{u}\mathbf{v}^{\mathsf{T}} = J$ and $\mathbf{v}^{\mathsf{T}}\mathbf{u} = n$, so $\det(I_n + J) = 1 + n$.

Solution to Exercise 5.5.48

Uses. Definition 5.1.1 (Determinant function), Lemma 5.5.8 (Replacing a row or a column), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Definition 2.4.26 (Trace), Definition 2.4.1 (Matrix product), Definition 5.3.6 (Determinant), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Column j of $A + tB$ is $\mathbf{a}_j + t\mathbf{b}_j$. Expand by multilinearity in every column, and group the terms by the number of columns in which $t\mathbf{b}_j$ was chosen.
Hint 2. The terms with exactly one such column j are $t \det A_j(\mathbf{b}_j)$. Evaluate $\det A_j(\mathbf{b}_j)$ with Lemma 5.5.8 and compare with the diagonal entries of $\operatorname{adj}(A)B$.
Sketch. Expanding, $\det(A + tB) = \sum_{S\subseteq[n]} t^{|S|} \det A_S$, where A_S takes $\mathbf{b}_j$ in the columns $j \in S$ and $\mathbf{a}_j$ elsewhere. So $e_k = \sum_{|S|=k} \det A_S$, with $e_0 = \det A$ and $e_n = \det B$. For $|S| = 1$, $\det A_j(\mathbf{b}_j) = \sum_i b_{ij}c_{ij}$, which is the (j,j) entry of $\operatorname{adj}(A)B$. For (b), take $B = I_n$: $\operatorname{tr}(\operatorname{adj} A) = \sum_i c_{ii}$. Expanding each $\det A_S$ by the Leibniz formula shows that these e_k are the coefficients of Proposition 5.3.19.

(a) Column j of $A + tB$ is $\mathbf{a}_j + t\mathbf{b}_j$. For $S \subseteq [n]$, let A_S be the matrix whose column j is $\mathbf{b}_j$ if $j \in S$ and $\mathbf{a}_j$ if $j \notin S$. Expanding by linearity in each column, one column at a time,

$$\det(A + tB) = \sum_{S\subseteq[n]} t^{|S|} \det A_S.$$

Grouping by $k = |S|$ gives the claim with $e_k = \sum_{|S|=k} \det A_S$, which does not depend on t. In particular $e_0 = \det A_\emptyset = \det A$ and $e_n = \det A_{[n]} = \det B$. The sets with one element give

$$e_1 = \sum_{j=1}^{n} \det A_j(\mathbf{b}_j) = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ij}c_{ij}$$

by part (b) of Lemma 5.5.8, where c_{ij} are the cofactors of A. The (j,j) entry of $\operatorname{adj}(A)B$ is $\sum_i (\operatorname{adj} A)_{ji} b_{ij} = \sum_i c_{ij}b_{ij}$, so $e_1 = \sum_j (\operatorname{adj}(A)B)_{jj} = \operatorname{tr}(\operatorname{adj}(A)B)$.

(b) With $B = I_n$, $e_1 = \operatorname{tr}(\operatorname{adj} A) = \sum_i c_{ii} = \sum_i \det A_{\hat{i}\hat{i}}$. It remains to see that these e_k are the coefficients c_k built in the proof of Proposition 5.3.19. (Comparing values of $t \mapsto \det(A + tI_n)$ is not enough, since over a small field a function does not determine its coefficients.) Now A_S has column $\mathbf{e}_j$ for $j \in S$, so in the Leibniz formula for $\det A_S$ the term of σ vanishes unless $\sigma(j) = j$ for all $j \in S$, that is, unless $S \subseteq \operatorname{Fix}(\sigma)$; in that case it is $\operatorname{sgn}(\sigma)\prod_{j\notin S} a_{\sigma(j)j}$. Writing $S' = \operatorname{Fix}(\sigma) \setminus S$,

$$\prod_{j\notin S} a_{\sigma(j)j} = \prod_{j\notin \operatorname{Fix}(\sigma)} a_{\sigma(j)j} \prod_{j\in S'} a_{jj}.$$

Summing over the S with $|S| = k$, that is, over the $S' \subseteq \operatorname{Fix}(\sigma)$ with $|S'| = |\operatorname{Fix}(\sigma)| - k$, gives exactly the coefficient $e_{\sigma,k}$ of that proof, and summing over σ gives $e_k = c_k$. (For $n = 2$, $\det A_{\hat{1}\hat{1}} + \det A_{\hat{2}\hat{2}} = a_{22} + a_{11} = \operatorname{tr} A$, in agreement with the coefficient of t^{n-1} found there.)

Solution to Exercise 5.5.49

Uses. (5.5.3), Theorem 5.5.14 (Cramer's rule), Corollary 5.5.12 (Integer inverses)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Recall (5.5.3): every solution satisfies $x_i \det B = \det B_i(\mathbf{y})$, whether or not B is invertible.
Hint 2. For (b), try a matrix all of whose columns are equal. For (c), try $n = 2$ and $B = 2I_2$.
Sketch. (a) True, by (5.5.3). (b) False: $B = J$ (all entries 1) with $n = 3$ and $\mathbf{y} = \mathbf{e}_1$; each $B_i(\mathbf{y})$ has two equal columns, but the equations contradict each other. (c) False: $2I_2\mathbf{x} = \mathbf{e}_1$ gives $x_1 = 1/2$. (d) True: Cramer gives $\mathbf{x} = \mathbf{0}$, so $\mathbf{y} = B\mathbf{0} = \mathbf{0}$.

(a) True. If $\mathbf{x}$ were a solution, then by (5.5.3), $\det B_i(\mathbf{y}) = x_i \det B = 0$, a contradiction.

(b) False. Let $n = 3$, let $B = J$ be the matrix with all entries 1, and let $\mathbf{y} = \mathbf{e}_1$. Replacing one column of J by $\mathbf{e}_1$ leaves two columns equal to $(1,1,1)$, so $\det B_i(\mathbf{y}) = 0$ for every i, and $\det J = 0$. But the equations $x_1 + x_2 + x_3 = 1$ and $x_1 + x_2 + x_3 = 0$ (rows 1 and 2) contradict each other, so there is no solution. (Exercise 5.5.35 gives another example.) For general $n \geq 2$, the same J and $\mathbf{y} = \mathbf{e}_1$ work when $n \geq 3$; for $n = 2$, take $B = 0$ and $\mathbf{y} = \mathbf{e}_1$.

(c) False. With $B = 2I_2$ and $\mathbf{y} = \mathbf{e}_1$, the solution is $\mathbf{x} = (1/2, 0)$. Integrality does hold when $\det B = \pm 1$, by Corollary 5.5.12.

(d) True. By Theorem 5.5.14, the unique solution has $x_i = \det B_i(\mathbf{y})/\det B = 0$ for every i, so $\mathbf{x} = \mathbf{0}$ and $\mathbf{y} = B\mathbf{0} = \mathbf{0}$.

Solution to Exercise 5.5.50

Uses. Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.10 (Multiplicativity), Exercise 5.1.5, Theorem 5.2.31 (Alternating functions and permutations), Definition 5.2.29 (Permutation matrix)
Hint 1. (a) List the six factors $c_j - c_i$ for $i < j$ with $c_0 = 1$, $c_1 = i$, $c_2 = -1$, $c_3 = -i$, and multiply them in pairs.
Hint 2. (b) The $(j+1, l+1)$ entry of V^2 is $\sum_{k=0}^{3} \omega^{k(j+l)}$. When is $\omega^{j+l} = 1$? If $z^4 = 1$ and $z \neq 1$, then $1 + z + z^2 + z^3 = 0$.
Sketch. (a) The factors are $i-1$, -2, $-i-1$, $-1-i$, $-2i$, $1-i$, with product $-16i$. (b) $\left(V^2\right)_{j+1,l+1} = \sum_k \omega^{k(j+l)}$ is 4 when $j + l \equiv 0 \bmod 4$ and 0 otherwise, so $V^2 = 4P$ with P the matrix of $j \mapsto -j \bmod 4$, a single transposition $(2\ 4)$ on the indices; $\det(V)^2 = 4^4(-1) = -256 = (-16i)^2$. (c) $(2-1)(4-1)(3-1)(4-2)(3-2)(3-4) = -12 = 3$.

(a) With $c_0 = 1$, $c_1 = i$, $c_2 = -1$, $c_3 = -i$, Theorem 5.5.21 gives the product of

$$c_1 - c_0 = i - 1, \quad c_2 - c_0 = -2, \quad c_3 - c_0 = -1 - i, \quad c_2 - c_1 = -1 - i, \quad c_3 - c_1 = -2i, \quad c_3 - c_2 = 1 - i.$$

Now $(i-1)(1-i) = -(1-i)^2 = -(-2i) = 2i$ and $(-1-i)^2 = (1+i)^2 = 2i$, so the product is

$$2i \cdot 2i \cdot (-2) \cdot (-2i) = (-4)(4i) = -16i.$$

(b) The $(j+1, l+1)$ entry of V^2 is $\sum_{k=0}^{3} \omega^{jk}\omega^{kl} = \sum_{k=0}^{3} z^k$ with $z = \omega^{j+l}$. If $j+l$ is divisible by 4, then $z = 1$ and the sum is 4. Otherwise $z \neq 1$ and $z^4 = \left(\omega^4\right)^{j+l} = 1$, so $(z-1)\left(1+z+z^2+z^3\right) = z^4 - 1 = 0$ gives $1 + z + z^2 + z^3 = 0$. So $V^2 = 4P$, where P has a 1 in position $(j+1, l+1)$ exactly when $l \equiv -j \bmod 4$. This is the permutation matrix E_π of the permutation π of $\{1,2,3,4\}$ that fixes 1 and 3 and swaps 2 and 4 (the indices $j+1$ for $j = 1, 3$), so $\det P = \operatorname{sgn}(\pi) = -1$ (Theorem 5.2.31). By Theorem 5.4.10 and Exercise 5.1.5,

$$(\det V)^2 = \det\left(V^2\right) = 4^4 \det P = -256.$$

Indeed $(-16i)^2 = 256i^2 = -256$. This determines $\det V$ up to sign, as $\pm 16i$.

(c) With $c_0, \ldots, c_3 = 1, 2, 4, 3$, the factors are $2-1=1$, $4-1=3$, $3-1=2$, $4-2=2$, $3-2=1$, and $3-4=-1$. Their product is $1 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot (-1) = -12 = 3$ in $\mathbb{F}_5$. (The argument of (b) works verbatim in $\mathbb{F}_5$ with $\omega = 2$ and gives $(\det V)^2 = 4^4 \cdot (-1) = -256 = 4$, and indeed $3^2 = 9 = 4$.)

Solution to Exercise 5.5.51

Uses. Theorem 5.5.9 (The adjugate and the inverse), Theorem 5.5.30 (Rank via minors), Theorem 4.2.11 (Rank–nullity), Theorem 4.2.7 (The image of T_A), Proposition 4.1.21 (Dimensions of subspaces)
Hint 1. The entries of $\operatorname{adj} A$ are, up to sign, the minors of order $n-1$. Use Theorem 5.5.30 for the last two cases.
Hint 2. If $\operatorname{rank} A = n-1$, then $A \operatorname{adj} A = 0$. Where do the columns of $\operatorname{adj} A$ lie, and what is the dimension of that subspace?
Sketch. If $\operatorname{rank} A = n$, $\operatorname{adj} A = (\det A)A^{-1}$ is invertible. If $\operatorname{rank} A \leq n-2$, all minors of order $n-1$ vanish (Theorem 5.5.30), so $\operatorname{adj} A = 0$. If $\operatorname{rank} A = n-1$, then $\det A = 0$, so $A \operatorname{adj} A = 0$ and $\operatorname{im} T_{\operatorname{adj} A} \subseteq \ker T_A$, which has dimension 1 by rank–nullity; and some minor of order $n-1$ is nonzero, so $\operatorname{adj} A \neq 0$.

If $\operatorname{rank} A = n$, then A is invertible and $\operatorname{adj} A = (\det A)A^{-1}$ is invertible (Theorem 5.5.9), so $\operatorname{rank}(\operatorname{adj} A) = n$.

The entries of $\operatorname{adj} A$ are $\pm \det A_{\hat{i}\hat{j}}$, the minors of order $n-1$ of A. If $\operatorname{rank} A \leq n-2$, these all vanish by Theorem 5.5.30 (if $A = 0$, trivially), so $\operatorname{adj} A = 0$ and $\operatorname{rank}(\operatorname{adj} A) = 0$.

Let $\operatorname{rank} A = n-1$. Then A is not invertible, so $\det A = 0$ and $A \operatorname{adj} A = 0$ by Theorem 5.5.9. So for every $\mathbf{x}$, $A(\operatorname{adj} A)\mathbf{x} = \mathbf{0}$, that is, $\operatorname{im} T_{\operatorname{adj} A} \subseteq \ker T_A$. By rank–nullity (Theorem 4.2.11), $\dim \ker T_A = n - (n-1) = 1$, so $\operatorname{rank}(\operatorname{adj} A) = \dim \operatorname{im} T_{\operatorname{adj} A} \leq 1$ (Theorem 4.2.7, Proposition 4.1.21). On the other hand, by Theorem 5.5.30, A has a nonzero minor of order $n-1$, so $\operatorname{adj} A \neq 0$ and $\operatorname{rank}(\operatorname{adj} A) \geq 1$. Hence $\operatorname{rank}(\operatorname{adj} A) = 1$.

Solution to Exercise 5.5.52

Uses. Definition 2.4.1 (Matrix product), Definition 5.1.1 (Determinant function), Theorem 5.2.31 (Alternating functions and permutations), Definition 5.3.6 (Determinant), Definition 5.5.29 (Minors of order k)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Column j of AB is $\sum_k b_{kj}\mathbf{a}_k$, where $\mathbf{a}_1, \ldots, \mathbf{a}_n \in F^m$ are the columns of A. Expand $\det(AB)$ by multilinearity in all m columns, as in Lemma 5.3.2.
Hint 2. The terms are indexed by functions $f : [m] \to [n]$. Which vanish? An injective f is determined by its image S and a permutation $\sigma \in S_m$.
Hint 3. For fixed $S = \{s_1 < \cdots < s_m\}$, collect the terms with image S and recognize the Leibniz formula for $B_{S[m]}$.
Sketch. Expanding, $\det(AB) = \sum_f \left(\prod_j b_{f(j)j}\right) \det\left(\mathbf{a}_{f(1)}, \ldots, \mathbf{a}_{f(m)}\right)$ over all $f : [m] \to [n]$. Non-injective f repeat a column. An injective f with image $S = \{s_1 < \cdots < s_m\}$ is $f(j) = s_{\sigma(j)}$ for a unique $\sigma \in S_m$, and $\det\left(\mathbf{a}_{s_{\sigma(1)}}, \ldots\right) = \operatorname{sgn}(\sigma) \det A_{[m]S}$. Summing over σ gives $\det A_{[m]S} \sum_\sigma \operatorname{sgn}(\sigma) \prod_j b_{s_{\sigma(j)}j} = \det A_{[m]S} \det B_{S[m]}$.

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n \in F^m$ be the columns of A. By Definition 2.4.1, column j of AB is $A\mathbf{b}_j = \sum_{k=1}^{n} b_{kj}\mathbf{a}_k$. Expanding $\det(AB)$ by linearity in each of its m columns (exactly as in the proof of Lemma 5.3.2, with $\mathbf{a}_k$ in place of $\mathbf{e}_k$),

$$\det(AB) = \sum_{f:[m]\to[n]} \left(\prod_{j=1}^{m} b_{f(j)j}\right) \det\left(\mathbf{a}_{f(1)}, \ldots, \mathbf{a}_{f(m)}\right),$$

the sum running over all functions $f : [m] \to [n]$. If f is not injective, the matrix $\left(\mathbf{a}_{f(1)}, \ldots, \mathbf{a}_{f(m)}\right)$ has two equal columns, so its determinant is 0.

An injective f has an image $S = f([m])$ with m elements; write $S = \{s_1 < s_2 < \cdots < s_m\}$. Then there is exactly one $\sigma \in S_m$ with $f(j) = s_{\sigma(j)}$ for all j, and conversely every pair (S, σ) gives an injective f in this way. The matrix with columns $\mathbf{a}_{s_1}, \ldots, \mathbf{a}_{s_m}$ is $A_{[m]S}$ (Definition 5.5.29), so by Theorem 5.2.31,

$$\det\left(\mathbf{a}_{s_{\sigma(1)}}, \ldots, \mathbf{a}_{s_{\sigma(m)}}\right) = \operatorname{sgn}(\sigma) \det A_{[m]S}.$$

Hence

$$\det(AB) = \sum_{|S|=m} \det A_{[m]S} \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \prod_{j=1}^{m} b_{s_{\sigma(j)} j}.$$

The (p, j) entry of $B_{S[m]}$ is $b_{s_p j}$, so the inner sum is the Leibniz formula (Definition 5.3.6) for $\det B_{S[m]}$. This proves the formula. If $m > n$, there are no injective f and no sets S, and both sides are 0.

Solution to Exercise 5.5.53

Uses. Corollary 5.4.15 (Rows), Lemma 5.1.8 (Column replacement), Theorem 5.4.26 (Block-triangular matrices), Theorem 5.4.7 (Invertibility and the determinant), induction on n

Techniques. Expose a factor with row and column operations (Technique 5.5.22), Induction on matrix size (Technique 5.5.17)

Hint 1. Subtract the last row from each other row. The difference $1/(x_i + y_j) - 1/(x_n + y_j)$ has numerator $x_n - x_i$, which does not depend on j.

Hint 2. Pull $x_n - x_i$ out of row i and $1/(x_n + y_j)$ out of column j; the last row becomes all 1s. Then subtract the last column from the others and do the same with the y's.

Hint 3. After both rounds, the matrix is block triangular with C_{n-1} in the corner and 1 in position (n, n).

Sketch. Subtracting row n from row $i < n$ gives entries $(x_n - x_i)/\big((x_i + y_j)(x_n + y_j)\big)$. Pulling out $x_n - x_i$ from rows $i < n$ and $1/(x_n + y_j)$ from every column makes row n all 1s and leaves $1/(x_i + y_j)$ elsewhere. Subtracting column n from columns $j < n$ and pulling out $y_n - y_j$ and $1/(x_i + y_n)$ leaves $\left(\begin{smallmatrix} C_{n-1} & 1 \\ 0 & 1 \end{smallmatrix}\right)$. So $\det C_n$ is $\det C_{n-1}$ times the factors involving index n, and induction finishes. For H_n, take $x_i = i$, $y_j = j - 1$; $\det H_3 = 1/2160$.

(a) We induct on n. For $n = 1$, both sides are $1/(x_1 + y_1)$. Let $n \geq 2$.

Rows. For each $i < n$, subtract row n from row i; the determinant does not change (Corollary 5.4.15). The new (i, j) entry, for $i < n$, is

$$\frac{1}{x_i + y_j} - \frac{1}{x_n + y_j} = \frac{x_n - x_i}{(x_i + y_j)(x_n + y_j)}.$$

Pull the factor $x_n - x_i$ out of row i, for each $i < n$, and then the factor $1/(x_n + y_j)$ out of column j, for each j (Corollary 5.4.15 and linearity in columns). The result is the matrix C' with entries $1/(x_i + y_j)$ for $i < n$ and with last row $(1, \ldots, 1)$, and

$$\det C_n = \frac{\prod_{i<n}(x_n - x_i)}{\prod_{j=1}^{n}(x_n + y_j)} \det C'.$$

Columns. For each $j < n$, subtract column n from column j of C' (Lemma 5.1.8). The last row becomes $(0, \ldots, 0, 1)$, and for $i < n$ the new (i, j) entry is

$$\frac{1}{x_i + y_j} - \frac{1}{x_i + y_n} = \frac{y_n - y_j}{(x_i + y_j)(x_i + y_n)}.$$

Pull $y_n - y_j$ out of column j, for each $j < n$, and $1/(x_i + y_n)$ out of row i, for each $i < n$. Row $i < n$ becomes $(1/(x_i + y_1), \ldots, 1/(x_i + y_{n-1}), 1)$, since its last entry $1/(x_i + y_n)$ is divided by $1/(x_i + y_n)$, and the last row stays $(0, \ldots, 0, 1)$. So the result is $\left(\begin{smallmatrix} C_{n-1} & \mathbf{1} \\ 0 & 1 \end{smallmatrix}\right)$, where C_{n-1} is built from $x_1, \ldots, x_{n-1}$ and $y_1, \ldots, y_{n-1}$, and its determinant is $\det C_{n-1}$ by Theorem 5.4.26. Altogether

$$\det C_n = \frac{\prod_{i<n}(x_n - x_i)\prod_{j<n}(y_n - y_j)}{\prod_{j=1}^{n}(x_n + y_j)\prod_{i<n}(x_i + y_n)} \det C_{n-1}.$$

The new numerator factors are those of the claimed numerator with larger index n, and the new denominator factors are the $x_i + y_j$ with $i = n$ or $j = n$. So the induction hypothesis for C_{n-1} gives the formula.

(b) With $x_i = i$ and $y_j = j - 1$, we have $x_i + y_j = i + j - 1 \geq 1$, so $C_n = H_n$. The x_i are distinct and the y_j are distinct, so every factor in the numerator of (a) is nonzero, and $\det H_n \neq 0$; H_n is invertible by Theorem 5.4.7. For $n = 3$, the numerator is $((2-1)(3-1)(3-2))^2 = 4$, and the denominator is the product of $i + j - 1$ over $i, j \in \{1, 2, 3\}$:

$$(1 \cdot 2 \cdot 3)(2 \cdot 3 \cdot 4)(3 \cdot 4 \cdot 5) = 6 \cdot 24 \cdot 60 = 8640.$$

So $\det H_3 = 4/8640 = 1/2160$.

Section 5.6

Solution to Exercise 5.6.6

Uses. Theorem 5.4.26 (Block-triangular matrices), Corollary 5.4.15 (Rows), Proposition 5.3.15 (Triangular matrices), Theorem 5.3.3 (The Leibniz formula), Theorem 5.2.31 (Alternating functions and permutations)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12), Expose a factor with row and column operations (Technique 5.5.22)
Hint 1. In (a), look at the upper right 2×3 corner. In (b), every row differs from row 1 in one entry. In (c), most columns have a single nonzero entry.
Hint 2. (a) is block lower triangular. In (b), subtract row 1 from the other rows. In (c), list the permutations that avoid the zeros, column by column.
Sketch. (a) Block lower triangular with diagonal blocks of determinants 3 and 25: $\det A = 75$. (b) Subtracting row 1 from rows 2, 3, 4 leaves an upper triangular matrix with diagonal $1, x-1, x-1, x-1$: $\det B(x) = (x-1)^3$. (c) Only $\sigma = (1\ 4)(2\ 3)$ survives, and it is even: $\det C = 1 \cdot 5 \cdot 2 \cdot 3 = 30$.

(a) The entries in rows 1, 2 and columns 3, 4, 5 are all 0, so $A = \begin{pmatrix} A_1 & 0 \\ X & A_2 \end{pmatrix}$ is block lower triangular, with

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix}.$$

By Theorem 5.4.26, $\det A = \det A_1 \det A_2$. Here $\det A_1 = 4 - 1 = 3$, and expanding A_2 along its first row, $\det A_2 = 1(1 - 0) - 2(0 - 12) + 0 = 25$. So $\det A = 75$. The block X never enters; elimination would have computed with it anyway.

(b) Subtract row 1 from each of rows 2, 3, 4. These replacements do not change the determinant (Corollary 5.4.15), and row i becomes $(x-1)\mathbf{e}_i^{\mathsf{T}}$. The result is upper triangular with diagonal entries $1, x-1, x-1, x-1$, so $\det B(x) = (x-1)^3$ by Proposition 5.3.15. The factor $x - 1$ was exposed by the replacements before anything was multiplied out.

(c) We find the permutations σ with $c_{\sigma(j)j} \neq 0$ for every j. Columns 1, 2, 3 have their only nonzero entries in rows 4, 3, 2, so $\sigma(1) = 4$, $\sigma(2) = 3$, $\sigma(3) = 2$. Column 4 has nonzero entries in rows 1 and 4, and row 4 is taken, so $\sigma(4) = 1$. Thus $\sigma = (1\ 4)(2\ 3)$, a product of two transpositions, which is even (Proposition 5.2.25). By Theorem 5.3.3,

$$\det C = c_{41}c_{32}c_{23}c_{14} = 1 \cdot 5 \cdot 2 \cdot 3 = 30.$$

The entry 7 belongs to no surviving term. Elimination would need swaps, and would carry the 7 along until it ended up above the diagonal.

Solution to Exercise 5.6.7

Uses. Corollary 5.4.15 (Rows), Lemma 5.1.8 (Column replacement), Proposition 5.3.17 (Block-diagonal matrices), Corollary 5.5.12 (Integer inverses), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate)
Techniques. Expose a factor with row and column operations (Technique 5.5.22), Induction on matrix size (Technique 5.5.17), Factor the matrix (Technique 5.6.1)
Hint 1. Subtract each row from the row below it, starting from the bottom. What does Pascal's rule do to the entries?
Hint 2. After the row operations, the first column is $\mathbf{e}_1$. Now subtract each column from the column to its right, starting from the right, and identify the lower right block.
Sketch. Replacing row i by row i minus row $i-1$ (for $i = n, \ldots, 2$) turns the entries into $\binom{i+j-3}{i-1}$, so column 1 becomes $\mathbf{e}_1$. Replacing column j by column j minus column $j-1$ (for $j = n, \ldots, 2$) turns row 1 into $\mathbf{e}_1^{\mathsf{T}}$ and the rest into $\binom{i+j-4}{i-2}$, which is P_{n-1}. So $\det P_n = \det P_{n-1} = \cdots = 1$, and $P_n^{-1} = \operatorname{adj} P_n$ is integral.

(a) We show $\det P_n = \det P_{n-1}$ for $n \geq 2$; since $P_1 = (1)$, induction on n then gives $\det P_n = 1$.

Rows. For $i = n, n-1, \ldots, 2$ in turn, subtract row $i-1$ from row i. Since we work from the bottom up, row $i-1$ is still the original row when it is used. For $i \geq 2$, Pascal's rule with $m = i+j-2 \geq 1$ and $k = i-1 \geq 1$ gives the new entry

$$\binom{i+j-2}{i-1} - \binom{i+j-3}{i-2} = \binom{i+j-3}{i-1}.$$

For $j = 1$ this is $\binom{i-2}{i-1} = 0$. So column 1 is now $\mathbf{e}_1$, and row 1 is still $(1, 1, \ldots, 1)$.

Columns. For $j = n, n-1, \ldots, 2$ in turn, subtract column $j-1$ from column j. In row 1 the new entries are $1 - 1 = 0$. In a row $i \geq 2$, Pascal's rule with $m = i+j-3 \geq 1$ and $k = i-1 \geq 1$ gives

$$\begin{aligned}\binom{i+j-3}{i-1} - \binom{i+j-4}{i-1} &= \binom{i+j-4}{i-2} \\ &= \binom{(i-1)+(j-1)-2}{(i-1)-1},\end{aligned}$$

the $(i-1, j-1)$ entry of P_{n-1}. Column 1 is unchanged.

The matrix is now $\begin{pmatrix} 1 & 0 \\ 0 & P_{n-1} \end{pmatrix}$, and all the operations were replacements, which do not change the determinant (Corollary 5.4.15, Lemma 5.1.8). By Proposition 5.3.17, $\det P_n = \det P_{n-1}$.

(b) Since $\det P_n = 1$ and P_n has integer entries, $P_n^{-1} = \operatorname{adj} P_n$ has integer entries (Corollary 5.5.12). For $P_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$ the cofactors are

$$\begin{aligned} c_{11} &= 3, \quad c_{12} = -3, c_{13} = 1, \\ c_{21} &= -3, c_{22} = 5, \quad c_{23} = -2, \\ c_{31} &= 1, \quad c_{32} = -2, c_{33} = 1, \end{aligned}$$

so

$$P_3^{-1} = \begin{pmatrix} 3 & -3 & 1 \\ -3 & 5 & -2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Multiplying out confirms $P_3 P_3^{-1} = I_3$.

Solution to Exercise 5.6.7

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Proposition 5.3.15 (Triangular matrices)
Techniques. Expose a factor with row and column operations (Technique 5.5.22), Induction on matrix size (Technique 5.5.17), Factor the matrix (Technique 5.6.1)
Hint 1. Look for a lower triangular L with $P_n = LL^{\mathsf{T}}$. Try L with entries $\binom{i-1}{k-1}$, the rows of Pascal's triangle.
Hint 2. The (i, j) entry of LL^{T} is $\sum_k \binom{i-1}{k-1}\binom{j-1}{k-1}$. Count the ways to choose $j-1$ objects from $i-1$ red and $j-1$ blue ones.

Sketch. With $l_{ik} = \binom{i-1}{k-1}$, L is lower triangular with 1s on the diagonal, and Vandermonde's identity gives $\left(LL^{\mathsf{T}}\right)_{ij} = \sum_k \binom{i-1}{k-1}\binom{j-1}{j-k} = \binom{i+j-2}{j-1} = p_{ij}$. So $\det P_n = (\det L)^2 = 1$.

Let $L \in \mathbb{Q}^{n\times n}$ have entries $l_{ik} = \binom{i-1}{k-1}$. Then $l_{ik} = 0$ for $k > i$ and $l_{ii} = 1$, so L is lower triangular with 1s on the diagonal, and $\det L = 1$ by Proposition 5.3.15. The terms with $k > \min(i,j)$ vanish, and for $k \le j$ we have $\binom{j-1}{k-1} = \binom{j-1}{j-k}$, so

$$\left(LL^{\mathsf{T}}\right)_{ij} = \sum_{k=1}^{\min(i,j)} \binom{i-1}{k-1}\binom{j-1}{j-k}.$$

The k-th term counts the ways to choose $j-1$ objects from a set of $i-1$ red and $j-1$ blue objects using exactly $k-1$ red and $j-k$ blue ones. Every choice has between 0 and $\min(i,j)-1$ red objects, so the sum is $\binom{i+j-2}{j-1} = \binom{i+j-2}{i-1} = p_{ij}$. Hence $P_n = LL^{\mathsf{T}}$, and $\det P_n = \det L \det L^{\mathsf{T}} = 1$ by Theorem 5.4.10 and Theorem 5.4.14.

The row operations of the first approach compute this factorization step by step. The second approach finds it at once, but only after guessing L.

Solution to Exercise 5.6.8

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Proposition 5.1.10 (Alternating implies antisymmetric), Definition 5.1.1 (Determinant function), Definition 2.4.26 (Trace)
Techniques. Recognize the determinant by its axioms (Technique 5.4.11), Expand by multilinearity (Technique 5.1.7)
Hint 1. Do not compute. Show that D is multilinear and alternating, and evaluate it at I_n.
Hint 2. For alternation, suppose columns k and l of X are equal. Only the summands $j = k$ and $j = l$ survive. How are their matrices related?
Hint 3. $D(I_n) = \sum_j \det(\mathbf{e}_1, \ldots, A\mathbf{e}_j, \ldots, \mathbf{e}_n)$. Expand $A\mathbf{e}_j = \sum_i a_{ij}\mathbf{e}_i$ in column j.
Sketch. Each summand is multilinear (in column j it is det composed with $\mathbf{x} \mapsto A\mathbf{x}$), so D is. If $\mathbf{x}_k = \mathbf{x}_l$ with $k \ne l$, the summands with $j \notin \{k,l\}$ have two equal columns, and the summands $j = k$ and $j = l$ differ by interchanging columns k and l, so they cancel. Hence $D = D(I_n)\det$, and $D(I_n) = \sum_j a_{jj} = \operatorname{tr} A$.

(a) Write $D = D_1 + \cdots + D_n$, where $D_j(X)$ is the j-th summand. We check the hypotheses of Theorem 5.3.4.

Multilinear. In column j, D_j is the composite of the linear map $\mathbf{x} \mapsto A\mathbf{x}$ with det as a function of column j, so it is linear in column j. In every other column D_j is linear because det is. So each D_j is multilinear, and so is their sum, since a sum of functions that are linear in a column is linear in that column.

Alternating. Suppose $\mathbf{x}_k = \mathbf{x}_l = \mathbf{x}$ for some $k < l$. If $j \notin \{k,l\}$, the matrix in $D_j(X)$ still has $\mathbf{x}$ in columns k and l, so $D_j(X) = 0$. The matrix in $D_k(X)$ has $A\mathbf{x}$ in column k and $\mathbf{x}$ in column l; the matrix in $D_l(X)$ has $\mathbf{x}$ in column k and $A\mathbf{x}$ in column l, and agrees with the first elsewhere. So it is the first matrix with columns k and l interchanged, and $D_l(X) = -D_k(X)$ by Proposition 5.1.10. Hence $D(X) = 0$.

By Theorem 5.3.4, $D(X) = D(I_n)\det X$. To find $D(I_n)$, note that $A\mathbf{e}_j = \sum_i a_{ij}\mathbf{e}_i$ and expand $D_j(I_n)$ in column j:

$$D_j(I_n) = \sum_{i=1}^{n} a_{ij} \det(\mathbf{e}_1, \ldots, \mathbf{e}_{j-1}, \mathbf{e}_i, \mathbf{e}_{j+1}, \ldots, \mathbf{e}_n).$$

For $i \ne j$ the vector $\mathbf{e}_i$ appears twice, so only $i = j$ survives, with value $a_{jj}\det I_n = a_{jj}$. So $D(I_n) = \sum_j a_{jj} = \operatorname{tr} A$, as desired.

(b) By Theorem 5.3.4, $E = E(I_n)\det$. Applying this to each summand and using (a),

$$\begin{aligned}\sum_j E(\mathbf{x}_1, \ldots, A\mathbf{x}_j, \ldots, \mathbf{x}_n) &= E(I_n)D(X) \\ &= E(I_n)(\operatorname{tr} A)\det X = (\operatorname{tr} A)E(X).\end{aligned}$$

The left side of (b) makes sense for any alternating form on a vector space, without a basis, so it can serve as a definition of the trace of an operator; compare Section 13.4. Over $\mathbb{R}$, $D(X)$ is the coefficient of t in

$\det(X + tAX) = \det(I_n + tA)\det X$, so (a) says that the trace is the rate at which $\det(I_n + tA)$ changes at $t = 0$ (Remark 5.5.28).

Solution to Exercise 5.6.9

Uses. Proposition 5.4.17 (Similar matrices have equal determinants), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Lemma 3.4.23 (Too many roots), Exercise 5.5.47, Exercise 5.1.5
Techniques. Polynomial identity arguments (Technique 5.5.24), det is a polynomial in the entries (Technique 5.3.18)
Hint 1. For (a), $BA = A^{-1}(AB)A$.
Hint 2. For (b), the entries of $tI_n - (A + sI_n)B$ have the form $\alpha + \beta s$. What does the Leibniz formula then say about its determinant? For which s does (a) apply to $A + sI_n$?
Hint 3. For (c), write $tI_n - AB = t\big(I_n - t^{-1}AB\big)$ and pull out t^n. The case $t = 0$ is multiplicativity.
Sketch. (a) $tI_n - BA = A^{-1}(tI_n - AB)A$ is similar to $tI_n - AB$. (b) Both determinants are sums of products of n affine functions of s, hence polynomial of degree at most n in s. By (a), $g(s) = 0$ whenever $A + sI_n$ is invertible, which fails for at most n values of s. An infinite field supplies $n + 1$ other values, so all $e_k = 0$ and $g(0) = 0$. (c) For $t \neq 0$, $\det(tI_n - AB) = t^n \det\big(I_n + \big(-t^{-1}A\big)B\big) = t^n \det\big(I_n + B\big(-t^{-1}A\big)\big) = \det(tI_n - BA)$. For $t = 0$, $\det(-AB) = (-1)^n \det A \det B = \det(-BA)$.

(a) Since A is invertible, $BA = A^{-1}(AB)A$, so

$$\begin{aligned} tI_n - BA &= A^{-1}(tI_n)A - A^{-1}(AB)A \\ &= A^{-1}(tI_n - AB)A. \end{aligned}$$

Similar matrices have equal determinants (Proposition 5.4.17), so $\det(tI_n - BA) = \det(tI_n - AB)$.

(b) Every entry of $tI_n - (A + sI_n)B = (tI_n - AB) - sB$ has the form $\alpha + \beta s$ with $\alpha, \beta \in F$ independent of s. Each term of the Leibniz formula is $\pm$ a product of n such entries, and multiplying out a product of n affine functions of s gives $c_0 + c_1 s + \cdots + c_n s^n$ for scalars c_k independent of s (as in the proof of Proposition 5.3.19). Summing over S_n, the first determinant has this form, and likewise the second, whose matrix $(tI_n - BA) - sB$ also has affine entries. Their difference g is therefore $e_0 + e_1 s + \cdots + e_n s^n$ for some $e_k \in F$.

By Proposition 5.3.19, $\det(A + sI_n)$ is the value at s of a polynomial of degree n with leading coefficient 1. That polynomial is not zero, so by Lemma 3.4.23 it has at most n roots, and $A + sI_n$ is invertible for all but at most n values of s. For each such s, (a) applied to $A + sI_n$ gives $g(s) = 0$.

Since F is infinite, there are $n + 1$ distinct values of s at which $A + sI_n$ is invertible. The polynomial $e_0 + e_1 x + \cdots + e_n x^n \in F[x]_{\leq n}$ vanishes at all of them, so it is zero by Lemma 3.4.23, and $g(s) = 0$ for every s. At $s = 0$ this is the identity for A.

(c) If $t \neq 0$, then $tI_n - AB = t\big(I_n + \big(-t^{-1}A\big)B\big)$, so by Exercise 5.1.5 and Exercise 5.5.47 (with $m = n$, and $-t^{-1}A$ in place of A),

$$\begin{aligned} \det(tI_n - AB) &= t^n \det\big(I_n + \big(-t^{-1}A\big)B\big) \\ &= t^n \det\big(I_n + B\big(-t^{-1}A\big)\big) = \det(tI_n - BA). \end{aligned}$$

If $t = 0$, then by Exercise 5.1.5 and Theorem 5.4.10, $\det(-AB) = (-1)^n \det A \det B = (-1)^n \det B \det A = \det(-BA)$.

The two proofs pay in different ways. The polynomial argument has to avoid the at most n values where $A + sI_n$ is singular and still find $n + 1$ others, so it needs more than $2n$ scalars, which an infinite field guarantees. Sylvester's identity holds over every field, but the reduction to it divides by t, so $t = 0$ is a separate case. In Section 7.3 the identity says that AB and BA have the same characteristic polynomial.

Solution to Exercise 5.6.10

Uses. Lemma 3.4.23 (Too many roots), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate)
Techniques. Polynomial identity arguments (Technique 5.5.24), Try small cases and small fields (Technique 1.2.38)
Hint 1. Over a finite field, how many values of t can be left after excluding up to n of them?

Hint 2. Polynomials of degree at most d that agree at $d+1$ distinct points are equal. What are the degrees here? For the example, try a diagonal matrix with diagonal entries 0 and 1.
Sketch. The last step needs more agreement points than the degree of the polynomials, which is up to $(n-1)^2$. Over a finite field, "all but at most n values" can be very few values, or none: for $A = \operatorname{diag}(0,1) \in \mathbb{F}_2^{2\times 2}$, both A and $A + I_2 = \operatorname{diag}(1,0)$ are singular. The argument is valid once $|F| \geq n^2 - n + 2$, in particular over every infinite field.

The error is the last step. Two polynomial functions of degree at most d are equal if they agree at $d+1$ distinct points (Lemma 3.4.23), and "all but finitely many points" supplies that many points only if the field has enough elements. Over a finite field it can supply none.

The example. Let $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$. Then $A + 0 \cdot I_2 = A$ and $A + 1 \cdot I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ are both singular, so there is no $t \in \mathbb{F}_2$ at which the two sides are known to agree, and the argument proves nothing about A.

When the argument works. The entries of $A + tI_n$ have degree at most 1 in t, so the cofactors of $A + tI_n$, which are determinants of size $n-1$, have degree at most $n-1$, and the entries of $\operatorname{adj}(\operatorname{adj}(A + tI_n))$ have degree at most $(n-1)^2$. On the other side, $\det(A + tI_n)^{n-2}$ has degree at most $n(n-2)$, so the entries of $\det(A + tI_n)^{n-2}(A + tI_n)$ have degree at most $n(n-2) + 1 = (n-1)^2$. The difference has degree at most $(n-1)^2$, and it vanishes at every t for which $A + tI_n$ is invertible, of which there are at least $|F| - n$ (Proposition 5.3.19). So the argument is correct when $|F| - n \geq (n-1)^2 + 1$, that is, $|F| \geq n^2 - n + 2$, and in particular over every infinite field.

Solution to Exercise 5.6.11

Uses. Proposition 5.4.1 (Determinants of elementary matrices), Corollary 5.4.4 (Products of elementary matrices), Exercise 5.4.48, Proposition 5.1.10 (Alternating implies antisymmetric)
Techniques. Reduce to elementary matrices (Technique 5.4.6)
Hint 1. One direction is multiplicativity. For the other, Exercise 5.4.48 writes every invertible matrix as replacements times one scaling.
Hint 2. For (b), the second proof of Proposition 5.1.10 performs a swap with three column replacements and one sign change.
Sketch. (a) Replacements have determinant 1 and E_{ic} has determinant $c > 0$, so such a product has positive determinant. Conversely, $A = E_1 \cdots E_t E_{n,\det A}$ with replacements E_k, and $E_{n,\det A}$ is a positive scaling when $\det A > 0$. (b) $E_{12} = E_{1,1,2}E_{2,-1,1}E_{1,1,2}E_{2,-1}$; a product of replacements and positive scalings has positive determinant, while $\det E_{12} = -1$.

(a) ($\Leftarrow$) By Proposition 5.4.1, a replacement has determinant 1 and E_{ic} has determinant c. So by Corollary 5.4.4, a product of replacements and scalings E_{ic} with $c > 0$ has determinant equal to the product of the scaling factors, which is positive.

($\Rightarrow$) By Exercise 5.4.48 (b), $A = E_1 \cdots E_t E_{n,\det A}$, where $E_1, \ldots, E_t$ are replacements. If $\det A > 0$, the last factor is a scaling by a positive number. (If $n = 1$, then $A = (\det A) = E_{1,\det A}$ directly.)

(b) In the second proof of Proposition 5.1.10, the columns $(\mathbf{x}, \mathbf{y})$ become $(\mathbf{x}+\mathbf{y}, \mathbf{y})$, then $(\mathbf{x}+\mathbf{y}, -\mathbf{x})$, then $(\mathbf{y}, -\mathbf{x})$, and finally $(\mathbf{y}, \mathbf{x})$ after scaling the second column by -1. Starting from I_2 and recording each column operation as a multiplication on the right (Proposition 3.2.28), this reads

$$\begin{aligned} E_{12} &= E_{1,1,2}E_{2,-1,1}E_{1,1,2}E_{2,-1} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned}$$

and multiplying out gives $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. A negative scaling is unavoidable: by (a), a product of replacements and positive scalings has positive determinant, while $\det E_{12} = -1$. Over $\mathbb{R}$, the sign of the determinant splits the invertible matrices into two classes, and the factors allowed in (a) never leave the class with positive determinant; Section 8.4 returns to this for rotations and reflections.

Solution to Exercise 5.6.12

Uses. Theorem 5.5.21 (The Vandermonde determinant), Exercise 5.5.26, Corollary 5.4.15 (Rows), Theorem 5.4.7 (Invertibility and the determinant), Corollary 5.4.8 (Vanishing determinants), Proposition 2.4.23 (Transpose of a product), Proposition 5.3.19 ($\det(A+tI_n)$ as a function of t), Lemma 3.4.23 (Too many roots)
Techniques. Reduce modulo p (Technique 5.3.21), det is a polynomial in the entries (Technique 5.3.18)
Hint 1. (a) Factor i out of row i. (b) Compute $\det M(x)$ as a polynomial in x and look at its parity. (c) Do not compute a determinant: look at the kernel. (d) How many roots can $\det(A+tI_n)$ have?
Hint 2. (b) The polynomial is x^3+x^2-2x-1, and $x^3+x^2=x^2(x+1)$ is even for integers. For real x, compare the values at 1 and 2. (c) If $\left(I_n+A^{\mathsf{T}}A\right)\mathbf{v}=\mathbf{0}$, multiply on the left by $\mathbf{v}^{\mathsf{T}}$.
Sketch. (a) Row i is $i\left(1,i,i^2,i^3\right)$, so $\det A = 4!\det V(1,2,3,4) = 24\cdot 12 = 288$. (b) $\det M(x)=x^3+x^2-2x-1$ is odd at every integer, but changes sign between 1 and 2. (c) $\mathbf{v}^{\mathsf{T}}\left(I_n+A^{\mathsf{T}}A\right)\mathbf{v}=\mathbf{v}^{\mathsf{T}}\mathbf{v}+(A\mathbf{v})^{\mathsf{T}}(A\mathbf{v})$ is a sum of squares, zero only for $\mathbf{v}=\mathbf{0}$. (d) $\det(A+tI_n)$ is a nonzero polynomial of degree n, so it cannot vanish at all of $0,1,\ldots,n$.

(a) Row i of A is $\left(i,i^2,i^3,i^4\right)=i\left(1,i,i^2,i^3\right)$, which is i times row i of the Vandermonde matrix $V(1,2,3,4)$. Pulling the factor i out of each row (Corollary 5.4.15) and using Exercise 5.5.26 (a),

$$\begin{aligned}\det A &= 1\cdot 2\cdot 3\cdot 4\cdot \det V(1,2,3,4)\\ &= 24\cdot(1!\cdot 2!\cdot 3!) = 288\neq 0.\end{aligned}$$

So A is invertible by Theorem 5.4.7. Recognizing the form made the computation unnecessary; even $\det V\neq 0$ follows at once from Theorem 5.5.21, since the nodes are distinct.

(b) Expanding along the first row,

$$\begin{aligned}\det M(x) &= x(x(x+1)-1)-1\cdot((x+1)-0)\\ &= x^3+x^2-2x-1.\end{aligned}$$

If x is an integer, $x^3+x^2=x^2(x+1)$ is a product with an even factor, and $2x$ is even, so $\det M(x)$ is odd, in particular nonzero, and $M(x)$ is invertible. (This is Technique 5.3.21 applied to a family: modulo 2 the determinant is 1 for every integer x.) On the other hand $\det M(1)=-1<0<7=\det M(2)$, and a polynomial function is continuous, so by the intermediate value theorem $\det M(x_0)=0$ for some real $x_0\in(1,2)$, and $M(x_0)$ is not invertible.

(c) Let $C=I_n+A^{\mathsf{T}}A$, and suppose $C\mathbf{v}=\mathbf{0}$. By Proposition 2.4.23, $\mathbf{v}^{\mathsf{T}}A^{\mathsf{T}}A\mathbf{v}=(A\mathbf{v})^{\mathsf{T}}(A\mathbf{v})$, so

$$\begin{aligned}0=\mathbf{v}^{\mathsf{T}}C\mathbf{v} &= \mathbf{v}^{\mathsf{T}}\mathbf{v}+(A\mathbf{v})^{\mathsf{T}}(A\mathbf{v})\\ &= \sum_{i=1}^{n}v_i^2+\sum_{k=1}^{m}(A\mathbf{v})_k^2.\end{aligned}$$

A sum of squares of real numbers is 0 only if every term is 0, so $\mathbf{v}=\mathbf{0}$. By Corollary 5.4.8, C is invertible. A determinant would be hopeless here: $\det C$ is a polynomial of degree up to $2n$ in the entries of A, with no evident sign.

(d) By Proposition 5.3.19, $\det(A+tI_n)=p(t)$ for the polynomial $p=x^n+c_{n-1}x^{n-1}+\cdots+c_0\in\mathbb{Q}[x]$. Since its coefficient of x^n is 1, p is not zero, so by Lemma 3.4.23 it does not vanish at all $n+1$ points $0,1,\ldots,n$. For a point t where it does not vanish, $A+tI_n$ is invertible by Theorem 5.4.7.

Solution to Exercise 5.6.13

Uses. Theorem 5.3.3 (The Leibniz formula), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.2.16 (Cycle decomposition), Proposition 5.2.8 (Reindexing sums over S_n)
Techniques. Cancel terms in pairs (Technique 5.3.5), Read structure off the Leibniz expansion (Technique 5.3.12)
Hint 1. In $t_{\sigma^{-1}}(A)$, reindex the product by $j=\sigma(i)$, and then use $a_{i\sigma(i)}=-a_{\sigma(i)i}$.
Hint 2. If $\sigma^2=\mathrm{id}$, every orbit of σ has one or two elements. Since n is odd, they cannot all have two.
Hint 3. Pair σ with σ^{-1}. The permutations equal to their own inverse are left over, and (b) disposes of them.
Sketch. (a) $\prod_j a_{\sigma^{-1}(j)j}=\prod_i a_{i\sigma(i)}=(-1)^n\prod_i a_{\sigma(i)i}$ and $\operatorname{sgn}\left(\sigma^{-1}\right)=\operatorname{sgn}(\sigma)$. (b) Orbits of an involution have size 1 or 2 and partition a set of odd size, so some j is fixed, and $a_{jj}=0$ is a factor. (c) $\sigma\mapsto\sigma^{-1}$ is an involution of S_n whose fixed points are the σ with $\sigma^2=\mathrm{id}$; pairs cancel by (a), and fixed points contribute 0 by (b). (d) $\det A=-\det A$ only gives $(1+1)\det A=0$; pairing never divides.

(a) By definition, $t_{\sigma^{-1}}(A) = \operatorname{sgn}(\sigma^{-1}) \prod_{j=1}^{n} a_{\sigma^{-1}(j)j}$. As i runs over $[n]$, so does $j = \sigma(i)$, so the product can be indexed by i instead:

$$\begin{aligned}\prod_{j=1}^{n} a_{\sigma^{-1}(j)j} &= \prod_{i=1}^{n} a_{i\sigma(i)} \\ &= \prod_{i=1}^{n} \bigl(-a_{\sigma(i)i}\bigr) = (-1)^n \prod_{i=1}^{n} a_{\sigma(i)i}.\end{aligned}$$

Since $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ (Proposition 5.2.25) and n is odd, $t_{\sigma^{-1}}(A) = -t_\sigma(A)$.

(b) Suppose $\sigma^2 = \mathrm{id}$, and let $i \in [n]$. The least $k \geq 1$ with $\sigma^k(i) = i$ is 1 or 2, so by Proposition 5.2.16 the orbit of i has one or two elements, and the orbits partition $[n]$. If every orbit had two elements, n would be even. So some orbit is $\{j\}$, that is, $\sigma(j) = j$, and $t_\sigma(A)$ contains the factor $a_{\sigma(j)j} = a_{jj} = 0$.

(c) The map $\varphi(\sigma) = \sigma^{-1}$ is a bijection of S_n with $\varphi(\varphi(\sigma)) = \sigma$, and its fixed points are the σ with $\sigma = \sigma^{-1}$, that is, $\sigma^2 = \mathrm{id}$. Let Q be the set of fixed points, and let P contain exactly one element of each pair $\{\sigma, \sigma^{-1}\}$ with $\sigma \neq \sigma^{-1}$. Every permutation lies in exactly one of Q, P, and $\varphi(P)$, so by Theorem 5.3.3, (a), and (b),

$$\begin{aligned}\det A &= \sum_{\sigma \in P} \bigl(t_\sigma(A) + t_{\sigma^{-1}}(A)\bigr) + \sum_{\sigma \in Q} t_\sigma(A) \\ &= 0 + 0 = 0.\end{aligned}$$

This is Technique 5.3.5 in its version with fixed points.

(d) The earlier argument shows $\det A = \det A^{\mathsf{T}} = \det(-A) = (-1)^n \det A = -\det A$, that is, $(1+1)\det A = 0$, and concludes by dividing by $1+1$. Over $\mathbb{F}_2$ this equation reads $0 = 0$. The pairing argument never divides: over $\mathbb{F}_2$, (a) says that $t_{\sigma^{-1}}(A) = t_\sigma(A)$, and the two terms of each pair cancel because $x + x = 0$ there. The hypothesis $a_{ii} = 0$ enters only in (b). When $1 + 1 \neq 0$ it follows from $a_{ii} = -a_{ii}$, but over $\mathbb{F}_2$ it is a genuine assumption: I_3 satisfies $a_{ji} = -a_{ij}$ over $\mathbb{F}_2$ and has determinant 1 (Exercise 5.4.33 (c)).

Solution to Exercise 5.6.14

Uses. Definition 5.2.21 (Inversion), Corollary 5.2.24 (Formulas for the sign), Proposition 5.2.26 (Sign of a cycle), Proposition 5.2.30 (Algebra of permutation matrices), Exercise 5.2.36, Corollary 5.4.27 (Several diagonal blocks), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Theorem 5.2.31 (Alternating functions and permutations)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Sort the pairs of positions $p < q$ by the parities of p and q. Only one kind of pair can be inverted.
Hint 2. For (c), rearrange the rows *and* the columns of M by π: consider $E_\pi^{\mathsf{T}} M E_\pi$, whose (p,q) entry is $m_{\pi(p)\pi(q)}$.
Sketch. (a) The odd positions carry $1, \ldots, n$ in order and the even positions $n+1, \ldots, 2n$ in order. A pair (position $2i$, position $2j-1$) with $j > i$ is inverted, since $n + i > j$, and no other pair is; there are $\sum_i (n-i) = n(n-1)/2$ of them. (b) $n = 3$: (2 4 5 3), sign -1; $n = 4$: (2 5 3)(4 6 7), sign 1. (c) $E_\pi^{\mathsf{T}} M E_\pi$ is block diagonal with blocks $\left(\begin{smallmatrix} a_i & b_i \\ c_i & d_i \end{smallmatrix}\right)$, and $\det\bigl(E_\pi^{\mathsf{T}} M E_\pi\bigr) = \operatorname{sgn}(\pi)^2 \det M = \det M$.

(a) The entry in position $2i-1$ is i, and the entry in position $2i$ is $n+i$. Consider positions $p < q$. If both are odd, or both even, their entries are in increasing order. If $p = 2i-1$ and $q = 2j$, then $i \leq j < n + j$, so the pair is in order. If $p = 2i$ and $q = 2j-1$, then $q > p$ means $j > i$, and the entries satisfy $n + i > n \geq j$, so the pair is an inversion. For fixed i there are $n - i$ such j, so

$$\operatorname{inv}(\pi) = \sum_{i=1}^{n} (n-i) = n\frac{n-1}{2},$$

and $\operatorname{sgn}(\pi) = (-1)^{n(n-1)/2}$ by Corollary 5.2.24.

(b) For $n = 3$ the bottom row is $1, 4, 2, 5, 3, 6$. Following $2 \mapsto 4 \mapsto 5 \mapsto 3 \mapsto 2$, and noting that 1 and 6 are fixed, $\pi = (2\ 4\ 5\ 3)$, a 4-cycle, with sign $-1 = (-1)^3$. For $n = 4$ the bottom row is $1, 5, 2, 6, 3, 7, 4, 8$, so $\pi = (2\ 5\ 3)(4\ 6\ 7)$, with sign $1 = (-1)^6$. Both agree with (a) by Proposition 5.2.26. The cycle type changes

irregularly with n, while the inversion count has a closed form, so inversions are the right tool for this permutation.

(c) Let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and $N = E_\pi^{\mathsf{T}} M E_\pi$. By Proposition 5.2.30 (a), column q of ME_π is column $\pi(q)$ of M. By Proposition 5.2.30 (c), $E_\pi^{\mathsf{T}} = E_{\pi^{-1}}$, and by Exercise 5.2.36 (a), row p of $E_{\pi^{-1}} X$ is row $\pi(p)$ of X. So $n_{pq} = m_{\pi(p)\pi(q)}$.

For $p, q \in \{2i-1, 2i\}$ the indices $\pi(p), \pi(q)$ lie in $\{i, n+i\}$, and these four entries of N form the block $\begin{pmatrix} m_{ii} & m_{i,n+i} \\ m_{n+i,i} & m_{n+i,n+i} \end{pmatrix} = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$. If instead $p \in \{2i-1, 2i\}$ and $q \in \{2j-1, 2j\}$ with $i \neq j$, then $m_{\pi(p)\pi(q)}$ is an off-diagonal entry of one of the diagonal matrices A, B, C, D, so it is 0. Hence N is block diagonal with the 2×2 blocks $\begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$, and by Corollary 5.4.27,

$$\det N = \prod_{i=1}^{n} (a_i d_i - b_i c_i).$$

Finally, $\det N = \det E_\pi^{\mathsf{T}} \det M \det E_\pi = \operatorname{sgn}(\pi)^2 \det M = \det M$ by Theorem 5.4.10, Theorem 5.4.14, and Theorem 5.2.31. So the sign does not enter: rearranging the rows and the columns in the same way costs $\operatorname{sgn}(\pi)^2 = 1$. Only rearranging the columns alone, as in $\det(ME_\pi) = \operatorname{sgn}(\pi) \det M$, costs a sign.

Solution to Exercise 5.6.15

Uses. Lemma 5.5.2 (A row or column with one nonzero entry), Theorem 5.2.31 (Alternating functions and permutations), Theorem 5.4.26 (Block-triangular matrices)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. Carry out the interchange of rows 1 and 3 for the given B, and compare the lower right 2×2 block with $B_{\widehat{31}}$.
Hint 2. After interchanging rows 1 and i, in which position is the old row 1? Deleting row i from B keeps the other rows in their original order.
Sketch. The lower right block is not $B_{\widehat{ij}}$: after interchanging rows 1 and i, the old row 1 sits in position i, behind rows $2, \ldots, i-1$. For the given B, the claim predicts $\det B = -1 \cdot \det I_2 = -1$, but B is the permutation matrix of the 3-cycle $(1\ 3\ 2)$, so $\det B = 1$, as $(-1)^{3+1}$ predicts. Moving row i to the top by $i-1$ adjacent interchanges keeps the other rows in order and gives the correct sign $(-1)^{i+j}$.

The error. The last step identifies the lower right block with $B_{\widehat{ij}}$. Interchanging rows 1 and i moves the old row 1 into position i, so after the first row is removed, the remaining rows appear in the order $2, 3, \ldots, i-1, 1, i+1, \ldots, n$. The minor $B_{\widehat{ij}}$ has them in the order $1, 2, \ldots, i-1, i+1, \ldots, n$. The two agree only when $i \leq 2$, and the same problem occurs with the columns when $j \geq 3$.

The counterexample. For the given B, row 3 is $(1, 0, 0)$, so $b_{31} = 1$ is the only nonzero entry, and $B_{\widehat{31}}$, obtained by deleting row 3 and column 1, is I_2. The claim, with $\varepsilon = -1$, predicts $\det B = -1$. But the columns of B are $\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2$, so $B = E_\sigma$ for the 3-cycle $\sigma = (1\ 3\ 2)$, which is even, and $\det B = 1$ (Theorem 5.2.31). In the argument, interchanging rows 1 and 3 gives $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, whose lower right block $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is $B_{\widehat{31}}$ with its rows interchanged.

The repair. Move row i to the top by $i-1$ interchanges of adjacent rows, which keeps the other rows in order, and likewise for column j. This gives the sign $(-1)^{(i-1)+(j-1)} = (-1)^{i+j}$ of Lemma 5.5.2, which for this B is $(-1)^4 = 1$.

Problems for Chapter 5

Solution to Problem 5.1

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Theorem 5.3.3 (The Leibniz formula), Proposition 5.2.25 (Multiplicativity of the sign), Proposition 5.1.11 (Antisymmetric implies alternating when $1 + 1 \neq 0$), Definition 5.1.1 (Determinant function)
Techniques. Cancel terms in pairs (Technique 5.3.5)

Hint 1. For (b), suppose columns $k \neq l$ are equal and let $\tau = (k\ l)$. Compare the terms of σ and $\sigma\tau$, and do not reindex the whole sum.
Hint 2. For (c), the product $a_{11}\cdots a_{1n}$ takes one entry from each column, and permuting the columns only permutes its factors. To break alternation, use a matrix whose columns are all equal but whose first row is not zero.
Sketch. (a) There is exactly one determinant function on $F^{n\times n}$, the Leibniz formula L, and every multilinear alternating D equals $D(I_n)L$. (b) If columns $k \neq l$ agree and $\tau = (k\ l)$, then $\sigma\tau$ uses the same entries as σ with the opposite sign, so $t_{\sigma\tau} = -t_\sigma$; grouping each even σ with $\sigma\tau$ writes $L(A)$ as a sum of zeros. (c) det is antisymmetric, and the product of the first-row entries is multilinear, invariant under column permutations, and 0 at I_n; over $\mathbb{F}_2$, invariance is antisymmetry. On the matrix with first row all 1s and other rows 0, $D = 0 + 1$. When $1 + 1 \neq 0$, antisymmetry implies alternation, and uniqueness leaves only det.

(a) *Statement* (Theorem 5.3.4). Let $n \geq 1$. The function

$$L(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

is a determinant function on $F^{n\times n}$, and it is the only one. Moreover, every multilinear alternating $D : F^{n\times n} \to F$ satisfies $D = D(I_n)L$.

(b) Let A have equal columns k and l, where $k \neq l$, and let $\tau = (k\ l)$. Write $t_\sigma(A) = \operatorname{sgn}(\sigma)\prod_j a_{\sigma(j)j}$. The permutation $\sigma\tau$ agrees with σ except that $(\sigma\tau)(k) = \sigma(l)$ and $(\sigma\tau)(l) = \sigma(k)$. So

$$\prod_j a_{(\sigma\tau)(j)j} = \left(\prod_{j\notin\{k,l\}} a_{\sigma(j)j}\right) a_{\sigma(l)k} a_{\sigma(k)l},$$

and since columns k and l are equal, $a_{\sigma(l)k} = a_{\sigma(l)l}$ and $a_{\sigma(k)l} = a_{\sigma(k)k}$. So the products of entries in $t_{\sigma\tau}(A)$ and $t_\sigma(A)$ agree, while $\operatorname{sgn}(\sigma\tau) = -\operatorname{sgn}(\sigma)$ by Proposition 5.2.25. Hence $t_{\sigma\tau}(A) = -t_\sigma(A)$.

The map $\sigma \mapsto \sigma\tau$ is its own inverse and changes the sign, so it maps the even permutations bijectively onto the odd ones. Therefore

$$L(A) = \sum_{\sigma \text{ even}} (t_\sigma(A) + t_{\sigma\tau}(A)) = \sum_{\sigma \text{ even}} 0 = 0.$$

No step divides by $1 + 1$: over $\mathbb{F}_2$ the two terms of a pair are equal and cancel because $x + x = 0$.

(c) Write $g(A) = a_{11}a_{12}\cdots a_{1n}$, so $D = \det + g$.

Normalized. $\det I_n = 1$, and $g(I_n)$ contains the factor $a_{12} = 0$ (as $n \geq 2$). So $D(I_n) = 1$.

Multilinear. g contains exactly one entry of each column, namely a_{1k} from column k. With the other columns fixed, g is a constant times a_{1k}, a linear function of column k. So g is multilinear, and so is $D = \det + g$.

Antisymmetric. Interchanging two columns negates det (Proposition 5.1.10) and permutes the factors of g, so it leaves g unchanged. In $\mathbb{F}_2$, $-1 = 1$, so both det and g are unchanged, and $D(A') = D(A) = -D(A)$ for the interchanged matrix A'.

Not alternating. Let A have first row $(1, \ldots, 1)$ and all other rows 0. All columns of A are equal to $\mathbf{e}_1$, so $\det A = 0$, but $g(A) = 1$, and $D(A) = 1 \neq 0$.

If $1 + 1 \neq 0$, an antisymmetric function is alternating (Proposition 5.1.11). So a normalized, multilinear, antisymmetric D would be a determinant function, hence equal to det by uniqueness, and det is alternating. So over such a field the three conditions determine det, while over $\mathbb{F}_2$ they allow at least two functions, for every $n \geq 2$. This is why the definition asks for alternation (compare Example 5.1.12, the case $n = 2$).

Solution to Problem 5.2

Uses. Theorem 5.3.4 (Existence and uniqueness of the determinant), Definition 5.1.1 (Determinant function), Theorem 5.4.10 (Multiplicativity), Exercise 5.1.5, Corollary 5.4.27 (Several diagonal blocks), Theorem 2.4.34 (Block multiplication)
Techniques. Recognize the determinant by its axioms (Technique 5.4.11)
Hint 1. For (b), fix A and regard $D(B) = \det(AB)$ as a function of the columns of B. Column j of AB is $A\mathbf{b}_j$.
Hint 2. For (c), take determinants: $(\det A)^2 = 2^n$. When is 2^n the square of a rational number? For even n, look for a 2×2 example and use blocks.
Sketch. (b) $D(B) = \det(A\mathbf{b}_1, \ldots, A\mathbf{b}_n)$ is linear in each column (a linear map followed by det) and vanishes when two columns of B agree, so $D(B) = D(I_n) \det B = \det A \det B$. Then $\det A \det A^{-1} = \det I_n = 1$. (c) $(\det A)^2 = 2^n$ with $\det A \in \mathbb{Q}$; for odd $n = 2k+1$ this makes $\det A/2^k$ a rational square root of 2. For even n, use $n/2$ diagonal blocks $\left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$. Over $\mathbb{R}$, $\sqrt{2}I_n$ works for every n.

(a) Theorem 5.4.7: $A \in F^{n\times n}$ is invertible if and only if $\det A \neq 0$. Theorem 5.4.10: $\det(AB) = \det A \det B$ for all $A, B \in F^{n\times n}$.

(b) Fix A, and let $D(B) = \det(AB)$. Column j of AB is $A\mathbf{b}_j$, so $D(B) = \det(A\mathbf{b}_1, \ldots, A\mathbf{b}_n)$. In column j, D is the linear map $\mathbf{x} \mapsto A\mathbf{x}$ followed by det, which is linear in column j; so D is multilinear. If two columns of B are equal, the corresponding columns of AB are equal, and $D(B) = 0$; so D is alternating. By Theorem 5.3.4, $D(B) = D(I_n) \det B = \det(A) \det B$.

If A is invertible, then $1 = \det I_n = \det\left(AA^{-1}\right) = \det A \det A^{-1}$, so $\det A \neq 0$.

(c) Suppose $A \in \mathbb{Q}^{n\times n}$ and $A^2 = 2I_n$. By (b) and Exercise 5.1.5, $(\det A)^2 = \det(2I_n) = 2^n$, and $\det A \in \mathbb{Q}$. If $n = 2k+1$ is odd, then $\left(\det A/2^k\right)^2 = 2^{2k+1}/2^{2k} = 2$, so $\sqrt{2}$ would be rational, which it is not. So n must be even. Conversely, $R = \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$ satisfies $R^2 = 2I_2$, and for even n the block-diagonal matrix with $n/2$ blocks R satisfies $A^2 = 2I_n$, since block-diagonal matrices multiply block by block (Theorem 2.4.34). So the answer over $\mathbb{Q}$ is: exactly the even n.

Over $\mathbb{R}$, $A = \sqrt{2}I_n$ works for every n. The obstruction over $\mathbb{Q}$ is arithmetic, the irrationality of $\sqrt{2}$, and the determinant is what detects it.

Solution to Problem 5.3

Uses. Theorem 5.5.14 (Cramer's rule), Proposition 2.4.8 (Four views of the product), Definition 5.1.1 (Determinant function), Theorem 5.4.7 (Invertibility and the determinant), Proposition 5.3.20 (Reduction modulo p), Corollary 5.5.12 (Integer inverses)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. For (b), write $\mathbf{y} = \sum_k x_k \mathbf{b}_k$ and expand $\det B_i(\mathbf{y})$ in column i.
Hint 2. For the second part of (c), if every integral $\mathbf{y}$ had an integral solution, what would that say about the columns of B^{-1}?
Sketch. (b) Linearity in column i gives $\det B_i(\mathbf{y}) = \sum_k x_k \det B_i(\mathbf{b}_k)$, and every term with $k \neq i$ has a repeated column, while $B_i(\mathbf{b}_i) = B$. (c) $\det B_i(\mathbf{y})$ is an integer. If all $B^{-1}\mathbf{e}_j$ were integral, B^{-1} would be, forcing $d = \pm 1$. For the example, $\mathbf{y} = \mathbf{e}_1$ gives $\mathbf{x} = (3/5, -1/5)$.

(a) (Theorem 5.5.14) If $B \in F^{n\times n}$ has $\det B \neq 0$ and $\mathbf{y} \in F^n$, then $B\mathbf{x} = \mathbf{y}$ has exactly one solution, and $x_i = \det B_i(\mathbf{y})/\det B$ for each i.

(b) Let $\mathbf{x}$ be a solution. Then $\mathbf{y} = B\mathbf{x} = \sum_k x_k \mathbf{b}_k$ (Proposition 2.4.8). By linearity in column i,

$$\det B_i(\mathbf{y}) = \sum_{k=1}^{n} x_k \det B_i(\mathbf{b}_k).$$

For $k \neq i$, the matrix $B_i(\mathbf{b}_k)$ has $\mathbf{b}_k$ in columns i and k, so its determinant is 0. For $k = i$, $B_i(\mathbf{b}_i) = B$. Hence $\det B_i(\mathbf{y}) = x_i \det B$.

If $\det B \neq 0$, then B is invertible (Theorem 5.4.7), so $\mathbf{x} = B^{-1}\mathbf{y}$ is the unique solution, and dividing by $\det B$ gives $x_i = \det B_i(\mathbf{y})/\det B$.

(c) If $\mathbf{y}$ is integral, then $B_i(\mathbf{y})$ has integer entries, so $\det B_i(\mathbf{y})$ is an integer (Proposition 5.3.20), and by (b), $dx_i = \det B_i(\mathbf{y}) \in \mathbb{Z}$.

Suppose that every integral $\mathbf{y}$ gives an integral solution. Then each column $B^{-1}\mathbf{e}_j$ of B^{-1}, the solution for $\mathbf{y} = \mathbf{e}_j$, is integral, so B^{-1} has integer entries, and $d = \pm 1$ by Corollary 5.5.12. So if $d \neq \pm 1$, some $\mathbf{e}_j$ has a non-integral solution.

For $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, $d = 5$. With $\mathbf{y} = \mathbf{e}_1$,

$$x_1 = \frac{1}{5}\det\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} = \frac{3}{5}, \quad x_2 = \frac{1}{5}\det\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = -\frac{1}{5}.$$

Check: $2 \cdot 3/5 - 1/5 = 1$ and $3/5 + 3 \cdot (-1/5) = 0$. The denominators are exactly d.

Solution to Problem 5.4

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Proposition 1.1.17 (No zero divisors), Corollary 5.4.8 (Vanishing determinants), Proposition 5.3.15 (Triangular matrices)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), try diagonal matrices. For (c), the two products have different sizes when $m \neq n$.
Hint 2. For (d) and (e), what is $A(1, \ldots, 1)$?
Sketch. (a) False: $I_n + \operatorname{diag}(-1, -1, 1, \ldots, 1) = \operatorname{diag}(0, 0, 2, \ldots, 2)$. (b) True: $d^3 = d$ gives $d(d-1)(d+1) = 0$. (c) False: $A = (1 \;\; 0)$ gives $\det\left(A^{\mathsf{T}}A\right) = 0$ and $\det\left(AA^{\mathsf{T}}\right) = 1$; true when $m = n$. (d) True: $(1, \ldots, 1)$ is in the kernel. (e) False: $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.

(a) False. Let $A = I_n$ and $B = \operatorname{diag}(-1, -1, 1, \ldots, 1)$. Both have determinant 1 (Proposition 5.3.15), but $A + B = \operatorname{diag}(0, 0, 2, \ldots, 2)$ has determinant 0.

(b) True. Let $d = \det A$. By Theorem 5.4.10, $d^3 = \det\left(A^3\right) = \det A = d$, so $d(d-1)(d+1) = d^3 - d = 0$. A field has no zero divisors (Proposition 1.1.17), so d is 0, 1, or -1. (Over $\mathbb{F}_2$ the set $\{-1, 0, 1\}$ is $\{0, 1\}$.)

(c) False in general. Let $A = (1 \;\; 0) \in \mathbb{R}^{1\times 2}$. Then $A^{\mathsf{T}}A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has determinant 0, while $AA^{\mathsf{T}} = (1)$ has determinant 1. The statement is true when $m = n$: then both sides equal $\det A^{\mathsf{T}} \det A = (\det A)^2$ by Theorem 5.4.10 and Theorem 5.4.14.

(d) True. The i-th entry of $A(1, \ldots, 1)$ is the sum of row i, so $A(1, \ldots, 1) = \mathbf{0}$ with $(1, \ldots, 1) \neq \mathbf{0}$, and $\det A = 0$ by Corollary 5.4.8.

(e) False for every $n \geq 2$ (for $n = 1$ it is true). $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ has row sums 1 and determinant 0, and adding a diagonal block I_{n-2} gives an example for each $n \geq 2$. The hypothesis says $A(1, \ldots, 1) = (1, \ldots, 1)$, a statement about one vector, which says nothing about $\det A$.

Solution to Problem 5.5

Uses. Definition 5.1.14 (The 2×2 determinant), Definition 2.4.26 (Trace), Theorem 5.4.10 (Multiplicativity), Proposition 1.1.17 (No zero divisors)
Hint 1. For (a), count the triples (a, b, d) with $ad = b^2$, splitting into the cases $b = 0$ and $b \neq 0$.
Hint 2. In (b), $d = -a$, and the condition becomes $b(-c) = a^2$. For (c), multiply out A^2; for the converse, take determinants of $A^2 = 0$ first.
Sketch. (a) $b = 0$: $ad = 0$ has $2p - 1$ solutions. $b \neq 0$: a is any of the $p - 1$ nonzero elements and $d = b^2a^{-1}$. Total $2p - 1 + (p-1)^2 = p^2$. (b) $d = -a$ and $b(-c) = a^2$ is the same equation $xy = z^2$, so again p^2. (c) Direct computation. If $A^2 = 0$, then $(\det A)^2 = 0$, so $A^2 = (\operatorname{tr} A)A$, forcing $\operatorname{tr} A = 0$. So there are p^2 such A.

(a) By Definition 5.1.14 we count the triples $(a, b, d) \in \mathbb{F}_p^3$ with $ad - b^2 = 0$. If $b = 0$, we need $ad = 0$, that is, $a = 0$ or $d = 0$ (Proposition 1.1.17): there are p triples with $a = 0$ and p with $d = 0$, and one triple has both, so $2p - 1$ in all. If $b \neq 0$ ($p - 1$ choices), then $ad = b^2 \neq 0$, so $a \neq 0$ ($p - 1$ choices), and then $d = b^2a^{-1}$ is determined. In all there are

$$2p - 1 + (p-1)^2 = p^2$$

such matrices.

(b) Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\operatorname{tr} A = 0$ means $d = -a$, and then $\det A = -a^2 - bc$, so the conditions say $b(-c) = a^2$, with a, b, c free otherwise. The map $(a, b, c) \mapsto (b, -c, a)$ is a bijection of $\mathbb{F}_p^3$ that turns these solutions into the solutions (x, y, z) of $xy = z^2$, and so does $(a, b, d) \mapsto (a, d, b)$ for the equation in (a). Both counts are the number of solutions of $xy = z^2$ in $\mathbb{F}_p^3$, which is p^2.

(c) With $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$A^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix},$$

$$(\operatorname{tr} A)A = \begin{pmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{pmatrix},$$

so $A^2 - (\operatorname{tr} A)A = \begin{pmatrix} bc-ad & 0 \\ 0 & bc-ad \end{pmatrix} = -(\det A) I_2$, as claimed.

If $\operatorname{tr} A = \det A = 0$, the identity gives $A^2 = 0$. Conversely, suppose $A^2 = 0$. Then $(\det A)^2 = \det(A^2) = 0$ by Theorem 5.4.10, so $\det A = 0$, and the identity becomes $(\operatorname{tr} A)A = A^2 = 0$. Either $\operatorname{tr} A = 0$, or $A = 0$, and the zero matrix also has trace 0. So $A^2 = 0$ exactly when $\operatorname{tr} A = \det A = 0$, and by (b) there are p^2 such matrices in $\mathbb{F}_p^{2\times 2}$, one of which is 0.

Solution to Problem 5.6

Uses. Definition 5.1.14 (The 2×2 determinant), Theorem 5.1.13 (The 2×2 determinant), Proposition 5.1.10 (Alternating implies antisymmetric), Example 5.1.15 (det is signed area)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. For (a), expand $\det(\mathbf{p}_k + \mathbf{c}, \mathbf{p}_{k+1} + \mathbf{c})$ into four terms, and sum over k. The sums $\sum_k \mathbf{p}_k$ and $\sum_k \mathbf{p}_{k+1}$ are equal.
Hint 2. For (b), apply (a) with $\mathbf{c} = -\mathbf{p}$. For (c), apply (b) to each triangle and then (a) with $-\mathbf{c}$.
Sketch. (a) The extra terms sum to $\det(\mathbf{s}, \mathbf{c}) + \det(\mathbf{c}, \mathbf{s})$ with $\mathbf{s} = \sum_k \mathbf{p}_k$, which is 0 by antisymmetry, and $\det(\mathbf{c}, \mathbf{c}) = 0$. (b) After translating by $-\mathbf{p}$, two of the three terms involve $\mathbf{0}$. (c) $\sum_k S(\mathbf{c}, \mathbf{p}_k, \mathbf{p}_{k+1}) = \frac{1}{2}\sum_k \det(\mathbf{p}_k - \mathbf{c}, \mathbf{p}_{k+1} - \mathbf{c}) = S$ by (a); for an interior $\mathbf{c}$ the triangles tile the polygon and are all counterclockwise. (d) $\frac{1}{2}(0 + 12 + 19 + 11 + 0) = 21$.

(a) By multilinearity, with $\det(\mathbf{c}, \mathbf{c}) = 0$,

$$\det(\mathbf{p}_k + \mathbf{c}, \mathbf{p}_{k+1} + \mathbf{c}) = \det(\mathbf{p}_k, \mathbf{p}_{k+1}) + \det(\mathbf{p}_k, \mathbf{c}) + \det(\mathbf{c}, \mathbf{p}_{k+1}).$$

Let $\mathbf{s} = \sum_{k=1}^m \mathbf{p}_k$. Since $\mathbf{p}_{m+1} = \mathbf{p}_1$, also $\sum_{k=1}^m \mathbf{p}_{k+1} = \mathbf{s}$. Summing over k and using linearity in each column, the extra terms add up to

$$\det(\mathbf{s}, \mathbf{c}) + \det(\mathbf{c}, \mathbf{s}) = 0$$

by antisymmetry (Proposition 5.1.10). So S is unchanged.

(b) By (a) with $\mathbf{c} = -\mathbf{p}$, and writing $\mathbf{q}' = \mathbf{q} - \mathbf{p}$ and $\mathbf{r}' = \mathbf{r} - \mathbf{p}$,

$$\begin{aligned} S(\mathbf{p}, \mathbf{q}, \mathbf{r}) &= S(\mathbf{0}, \mathbf{q}', \mathbf{r}') \\ &= \frac{1}{2}(\det(\mathbf{0}, \mathbf{q}') + \det(\mathbf{q}', \mathbf{r}') + \det(\mathbf{r}', \mathbf{0})) \\ &= \frac{1}{2}\det(\mathbf{q}', \mathbf{r}'), \end{aligned}$$

since a matrix with a zero column has determinant 0. The triangle $\mathbf{pqr}$ is half of the parallelogram spanned by $\mathbf{q}'$ and $\mathbf{r}'$, translated to $\mathbf{p}$. By Example 5.1.15, $\det(\mathbf{q}', \mathbf{r}')$ is the signed area of that parallelogram, positive exactly when the turn from $\mathbf{q}'$ to $\mathbf{r}'$ is counterclockwise, that is, when $\mathbf{p}, \mathbf{q}, \mathbf{r}$ run counterclockwise.

(c) By (b), $S(\mathbf{c}, \mathbf{p}_k, \mathbf{p}_{k+1}) = \frac{1}{2}\det(\mathbf{p}_k - \mathbf{c}, \mathbf{p}_{k+1} - \mathbf{c})$. Summing over k gives $S(\mathbf{p}_1 - \mathbf{c}, \ldots, \mathbf{p}_m - \mathbf{c})$, which is $S(\mathbf{p}_1, \ldots, \mathbf{p}_m)$ by (a).

Now let the polygon be convex with counterclockwise vertices, and let $\mathbf{c}$ be a point inside it. The segments from $\mathbf{c}$ to the vertices cut the polygon into the triangles $\mathbf{c}\mathbf{p}_k\mathbf{p}_{k+1}$, and each of these runs counterclockwise, since the boundary of the polygon does. By (b), each $S(\mathbf{c}, \mathbf{p}_k, \mathbf{p}_{k+1})$ is the area of its triangle, and their sum is the area of the polygon.

(d) The terms $\det(\mathbf{p}_k, \mathbf{p}_{k+1})$ are

$$\det\begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix} = 0, \quad \det\begin{pmatrix} 4 & 5 \\ 0 & 3 \end{pmatrix} = 12,$$

$$\det\begin{pmatrix} 5 & 2 \\ 3 & 5 \end{pmatrix} = 19, \quad \det\begin{pmatrix} 2 & -1 \\ 5 & 3 \end{pmatrix} = 11,$$

$$\det\begin{pmatrix} -1 & 0 \\ 3 & 0 \end{pmatrix} = 0,$$

so $S = \frac{1}{2}(0 + 12 + 19 + 11 + 0) = 21$. The pentagon is convex and its vertices run counterclockwise: the edge vectors $(4, 0), (1, 3), (-3, 2), (-3, -2), (1, -3)$ point in the directions of about $0°$, $72°$, $146°$, $214°$, and $288°$, so the boundary turns left at every vertex and makes exactly one full turn in all. Hence the area is 21. (Left turns alone are not enough: the edges of a five-pointed star also turn left at every vertex, but they make two full turns.)

Solution to Problem 5.7

> *Uses.* Lemma 5.1.8 (Column replacement), Lemma 5.5.2 (A row or column with one nonzero entry), Corollary 5.4.8 (Vanishing determinants), Theorem 5.5.14 (Cramer's rule), Proposition 2.4.8 (Four views of the product)
> *Hint 1.* For (a), subtract column 1 from the other columns, and expand along the first row.
> *Hint 2.* For (c), the two conditions on λ are one linear system $M\boldsymbol{\lambda} = (1, \mathbf{x})$. For (d), apply (a) to M_i, which is the matrix of the same kind built from the points with $\mathbf{p}_i$ replaced by $\mathbf{x}$.
> *Sketch.* (a) Column replacements turn column $i + 1$ into $(0, \mathbf{p}_i - \mathbf{p}_0)$, after which row 1 is $\mathbf{e}_1^{\mathsf{T}}$ and the minor is the matrix with columns $\mathbf{p}_i - \mathbf{p}_0$. (b) Corollary 5.4.8. (c) The conditions read $M\boldsymbol{\lambda} = (1, \mathbf{x})$; apply Cramer's rule. (d) By (a) and Problem 5.6 (b), $\det M_i$ is twice the signed area of the triangle with $\mathbf{p}_i$ replaced by $\mathbf{x}$, and $\det M$ is twice that of $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$. For the example, $\det M = 16$ and $\lambda = (1/4, 1/4, 1/2)$.

(a) For $i = 1, \ldots, n$, subtract column 1 from column $i + 1$. These replacements do not change the determinant (Lemma 5.1.8), and column $i + 1$ becomes $(0, \mathbf{p}_i - \mathbf{p}_0)$. Now row 1 is $(1, 0, \ldots, 0)$, so by Lemma 5.5.2 the determinant is $(-1)^{1+1} \cdot 1$ times the determinant of the matrix left after deleting row 1 and column 1, whose columns are $\mathbf{p}_1 - \mathbf{p}_0, \ldots, \mathbf{p}_n - \mathbf{p}_0$.

(b) By (a) and Corollary 5.4.8, $\det M \neq 0$ exactly when the columns $\mathbf{p}_i - \mathbf{p}_0$ are independent. For $n = 2$ this says that $\mathbf{p}_1 - \mathbf{p}_0$ and $\mathbf{p}_2 - \mathbf{p}_0$ do not lie on one line through $\mathbf{0}$, that is, the three points do not lie on one line.

(c) Let $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_n)$. By Proposition 2.4.8, $M\boldsymbol{\lambda} = \sum_i \lambda_i (1, \mathbf{p}_i)$, whose first entry is $\sum_i \lambda_i$ and whose other entries form $\sum_i \lambda_i \mathbf{p}_i$. So the two conditions say exactly $M\boldsymbol{\lambda} = (1, \mathbf{x})$. Since $\det M \neq 0$, Cramer's rule (Theorem 5.5.14) gives a unique solution, with $\lambda_i = \det M_i / \det M$.

(d) M_i is the matrix M built from the points $\mathbf{p}_0, \ldots, \mathbf{x}, \ldots, \mathbf{p}_n$, with $\mathbf{x}$ in place of $\mathbf{p}_i$. By (a) and Problem 5.6 (b), $\det M = 2S(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$ and

$$\begin{aligned} \det M_0 &= 2S(\mathbf{x}, \mathbf{p}_1, \mathbf{p}_2), \\ \det M_1 &= 2S(\mathbf{p}_0, \mathbf{x}, \mathbf{p}_2), \\ \det M_2 &= 2S(\mathbf{p}_0, \mathbf{p}_1, \mathbf{x}). \end{aligned}$$

So λ_i is the signed area of the triangle with $\mathbf{p}_i$ replaced by $\mathbf{x}$, divided by the signed area of $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$. It is positive exactly when the two signed areas have the same sign, that is, the two triangles have the same orientation.

For $\mathbf{p}_0 = (0, 0)$, $\mathbf{p}_1 = (4, 0)$, $\mathbf{p}_2 = (0, 4)$, and $\mathbf{x} = (1, 2)$:

$$\det M = \det\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = 16,$$

$$\det M_1 = \det\begin{pmatrix} 1 & 0 \\ 2 & 4 \end{pmatrix} = 4,$$

$$\det M_2 = \det\begin{pmatrix} 4 & 1 \\ 0 & 2 \end{pmatrix} = 8,$$

and $\det M_0 = \det(\mathbf{p}_1 - \mathbf{x}, \mathbf{p}_2 - \mathbf{x}) = \det\begin{pmatrix} 3 & -1 \\ -2 & 2 \end{pmatrix} = 4$. So $\lambda = (1/4, 1/4, 1/2)$, and indeed $\frac{1}{4}(0,0) + \frac{1}{4}(4,0) + \frac{1}{2}(0,4) = (1,2)$. All three are positive, so $(1,2)$ lies inside the triangle in this sense. These λ_i are the barycentric coordinates of Section 4.7, and the sign test is the standard way to decide whether a point lies in a triangle.

Solution to Problem 5.8

Uses. Definition 5.4.18 (Determinant of an operator), Definition 1.6.1 (Basis), Proposition 5.3.15 (Triangular matrices), Proposition 5.4.20 (Properties of the determinant of an operator), Exercise 5.4.40

Hint 1. Symmetric matrices are fixed by the transpose and skew-symmetric ones are negated. What does T do to each kind?

Hint 2. For the basis in (a), $U_{ij} = \frac{1}{2}\big((U_{ij} + U_{ji}) + (U_{ij} - U_{ji})\big)$, and there are $n + 2 \cdot n(n-1)/2 = n^2$ matrices in the list.

Sketch. (a) T multiplies U_{ii} and $U_{ij} + U_{ji}$ by $\alpha + \beta$, and $U_{ij} - U_{ji}$ by $\alpha - \beta$. There are $n + n(n-1)/2 = n(n+1)/2$ basis vectors of the first kind. (b) $\alpha = 0$, $\beta = 1$ gives $(-1)^{n(n-1)/2}$. (c) The matrix is block diagonal with blocks $\alpha + \beta$, $\begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$, $\alpha + \beta$, so $\det T = (\alpha+\beta)^2(\alpha^2 - \beta^2) = (\alpha+\beta)^3(\alpha - \beta)$. (d) Invertible iff $\alpha + \beta \neq 0$ and, for $n \geq 2$, $\alpha - \beta \neq 0$; for $\alpha = \beta \neq 0$ the kernel is the space of skew-symmetric matrices.

(a) Let $S_{ij} = U_{ij} + U_{ji}$ and $K_{ij} = U_{ij} - U_{ji}$ for $i < j$. The list $U_{11}, \dots, U_{nn}, S_{ij}, K_{ij}$ has $n + 2 \cdot n(n-1)/2 = n^2$ entries.

It spans. Each U_{ii} is in the list, and for $i < j$, $U_{ij} = (1+1)^{-1}\big(S_{ij} + K_{ij}\big)$ and $U_{ji} = (1+1)^{-1}\big(S_{ij} - K_{ij}\big)$. So the span contains every matrix unit, hence all of $F^{n \times n}$.

It is independent. Suppose $\sum_i a_i U_{ii} + \sum_{i<j}\big(b_{ij} S_{ij} + c_{ij} K_{ij}\big) = 0$. The (i,i) entry gives $a_i = 0$. For $i < j$, the (i,j) and (j,i) entries give $b_{ij} + c_{ij} = 0$ and $b_{ij} - c_{ij} = 0$, so $(1+1)b_{ij} = 0$, hence $b_{ij} = 0$ and $c_{ij} = 0$.

So the list is a basis (Definition 1.6.1). Since $U_{ii}^{\mathsf{T}} = U_{ii}$, $S_{ij}^{\mathsf{T}} = S_{ij}$, and $K_{ij}^{\mathsf{T}} = -K_{ij}$,

$$T(U_{ii}) = (\alpha + \beta) U_{ii},$$

$$T\big(S_{ij}\big) = (\alpha + \beta) S_{ij},$$

$$T\big(K_{ij}\big) = (\alpha - \beta) K_{ij}.$$

So the matrix of T in this basis is diagonal, with $\alpha + \beta$ appearing $n + n(n-1)/2 = n(n+1)/2$ times and $\alpha - \beta$ appearing $n(n-1)/2$ times. By Definition 5.4.18 and Proposition 5.3.15, $\det T$ is the product of these diagonal entries, as claimed.

(b) The transpose operator is the case $\alpha = 0$, $\beta = 1$, and the formula gives $1 \cdot (-1)^{n(n-1)/2}$, which is the answer of Exercise 5.4.40.

(c) For $n = 2$, $T(U_{11}) = (\alpha+\beta)U_{11}$, $T(U_{22}) = (\alpha+\beta)U_{22}$, $T(U_{12}) = \alpha U_{12} + \beta U_{21}$, and $T(U_{21}) = \beta U_{12} + \alpha U_{21}$. In the basis $U_{11}, U_{12}, U_{21}, U_{22}$ the matrix of T is

$$\begin{pmatrix} \alpha+\beta & 0 & 0 & 0 \\ 0 & \alpha & \beta & 0 \\ 0 & \beta & \alpha & 0 \\ 0 & 0 & 0 & \alpha+\beta \end{pmatrix},$$

which is block diagonal, so $\det T = (\alpha+\beta)^2(\alpha^2 - \beta^2) = (\alpha+\beta)^3(\alpha - \beta)$ (Corollary 5.4.27). This is the formula of (a) for $n = 2$, and the computation used nothing about $1 + 1$. (Over $\mathbb{F}_2$ it reads $(\alpha+\beta)^4$.)

(d) By Proposition 5.4.20, T is invertible exactly when $\det T \neq 0$, that is, when $\alpha+\beta \neq 0$ and, if $n \geq 2$, also $\alpha-\beta \neq 0$. If $\alpha=\beta \neq 0$, then in the basis of (a) the diagonal entries are $1+1$ times α for U_{ii} and S_{ij}, which is nonzero, and 0 for K_{ij}. So $\ker T = \operatorname{span}\left(K_{ij} \mid i<j\right)$, the space of skew-symmetric matrices, of dimension $n(n-1)/2$. Directly: $\alpha\left(X+X^{\mathsf{T}}\right)=0$ exactly when $X^{\mathsf{T}}=-X$.

Solution to Problem 5.9

Uses. Definition 3.4.22 (Vandermonde matrix), Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Definition 2.4.1 (Matrix product)
Techniques. Factor the matrix (Technique 5.6.1)
Hint 1. The (l,j) entry of V is x_l^{j-1}. Compute the (i,j) entry of $V^{\mathsf{T}}V$ as a sum over l.
Hint 2. For (c), a real square is nonnegative. Over $\mathbb{C}$, try $x=(i,-i)$. For (d), you do not need the x_i: compute $\det H$ from the power sums.
Sketch. (a) $\left(V^{\mathsf{T}}V\right)_{ij}=\sum_l x_l^{i-1}x_l^{j-1}=s_{i+j-2}$, so $\det H=(\det V)^2$. (b) $2\left(x_1^2+x_2^2\right)-(x_1+x_2)^2=(x_1-x_2)^2$. (c) Squares of reals; over $\mathbb{C}$, $x=(i,-i)$ gives $H=\left(\begin{smallmatrix}2&0\\0&-2\end{smallmatrix}\right)$. (d) $\det\left(\begin{smallmatrix}3&0&6\\0&6&0\\6&0&18\end{smallmatrix}\right)=108>0$, so yes.

(a) The (l,j) entry of V is x_l^{j-1} (Definition 3.4.22), so the (i,l) entry of V^{T} is x_l^{i-1}, and

$$\left(V^{\mathsf{T}}V\right)_{ij}=\sum_{l=1}^{n}x_l^{i-1}x_l^{j-1}=\sum_{l=1}^{n}x_l^{i+j-2}=s_{i+j-2}.$$

So $H=V^{\mathsf{T}}V$, and by Theorem 5.4.10, Theorem 5.4.14, and Theorem 5.5.21,

$$\det H=\det V^{\mathsf{T}}\det V=(\det V)^2=\prod_{1\leq i<j\leq n}\left(x_j-x_i\right)^2.$$

(b) For $n=2$, $H=\left(\begin{smallmatrix}2&x_1+x_2\\x_1+x_2&x_1^2+x_2^2\end{smallmatrix}\right)$, and

$$\begin{aligned}\det H&=2x_1^2+2x_2^2-(x_1+x_2)^2\\&=x_1^2-2x_1x_2+x_2^2=(x_2-x_1)^2.\end{aligned}$$

(c) Over $\mathbb{R}$, $\det H=(\det V)^2\geq 0$, with equality exactly when $\det V=0$, that is, when some factor x_j-x_i is 0 (Proposition 1.1.17). Over $\mathbb{C}$, take $n=2$, $x_1=i$, and $x_2=-i$, where i is the imaginary unit: then $s_0=2$, $s_1=0$, $s_2=i^2+(-i)^2=-2$, and $\det H=\det\left(\begin{smallmatrix}2&0\\0&-2\end{smallmatrix}\right)=-4$, which is $(x_2-x_1)^2=(-2i)^2$.

(d) Here $s_0=3$, and

$$\det H=\det\begin{pmatrix}3&0&6\\0&6&0\\6&0&18\end{pmatrix}=6(3\cdot 18-6\cdot 6)=108>0,$$

expanding along the middle row. By (c), the x_i are distinct. (Indeed $x=\left(-\sqrt{3},0,\sqrt{3}\right)$ has these power sums, and $\left(\sqrt{3}\right)^2\left(2\sqrt{3}\right)^2\left(\sqrt{3}\right)^2=108$.)

Solution to Problem 5.10

Uses. Corollary 5.4.15 (Rows), Proposition 5.3.15 (Triangular matrices), Definition 5.5.6 (Cofactors, cofactor matrix, adjugate), Theorem 5.5.9 (The adjugate and the inverse), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Expose a factor with row and column operations (Technique 5.5.22)
Hint 1. Subtract a times row $i-1$ from row i, for $i=n,\ldots,2$. What happens to the entries left of the diagonal?
Hint 2. For (c), the inverse of K_3 is tridiagonal. Guess a tridiagonal T with $-a$ next to the diagonal, and check $K_nT=\left(1-a^2\right)I_n$ column by column.
Sketch. (a) The row operations make the entries left of the diagonal 0 and multiply those on and right of it by $1-a^2$ (except in row 1), leaving an upper triangular matrix with diagonal $1,1-a^2,\ldots,1-a^2$. (b) $\operatorname{adj}K_3=\left(1-a^2\right)\left(\begin{smallmatrix}1&-a&0\\-a&1+a^2&-a\\0&-a&1\end{smallmatrix}\right)$. (c) $K_n^{-1}=\left(1-a^2\right)^{-1}T$, where T is tridiagonal with diagonal $1,1+a^2,\ldots,1+a^2,1$ and $-a$ beside the diagonal.

(a) For $i = n, n-1, \ldots, 2$ in turn, subtract a times row $i-1$ from row i; working from the bottom up, row $i-1$ is still the original row when it is used. These replacements do not change the determinant (Corollary 5.4.15). For $i \geq 2$ the new (i,j) entry is $a^{|i-j|} - a \cdot a^{|i-1-j|}$. If $j \leq i-1$, this is $a^{i-j} - a \cdot a^{i-1-j} = 0$. If $j \geq i$, it is $a^{j-i} - a \cdot a^{j-i+1} = a^{j-i}\left(1-a^2\right)$. So the new matrix is upper triangular with diagonal entries $1, 1-a^2, \ldots, 1-a^2$, and $\det K_n = \left(1-a^2\right)^{n-1}$ by Proposition 5.3.15.

(b) The cofactors of K_3 are

$$c_{11} = 1-a^2, c_{12} = -\left(a-a^3\right), c_{13} = a^2 - a^2 = 0,$$
$$c_{22} = 1-a^4, c_{23} = -\left(a-a^3\right), c_{33} = 1-a^2,$$

and $c_{ji} = c_{ij}$ since K_3 is symmetric. Factoring out $1-a^2$,

$$\operatorname{adj} K_3 = \left(1-a^2\right)\begin{pmatrix} 1 & -a & 0 \\ -a & 1+a^2 & -a \\ 0 & -a & 1 \end{pmatrix}.$$

By (a), $\det K_3 = \left(1-a^2\right)^2$, so when $a^2 \neq 1$, Theorem 5.5.9 gives

$$K_3^{-1} = \frac{1}{1-a^2}\begin{pmatrix} 1 & -a & 0 \\ -a & 1+a^2 & -a \\ 0 & -a & 1 \end{pmatrix}.$$

(c) Let $T \in F^{n\times n}$ be tridiagonal with diagonal entries $1, 1+a^2, \ldots, 1+a^2, 1$ (the first and last are 1) and every entry next to the diagonal equal to $-a$. We check $K_n T = \left(1-a^2\right) I_n$ one column at a time.

Column 1 of T is $(1, -a, 0, \ldots, 0)$, so the $(i,1)$ entry of $K_n T$ is $a^{|i-1|} - a \cdot a^{|i-2|}$. For $i = 1$ it is $1-a^2$, and for $i \geq 2$ it is $a^{i-1} - a^{i-1} = 0$. *Column* n is the mirror image of column 1, with the same result.

Column j *with* $1 < j < n$ has entries $-a, 1+a^2, -a$ in rows $j-1, j, j+1$. The (i,j) entry of $K_n T$ is

$$-a \cdot a^{|i-j+1|} + \left(1+a^2\right)a^{|i-j|} - a \cdot a^{|i-j-1|}.$$

If $i = j$, this is $-a^2 + 1 + a^2 - a^2 = 1-a^2$. If $i < j$, put $d = j-i \geq 1$; the entry is $-a^d + a^d + a^{d+2} - a^{d+2} = 0$. If $i > j$, put $d = i-j \geq 1$; it is $-a^{d+2} + a^d + a^{d+2} - a^d = 0$.

So $K_n T = \left(1-a^2\right) I_n$. Taking determinants, $\det K_n \det T = \left(1-a^2\right)^n \neq 0$, so K_n is invertible (Theorem 5.4.7), and multiplying by K_n^{-1} gives $K_n^{-1} = \left(1-a^2\right)^{-1} T$.

Solution to Problem 5.11

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.7 (Invertibility and the determinant), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Theorem 7.1.32 (Fundamental Theorem of Algebra)
Techniques. det is a polynomial in the entries (Technique 5.3.18)
Hint 1. The product computation is fine. Test the sentence "so is $M = A^{-1}N$" on 2×2 matrices, with A a swap.
Hint 2. For (b), commuting lets you regroup $\left(A^{-1}N\right)^k$. For (c), the polynomial $\det(I_n + tN)$ never vanishes; over $\mathbb{C}$, what does that force?
Sketch. (a) $A^{-1}N$ need not be nilpotent: $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $N = \left(\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right)$ give $A+N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, singular, and $A^{-1}N = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. (b) If $AN = NA$, then $\left(A^{-1}N\right)^k = A^{-k}N^k = 0$. (c) $g(t) = \det(I_n + tN)$ is $e_0 + e_1 t + \cdots + e_n t^n$ and has no zeros, since tN is nilpotent and commutes with I_n. A nonconstant complex polynomial has a root, so g is constant, equal to $g(0) = 1$. (d) $\det(A+N) = \det A \det\left(I_n + A^{-1}N\right) = \det A$.

(a) The error is the claim that $M = A^{-1}N$ is nilpotent. (The displayed product is correct: it telescopes to $I_n - (-M)^k$.) Let

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

Then A is invertible ($A^{-1} = A$), $N^2 = 0$, and $A + N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has a zero row, so it is not invertible. Here $A^{-1}N = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$, whose powers $\begin{pmatrix} (-1)^k & 0 \\ 0 & 0 \end{pmatrix}$ are never 0.

(b) Suppose $AN = NA$. Multiplying by A^{-1} on both sides gives $NA^{-1} = A^{-1}N$. So A^{-1} and N commute, and by induction on k, $\left(A^{-1}N\right)^k = A^{-k}N^k$ (move each A^{-1} to the left past the factors N). If $N^k = 0$, then $\left(A^{-1}N\right)^k = 0$, $M = A^{-1}N$ is nilpotent, and the argument of (a) is valid.

(c) Let $N^k = 0$ and $g(t) = \det(I_n + tN)$ for $t \in \mathbb{C}$. Since $(tN)^k = t^k N^k = 0$ and tN commutes with I_n, (b) with $A = I_n$ shows that $I_n + tN$ is invertible, so $g(t) \neq 0$ for every t (Theorem 5.4.7). Every entry of $I_n + tN$ has the form $\alpha + \beta t$, so, as in the proof of Proposition 5.3.19, $g(t) = e_0 + e_1 t + \cdots + e_n t^n$ for some $e_0, \ldots, e_n \in \mathbb{C}$ and all t, with $e_0 = g(0) = \det I_n = 1$. If $e_m \neq 0$ for some $m \geq 1$, the polynomial $e_0 + e_1 x + \cdots + e_n x^n \in \mathbb{C}[x]$ has degree at least 1, so it has a root $t_0 \in \mathbb{C}$ (Theorem 7.1.32), and $g(t_0) = 0$, which is impossible. So $e_1 = \cdots = e_n = 0$, and $g(t) = 1$ for every t. In particular $\det(I_n + N) = g(1) = 1$.

(d) By (b), $A^{-1}N$ is nilpotent, so by (c) and Theorem 5.4.10,

$$\begin{aligned} \det(A + N) &= \det\left(A\left(I_n + A^{-1}N\right)\right) \\ &= \det A \det\left(I_n + A^{-1}N\right) = \det A. \end{aligned}$$

A real matrix can be regarded as a complex one with the same determinant, since the Leibniz formula uses only sums and products of the entries, so (c) and (d) also hold over $\mathbb{R}$.

Solution to Problem 5.12

> *Uses.* Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.10 (Multiplicativity), Proposition 2.4.8 (Four views of the product), Proposition 5.3.15 (Triangular matrices), Lemma 5.5.23 (A polynomial with prescribed roots)
> *Techniques.* Factor the matrix (Technique 5.6.1)
> *Hint 1.* In $(C\mathbf{v}_m)_j = \sum_k c_{(k-j) \bmod n}\omega^{mk}$, substitute $l = (k-j) \bmod n$ and use $\omega^n = 1$.
> *Hint 2.* Put the $\mathbf{v}_m$ side by side as the columns of a matrix W. Then $CW = WD$ with D diagonal, and W is a Vandermonde matrix.
> *Hint 3.* For (d), $f(z) = 1 + z$. For the value, evaluate $t^n - 1 = \prod_m (t - \omega^m)$ at $t = -1$.
> *Sketch.* (a) $(C\mathbf{v}_m)_j = \sum_l c_l \omega^{m(j+l)} = \omega^{mj} f(\omega^m)$. (b) W with columns $\mathbf{v}_m$ is $V\left(1, \omega, \ldots, \omega^{n-1}\right)$, invertible since the nodes are distinct, and $CW = W \operatorname{diag}(f(\omega^m))$; cancel $\det W$. (c) $n = 3$ and $\omega^4 = \omega$. (d) $f(\omega^m) = 1 + \omega^m = 0$ iff $\omega^m = -1$, possible iff n is even. For odd n, $t^n - 1 = \prod_m (t - \omega^m)$ at $t = -1$ gives $-2 = -\prod_m (1 + \omega^m)$.

(a) By definition, $(C\mathbf{v}_m)_j = \sum_{k=0}^{n-1} c_{(k-j) \bmod n}\omega^{mk}$. As k runs over $0, \ldots, n-1$, so does $l = (k-j) \bmod n$, and $k = j + l$ or $k = j + l - n$. Since $\omega^n = 1$, in both cases $\omega^{mk} = \omega^{m(j+l)}$. So

$$(C\mathbf{v}_m)_j = \sum_{l=0}^{n-1} c_l \omega^{mj}\omega^{ml} = \omega^{mj} f(\omega^m),$$

which is the j-th entry of $f(\omega^m)\mathbf{v}_m$.

(b) Let W be the matrix with columns $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$. Its (k, m) entry is $\omega^{km} = \left(\omega^k\right)^m$, so W is the Vandermonde matrix $V\left(1, \omega, \ldots, \omega^{n-1}\right)$, and $\det W \neq 0$ by Theorem 5.5.21, since the nodes are distinct. By (a) and Proposition 2.4.8, column m of CW is $f(\omega^m)\mathbf{v}_m$, which is column m of WD for $D = \operatorname{diag}\left(f(1), f(\omega), \ldots, f\left(\omega^{n-1}\right)\right)$. So $CW = WD$, and by Theorem 5.4.10 and Proposition 5.3.15,

$$\det C \det W = \det W \det D = \det W \prod_{m=0}^{n-1} f(\omega^m).$$

Dividing by $\det W \neq 0$ gives the claim.

(c) For $n = 3$ and $(c_0, c_1, c_2) = (a, b, c)$, the matrix is $\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ and $f(z) = a + bz + cz^2$. So $\det = f(1)f(\omega)f\left(\omega^2\right)$ with $f\left(\omega^2\right) = a + b\omega^2 + c\omega^4 = a + b\omega^2 + c\omega$, since $\omega^3 = 1$.

(d) Here $f(z) = 1 + z$, so by (b), $\det C = 0$ exactly when $\omega^m = -1$ for some m. If n is even, $\omega^{n/2} = -1$. If n is odd, then $(-1)^n = -1$, while $(\omega^m)^n = 1$, so no ω^m equals -1. Hence C is singular exactly when n is even.

For odd n, the polynomial $x^n - 1$ has degree n, leading coefficient 1, and the n distinct roots $1, \omega, \dots, \omega^{n-1}$, so by Lemma 5.5.23, $t^n - 1 = \prod_{m=0}^{n-1}(t - \omega^m)$ for every $t \in \mathbb{C}$. At $t = -1$,

$$\begin{aligned} -2 = (-1)^n - 1 &= \prod_m (-1 - \omega^m) \\ &= (-1)^n \prod_m (1 + \omega^m) = -\det C, \end{aligned}$$

so $\det C = 2$. For $n = 3$ this is the matrix of Exercise 5.3.38 (a).

Solution to Problem 5.13

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.14 (The transpose), Proposition 5.3.15 (Triangular matrices), Proposition 5.1.10 (Alternating implies antisymmetric), Exercise 5.2.34
Techniques. Factor the matrix (Technique 5.6.1)
Hint 1. Expand $(x_i + y_j)^{n-1}$ by the binomial theorem and read the result as a product of three matrices.
Hint 2. $M = XDY$, where X has entries x_i^k, D is diagonal with the binomial coefficients, and Y has entries y_j^{n-1-k}. Up to the order of its rows, Y is $V(y_1, \dots, y_n)^{\mathsf{T}}$.
Sketch. By the binomial theorem, $M = XDY$ with $X = V(x_1, \dots, x_n)$, $D = \operatorname{diag}\left(\binom{n-1}{0}, \dots, \binom{n-1}{n-1}\right)$, and Y equal to $V(y_1, \dots, y_n)^{\mathsf{T}}$ with its rows reversed. Reversing n rows takes $\lfloor n/2 \rfloor$ interchanges, and $(-1)^{\lfloor n/2 \rfloor} = (-1)^{n(n-1)/2}$, which turns $\prod_{i<j}(y_j - y_i)$ into $\prod_{i<j}(y_i - y_j)$. (c) Over $\mathbb{Q}$: invertible iff the x_i are distinct and the y_j are distinct. Over $\mathbb{F}_2$: never, since $\binom{2}{1} = 0$.

(a) Index rows and columns of auxiliary matrices by $k = 0, \dots, n-1$ where needed. By the binomial theorem,

$$(x_i + y_j)^{n-1} = \sum_{k=0}^{n-1} x_i^k \binom{n-1}{k} y_j^{n-1-k}.$$

So $M = XDY$, where X has (i,k) entry x_i^k, $D = \operatorname{diag}\left(\binom{n-1}{0}, \dots, \binom{n-1}{n-1}\right)$, and Y has (k,j) entry y_j^{n-1-k}. Now $X = V(x_1, \dots, x_n)$. The matrix $V(y_1, \dots, y_n)^{\mathsf{T}}$ has (k,j) entry y_j^k, so Y is $V(y_1, \dots, y_n)^{\mathsf{T}}$ with its rows in reverse order. Reversing the order of n rows takes $\lfloor n/2 \rfloor$ interchanges (the first row with the last, and so on), each of which negates the determinant (Proposition 5.1.10 with Theorem 5.4.14), and $(-1)^{\lfloor n/2 \rfloor} = (-1)^{n(n-1)/2}$ by Exercise 5.2.34 (c). So, by Theorem 5.5.21 and Theorem 5.4.14,

$$\det Y = (-1)^{n(n-1)/2} \prod_{1 \le i < j \le n} (y_j - y_i) = \prod_{1 \le i < j \le n} (y_i - y_j),$$

because the product has $n(n-1)/2$ factors and each sign change flips one of them. By Theorem 5.4.10 and Proposition 5.3.15,

$$\begin{aligned} \det M &= \det X \det D \det Y \\ &= \prod_{i<j} (x_j - x_i) \cdot \prod_{k=0}^{n-1} \binom{n-1}{k} \cdot \prod_{i<j} (y_i - y_j). \end{aligned}$$

(b) For $n = 2$ the binomial coefficients are $1, 1$, and

$$\begin{aligned} \det M &= (x_1 + y_1)(x_2 + y_2) - (x_1 + y_2)(x_2 + y_1) \\ &= x_1 y_2 + x_2 y_1 - x_1 y_1 - x_2 y_2 \\ &= (x_2 - x_1)(y_1 - y_2). \end{aligned}$$

(c) For $n = 3$ the coefficients are $\binom{2}{0}, \binom{2}{1}, \binom{2}{2} = 1, 2, 1$. Over $\mathbb{Q}$ they are nonzero, so M is invertible exactly when the x_i are distinct and the y_j are distinct (Proposition 1.1.17, Theorem 5.4.7). Over $\mathbb{F}_2$, $\binom{2}{1} = 2 = 0$, so $\det M = 0$ and M is never invertible. This agrees with $(x + y)^2 = x^2 + y^2$ over $\mathbb{F}_2$: there M is the sum of

the matrix with all columns $\left(x_1^2, x_2^2, x_3^2\right)$ and the matrix with all rows $\left(y_1^2, y_2^2, y_3^2\right)$, and both have rank at most 1.

Solution to Problem 5.14

Uses. Theorem 5.3.3 (The Leibniz formula), Proposition 5.2.25 (Multiplicativity of the sign), Definition 5.2.10 (Transposition), Proposition 5.5.18 (Tridiagonal determinants)
Techniques. Read structure off the Leibniz expansion (Technique 5.3.12), Induction on matrix size (Technique 5.5.17)
Hint 1. For (a), induct on n and look at $\sigma(n)$, which is n or $n-1$. If $\sigma(n) = n-1$, which j can have $\sigma(j) = n$?
Hint 2. The same case split counts the permutations in (b). In (c), the term of σ is nonzero only if $|\sigma(j) - j| \leq 1$ for every j.
Sketch. (a) Induct on n. If $\sigma(n) = n$, restrict to $[n-1]$. If $\sigma(n) = n-1$, then $\sigma^{-1}(n) \in \{n-1, n\}$ forces $\sigma(n-1) = n$, and $(n-1\ n)\sigma$ fixes $n-1$ and n. (b) The two cases give $N_n = N_{n-1} + N_{n-2}$ with $N_0 = N_1 = 1$. (c) For the product σ of the transpositions of P, $\operatorname{sgn}(\sigma) = (-1)^{|P|}$, and each pair contributes $t_{k+1,k}t_{k,k+1} = c_k b_k$; all other terms of the Leibniz formula vanish. (d) Split by whether n is matched. In the example, every term equals 1.

(a) A product of disjoint adjacent transpositions moves each point by at most 1. For the converse we induct on n; for $n \leq 1$ the only permutation is the identity, the empty product. Let $n \geq 2$ and let σ satisfy $|\sigma(j) - j| \leq 1$ for every j. Then $\sigma(n) \in \{n-1, n\}$.

If $\sigma(n) = n$, then σ maps $[n-1]$ onto itself, and its restriction has the same property, so by induction it is a product of disjoint adjacent transpositions of $[n-1]$, and so is σ.

If $\sigma(n) = n-1$, let $j = \sigma^{-1}(n)$. Then $|n-j| \leq 1$ and $j \neq n$, so $j = n-1$, that is, $\sigma(n-1) = n$. For $j \leq n-2$ we have $\sigma(j) \leq j+1 \leq n-1$ and $\sigma(j) \neq n-1$, so σ maps $[n-2]$ onto itself. Its restriction to $[n-2]$ has the property, so by induction it is a product ρ of disjoint adjacent transpositions of $[n-2]$, and $\sigma = \rho(n-1\ n)$, where $(n-1\ n)$ is disjoint from the factors of ρ.

(b) Let N_n be the number of such permutations of $[n]$, with $N_0 = N_1 = 1$. The two cases of (a) give $N_n = N_{n-1} + N_{n-2}$ for $n \geq 2$, since in each case the restriction can be any of the permutations counted for $[n-1]$, respectively $[n-2]$. So $N_n = F_{n+1}$, by induction from $N_0 = F_1$ and $N_1 = F_2$.

(c) The term of σ in the Leibniz formula (Theorem 5.3.3) contains $t_{\sigma(j)j}$, which is 0 unless $|\sigma(j) - j| \leq 1$. So only the permutations of (a) contribute. Such a σ is the product of the transpositions $(k\ k+1)$ for the pairs of a matching P, and different matchings give different permutations, so the surviving terms correspond to the matchings. For the σ of P, $\operatorname{sgn}(\sigma) = (-1)^{|P|}$ by Proposition 5.2.25. Each unmatched i contributes $t_{ii} = a_i$, and each pair $\{k, k+1\} \in P$ contributes $t_{\sigma(k)k}t_{\sigma(k+1),k+1} = t_{k+1,k}t_{k,k+1} = c_k b_k$. So the term of σ is $\prod_P(-b_k c_k)\prod_{\text{unmatched}} a_i$, as claimed.

(d) Let d_m be the sum in (c) for T_m, the upper left $m \times m$ corner. Split the matchings of $[m]$ according to whether m is matched. If it is unmatched, the rest is any matching of $[m-1]$, and these terms add up to $a_m d_{m-1}$. If it is matched, its pair is $\{m-1, m\}$, the rest is any matching of $[m-2]$, and these terms add up to $-b_{m-1}c_{m-1}d_{m-2}$. So $d_m = a_m d_{m-1} - b_{m-1}c_{m-1}d_{m-2}$, the recurrence of Proposition 5.5.18.

In Example 5.5.19 (b), $a_k = 1$ and $-b_k c_k = -i \cdot i = 1$, so every term in (c) equals 1, and $\det T$ is the number of matchings of $[n]$, which is F_{n+1} by (b). The cofactor recurrence computes the determinants; the matching formula explains what they count.

Solution to Problem 5.15

Uses. Proposition 5.3.20 (Reduction modulo p), Exercise 5.6.13, Corollary 5.4.27 (Several diagonal blocks), Example 5.4.24 (A matrix with constant row sums)
Techniques. Reduce modulo p (Technique 5.3.21), Cancel terms in pairs (Technique 5.3.5)
Hint 1. Reduce modulo 2. What do the hypotheses say about the reduction $\overline{A}$ over $\mathbb{F}_2$?
Hint 2. $\overline{A}$ is alternating in the sense of Exercise 5.6.13. For (b), start with $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.
Sketch. (a) Over $\mathbb{F}_2$, $\overline{a}_{ii} = 0$ and $\overline{a}_{ji} = \overline{a}_{ij} = -\overline{a}_{ij}$, so $\overline{A}$ is alternating of odd size and $\det \overline{A} = 0$; hence $\det A$ is even. (b) Block-diagonal matrices with blocks $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ satisfy the hypotheses and have determinant ± 1. (c) I_n has odd determinant; $J + I_n$ has determinant $n+1$.

(a) Let $\overline{A} \in \mathbb{F}_2^{n\times n}$ be the reduction of A modulo 2. Since a_{ii} is even, $\bar{a}_{ii} = 0$. Since $a_{ij} + a_{ji}$ is even, $\bar{a}_{ji} = -\bar{a}_{ij}$ in $\mathbb{F}_2$. So $\overline{A}$ is alternating in the sense of Exercise 5.6.13, and since n is odd, $\det \overline{A} = 0$ there. By Proposition 5.3.20, $(\det A) \bmod 2 = \det \overline{A} = 0$, so $\det A$ is even.

The argument must work in characteristic 2, which is why it uses the pairing of Exercise 5.6.13 and not the identity $\det A = -\det A$: over $\mathbb{F}_2$, $-\det A = \det A$ carries no information.

(b) Let $n = 2k$ and let A be block diagonal with k blocks $R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Its diagonal entries are 0, and $a_{ij} + a_{ji}$ is 0 or 2, so A satisfies the hypotheses. By Corollary 5.4.27, $\det A = (\det R)^k = (-1)^k$, which is odd.

(c) For odd n, I_n has $a_{ij} + a_{ji} = 0$ for $i \neq j$ but odd diagonal entries, and $\det I_n = 1$ is odd. So the diagonal hypothesis is needed (it is not implied by the other one, which for $i = j$ only says that $2a_{ii}$ is even). The matrix $J + I_n$ has diagonal entries 2 and $a_{ij} + a_{ji} = 2$, and by Example 5.4.24 with $x = 2$, $\det(J + I_n) = (n + 1) \cdot 1^{n-1} = n + 1$, which is even for odd n, as (a) predicts.

Solution to Problem 5.16

Uses. Proposition 2.4.23 (Transpose of a product), Corollary 5.4.8 (Vanishing determinants), Theorem 5.4.7 (Invertibility and the determinant), Proposition 5.3.19 ($\det(A + tI_n)$ as a function of t), Proposition 5.3.15 (Triangular matrices)
Techniques. det is a polynomial in the entries (Technique 5.3.18)
Hint 1. A 1×1 matrix equals its own transpose. For invertibility, if $(I_n + tA)\mathbf{x} = \mathbf{0}$, multiply on the left by $\mathbf{x}^\mathsf{T}$.
Hint 2. $g(t) = \det(I_n + tA)$ is a polynomial in t, so it is continuous, it is 1 at $t = 0$, and by (a) it never vanishes.
Sketch. (a) $\mathbf{x}^\mathsf{T}A\mathbf{x} = \left(\mathbf{x}^\mathsf{T}A\mathbf{x}\right)^\mathsf{T} = -\mathbf{x}^\mathsf{T}A\mathbf{x}$, so it is 0; then $\mathbf{x}^\mathsf{T}(I_n + tA)\mathbf{x} = \sum x_i^2$ forces $\mathbf{x} = \mathbf{0}$. (b) A continuous function on $[0, 1]$ with $g(0) = 1$ and no zeros is positive at 1. (c) $1 + p^2$, and $1 + p^2 + q^2 + r^2$. (d) $A = \operatorname{diag}(-2, 2)$.

(a) The 1×1 matrix $\mathbf{x}^\mathsf{T}A\mathbf{x}$ equals its transpose, which by Proposition 2.4.23 is $\mathbf{x}^\mathsf{T}A^\mathsf{T}\mathbf{x} = -\mathbf{x}^\mathsf{T}A\mathbf{x}$. So $2\mathbf{x}^\mathsf{T}A\mathbf{x} = 0$, and $\mathbf{x}^\mathsf{T}A\mathbf{x} = 0$. Now suppose $(I_n + tA)\mathbf{x} = \mathbf{0}$. Then

$$0 = \mathbf{x}^\mathsf{T}(I_n + tA)\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{x} + t\mathbf{x}^\mathsf{T}A\mathbf{x} = x_1^2 + \cdots + x_n^2,$$

so $\mathbf{x} = \mathbf{0}$, and $I_n + tA$ is invertible by Corollary 5.4.8.

(b) Let $g(t) = \det(I_n + tA)$. Its entries are affine in t, so as in the proof of Proposition 5.3.19, $g(t) = e_0 + e_1 t + \cdots + e_n t^n$ for real constants e_k. In particular g is continuous, $g(0) = \det I_n = 1$, and by (a) and Theorem 5.4.7, $g(t) \neq 0$ for every t. If $g(1) < 0$, the intermediate value theorem would give $t \in (0, 1)$ with $g(t) = 0$. So $\det(I_n + A) = g(1) > 0$.

(c) For $n = 2$, $A = \begin{pmatrix} 0 & p \\ -p & 0 \end{pmatrix}$ and $\det(I_2 + A) = \det\begin{pmatrix} 1 & p \\ -p & 1 \end{pmatrix} = 1 + p^2$. For $n = 3$, $A = \begin{pmatrix} 0 & p & q \\ -p & 0 & r \\ -q & -r & 0 \end{pmatrix}$, and expanding along the first row,

$$\det\begin{pmatrix} 1 & p & q \\ -p & 1 & r \\ -q & -r & 1 \end{pmatrix} = \left(1 + r^2\right) - p(-p + qr) + q(pr + q)$$
$$= 1 + p^2 + q^2 + r^2.$$

Both are at least 1. In fact $\det(I_n + A) \geq 1$ for every real skew-symmetric A, but the proof needs the spectral theory of Chapter 8.

(d) $A = \operatorname{diag}(-2, 2)$ has trace 0, and $\det(I_2 + A) = (-1) \cdot 3 = -3 < 0$ (Proposition 5.3.15). Here $I_2 + tA$ is singular at $t = 1/2$, and the argument of (b) breaks at exactly that point.

Solution to Problem 5.17

Uses. Theorem 4.1.10 (Independent lists are short), Corollary 5.4.8 (Vanishing determinants), Exercise 5.5.48, Theorem 7.1.32 (Fundamental Theorem of Algebra), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. det is a polynomial in the entries (Technique 5.3.18)
Hint 1. For (a), if $\dim W > n$, a basis $X_1, \ldots, X_m$ gives more than n vectors $X_k\mathbf{e}_1$ in F^n. A relation among them produces a nonzero singular matrix in W.

Hint 2. For (c), $\det B \neq 0$, and $\det(A + tB)$ is a polynomial in t of degree n with leading coefficient $\det B$. Over $\mathbb{C}$ it has a root. For (d), use that a real polynomial of odd degree has a real root.
Sketch. (a) If $m = \dim W > n$, then $\sum_k c_k X_k \mathbf{e}_1 = \mathbf{0}$ nontrivially, and $X = \sum_k c_k X_k$ is a nonzero matrix of W with $X\mathbf{e}_1 = \mathbf{0}$. (b) $\det = a^2 + b^2$. (c) For independent $A, B \in W$, $B \neq 0$, so $\det B \neq 0$, and $\det(A + tB)$ has degree $n \geq 1$ in t; a root $t_0 \in \mathbb{C}$ gives the nonzero singular $A + t_0 B \in W$. (d) The same with a real root of an odd-degree polynomial.

(a) Suppose $m = \dim W > n$, and let $X_1, \ldots, X_m$ be a basis of W. The m vectors $X_1\mathbf{e}_1, \ldots, X_m\mathbf{e}_1$ lie in F^n, so they are dependent (Theorem 4.1.10): $\sum_k c_k X_k \mathbf{e}_1 = \mathbf{0}$ with not all $c_k = 0$. Let $X = \sum_k c_k X_k \in W$. Since the X_k are independent, $X \neq 0$; but $X\mathbf{e}_1 = \mathbf{0}$, so X is not invertible (Corollary 5.4.8). This contradicts the hypothesis, so $\dim W \leq n$.

(b) $W = \operatorname{span}(I_2, R)$ with $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so it is a subspace, and I_2, R are independent, so $\dim W = 2$. For $X = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, $\det X = a^2 + b^2$, which is positive unless $a = b = 0$. So every nonzero $X \in W$ is invertible.

(c) Suppose $\dim W \geq 2$, and let $A, B \in W$ be independent. Then $B \neq 0$, so B is invertible and $\det B \neq 0$. By Exercise 5.5.48 (a), there are $e_0, \ldots, e_n \in \mathbb{C}$ with $\det(A + tB) = e_0 + e_1 t + \cdots + e_n t^n$ for all t, and $e_n = \det B \neq 0$. So $e_0 + e_1 x + \cdots + e_n x^n$ has degree $n \geq 2$, and it has a root $t_0 \in \mathbb{C}$ (Theorem 7.1.32). Then $A + t_0 B \in W$ is singular, and it is nonzero because A and B are independent. This contradicts the hypothesis, so $\dim W \leq 1$. The bound is attained by $W = \operatorname{span}(I_n)$.

(d) For $n = 1$ this is (a). For odd $n \geq 3$, argue as in (c): with A, B independent in W, $g(t) = \det(A + tB) = e_0 + \cdots + e_n t^n$ with real e_k and $e_n = \det B \neq 0$. Since n is odd, $g(t)/t^n \to e_n$ as $t \to \pm\infty$ and t^n changes sign, so g takes both signs, and by the intermediate value theorem $g(t_0) = 0$ for some real t_0. Then $A + t_0 B$ is a nonzero singular matrix in W, a contradiction.

So the largest possible $\dim W$ depends on the field: 1 over $\mathbb{C}$, 1 over $\mathbb{R}$ for odd n, but 2 over $\mathbb{R}$ for $n = 2$, by (b). (For $n = 4$ over $\mathbb{R}$, a construction from the quaternions reaches $\dim W = 4$.)

Solution to Problem 5.18

Uses. Lemma 5.1.8 (Column replacement), Definition 5.1.1 (Determinant function), Proposition 5.3.15 (Triangular matrices)
Techniques. Polynomial identity arguments (Technique 5.5.24), det is a polynomial in the entries (Technique 5.3.18)
Hint 1. Every column of $T_n + tJ$ is a column of T_n plus $t(1, \ldots, 1)$. Subtract column 1 from the others.
Hint 2. An affine function $\alpha + \beta t$ is determined by two of its values. At $t = -a$ and at $t = -b$ the matrix $T_n + tJ$ is triangular.
Sketch. After subtracting column 1 from the others, only column 1 involves t, linearly, so $f(t) = \det(T_n + tJ) = \alpha + \beta t$. $T_n - aJ$ is lower triangular with diagonal $x - a$, and $T_n - bJ$ is upper triangular with diagonal $x - b$, so $f(-a) = (x - a)^n$ and $f(-b) = (x - b)^n$. Solve for $f(0) = \alpha$.

Let $\mathbf{1} = (1, \ldots, 1)$ and let $\mathbf{t}_1, \ldots, \mathbf{t}_n$ be the columns of T_n, so the columns of $T_n + tJ$ are $\mathbf{t}_j + t\mathbf{1}$. Subtracting column 1 from each other column does not change the determinant (Lemma 5.1.8), and leaves the columns $\mathbf{t}_1 + t\mathbf{1}, \mathbf{t}_2 - \mathbf{t}_1, \ldots, \mathbf{t}_n - \mathbf{t}_1$, in which only the first involves t. By linearity in column 1,

$$\begin{aligned} f(t) &= \det(T_n + tJ) \\ &= \det(\mathbf{t}_1, \mathbf{t}_2 - \mathbf{t}_1, \ldots) + t \det(\mathbf{1}, \mathbf{t}_2 - \mathbf{t}_1, \ldots) \\ &= \alpha + \beta t \end{aligned}$$

for constants $\alpha, \beta \in F$. Note $f(0) = \det T_n = \alpha$.

The matrix $T_n - aJ$ has diagonal entries $x - a$, entries 0 above the diagonal, and $b - a$ below it, so it is lower triangular and $f(-a) = (x - a)^n$ (Proposition 5.3.15). Likewise $T_n - bJ$ is upper triangular with diagonal entries $x - b$, so $f(-b) = (x - b)^n$. So

$$\alpha - a\beta = (x - a)^n, \quad \alpha - b\beta = (x - b)^n.$$

Multiplying the first equation by b, the second by a, and subtracting, $(b - a)\alpha = b(x - a)^n - a(x - b)^n$. Since $a \neq b$,

$$\det T_n = \alpha = \frac{b(x-a)^n - a(x-b)^n}{b-a}.$$

For $n = 2$ the formula gives

$$\begin{aligned}\frac{b(x-a)^2 - a(x-b)^2}{b-a} &= \frac{bx^2 + a^2 b - ax^2 - ab^2}{b-a} \\ &= \frac{(b-a)x^2 - ab(b-a)}{b-a} = x^2 - ab,\end{aligned}$$

which is $\det\begin{pmatrix} x & a \\ b & x \end{pmatrix}$. The proof never multiplied anything out: it located two values of t at which the matrix becomes triangular and used that f is determined by two values. When $a = b$ the formula breaks down, and the answer is $(x-a)^{n-1}(x + (n-1)a)$ (compare Example 5.4.24).

Solution to Problem 5.19

Uses. Theorem 5.5.21 (The Vandermonde determinant), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.15 (Triangular matrices), Definition 1.2.17 (Evaluation), Lemma 5.5.23 (A polynomial with prescribed roots), Proposition 5.3.20 (Reduction modulo p), Exercise 5.5.26
Techniques. Expose a factor with row and column operations (Technique 5.5.22), Factor the matrix (Technique 5.6.1)
Hint 1. For (a), column $k+1$ is column $k+1$ of the Vandermonde matrix plus a combination of earlier columns. Write the matrix as VU with U upper triangular.
Hint 2. For (b), take $p_k(t) = t(t-1)\cdots(t-k+1)$, whose values at integers $x \geq 0$ are $k!\binom{x}{k}$. Translating all x_i by the same integer does not change the product, so you may assume $x_i \geq 0$.
Sketch. (a) $p_k = x^k + \sum_{l<k} u_{lk} x^l$, so the matrix is VU with U upper triangular with 1s on the diagonal; $\det = \det V$. (b) Shift so that all $x_i \geq 0$ and use the falling factorials: the matrix has entries $k!\binom{x_i}{k}$, so $\det V = \left(\prod_k k!\right)\det\left(\binom{x_i}{k}\right)$, and the last determinant is an integer. (c) $\prod_{i<j}(j-i) = 1!2!\cdots(n-1)!$.

(a) Write $p_k = \sum_{l=0}^{k} u_{lk} x^l$ with $u_{kk} = 1$, and set $u_{lk} = 0$ for $l > k$. Let $U \in F^{n\times n}$ have $(l+1, k+1)$ entry u_{lk}; it is upper triangular with 1s on the diagonal, so $\det U = 1$ (Proposition 5.3.15). Let $V = V(x_1, \ldots, x_n)$, with $(i, l+1)$ entry x_i^l. By Definition 1.2.17,

$$p_k(x_i) = \sum_{l=0}^{n-1} x_i^l u_{lk} = (VU)_{i,k+1}.$$

So the matrix in question is VU, and its determinant is $\det V \det U = \det V = \prod_{i<j}(x_j - x_i)$ by Theorem 5.4.10 and Theorem 5.5.21.

(b) Replacing every x_i by $x_i + c$, for an integer c, does not change any difference $x_j - x_i$. So we may assume that every $x_i \geq 0$. For $k \geq 0$, multiplying out the product one factor at a time (as in the proof of Lemma 5.5.23) shows that there is a polynomial $p_k \in \mathbb{Q}[x]$ of degree k, with coefficient 1 of x^k, such that

$$p_k(t) = t(t-1)\cdots(t-k+1) \quad \text{for every } t,$$

with $p_0 = 1$. For an integer $t \geq 0$, $p_k(t) = k!\binom{t}{k}$: if $t \geq k$ this is the formula for $\binom{t}{k}$, and if $t < k$ both sides are 0. By (a),

$$\begin{aligned}\prod_{i<j}(x_j - x_i) &= \det(p_k(x_i)) = \det\left(k!\binom{x_i}{k}\right) \\ &= \left(\prod_{k=0}^{n-1} k!\right)\det\left(\binom{x_i}{k}\right),\end{aligned}$$

pulling the factor $k!$ out of column $k+1$. The matrix $\left(\binom{x_i}{k}\right)$ has integer entries, so its determinant is an integer (Proposition 5.3.20), and $0! \cdot 1! \cdots (n-1)!$ divides the product.

(c) By Exercise 5.5.26 (a), $\prod_{1\le i<j\le n}(j-i) = \det V(1,\dots,n) = 1!\cdot 2!\cdots(n-1)!$. So the quotient in (c) is $\prod_{i<j}(x_j - x_i)$ divided by $0!\cdot 1!\cdots(n-1)!$, which is an integer by (b). For $n=3$ this says that among three integers some two have the same parity; for $n=4$, that the six differences of four integers have a product divisible by 12.

Solution to Problem 5.20

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Proposition 5.3.15 (Triangular matrices), Corollary 5.5.12 (Integer inverses), Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Factor the matrix (Technique 5.6.1)
Hint 1. Apply the identity to $m = \gcd(i,j)$. The divisors of $\gcd(i,j)$ are exactly the common divisors of i and j.
Hint 2. L is lower triangular with 1s on the diagonal. For (c), when is $\prod_k \varphi(k) = 1$?
Sketch. (a) $\left(L\Phi L^{\mathsf T}\right)_{ij} = \sum_{d\mid i,\, d\mid j}\varphi(d) = \sum_{d\mid \gcd(i,j)}\varphi(d) = \gcd(i,j)$. (b) $\det L = 1$, so $\det G_n = \det\Phi$; $\det G_6 = 1\cdot 1\cdot 2\cdot 2\cdot 4\cdot 2 = 32$. (c) $\varphi(d)\ge 1$, so $\det G_n \neq 0$; the inverse is integral iff $\det G_n = \pm 1$, iff $\varphi(k)=1$ for all $k\le n$, iff $n\le 2$.

(a) The (i,j) entry of $L\Phi L^{\mathsf T}$ is

$$\sum_{d=1}^{n} l_{id}\varphi(d)l_{jd} = \sum_{d\le n,\, d\mid i,\, d\mid j}\varphi(d).$$

A common divisor of i and j is at most $i\le n$, so the condition $d\le n$ is automatic, and the common divisors of i and j are exactly the divisors of $\gcd(i,j)$. By the given identity with $m=\gcd(i,j)$, the sum is $\gcd(i,j) = g_{ij}$.

(b) If $d\mid i$ then $d\le i$, so L is lower triangular, and $l_{ii}=1$. So $\det L = \det L^{\mathsf T} = 1$ (Proposition 5.3.15, Theorem 5.4.14), and by Theorem 5.4.10,

$$\det G_n = \det L\det\Phi\det L^{\mathsf T} = \varphi(1)\varphi(2)\cdots\varphi(n).$$

Since $\varphi(1),\dots,\varphi(6) = 1,1,2,2,4,2$, we get $\det G_6 = 32$.

(c) Every $\varphi(d)\ge 1$, since $k=1$ is counted, so $\det G_n \ge 1$ and G_n is invertible (Theorem 5.4.7). By Corollary 5.5.12, G_n^{-1} has integer entries exactly when $\det G_n = \pm 1$, that is, when $\varphi(k)=1$ for every $k\le n$. Since $\varphi(1)=\varphi(2)=1$ and $\varphi(3)=2$, this happens exactly for $n\le 2$.

The factorization is the whole proof. Computing $\det G_n$ by elimination for each n gives no hint of the pattern, while the identity $m=\sum_{d\mid m}\varphi(d)$ says that each entry is a sum of products, which is what Technique 5.6.1 asks for. (This is H. J. S. Smith's determinant, from 1875.)

Section 6.1

Solution to Exercise 6.1.12

Uses. Definition 6.1.8 (Dot product), Proposition 6.1.10 (The dot product is an inner product)
Hint 1. Conjugate the coordinates of the second vector, multiply coordinatewise, and add.
Hint 2. For the last part, expand $(\mathbf{v}-c\mathbf{w})\cdot\mathbf{w}$ by linearity in the first slot; it is a linear equation in c.
Sketch. $\mathbf{v}\cdot\mathbf{w} = -i + (1+i) + (2+2i) = 3+2i$, and $\mathbf{w}\cdot\mathbf{v} = 3-2i$. The squared lengths are $1+1+4=6$ and $1+2+2=5$. Linearity gives $(\mathbf{v}-c\mathbf{w})\cdot\mathbf{w} = (3+2i) - 5c$, so $c = (3+2i)/5$.

By Definition 6.1.8,

$$\mathbf{v}\cdot\mathbf{w} = 1\cdot\bar{i} + i\cdot\overline{(1+i)} + 2\cdot\overline{(1-i)} = -i + i(1-i) + 2(1+i) = -i + (i+1) + (2+2i) = 3+2i,$$

and by conjugate symmetry $\mathbf{w}\cdot\mathbf{v} = \overline{3+2i} = 3-2i$. Directly,

$$\mathbf{w}\cdot\mathbf{v} = i\cdot 1 + (1+i)(-i) + (1-i)2 = i + (1-i) + (2-2i) = 3-2i.$$

By Proposition 6.1.10, $\mathbf{v}\cdot\mathbf{v} = 1+1+4 = 6$ and $\mathbf{w}\cdot\mathbf{w} = 1 + |1+i|^2 + |1-i|^2 = 1+2+2 = 5$.

By linearity in the first slot, $(\mathbf{v}-c\mathbf{w})\cdot\mathbf{w} = \mathbf{v}\cdot\mathbf{w} - c(\mathbf{w}\cdot\mathbf{w}) = (3+2i) - 5c$. This vanishes exactly for $c = (3+2i)/5$.

Solution to Exercise 6.1.20

Uses. Corollary 6.1.19 ($A^*A = 0$ forces $A = 0$)
Techniques. The trace of A^*A (Technique 6.1.18)
Hint 1. Rewrite A^2 using the hypothesis $A^* = A$.
Hint 2. $A^*A = AA = 0$. Now use Corollary 6.1.19.
Sketch. $A^*A = A^2 = 0$, so $A = 0$ by Corollary 6.1.19. Without $A^* = A$ the claim fails for $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Since $A^* = A$, we have $A^*A = AA = A^2 = 0$, so $A = 0$ by Corollary 6.1.19. The hypothesis $A^* = A$ cannot be dropped: $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ satisfies $N^2 = 0$ but $N \neq 0$. (Indeed $N^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq N$.)

Solution to Exercise 6.1.28

Uses. Proposition 6.1.27 (Basic properties of the norm)
Hint 1. By Proposition 6.1.27(b), the condition only involves $|a|$.
Hint 2. $\|a\mathbf{v}\| = 1$ if and only if $|a| = 1/\|\mathbf{v}\|$. How many real numbers, and which complex numbers, have a given modulus?
Sketch. $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$, so the condition is $|a| = 1/\|\mathbf{v}\|$. Over $\mathbb{R}$ this gives $a = \pm 1/\|\mathbf{v}\|$; over $\mathbb{C}$ it gives the circle of radius $1/\|\mathbf{v}\|$ centred at 0.

By Proposition 6.1.27, $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$, and $\|\mathbf{v}\| > 0$ since $\mathbf{v} \neq \mathbf{0}$. So $\|a\mathbf{v}\| = 1$ if and only if $|a| = 1/\|\mathbf{v}\|$. Over $\mathbb{R}$ there are exactly two such scalars, $a = 1/\|\mathbf{v}\|$ and $a = -1/\|\mathbf{v}\|$. Over $\mathbb{C}$ there are infinitely many: all $a = s + ti$ with $s^2 + t^2 = 1/\|\mathbf{v}\|^2$, a circle of radius $1/\|\mathbf{v}\|$ about 0 in the complex plane. So a nonzero vector in a complex space has a whole circle of normalizations, which differ by factors of modulus 1.

Solution to Exercise 6.1.31

Uses. Proposition 6.1.30 (Parallelogram law), Proposition 6.1.29 (Expanding a norm)
Hint 1. The parallelogram law relates the four given or unknown norms.
Hint 2. For the real part, solve the expansion of $\|\mathbf{x} + \mathbf{y}\|^2$ for $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$.
Sketch. $\|\mathbf{x} - \mathbf{y}\|^2 = 2 \cdot 9 + 2 \cdot 16 - 36 = 14$, and $36 = 9 + 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle + 16$ gives $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle = 11/2$.

By Proposition 6.1.30,

$$\|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 - \|\mathbf{x} + \mathbf{y}\|^2 = 18 + 32 - 36 = 14,$$

so $\|\mathbf{x} - \mathbf{y}\| = \sqrt{14}$. By Proposition 6.1.29, $36 = 9 + 2\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle + 16$, so $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle = 11/2$. (As a check, Proposition 6.1.29 with the minus sign gives $9 - 11 + 16 = 14$.)

Solution to Exercise 6.1.35

Uses. Proposition 6.1.32 (Polarization identities), Proposition 6.1.30 (Parallelogram law)
Techniques. Polarization (Technique 6.1.33)
Hint 1. Substitute into Proposition 6.1.32(b).
Hint 2. For $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, use the parallelogram law, for the pair $\mathbf{x}, \mathbf{y}$ or for the pair $\mathbf{x}, i\mathbf{y}$.
Sketch. $\langle \mathbf{x}, \mathbf{y}\rangle = (10 - 2 + 4i - 8i)/4 = 2 - i$. The parallelogram law gives $2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right) = 10 + 2 = 12$, so the sum is 6.

By Proposition 6.1.32(b),

$$\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}(10 - 2 + 4i - 8i) = \frac{1}{4}(8 - 4i) = 2 - i.$$

By Proposition 6.1.30, $2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 = 10 + 2 = 12$, so $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = 6$. The data are consistent: the parallelogram law for $\mathbf{x}$ and $i\mathbf{y}$, where $\|i\mathbf{y}\| = \|\mathbf{y}\|$, gives $4 + 8 = 12$ as well.

Solution to Exercise 6.1.42

Uses. Definition 6.1.2 (Inner product), Proposition 6.1.22 (Inner products on $\mathbb{F}^n$), Proposition 6.1.4 (Sesquilinearity), Definition 1.2.17 (Evaluation)
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. (a) and (e) have the form $\mathbf{y}^* M \mathbf{x}$; write down M and use Proposition 6.1.22. To show $\mathbf{x}^* M \mathbf{x} > 0$, complete squares.
Hint 2. (b) Look for a nonzero matrix with $\langle A, A\rangle = 0$. (c) Compute $\langle \mathbf{e}_2, \mathbf{e}_2\rangle$.
Hint 3. (d) A polynomial $a + bx + cx^2$ that vanishes at -1, 0, and 1: solve for a, b, c. (e) Try to write $\mathbf{x}^* M \mathbf{x}$ as $|x_1|^2 + |x_2 + ix_1|^2$.

Sketch. (a) Yes: M is symmetric and $\mathbf{x}^{\mathsf{T}}M\mathbf{x} = (x_1+x_3)^2 + 2x_2^2 + 2x_3^2$. (b) No: $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \neq 0$ has $\langle A, A\rangle = 0$. (c) No: $\langle \mathbf{e}_2, \mathbf{e}_2\rangle = i$ is not real. (d) Yes: it is symmetric, bilinear, a sum of squares on the diagonal, and a polynomial of degree at most 2 vanishing at three points is 0. (e) Yes: $M = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix}$ is self-conjugate-transpose and $\mathbf{x}^*M\mathbf{x} = |x_1|^2 + |x_2 + ix_1|^2$.

(a) Yes. The formula is $\mathbf{y}^{\mathsf{T}}M\mathbf{x}$ with $M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$, which is symmetric, and

$$\mathbf{x}^{\mathsf{T}}M\mathbf{x} = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_3 = (x_1+x_3)^2 + 2x_2^2 + 2x_3^2.$$

This is a sum of nonnegative terms, and it is 0 only if $x_2 = x_3 = 0$ and $x_1 + x_3 = 0$, that is, $\mathbf{x} = \mathbf{0}$. By Proposition 6.1.22(b), it is an inner product.

(b) No. Definiteness fails: $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is nonzero, but $\langle A, A\rangle = \operatorname{tr}(A)^2 = 0$. (The formula is symmetric, bilinear, and positive.)

(c) No. $\langle \mathbf{e}_2, \mathbf{e}_2\rangle = i$ is not a nonnegative real number, so positivity fails. (Conjugate symmetry fails as well, since $\bar{i} \neq i$.)

(d) Yes. Evaluation at a point is linear (Definition 1.2.17), so the formula is linear in p, and it is visibly symmetric. Also $\langle p, p\rangle = p(-1)^2 + p(0)^2 + p(1)^2 \geq 0$. If $\langle p, p\rangle = 0$, each term vanishes, so $p = a + bx + cx^2$ satisfies $p(0) = a = 0$, $p(1) = a + b + c = 0$, and $p(-1) = a - b + c = 0$. Then $b + c = 0$ and $-b + c = 0$, so $b = c = 0$ and $p = \mathbf{0}$.

(e) Yes. Reading off the coefficient of $\overline{y_i}x_j$ as m_{ij}, the formula is $\mathbf{y}^*M\mathbf{x}$ with $M = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix}$. The conjugate transpose of M is $\begin{pmatrix} \overline{2} & \overline{i} \\ \overline{-i} & \overline{1} \end{pmatrix} = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix} = M$. Also,

$$\mathbf{x}^*M\mathbf{x} = 2|x_1|^2 - i\overline{x_1}x_2 + ix_1\overline{x_2} + |x_2|^2.$$

On the other hand, by Proposition 6.1.29 in $\mathbb{C}^1$ (or by direct expansion), $|x_2 + ix_1|^2 = |x_2|^2 + ix_1\overline{x_2} - i\overline{x_1}x_2 + |x_1|^2$. Hence $\mathbf{x}^*M\mathbf{x} = |x_1|^2 + |x_2 + ix_1|^2$, which is 0 only if $x_1 = 0$ and then $x_2 = 0$. By Proposition 6.1.22(b), this is an inner product.

Solution to Exercise 6.1.43

Uses. Proposition 6.1.13 (Integral inner products), Lemma 6.1.14 (Nonnegative functions with zero integral)
Hint 1. In each case symmetry and bilinearity come from the integral. Only positivity and definiteness need thought.
Hint 2. (b) Find f that is large where $t < 0$ and small where $t > 0$. (c) Find a nonzero continuous function that vanishes on $[0, 1/2]$.
Hint 3. For (a) and (d), use Lemma 6.1.14 (on the relevant interval), then continuity or the mean value theorem.
Sketch. (a) Yes: $t^2 f(t)^2 = 0$ for all t forces $f = 0$ off 0, and then at 0 by continuity. (b) No: $f(t) = 1 - t$ gives $\langle f, f\rangle = -4/3$. (c) No: $f(t) = \max(0, t - 1/2)$ is nonzero with $\langle f, f\rangle = 0$. (d) Yes: $\langle f, f\rangle = 0$ gives $f(0) = 0$ and $f' = 0$, so $f = 0$.

In all four cases the formula is symmetric and bilinear, by linearity of the integral and of $f \mapsto f'$. So only positivity and definiteness are at stake.

(a) Yes. $\langle f, f\rangle = \int_{-1}^{1} t^2 f(t)^2 \, dt \geq 0$, as the integrand is continuous and nonnegative. If $\langle f, f\rangle = 0$, then Lemma 6.1.14 (on $[-1, 1]$) gives $t^2 f(t)^2 = 0$ for all t, so $f(t) = 0$ for all $t \neq 0$. By continuity, $f(0) = \lim_{t\to 0} f(t) = 0$ as well. So $f = 0$.

(b) No. Positivity fails for $f(t) = 1 - t$:

$$\langle f, f\rangle = \int_{-1}^{1} t(1-t)^2 \, dt = \int_{-1}^{1} \left(t - 2t^2 + t^3\right) dt = 0 - \frac{4}{3} + 0 = -\frac{4}{3} < 0.$$

(c) No. Definiteness fails: $f(t) = \max(0, t - 1/2)$ is continuous and nonzero, but $f = 0$ on $[0, 1/2]$, so $\langle f, f\rangle = 0$. (Positivity does hold.)

(d) Yes. $\langle f, f\rangle = f(0)^2 + \int_0^1 f'(t)^2\,dt \geq 0$. If $\langle f, f\rangle = 0$, both nonnegative terms vanish. So $f(0) = 0$, and Lemma 6.1.14 applied to the continuous function $(f')^2$ gives $f' = 0$ on $[0, 1]$. By the mean value theorem f is constant, so $f = f(0) = 0$. Without the term $f(0)g(0)$, definiteness would fail for the constant function 1.

Solution to Exercise 6.1.44

Uses. Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.29 (Expanding a norm)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Expand each inner product into the four values $\langle \mathbf{x}, \mathbf{x}\rangle = 1$, $\langle \mathbf{x}, \mathbf{y}\rangle = 1 - i$, $\langle \mathbf{y}, \mathbf{x}\rangle = 1 + i$, $\langle \mathbf{y}, \mathbf{y}\rangle = 4$. Scalars leave the second slot conjugated.
Hint 2. For (b) and (c), Proposition 6.1.29 saves work: $\|\mathbf{x} + a\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\mathrm{Re}(\bar{a}\langle \mathbf{x}, \mathbf{y}\rangle) + |a|^2\|\mathbf{y}\|^2$.
Sketch. (a) $2 - 2(1 - i) + i(1 + i) - 4i = -1 - i$. (b) $1 + 2\,\mathrm{Re}(i(1 - i)) + 4 = 7$. (c) $|1 + i|^2 + 2\,\mathrm{Re}((1 + i)(1 - i)) + 4 = 10$. (d) $c = (1 - i)/4$ and $\|\mathbf{x} - c\mathbf{y}\|^2 = 1 - 1 + 1/2 = 1/2$.

We use $\langle \mathbf{x}, \mathbf{x}\rangle = 1$, $\langle \mathbf{y}, \mathbf{y}\rangle = 4$, $\langle \mathbf{x}, \mathbf{y}\rangle = 1 - i$, and $\langle \mathbf{y}, \mathbf{x}\rangle = \overline{1 - i} = 1 + i$.

(a) Expanding,

$$\begin{aligned}\langle 2\mathbf{x} + i\mathbf{y}, \mathbf{x} - \mathbf{y}\rangle &= 2\langle \mathbf{x}, \mathbf{x}\rangle - 2\langle \mathbf{x}, \mathbf{y}\rangle + i\langle \mathbf{y}, \mathbf{x}\rangle - i\langle \mathbf{y}, \mathbf{y}\rangle \\ &= 2 - (2 - 2i) + (i - 1) - 4i = -1 - i.\end{aligned}$$

(b) By Proposition 6.1.29 with $a = -i$, $\bar{a} = i$, and $|a| = 1$,

$$\|\mathbf{x} - i\mathbf{y}\|^2 = 1 + 2\,\mathrm{Re}(i(1 - i)) + 4 = 1 + 2\,\mathrm{Re}(1 + i) + 4 = 7.$$

(c) Here the scalar sits on $\mathbf{x}$, so we expand directly:

$$\|(1 + i)\mathbf{x} + \mathbf{y}\|^2 = |1 + i|^2\langle \mathbf{x}, \mathbf{x}\rangle + (1 + i)\langle \mathbf{x}, \mathbf{y}\rangle + \overline{(1 + i)}\langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle,$$

which is

$$2 + (1 + i)(1 - i) + (1 - i)(1 + i) + 4 = 2 + 2 + 2 + 4 = 10.$$

(d) By linearity in the first slot, $\langle \mathbf{x} - c\mathbf{y}, \mathbf{y}\rangle = (1 - i) - 4c$, which vanishes exactly for $c = (1 - i)/4$. Then $\bar{c}\langle \mathbf{x}, \mathbf{y}\rangle = (1 + i)(1 - i)/4 = 1/2$ and $|c|^2 = 2/16 = 1/8$, so by Proposition 6.1.29 with $a = -c$,

$$\|\mathbf{x} - c\mathbf{y}\|^2 = 1 - 2 \cdot \frac{1}{2} + \frac{1}{8} \cdot 4 = \frac{1}{2}.$$

Solution to Exercise 6.1.45

Uses. Proposition 6.1.22 (Inner products on $\mathbb{F}^n$), Proposition 6.1.29 (Expanding a norm), Proposition 6.1.30 (Parallelogram law)
Techniques. Polarization (Technique 6.1.33)
Hint 1. By Proposition 6.1.22(a), an inner product on $\mathbb{R}^2$ is determined by $\langle \mathbf{e}_1, \mathbf{e}_1\rangle$, $\langle \mathbf{e}_2, \mathbf{e}_2\rangle$, and $c = \langle \mathbf{e}_1, \mathbf{e}_2\rangle$. Find c from $\|\mathbf{e}_1 + \mathbf{e}_2\|^2$.
Hint 2. In (a), you must also check that the matrix you find gives an inner product: complete the square.
Hint 3. In (c), the same computation forces a value of c. Find a vector $\mathbf{x}$ with $\langle \mathbf{x}, \mathbf{x}\rangle < 0$ for that value.
Sketch. (a) $4 = 1 + 2c + 4$ forces $c = -1/2$, so $\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 - (x_1y_2 + x_2y_1)/2 + 4x_2y_2$, which is an inner product because $x_1^2 - x_1x_2 + 4x_2^2 = (x_1 - x_2/2)^2 + 15x_2^2/4$. (b) Both give $\|\mathbf{e}_1 - \mathbf{e}_2\|^2 = 6$. (c) $16 = 5 + 2c$ forces $c = 11/2$, and then $(2, -1)$ would have $\langle \mathbf{x}, \mathbf{x}\rangle = 4 - 22 + 4 < 0$.

(a) Suppose $\langle \cdot, \cdot\rangle$ is such an inner product, and let $c = \langle \mathbf{e}_1, \mathbf{e}_2\rangle = \langle \mathbf{e}_2, \mathbf{e}_1\rangle$ (it is real and symmetric). By Proposition 6.1.29, $\|\mathbf{e}_1 + \mathbf{e}_2\|^2 = 1 + 2c + 4$, so $4 = 5 + 2c$ and $c = -1/2$. By Proposition 6.1.22(a), $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}}M\mathbf{x}$ with $M = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 4 \end{pmatrix}$, that is,

$$\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 - \frac{1}{2}(x_1y_2 + x_2y_1) + 4x_2y_2.$$

So there is at most one such inner product. Conversely, M is symmetric and

$$\mathbf{x}^{\mathsf{T}} M\mathbf{x} = x_1^2 - x_1x_2 + 4x_2^2 = \left(x_1 - \frac{x_2}{2}\right)^2 + \frac{15}{4}x_2^2,$$

which is positive unless $x_2 = 0$ and $x_1 = x_2/2 = 0$. By Proposition 6.1.22(b), this is an inner product, and it has $\|\mathbf{e}_1\|^2 = 1$, $\|\mathbf{e}_2\|^2 = 4$, and $\|\mathbf{e}_1 + \mathbf{e}_2\|^2 = 1 - 1 + 4 = 4$, as required.

(b) From the formula, $\|\mathbf{e}_1 - \mathbf{e}_2\|^2 = 1 - 2c + 4 = 1 + 1 + 4 = 6$. From Proposition 6.1.30, $\|\mathbf{e}_1 - \mathbf{e}_2\|^2 = 2 \cdot 1 + 2 \cdot 4 - 4 = 6$. So $\|\mathbf{e}_1 - \mathbf{e}_2\| = \sqrt{6}$.

(c) Suppose such an inner product exists. The argument of (a) gives $16 = 5 + 2c$, so $c = 11/2$, and $\langle \mathbf{x}, \mathbf{x}\rangle = x_1^2 + 11x_1x_2 + 4x_2^2$ for all $\mathbf{x}$. For $\mathbf{x} = (2, -1)$ this is $4 - 22 + 4 = -14 < 0$, contradicting positivity. (In Section 6.2 this also follows from the triangle inequality, as $4 > 1 + 2$.)

Solution to Exercise 6.1.46

Uses. Proposition 6.1.22 (Inner products on $\mathbb{F}^n$)
Techniques. Plug in clever test vectors (Technique 6.1.40), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. The formula is $\mathbf{y}^{\mathsf{T}} M\mathbf{x}$ for the symmetric matrix $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, so by Proposition 6.1.22 the question is when $q(\mathbf{x}) = ax_1^2 + 2bx_1x_2 + cx_2^2 > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
Hint 2. Necessary conditions come from test vectors: try $\mathbf{e}_1$, and then a vector that makes the cross term as harmful as possible, such as $(-b, a)$.
Hint 3. For sufficiency, when $a > 0$, complete the square in $aq(\mathbf{x})$.
Sketch. Exactly when $a > 0$ and $ac - b^2 > 0$. Necessity: $q(\mathbf{e}_1) = a$ and $q(-b, a) = a\big(ac - b^2\big)$. Sufficiency: $aq(\mathbf{x}) = (ax_1 + bx_2)^2 + \big(ac - b^2\big)x_2^2$.

The formula is $\mathbf{y}^{\mathsf{T}} M\mathbf{x}$ with $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, which is symmetric. By Proposition 6.1.22(b), it is an inner product if and only if $q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}} M\mathbf{x} = ax_1^2 + 2bx_1x_2 + cx_2^2$ is positive for all $\mathbf{x} \neq \mathbf{0}$. We show that this happens exactly when $a > 0$ and $ac - b^2 > 0$.

Suppose $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. Then $a = q(\mathbf{e}_1) > 0$. The vector $(-b, a)$ is nonzero since $a \neq 0$, and

$$q(-b, a) = ab^2 - 2b \cdot ba + ca^2 = a\big(ac - b^2\big),$$

which is positive; dividing by $a > 0$ gives $ac - b^2 > 0$.

Conversely, suppose $a > 0$ and $ac - b^2 > 0$. Completing the square,

$$aq(\mathbf{x}) = a^2x_1^2 + 2abx_1x_2 + acx_2^2 = (ax_1 + bx_2)^2 + \big(ac - b^2\big)x_2^2.$$

Both terms are nonnegative, so $aq(\mathbf{x}) \geq 0$, and it is 0 only if $x_2 = 0$ and $ax_1 + bx_2 = 0$, that is, $\mathbf{x} = \mathbf{0}$. Since $a > 0$, $q(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$.

The number $ac - b^2$ is $\det M$; Section 12.3 extends this test to $n \times n$ matrices.

Solution to Exercise 6.1.47

Uses. Notation 6.1.7 (Conjugate transpose), Notation 6.1.1 (Real and complex scalars)
Techniques. Plug in clever test vectors (Technique 6.1.40)
Hint 1. Compute $\mathbf{x}^* M\mathbf{x}$ for $\mathbf{x} = \mathbf{e}_j$, $\mathbf{e}_j + \mathbf{e}_k$, and $\mathbf{e}_j + i\mathbf{e}_k$.
Hint 2. With $\mathbf{x} = \mathbf{e}_j + \mathbf{e}_k$ you learn that $m_{jk} + m_{kj}$ is real; with $\mathbf{e}_j + i\mathbf{e}_k$, that $i\big(m_{jk} - m_{kj}\big)$ is real. What do these say about real and imaginary parts?
Sketch. $\mathbf{e}_j^* M\mathbf{e}_j = m_{jj}$ is real. For $j \neq k$, $\mathbf{x} = \mathbf{e}_j + \mathbf{e}_k$ gives $m_{jk} + m_{kj} \in \mathbb{R}$, so $\operatorname{Im} m_{jk} = -\operatorname{Im} m_{kj}$; and $\mathbf{x} = \mathbf{e}_j + i\mathbf{e}_k$ gives $i\big(m_{jk} - m_{kj}\big) \in \mathbb{R}$, so $\operatorname{Re} m_{jk} = \operatorname{Re} m_{kj}$. Together, $m_{jk} = \overline{m_{kj}}$.

For $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x}^* M\mathbf{x} = \sum_{p,q} \overline{x_p} m_{pq} x_q$. Fix j. For $\mathbf{x} = \mathbf{e}_j$ this is m_{jj}, so $m_{jj} \in \mathbb{R}$, that is, $m_{jj} = \overline{m_{jj}}$.

Now fix $j \neq k$. For $\mathbf{x} = \mathbf{e}_j + \mathbf{e}_k$, only the terms with $p, q \in \{j, k\}$ survive, and

$$\mathbf{x}^* M\mathbf{x} = m_{jj} + m_{jk} + m_{kj} + m_{kk}.$$

Since this and m_{jj}, m_{kk} are real, $m_{jk} + m_{kj} \in \mathbb{R}$, so $\operatorname{Im} m_{jk} = -\operatorname{Im} m_{kj}$. For $\mathbf{x} = \mathbf{e}_j + i\mathbf{e}_k$ we have $x_j = 1$ and $x_k = i$, so

$$\mathbf{x}^* M\mathbf{x} = m_{jj} + m_{jk}i + \bar{i}m_{kj} + \bar{i}m_{kk}i = m_{jj} + i\big(m_{jk} - m_{kj}\big) + m_{kk}.$$

So $i\big(m_{jk} - m_{kj}\big)$ is real, which means $m_{jk} - m_{kj}$ has real part 0: $\operatorname{Re} m_{jk} = \operatorname{Re} m_{kj}$. Together, $m_{jk} = \operatorname{Re} m_{kj} - i \operatorname{Im} m_{kj} = \overline{m_{kj}}$. Hence the (k, j) entry of M^*, which is $\overline{m_{jk}}$, equals m_{kj}, and $M^* = M$.

If $\mathbf{x}^* M\mathbf{x} > 0$ for every $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{x}^* M\mathbf{x}$ is real for every $\mathbf{x}$ (it is 0 at $\mathbf{0}$), so $M^* = M$ by what we just proved.

Solution to Exercise 6.1.48

Uses. Proposition 6.1.30 (Parallelogram law), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. Write $\mathbf{z} - \mathbf{x}$ and $\mathbf{z} - \mathbf{y}$ as $\mathbf{u} + \mathbf{w}$ and $\mathbf{u} - \mathbf{w}$ for suitable $\mathbf{u}, \mathbf{w}$.
Hint 2. Take $\mathbf{u} = \mathbf{z} - \mathbf{m}$ and $\mathbf{w} = (\mathbf{y} - \mathbf{x})/2$, then apply the parallelogram law.
Sketch. With $\mathbf{u} = \mathbf{z} - \mathbf{m}$ and $\mathbf{w} = (\mathbf{y} - \mathbf{x})/2$, we have $\mathbf{z} - \mathbf{x} = \mathbf{u} + \mathbf{w}$ and $\mathbf{z} - \mathbf{y} = \mathbf{u} - \mathbf{w}$. The parallelogram law gives $2\|\mathbf{u}\|^2 + 2\|\mathbf{w}\|^2$, and $2\|\mathbf{w}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2/2$.

Let $\mathbf{u} = \mathbf{z} - \mathbf{m}$ and $\mathbf{w} = (\mathbf{y} - \mathbf{x})/2$. Since $\mathbf{m} - \mathbf{x} = (\mathbf{y} - \mathbf{x})/2 = \mathbf{w}$ and $\mathbf{m} - \mathbf{y} = (\mathbf{x} - \mathbf{y})/2 = -\mathbf{w}$,

$$\mathbf{z} - \mathbf{x} = (\mathbf{z} - \mathbf{m}) + (\mathbf{m} - \mathbf{x}) = \mathbf{u} + \mathbf{w}, \quad \mathbf{z} - \mathbf{y} = \mathbf{u} - \mathbf{w}.$$

By Proposition 6.1.30 and Proposition 6.1.27(b),

$$\|\mathbf{z} - \mathbf{x}\|^2 + \|\mathbf{z} - \mathbf{y}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{w}\|^2 = 2\|\mathbf{z} - \mathbf{m}\|^2 + 2 \cdot \frac{1}{4}\|\mathbf{y} - \mathbf{x}\|^2,$$

which is the identity, since $\|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$.

In $\mathbb{R}^2$, $\mathbf{m}$ is the midpoint of the side from $\mathbf{x}$ to $\mathbf{y}$, and $\|\mathbf{z} - \mathbf{m}\|$ is the length of the median from $\mathbf{z}$. The identity expresses the length of a median through the three side lengths:

$$\|\mathbf{z} - \mathbf{m}\|^2 = \frac{1}{4}\Big(2\|\mathbf{z} - \mathbf{x}\|^2 + 2\|\mathbf{z} - \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\Big).$$

Solution to Exercise 6.1.49

Uses. Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)
Hint 1. Expand $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle$ into four terms.
Hint 2. The two cross terms are $-\langle \mathbf{x}, \mathbf{y} \rangle$ and $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle}$, whose sum is $-2i \operatorname{Im}\langle \mathbf{x}, \mathbf{y} \rangle$.
Sketch. $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \big(\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\big) - 2i \operatorname{Im}\langle \mathbf{x}, \mathbf{y} \rangle$. Its real part is $\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$, and it vanishes exactly when both the real part and $\operatorname{Im}\langle \mathbf{x}, \mathbf{y} \rangle$ do. In $\mathbb{R}^2$: a parallelogram is a rhombus if and only if its diagonals are perpendicular.

Expanding with linearity in the first slot and Proposition 6.1.4,

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \|\mathbf{x}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle - \|\mathbf{y}\|^2.$$

With $z = \langle \mathbf{x}, \mathbf{y} \rangle$, conjugate symmetry gives $-z + \bar{z} = -2i \operatorname{Im} z$, so

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \big(\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\big) - 2i \operatorname{Im}\langle \mathbf{x}, \mathbf{y} \rangle.$$

(a) The real part of the right side is $\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$, which is 0 if and only if $\|\mathbf{x}\| = \|\mathbf{y}\|$ (norms are nonnegative). In $\mathbb{R}^2$ with the dot product, $\mathbf{x} + \mathbf{y}$ and $\mathbf{x} - \mathbf{y}$ are the diagonals of the parallelogram on $\mathbf{x}$ and $\mathbf{y}$, so (a) says that the parallelogram is a rhombus (all sides equal) if and only if its diagonals are perpendicular, provided $\mathbf{x}, \mathbf{y}$ are linearly independent so that the parallelogram is not degenerate.

(b) A complex number is 0 if and only if its real and imaginary parts are. So $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = 0$ if and only if $\|\mathbf{x}\| = \|\mathbf{y}\|$ and $\operatorname{Im}\langle \mathbf{x}, \mathbf{y} \rangle = 0$, that is, $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{R}$. For instance, $\mathbf{x} = (1, 0)$ and $\mathbf{y} = (0, 1)$ in $\mathbb{C}^2$ qualify, but $\mathbf{x} = (1)$ and $\mathbf{y} = (i)$ in $\mathbb{C}^1$ do not: they have equal norms, while $\langle \mathbf{x} + \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = (1 + i)\overline{(1 - i)} = (1 + i)^2 = 2i$.

Solution to Exercise 6.1.50

Uses. Proposition 6.1.37 (Inner products distinguish vectors), Proposition 6.1.4 (Sesquilinearity), Definition 1.3.10 (Span)
Techniques. Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. By Proposition 6.1.37, it suffices to show $\langle \mathbf{x} - \mathbf{y}, \mathbf{w} \rangle = 0$ for every $\mathbf{w} \in V$.
Hint 2. Write $\mathbf{w} = \sum_j c_j \mathbf{v}_j$ and expand the second slot with Proposition 6.1.4.

Sketch. Every $\mathbf{w} = \sum_j c_j \mathbf{v}_j$, so $\langle \mathbf{x}-\mathbf{y}, \mathbf{w}\rangle = \sum_j \overline{c_j}\langle \mathbf{x}-\mathbf{y}, \mathbf{v}_j\rangle = 0$, and Proposition 6.1.37 gives $\mathbf{x} = \mathbf{y}$. In $\mathbb{R}^2$ with the list $\mathbf{e}_1$, the vectors $\mathbf{e}_2$ and $\mathbf{0}$ have the same inner product with $\mathbf{e}_1$.

Let $\mathbf{u} = \mathbf{x} - \mathbf{y}$, so that $\langle \mathbf{u}, \mathbf{v}_j\rangle = 0$ for every j. Let $\mathbf{w} \in V$. Since the list spans V, $\mathbf{w} = \sum_j c_j \mathbf{v}_j$ for some scalars c_j, and by Proposition 6.1.4,

$$\langle \mathbf{u}, \mathbf{w}\rangle = \sum_{j=1}^{k} \overline{c_j}\langle \mathbf{u}, \mathbf{v}_j\rangle = 0 = \langle \mathbf{0}, \mathbf{w}\rangle.$$

By Proposition 6.1.37, $\mathbf{u} = \mathbf{0}$, that is, $\mathbf{x} = \mathbf{y}$.

Without the spanning hypothesis the conclusion fails: in $\mathbb{R}^2$ with the list consisting of $\mathbf{e}_1$ alone, $\langle \mathbf{e}_2, \mathbf{e}_1\rangle = 0 = \langle \mathbf{0}, \mathbf{e}_1\rangle$, but $\mathbf{e}_2 \neq \mathbf{0}$.

Solution to Exercise 6.1.51

Uses. Proposition 6.1.32 (Polarization identities), Proposition 6.1.27 (Basic properties of the norm), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Polarization (Technique 6.1.33)
Hint 1. Express $\langle T\mathbf{v}, T\mathbf{w}\rangle$ through norms with Proposition 6.1.32.
Hint 2. Each norm that appears is $\|T\mathbf{v} + cT\mathbf{w}\| = \|T(\mathbf{v} + c\mathbf{w})\|$. For injectivity, look at the kernel.
Sketch. By polarization in W and linearity, $\langle T\mathbf{v}, T\mathbf{w}\rangle$ is a combination of the numbers $\|T(\mathbf{v} + c\mathbf{w})\|^2 = \|\mathbf{v} + c\mathbf{w}\|^2$, which is the same combination that polarization in V gives for $\langle \mathbf{v}, \mathbf{w}\rangle$. If $T\mathbf{v} = \mathbf{0}$, then $\|\mathbf{v}\| = \|T\mathbf{v}\| = 0$.

Suppose $\mathbb{F} = \mathbb{C}$; the real case is the same with c ranging over ± 1 and the coefficients of Proposition 6.1.32(a). For $c \in \mathbb{C}$, linearity gives $T\mathbf{v} + cT\mathbf{w} = T(\mathbf{v} + c\mathbf{w})$, so by hypothesis $\|T\mathbf{v} + cT\mathbf{w}\| = \|\mathbf{v} + c\mathbf{w}\|$. By Proposition 6.1.32 in W and then in V,

$$\langle T\mathbf{v}, T\mathbf{w}\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \left\|T\mathbf{v} + i^k T\mathbf{w}\right\|^2 = \frac{1}{4}\sum_{k=0}^{3} i^k \left\|\mathbf{v} + i^k \mathbf{w}\right\|^2 = \langle \mathbf{v}, \mathbf{w}\rangle.$$

For injectivity, if $T\mathbf{v} = \mathbf{0}$, then $\|\mathbf{v}\| = \|T\mathbf{v}\| = 0$, so $\mathbf{v} = \mathbf{0}$ by Proposition 6.1.27. So $\ker T = \{\mathbf{0}\}$, and T is injective by Theorem 2.6.11.

Solution to Exercise 6.1.52

Uses. Example 6.1.39 (A rotation), Corollary 6.1.38 (Linear maps are determined by inner products), Proposition 6.1.4 (Sesquilinearity)
Techniques. Plug in clever test vectors (Technique 6.1.40)
Hint 1. (a) Look at Example 6.1.39.
Hint 2. (b) Apply the hypothesis to $\mathbf{v} + \mathbf{w}$, and use that a real inner product is symmetric.
Hint 3. (c) Combine (b) with the hypothesis to get $\langle T\mathbf{v}, \mathbf{w}\rangle = -\langle T\mathbf{v}, \mathbf{w}\rangle$.
Sketch. (a) False: the rotation by $\pi/2$ on $\mathbb{R}^2$. (b) Expanding $\langle T(\mathbf{v}+\mathbf{w}), \mathbf{v}+\mathbf{w}\rangle = 0$ gives $\langle T\mathbf{v}, \mathbf{w}\rangle + \langle T\mathbf{w}, \mathbf{v}\rangle = 0$, and $\langle T\mathbf{w}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$ by symmetry; conversely put $\mathbf{w} = \mathbf{v}$. (c) Then $\langle T\mathbf{v}, \mathbf{w}\rangle = -\langle T\mathbf{v}, \mathbf{w}\rangle$, so it is 0 and $T = 0$.

(a) False. The rotation $T\mathbf{x} = (-x_2, x_1)$ of $\mathbb{R}^2$ satisfies $\langle T\mathbf{x}, \mathbf{x}\rangle = 0$ for every $\mathbf{x}$ but is not 0 (Example 6.1.39).

(b) Suppose $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$. For $\mathbf{v}, \mathbf{w} \in V$, expanding as in the proof of Proposition 6.1.41,

$$0 = \langle T(\mathbf{v}+\mathbf{w}), \mathbf{v}+\mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle + \langle T\mathbf{w}, \mathbf{v}\rangle.$$

A real inner product is symmetric, so $\langle T\mathbf{w}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$, and hence $\langle T\mathbf{v}, \mathbf{w}\rangle = -\langle \mathbf{v}, T\mathbf{w}\rangle$. Conversely, if this holds for all $\mathbf{v}, \mathbf{w}$, then taking $\mathbf{w} = \mathbf{v}$ gives $\langle T\mathbf{v}, \mathbf{v}\rangle = -\langle \mathbf{v}, T\mathbf{v}\rangle = -\langle T\mathbf{v}, \mathbf{v}\rangle$, so $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$.

(c) By (b) and the hypothesis on T, $\langle T\mathbf{v}, \mathbf{w}\rangle = -\langle \mathbf{v}, T\mathbf{w}\rangle = -\langle T\mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w}$. So $\langle T\mathbf{v}, \mathbf{w}\rangle = 0 = \langle 0\mathbf{v}, \mathbf{w}\rangle$, and $T = 0$ by Corollary 6.1.38.

Solution to Exercise 6.1.53

Uses. Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v}, \mathbf{v}\rangle$ determine T), Proposition 6.1.4 (Sesquilinearity), Notation 6.1.1 (Real and complex scalars)
Techniques. Polarization (Technique 6.1.33)
Hint 1. One direction is a one-line computation with conjugate symmetry.

Hint 2. For the other, consider $B(\mathbf{v},\mathbf{w}) = \langle T\mathbf{v},\mathbf{w}\rangle - \langle \mathbf{v}, T\mathbf{w}\rangle$. How does B treat scalars in each slot, and what is $B(\mathbf{v},\mathbf{v})$?
Hint 3. B is linear in the first slot, conjugate-linear in the second, and $B(\mathbf{v},\mathbf{v}) = 2i\operatorname{Im}\langle T\mathbf{v},\mathbf{v}\rangle$. Use (6.1.4).
Sketch. If T satisfies the identity, then $\langle T\mathbf{v},\mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v},\mathbf{v}\rangle}$ is real. Conversely, $B(\mathbf{v},\mathbf{w}) = \langle T\mathbf{v},\mathbf{w}\rangle - \langle \mathbf{v}, T\mathbf{w}\rangle$ is linear in $\mathbf{v}$ and conjugate-linear in $\mathbf{w}$, and $B(\mathbf{v},\mathbf{v}) = 2i\operatorname{Im}\langle T\mathbf{v},\mathbf{v}\rangle = 0$. By (6.1.4), $B = 0$.

Suppose $\langle T\mathbf{v},\mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle$ for all $\mathbf{v},\mathbf{w}$. Then $\langle T\mathbf{v},\mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \overline{\langle T\mathbf{v},\mathbf{v}\rangle}$ by conjugate symmetry, so $\langle T\mathbf{v},\mathbf{v}\rangle$ is real.

Conversely, suppose $\langle T\mathbf{v},\mathbf{v}\rangle \in \mathbb{R}$ for all $\mathbf{v}$, and define

$$B(\mathbf{v},\mathbf{w}) = \langle T\mathbf{v},\mathbf{w}\rangle - \langle \mathbf{v}, T\mathbf{w}\rangle.$$

The first term is linear in $\mathbf{v}$ and conjugate-linear in $\mathbf{w}$, as in the second proof of Proposition 6.1.41. So is the second term: it is linear in $\mathbf{v}$ by linearity in the first slot, and by Proposition 6.1.4,

$$\langle \mathbf{v}, T(a\mathbf{w}+b\mathbf{z})\rangle = \langle \mathbf{v}, aT\mathbf{w}+bT\mathbf{z}\rangle = \overline{a}\langle \mathbf{v},T\mathbf{w}\rangle + \overline{b}\langle \mathbf{v},T\mathbf{z}\rangle.$$

Hence B is linear in the first slot and conjugate-linear in the second. Its diagonal values are

$$B(\mathbf{v},\mathbf{v}) = \langle T\mathbf{v},\mathbf{v}\rangle - \overline{\langle T\mathbf{v},\mathbf{v}\rangle} = 2i\operatorname{Im}\langle T\mathbf{v},\mathbf{v}\rangle = 0.$$

The identity (6.1.4) holds for every such B, as its derivation used only how B treats scalars in each slot. So $B(\mathbf{v},\mathbf{w}) = \frac{1}{4}\sum_k i^k B\big(\mathbf{v}+i^k\mathbf{w},\mathbf{v}+i^k\mathbf{w}\big) = 0$ for all $\mathbf{v},\mathbf{w}$, which is the claim.

Solution to Exercise 6.1.54

Uses. Corollary 6.1.19 ($A^*A = 0$ forces $A = 0$), Proposition 2.4.23 (Transpose of a product), Definition 2.4.1 (Matrix product)
Techniques. The trace of A^*A (Technique 6.1.18)
Hint 1. Show $(BC)^* = C^*B^*$ entrywise, and deduce $(A^m)^* = A^m$ for every $m \geq 1$.
Hint 2. If $A^{2m} = 0$, then $(A^m)^*A^m = A^{2m} = 0$. What does Corollary 6.1.19 say?
Hint 3. Choose j with $2^j \geq k$ and halve the exponent j times.
Sketch. $(A^m)^* = (A^*)^m = A^m$, so $A^{2m} = 0$ gives $(A^m)^*A^m = 0$ and $A^m = 0$. Starting from $A^{2^j} = 0$ with $2^j \geq k$, halving j times gives $A = 0$.

First, for $B, C \in \mathbb{C}^{n\times n}$, the (i,j) entry of $(BC)^*$ is

$$\overline{(BC)_{ji}} = \overline{\sum_l b_{jl}c_{li}} = \sum_l \overline{c_{li}}\,\overline{b_{jl}} = \sum_l (C^*)_{il}(B^*)_{lj},$$

which is the (i,j) entry of C^*B^* (Definition 2.4.1). So $(BC)^* = C^*B^*$. By induction on m, using $A^* = A$,

$$(A^m)^* = \big(A^{m-1}A\big)^* = A^*\big(A^{m-1}\big)^* = AA^{m-1} = A^m \quad \text{for all } m \geq 1.$$

Next, if $m \geq 1$ and $A^{2m} = 0$, then $(A^m)^*A^m = A^mA^m = A^{2m} = 0$, so $A^m = 0$ by Corollary 6.1.19.

Now choose $j \geq 0$ with $2^j \geq k$. Then $A^{2^j} = A^{2^j-k}A^k = 0$ (with $A^0 = I$). If $j = 0$, this says $A = 0$. Otherwise, applying the previous paragraph with $m = 2^{j-1}, 2^{j-2}, \ldots, 1$ in turn gives $A^{2^{j-1}} = 0$, then $A^{2^{j-2}} = 0$, and finally $A^1 = 0$.

Solution to Exercise 6.1.55

Uses. Proposition 6.1.16 (Frobenius inner product)
Techniques. The trace of A^*A (Technique 6.1.18)
Hint 1. Rewrite A^2 using $A = -A^{\mathsf{T}}$.
Hint 2. $A^2 = -A^{\mathsf{T}}A$, and $\operatorname{tr}\big(A^{\mathsf{T}}A\big)$ is a sum of squares.
Sketch. $\operatorname{tr}\big(A^2\big) = -\operatorname{tr}\big(A^{\mathsf{T}}A\big) = -\sum_{i,j} a_{ij}^2 \leq 0$, with equality only for $A = 0$. If $A^2 = I_n$, then $\operatorname{tr}\big(A^2\big) = n > 0$. The rotation $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has $J^2 = -I_2$.

(a) Since $A = -A^{\mathsf{T}}$, $A^2 = \big(-A^{\mathsf{T}}\big)A = -A^{\mathsf{T}}A$. By Proposition 6.1.16 (over $\mathbb{R}$, $A^* = A^{\mathsf{T}}$),

$$\operatorname{tr}\big(A^2\big) = -\operatorname{tr}\big(A^{\mathsf{T}}A\big) = -\sum_{i,j} a_{ij}^2 \leq 0,$$

with equality only if every $a_{ij} = 0$.

(b) If $A^2 = I_n$, then $\operatorname{tr}(A^2) = n > 0$, contradicting (a). On the other hand, $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is skew-symmetric with $J^2 = -I_2$, and indeed $\operatorname{tr}(J^2) = -2 = -\sum_{i,j} j_{ij}^2$.

Solution to Exercise 6.1.56

Uses. Proposition 6.1.16 (Frobenius inner product), Definition 6.1.8 (Dot product), Definition 2.4.1 (Matrix product)
Hint 1. The (i, j) entry of $\mathbf{u}\mathbf{v}^*$ is $u_i\overline{v_j}$.
Hint 2. Use the entrywise formula of Proposition 6.1.16; the double sum factors into a product of two single sums.
Sketch. By Proposition 6.1.16 the inner product is $\sum_{i,j} u_i\overline{v_j}\,\overline{x_i}\overline{y_j}$, and this double sum factors as $\left(\sum_i u_i\overline{x_i}\right)\left(\sum_j y_j\overline{v_j}\right)$. With $\mathbf{x} = \mathbf{u}$ and $\mathbf{y} = \mathbf{v}$ this is $\|\mathbf{u}\|^2\|\mathbf{v}\|^2$.

By Definition 2.4.1, the (i, j) entry of $\mathbf{u}\mathbf{v}^*$ is $u_i\overline{v_j}$, and that of $\mathbf{x}\mathbf{y}^*$ is $x_i\overline{y_j}$. By Proposition 6.1.16,

$$\langle \mathbf{u}\mathbf{v}^*, \mathbf{x}\mathbf{y}^* \rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} u_i\overline{v_j}\,\overline{x_i\overline{y_j}} = \sum_{i,j} u_i\overline{x_i}y_j\overline{v_j} = \left(\sum_i u_i\overline{x_i}\right)\left(\sum_j y_j\overline{v_j}\right),$$

which is $(\mathbf{u}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{v})$. Taking $\mathbf{x} = \mathbf{u}$ and $\mathbf{y} = \mathbf{v}$ gives $\|\mathbf{u}\mathbf{v}^*\|^2 = (\mathbf{u}\cdot\mathbf{u})(\mathbf{v}\cdot\mathbf{v}) = \|\mathbf{u}\|^2\|\mathbf{v}\|^2$, and taking nonnegative square roots, $\|\mathbf{u}\mathbf{v}^*\| = \|\mathbf{u}\|\|\mathbf{v}\|$.

Solution to Exercise 6.1.57

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.1.4 (Sesquilinearity)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Sum over all ordered pairs (i, j) instead; the terms with $i = j$ are 0, and each unordered pair is counted twice.
Hint 2. Expand $\|\mathbf{x}_i - \mathbf{x}_j\|^2 = 2 - 2\operatorname{Re}\langle \mathbf{x}_i, \mathbf{x}_j\rangle$, and recognize $\sum_{i,j}\langle \mathbf{x}_i, \mathbf{x}_j\rangle$ as $\langle \mathbf{s}, \mathbf{s}\rangle$ for $\mathbf{s} = \sum_i \mathbf{x}_i$.
Sketch. Each term is $2 - 2\operatorname{Re}\langle \mathbf{x}_i, \mathbf{x}_j\rangle$, and $\sum_{i,j}\langle \mathbf{x}_i, \mathbf{x}_j\rangle = \|\mathbf{s}\|^2$, so the sum over all ordered pairs is $2n^2 - 2\|\mathbf{s}\|^2$; the sum over $i < j$ is half of this. (b) is immediate, and in (c) the maximum 9 is attained by three unit vectors at angles $2\pi/3$, which sum to $\mathbf{0}$.

(a) Let $\mathbf{s} = \mathbf{x}_1 + \cdots + \mathbf{x}_n$. By sesquilinearity, $\sum_{i,j}\langle \mathbf{x}_i, \mathbf{x}_j\rangle = \langle \mathbf{s}, \mathbf{s}\rangle = \|\mathbf{s}\|^2$, the sum running over all ordered pairs. By Proposition 6.1.29 and $\|\mathbf{x}_i\| = 1$, $\|\mathbf{x}_i - \mathbf{x}_j\|^2 = 2 - 2\operatorname{Re}\langle \mathbf{x}_i, \mathbf{x}_j\rangle$. Summing over all ordered pairs,

$$\sum_{i,j}\|\mathbf{x}_i - \mathbf{x}_j\|^2 = 2n^2 - 2\operatorname{Re}\sum_{i,j}\langle \mathbf{x}_i, \mathbf{x}_j\rangle = 2n^2 - 2\|\mathbf{s}\|^2.$$

The terms with $i = j$ vanish, and each pair $i < j$ occurs twice, as (i, j) and (j, i), with the same value. So the sum over $i < j$ is half of this, $n^2 - \|\mathbf{s}\|^2$.

(b) Since $\|\mathbf{s}\|^2 \geq 0$, the sum is at most n^2, with equality if and only if $\|\mathbf{s}\| = 0$, that is, $\mathbf{s} = \mathbf{0}$ by Proposition 6.1.27.

(c) By (b) the sum is at most 9, with equality exactly when $\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 = \mathbf{0}$. This happens for $\mathbf{x}_1 = (1, 0)$, $\mathbf{x}_2 = \left(-1/2, \sqrt{3}/2\right)$, and $\mathbf{x}_3 = \left(-1/2, -\sqrt{3}/2\right)$, the vertices of an equilateral triangle inscribed in the unit circle. Each side has squared length $(3/2)^2 + \left(\sqrt{3}/2\right)^2 = 3$, and $3 + 3 + 3 = 9$.

Solution to Exercise 6.1.58

Uses. Proposition 1.1.24 ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime), Proposition 1.1.17 (No zero divisors)
Hint 1. (a) List the squares in $\mathbb{F}_5$.
Hint 2. (b) For odd p, it suffices to find $a, b \in \mathbb{F}_p$ with $a^2 + b^2 = -1$, and then take $\mathbf{x} = (a, b, 1)$. Compare the sets $\{a^2\}$ and $\{-1 - b^2\}$.
Hint 3. Count: $a^2 = c^2$ if and only if $a = \pm c$, so there are exactly $(p + 1)/2$ squares in $\mathbb{F}_p$. Two subsets of $\mathbb{F}_p$ of that size must meet.
Sketch. (a) $(1, 2)$, since $1 + 4 = 5 = 0$. (b) For $p = 2$ take $(1, 1, 0)$. For odd p, the set S of squares has $(p + 1)/2$ elements, and so does $-1 - S$; since $(p + 1)/2 + (p + 1)/2 > p$, they share an element $a^2 = -1 - b^2$, and $(a, b, 1)$ works.

(a) The squares in $\mathbb{F}_5$ are $0, 1, 4, 4, 1$ (of $0, 1, 2, 3, 4$). Since $1 + 4 = 5 = 0$, the vector $\mathbf{x} = (1, 2)$ works.

(b) For $p = 2$, $\mathbf{x} = (1, 1, 0)$ works, since $1 + 1 = 0$ in $\mathbb{F}_2$. Let p be odd. It suffices to find $a, b \in \mathbb{F}_p$ with $a^2 + b^2 + 1 = 0$, because then $\mathbf{x} = (a, b, 1) \neq \mathbf{0}$ works.

Let $S = \left\{a^2 \;\middle|\; a \in \mathbb{F}_p\right\}$. For $a, c \in \mathbb{F}_p$, $a^2 = c^2$ if and only if $(a-c)(a+c) = 0$, which by Proposition 1.1.17 means $a = c$ or $a = -c$. For $a \neq 0$ we have $a \neq -a$, since $a = -a$ would give $2a = 0$, and $2 \neq 0$ in $\mathbb{F}_p$ for odd p. So the $p-1$ nonzero elements fall into $(p-1)/2$ pairs $\{a, -a\}$, each pair having one square, and distinct pairs having distinct squares. Together with $0^2 = 0$, S has exactly $(p+1)/2$ elements.

The map $s \mapsto -1 - s$ is a bijection of $\mathbb{F}_p$, so $T = \left\{-1 - b^2 \;\middle|\; b \in \mathbb{F}_p\right\}$ also has $(p+1)/2$ elements. Since $|S| + |T| = p + 1 > p = \left|\mathbb{F}_p\right|$, the sets S and T intersect: $a^2 = -1 - b^2$ for some $a, b \in \mathbb{F}_p$. Then $a^2 + b^2 + 1 = 0$, as desired. For $n > 3$, pad with zeros: $(x_1, x_2, x_3, 0, \ldots, 0)$ is nonzero with $\sum_k x_k^2 = 0$.

Solution to Exercise 6.1.59

Uses. Proposition 6.1.32 (Polarization identities), Proposition 5.1.10 (Alternating implies antisymmetric)
Techniques. Polarization (Technique 6.1.33)
Hint 1. Which properties of an inner product did the proof of Proposition 6.1.32(a) use? Does B have them?
Hint 2. Expand $B(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - B(\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y})$ for a B that is only bilinear.
Hint 3. Look for a counterexample among 2×2 determinants.
Sketch. The real polarization identity needs symmetry: for a bilinear B, the right side equals $(B(\mathbf{x}, \mathbf{y}) + B(\mathbf{y}, \mathbf{x}))/2$. The claim is false: $B(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1$ vanishes on the diagonal but $B(\mathbf{e}_1, \mathbf{e}_2) = 1$. What is true is $B(\mathbf{y}, \mathbf{x}) = -B(\mathbf{x}, \mathbf{y})$.

The error is the first equality. The real polarization identity was derived for inner products, and its proof used symmetry: the two cross terms of $\|\mathbf{x}+\mathbf{y}\|^2$ were combined into $2\langle \mathbf{x}, \mathbf{y}\rangle$. For a bilinear B without symmetry, expanding gives

$$B(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) - B(\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}) = 2B(\mathbf{x}, \mathbf{y}) + 2B(\mathbf{y}, \mathbf{x}),$$

so the right side of the displayed formula is $(B(\mathbf{x}, \mathbf{y}) + B(\mathbf{y}, \mathbf{x}))/2$, not $B(\mathbf{x}, \mathbf{y})$.

The claim is false. On $\mathbb{R}^2$, $B(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1 = \det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$ is linear in each slot and $B(\mathbf{v}, \mathbf{v}) = 0$ for every $\mathbf{v}$, but $B(\mathbf{e}_1, \mathbf{e}_2) = 1$. What the argument does prove, correctly read, is that $B(\mathbf{x}, \mathbf{y}) + B(\mathbf{y}, \mathbf{x}) = 0$: a bilinear function that vanishes on the diagonal is antisymmetric. This is the argument of Proposition 5.1.10, which applies to any bilinear function of two vectors (Remark 5.1.6).

Solution to Exercise 6.1.60

Uses. Lemma 6.1.14 (Nonnegative functions with zero integral), Proposition 6.1.13 (Integral inner products)
Hint 1. Which hypothesis of Lemma 6.1.14 is missing for an integrable f?
Hint 2. Find a nonzero integrable function whose square has integral 0. Change a continuous function at one point.
Sketch. The step "$\int f^2 = 0$ and $f^2 \geq 0$ imply $f^2 = 0$" needs continuity (Lemma 6.1.14). The function equal to 1 at $t = 1/2$ and 0 elsewhere is integrable, nonzero, and has $\int f^2 = 0$. So the claim is false: definiteness fails on $R[0,1]$.

The error is in the last step. A nonnegative function with integral 0 must vanish if it is continuous (Lemma 6.1.14), but not in general. Let $f(1/2) = 1$ and $f(t) = 0$ for $t \neq 1/2$. Then f is Riemann-integrable and $f^2 = f$, and $\int_0^1 f^2 = 0$, since every Riemann sum over a partition of mesh δ is at most 2δ (the point $1/2$ lies in at most two subintervals). But $f \neq 0$.

So the claim is false: the formula is symmetric, bilinear, and positive on $R[0,1]$, but not definite. This is why Proposition 6.1.13 is stated for continuous functions.

Solution to Exercise 6.1.61 BEYOND 21-242

Uses. Proposition 6.1.32 (Polarization identities), Proposition 6.1.30 (Parallelogram law), Exercise 2.1.51
Hint 1. (a) uses only $\|-\mathbf{v}\| = \|\mathbf{v}\|$ and $\|2\mathbf{x}\| = 2\|\mathbf{x}\|$.
Hint 2. (b) Apply the parallelogram law to the pairs $(\mathbf{x}+\mathbf{y}, \mathbf{z})$ and $(\mathbf{x}-\mathbf{y}, \mathbf{z})$, and subtract.
Hint 3. (c) Put $\mathbf{z} = \mathbf{x}$ in (b) to get $\langle 2\mathbf{x}, \mathbf{y}\rangle = 2\langle \mathbf{x}, \mathbf{y}\rangle$. Then choose $\mathbf{x}$ and $\mathbf{z}$ with $\mathbf{x}+\mathbf{z} = \mathbf{u}$ and $\mathbf{x}-\mathbf{z} = \mathbf{w}$.
Hint 4. (d) Additivity gives homogeneity for integers, then for rationals. For real t, show $t \mapsto \langle t\mathbf{x}, \mathbf{y}\rangle$ is continuous, using $|\|\mathbf{a}\| - \|\mathbf{b}\|| \leq \|\mathbf{a}-\mathbf{b}\|$.
Sketch. (a) is immediate. (b) Subtracting the parallelogram laws for $(\mathbf{x}+\mathbf{y}, \mathbf{z})$ and $(\mathbf{x}-\mathbf{y}, \mathbf{z})$ gives $4\langle \mathbf{x}+\mathbf{z}, \mathbf{y}\rangle + 4\langle \mathbf{x}-\mathbf{z}, \mathbf{y}\rangle = 8\langle \mathbf{x}, \mathbf{y}\rangle$. (c) $\mathbf{z} = \mathbf{x}$ gives $\langle 2\mathbf{x}, \mathbf{y}\rangle = 2\langle \mathbf{x}, \mathbf{y}\rangle$, and then $\mathbf{x} = (\mathbf{u}+\mathbf{w})/2$, $\mathbf{z} = (\mathbf{u}-\mathbf{w})/2$ give additivity. (d) $f(t) = \langle t\mathbf{x}, \mathbf{y}\rangle$ is additive and continuous, so $f(t) = tf(1)$. Symmetry, linearity in the first slot, and $\langle \mathbf{x}, \mathbf{x}\rangle = \|\mathbf{x}\|^2$ are the inner product axioms.

(a) Since $\|-\mathbf{v}\| = |-1|\|\mathbf{v}\| = \|\mathbf{v}\|$, we have $\|\mathbf{y}-\mathbf{x}\| = \|\mathbf{x}-\mathbf{y}\|$, which gives symmetry, and $\|-\mathbf{x}-\mathbf{y}\| = \|\mathbf{x}+\mathbf{y}\|$ and $\|-\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}\|$, which give $\langle -\mathbf{x},\mathbf{y}\rangle = -\langle \mathbf{x},\mathbf{y}\rangle$. Also $\langle \mathbf{x},\mathbf{x}\rangle = \big(\|2\mathbf{x}\|^2 - \|\mathbf{0}\|^2\big)/4 = \|\mathbf{x}\|^2$.

(b) The parallelogram law for the pairs $(\mathbf{x}+\mathbf{y},\mathbf{z})$ and $(\mathbf{x}-\mathbf{y},\mathbf{z})$ reads

$$\|\mathbf{x}+\mathbf{y}+\mathbf{z}\|^2 + \|\mathbf{x}+\mathbf{y}-\mathbf{z}\|^2 = 2\|\mathbf{x}+\mathbf{y}\|^2 + 2\|\mathbf{z}\|^2,$$
$$\|\mathbf{x}-\mathbf{y}+\mathbf{z}\|^2 + \|\mathbf{x}-\mathbf{y}-\mathbf{z}\|^2 = 2\|\mathbf{x}-\mathbf{y}\|^2 + 2\|\mathbf{z}\|^2.$$

Subtract the second from the first and regroup the left side:

$$\big(\|(\mathbf{x}+\mathbf{z})+\mathbf{y}\|^2 - \|(\mathbf{x}+\mathbf{z})-\mathbf{y}\|^2\big) + \big(\|(\mathbf{x}-\mathbf{z})+\mathbf{y}\|^2 - \|(\mathbf{x}-\mathbf{z})-\mathbf{y}\|^2\big)$$
$$= 2\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\big).$$

By the definition of $\langle\cdot,\cdot\rangle$, this says $4\langle \mathbf{x}+\mathbf{z},\mathbf{y}\rangle + 4\langle \mathbf{x}-\mathbf{z},\mathbf{y}\rangle = 8\langle \mathbf{x},\mathbf{y}\rangle$.

(c) First, $\langle \mathbf{0},\mathbf{y}\rangle = \big(\|\mathbf{y}\|^2 - \|-\mathbf{y}\|^2\big)/4 = 0$. Taking $\mathbf{z} = \mathbf{x}$ in (b) gives $\langle 2\mathbf{x},\mathbf{y}\rangle + \langle \mathbf{0},\mathbf{y}\rangle = 2\langle \mathbf{x},\mathbf{y}\rangle$, so $\langle 2\mathbf{x},\mathbf{y}\rangle = 2\langle \mathbf{x},\mathbf{y}\rangle$ for all $\mathbf{x},\mathbf{y}$. Now let $\mathbf{u},\mathbf{w}\in V$, and apply (b) with $\mathbf{x} = (\mathbf{u}+\mathbf{w})/2$ and $\mathbf{z} = (\mathbf{u}-\mathbf{w})/2$, so that $\mathbf{x}+\mathbf{z} = \mathbf{u}$ and $\mathbf{x}-\mathbf{z} = \mathbf{w}$:

$$\langle \mathbf{u},\mathbf{y}\rangle + \langle \mathbf{w},\mathbf{y}\rangle = 2\left\langle \frac{\mathbf{u}+\mathbf{w}}{2},\mathbf{y}\right\rangle = \langle \mathbf{u}+\mathbf{w},\mathbf{y}\rangle,$$

the last step by the doubling rule applied to $(\mathbf{u}+\mathbf{w})/2$.

(d) Fix $\mathbf{x},\mathbf{y}$ and let $f(t) = \langle t\mathbf{x},\mathbf{y}\rangle$ for $t\in\mathbb{R}$. By (c), $f(s+t) = \langle s\mathbf{x}+t\mathbf{x},\mathbf{y}\rangle = f(s)+f(t)$, so f is additive. We show that f is continuous. The triangle inequality gives $\|\mathbf{a}\| \le \|\mathbf{a}-\mathbf{b}\| + \|\mathbf{b}\|$ and the same with $\mathbf{a},\mathbf{b}$ exchanged, so $|\|\mathbf{a}\| - \|\mathbf{b}\|| \le \|\mathbf{a}-\mathbf{b}\|$. Hence

$$|\|t\mathbf{x}\pm\mathbf{y}\| - \|s\mathbf{x}\pm\mathbf{y}\|| \le \|(t-s)\mathbf{x}\| = |t-s|\|\mathbf{x}\|,$$

so $t\mapsto \|t\mathbf{x}\pm\mathbf{y}\|$ is continuous, and so is $f(t) = \big(\|t\mathbf{x}+\mathbf{y}\|^2 - \|t\mathbf{x}-\mathbf{y}\|^2\big)/4$. A continuous additive function $\mathbb{R}\to\mathbb{R}$ satisfies $f(t) = tf(1)$ for all t (Exercise 2.1.51). Directly: additivity gives $f(mt) = mf(t)$ for positive integers m by induction, $f(0) = 0$, and $f(-t) = -f(t)$ by (a); so for $q = m/n$ with $n>0$, $nf(q) = f(m) = mf(1)$, that is, $f(q) = qf(1)$. Every real t is a limit of rationals q_k, and continuity gives $f(t) = \lim f(q_k) = \lim q_k f(1) = tf(1)$. So $\langle t\mathbf{x},\mathbf{y}\rangle = t\langle \mathbf{x},\mathbf{y}\rangle$.

Together with (c), $\langle\cdot,\cdot\rangle$ is linear in the first slot. It is symmetric by (a), which over $\mathbb{R}$ is conjugate symmetry, and $\langle \mathbf{x},\mathbf{x}\rangle = \|\mathbf{x}\|^2$ is nonnegative and vanishes only for $\mathbf{x} = \mathbf{0}$. So $\langle\cdot,\cdot\rangle$ is an inner product, and its norm $\sqrt{\langle \mathbf{x},\mathbf{x}\rangle}$ is $\|\mathbf{x}\|$.

Section 6.2

Solution to Exercise 6.2.10

Uses. Definition 6.1.8 (Dot product), (6.2.1)
Techniques. Orthogonal decomposition (Technique 6.2.4)
Hint 1. $c = (\mathbf{x}\cdot\mathbf{y})/\|\mathbf{y}\|^2$.
Hint 2. Compute $c\mathbf{y}$ coordinate by coordinate: $(3+i)i = -1+3i$ and $(3+i)(1-i) = 4-2i$.
Sketch. $c = (3+i)/3$, $c\mathbf{y} = ((-1+3i)/3, (4-2i)/3)$, and $\mathbf{w} = (4/3, (2+2i)/3)$. Then $\|\mathbf{w}\|^2 = 8/3$ and $|3+i|^2/3 = 10/3$, which add to 6.

By Technique 6.2.4, $c = (\mathbf{x}\cdot\mathbf{y})/\|\mathbf{y}\|^2 = (3+i)/3$. Since $(3+i)i = -1+3i$ and $(3+i)(1-i) = 4-2i$,

$$c\mathbf{y} = \left(\frac{-1+3i}{3}, \frac{4-2i}{3}\right), \quad \mathbf{w} = \mathbf{x} - c\mathbf{y} = \left(\frac{4}{3}, \frac{2+2i}{3}\right).$$

Check:

$$\mathbf{w}\cdot\mathbf{y} = \frac{4}{3}\bar{i} + \frac{2+2i}{3}\overline{(1-i)} = \frac{-4i}{3} + \frac{4i}{3} = 0,$$

using $(2+2i)(1+i) = 4i$. Now $\|\mathbf{w}\|^2 = 16/9 + 8/9 = 8/3$ (as $|2+2i|^2 = 8$), and $|\mathbf{x}\cdot\mathbf{y}|^2/\|\mathbf{y}\|^2 = 10/3$. Their sum is $18/3 = 6 = \|\mathbf{x}\|^2$, as (6.2.1) predicts. In particular $|\mathbf{x}\cdot\mathbf{y}|^2 = 10 \le 18 = \|\mathbf{x}\|^2\|\mathbf{y}\|^2$.

Solution to Exercise 6.2.12

Uses. Proposition 6.1.27 (Basic properties of the norm), Theorem 6.2.11 (Triangle inequality), Notation 6.1.1 (Real and complex scalars)
Hint 1. By Proposition 6.1.27, both conditions are equations between real multiples of $\|\mathbf{x}\|$. Divide by $\|\mathbf{x}\|$.
Hint 2. The first condition becomes $|1+a| = 1+|a|$, the triangle inequality in $\mathbb{C} = \mathbb{C}^1$ itself. For the second, square $|1+a| = |1-|a||$ and expand $|1+a|^2 = 1 + 2\operatorname{Re} a + |a|^2$.
Sketch. The first condition is $|1+a| = 1+|a|$, which by Theorem 6.2.11(b) in $\mathbb{C}$ means $a \ge 0$ is real. Squaring the second gives $2\operatorname{Re} a = -2|a|$, so a is a real number ≤ 0.

By Proposition 6.1.27, $\|\mathbf{x}+a\mathbf{x}\| = |1+a|\|\mathbf{x}\|$ and $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$, and $\|\mathbf{x}\| > 0$.

The first condition is $|1+a| = 1+|a|$. This is the equality case of Theorem 6.2.11 in $\mathbb{C}$ with the dot product $z\overline{w}$, for the vectors 1 and a. By (b), it holds if and only if one of $1, a$ is a nonnegative real multiple of the other, that is, if and only if a is a real number with $a \ge 0$. (If $1 = ba$ with $b \ge 0$, then $b > 0$ and $a = 1/b > 0$.)

The second condition is $|1+a| = |1-|a||$. Both sides are nonnegative, so it is equivalent to the equality of squares $1 + 2\operatorname{Re} a + |a|^2 = 1 - 2|a| + |a|^2$, that is, $\operatorname{Re} a = -|a|$. Writing $a = s + ti$, this says $s = -\sqrt{s^2+t^2}$, which holds if and only if $t = 0$ and $s \le 0$. So the second condition holds exactly for real $a \le 0$; for $a < 0$, the vectors $\mathbf{x}$ and $a\mathbf{x}$ point in opposite directions.

Solution to Exercise 6.2.20

Uses. Definition 6.2.16 (Angle), Proposition 6.1.13 (Integral inner products)
Hint 1. $\langle t, t^2\rangle = \int_0^1 t^3\,dt$, and $\|t^k\|^2 = \int_0^1 t^{2k}\,dt$.
Hint 2. For the second part, the condition is $\langle 1, t+a\rangle = (1/2)\|1\|\|t+a\|$. The left side must be positive. Square and solve.
Sketch. $\cos\theta = (1/4)/\left(\sqrt{1/3}\sqrt{1/5}\right) = \sqrt{15}/4$. For the second part, $1/2 + a = (1/2)\sqrt{1/3 + a + a^2}$ with $1/2 + a > 0$; squaring gives $9a^2 + 9a + 2 = 0$, so $a = -1/3$ or $a = -2/3$, and only $a = -1/3$ has $1/2 + a > 0$.

We have $\langle t, t^2\rangle = \int_0^1 t^3\,dt = 1/4$, $\|t\|^2 = 1/3$, and $\|t^2\|^2 = \int_0^1 t^4\,dt = 1/5$. So

$$\cos\theta = \frac{1/4}{\sqrt{\frac{1}{3}\cdot\frac{1}{5}}} = \frac{\sqrt{15}}{4} \approx 0.968.$$

For the second part, $t + a \ne 0$ for every a, $\|1\| = 1$, $\langle 1, t+a\rangle = 1/2 + a$, and

$$\|t+a\|^2 = \int_0^1 \left(t^2 + 2at + a^2\right)dt = \frac{1}{3} + a + a^2.$$

The angle is $\pi/3$ if and only if $1/2 + a = (1/2)\sqrt{1/3 + a + a^2}$. This requires $1/2 + a \ge 0$, and then it is equivalent to the squared equation $4(1/2+a)^2 = 1/3 + a + a^2$, that is, $3a^2 + 3a + 2/3 = 0$, or $9a^2 + 9a + 2 = 0$. Its roots are $a = -1/3$ and $a = -2/3$. Only $a = -1/3$ satisfies $1/2 + a \ge 0$; for $a = -2/3$ the angle is $2\pi/3$. Check: $\|t - 1/3\|^2 = 1/3 - 1/3 + 1/9 = 1/9$, so $\cos\theta = (1/6)/(1/3) = 1/2$.

Solution to Exercise 6.2.27

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), (6.2.1), Definition 6.2.16 (Angle)
Techniques. Orthogonal decomposition (Technique 6.2.4)
Hint 1. Every inner product in this exercise carries the weights $1, 2, 3$. For instance $\|\mathbf{y}\|^2 = 1 + 2 + 3\cdot 4$.
Hint 2. $c = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$.
Sketch. $\langle \mathbf{x}, \mathbf{y}\rangle = 5$, $\|\mathbf{x}\|^2 = 6$, $\|\mathbf{y}\|^2 = 15$, and $25 \le 90$. Then $c = 1/3$, $\mathbf{w} = (2/3, 4/3, 1/3)$, $\|\mathbf{w}\|^2 = 13/3$, and $13/3 + 25/15 = 6$. The cosine is $5/\sqrt{90} = \sqrt{10}/6$; for the dot product it is $2/\sqrt{18} = \sqrt{2}/3$.

(a) $\langle \mathbf{x}, \mathbf{y}\rangle = 1 - 2 + 6 = 5$, $\|\mathbf{x}\|^2 = 1 + 2 + 3 = 6$, and $\|\mathbf{y}\|^2 = 1 + 2 + 12 = 15$. Indeed $5^2 = 25 \le 6 \cdot 15 = 90$.

(b) By Technique 6.2.4, $c = 5/15 = 1/3$ and $\mathbf{w} = \mathbf{x} - \mathbf{y}/3 = (2/3, 4/3, 1/3)$. Check:

$$\langle \mathbf{w}, \mathbf{y} \rangle = \frac{2}{3} - 2 \cdot \frac{4}{3} + 3 \cdot \frac{2}{3} = 0.$$

Also

$$\|\mathbf{w}\|^2 = \frac{4}{9} + 2 \cdot \frac{16}{9} + 3 \cdot \frac{1}{9} = \frac{39}{9} = \frac{13}{3},$$

and

$$\|\mathbf{w}\|^2 + \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\|\mathbf{y}\|^2} = \frac{13}{3} + \frac{25}{15} = 6 = \|\mathbf{x}\|^2,$$

as (6.2.1) predicts.

(c) $\cos\theta = 5/\sqrt{90} = 5/\left(3\sqrt{10}\right) = \sqrt{10}/6 \approx 0.527$. For the dot product, $\mathbf{x} \cdot \mathbf{y} = 2$, $\|\mathbf{x}\|^2 = 3$, $\|\mathbf{y}\|^2 = 6$, and the cosine is $2/\sqrt{18} = \sqrt{2}/3 \approx 0.471$. The angle depends on the inner product.

Solution to Exercise 6.2.28

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 6.2.11 (Triangle inequality)
Hint 1. Both vectors are nonzero for every z. By Theorem 6.2.1(b), equality means $\mathbf{y} = a\mathbf{x}$ for some $a \in \mathbb{C}$. Compare second coordinates first.
Hint 2. For (b), Theorem 6.2.11(b) requires a nonnegative real multiple. Is the multiple from (a) real?
Sketch. (a) $\mathbf{y} = a\mathbf{x}$ forces $1 = ai$, so $a = -i$ and $z = -i$. (b) The only dependent case has $\mathbf{y} = -i\mathbf{x}$, which is not a nonnegative real multiple, so there is no such z; indeed $\|\mathbf{x} + \mathbf{y}\| = 2 < 2\sqrt{2}$ for $z = -i$.

(a) Both $\mathbf{x}$ and $\mathbf{y}$ are nonzero, so by Theorem 6.2.1(b) equality holds if and only if $\mathbf{y} = a\mathbf{x}$ for some $a \in \mathbb{C}$ (a relation $\mathbf{x} = b\mathbf{y}$ has $b \neq 0$ and can be inverted). Comparing second coordinates, $1 = ai$, so $a = -i$, and then $z = a \cdot 1 = -i$. Conversely, $z = -i$ gives $\mathbf{y} = -i\mathbf{x}$. Check: $\mathbf{x} \cdot \mathbf{y} = 1 \cdot \overline{-i} + i \cdot 1 = 2i$, and $|2i| = 2 = \sqrt{2}\sqrt{2} = \|\mathbf{x}\|\|\mathbf{y}\|$.

(b) By Theorem 6.2.11(b), equality requires one of $\mathbf{x}, \mathbf{y}$ to be a nonnegative real multiple of the other, so in particular equality in Cauchy–Schwarz. By (a), this forces $z = -i$ and $\mathbf{y} = -i\mathbf{x}$, equivalently $\mathbf{x} = i\mathbf{y}$, and neither $-i$ nor i is a nonnegative real number. So there is no such z. For $z = -i$ directly: $\mathbf{x} + \mathbf{y} = (1 - i, 1 + i)$ has norm $\sqrt{2 + 2} = 2 < 2\sqrt{2}$.

Solution to Exercise 6.2.29

Uses. Definition 6.1.8 (Dot product)
Techniques. Rotate the phase (Technique 6.2.8)
Hint 1. $\langle \mathbf{x}, \mathbf{y} \rangle = 1 \cdot \overline{1} + 1 \cdot \overline{i}$.
Hint 2. For (b), $\operatorname{Re}\langle \lambda\mathbf{x}, \mathbf{y} \rangle = \operatorname{Re}(\lambda\langle \mathbf{x}, \mathbf{y} \rangle)$ is largest when λ rotates $\langle \mathbf{x}, \mathbf{y} \rangle$ onto the positive real axis.
Sketch. $\langle \mathbf{x}, \mathbf{y} \rangle = 1 - i$ and $\|\mathbf{x}\| = \|\mathbf{y}\| = \sqrt{2}$. The two ratios are $1/2$ and $\sqrt{2}/2$, giving angles $\pi/3$ and $\pi/4$. In (b), $\lambda = (1 + i)/\sqrt{2}$ and the maximum is $\sqrt{2} = |1 - i|$.

(a) $\langle \mathbf{x}, \mathbf{y} \rangle = 1 + \overline{i} = 1 - i$, and $\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 = 2$. So

$$\frac{\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{1}{2}, \quad \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{\sqrt{2}}{2},$$

with angles $\pi/3$ and $\pi/4$. The first is the angle between $\mathbf{x}$ and $\mathbf{y}$ regarded as vectors $(1, 0, 1, 0)$ and $(1, 0, 0, 1)$ of $\mathbb{R}^4$.

(b) $\operatorname{Re}\langle \lambda\mathbf{x}, \mathbf{y} \rangle = \operatorname{Re}(\lambda(1 - i)) \leq |\lambda(1 - i)| = \sqrt{2}$, with equality if and only if $\lambda(1 - i)$ is a nonnegative real number. With $1 - i = \sqrt{2}\mu$, $\mu = (1 - i)/\sqrt{2}$, this happens for $\lambda = \overline{\mu} = (1 + i)/\sqrt{2}$, and indeed $\lambda(1 - i) = (1 + i)(1 - i)/\sqrt{2} = \sqrt{2}$. So the maximum is $\sqrt{2} = |\langle \mathbf{x}, \mathbf{y} \rangle|$, and the smallest real angle between $\mathbf{y}$ and a unit-modulus multiple of $\mathbf{x}$ is $\pi/4$, as Technique 6.2.8 predicts.

Solution to Exercise 6.2.30

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Definition 6.1.8 (Dot product), Remark 6.2.2

Hint 1. The dot product on $\mathbb{C}^n$ conjugates the second vector. Which vector should play the role of $\mathbf{y}$ so that $\mathbf{z}\cdot\mathbf{y} = \sum_k z_k w_k$?
Hint 2. For the example, try $\mathbf{z} = \mathbf{w} = (1, i)$.
Sketch. With $\overline{\mathbf{w}} = (\overline{w_1}, \ldots, \overline{w_n})$, $\mathbf{z}\cdot\overline{\mathbf{w}} = \sum_k z_k w_k$ and $\|\overline{\mathbf{w}}\| = \|\mathbf{w}\|$, so this is Cauchy–Schwarz for $\mathbf{z}, \overline{\mathbf{w}}$, with equality iff they are dependent. For $\mathbf{z} = \mathbf{w} = (1, i)$, the left side is $\left|1 + i^2\right|^2 = 0 < 4$.

Let $\mathbf{z} = (z_1, \ldots, z_n)$ and $\overline{\mathbf{w}} = (\overline{w_1}, \ldots, \overline{w_n})$. Then $\mathbf{z}\cdot\overline{\mathbf{w}} = \sum_k z_k\overline{\overline{w_k}} = \sum_k z_k w_k$, and $\|\overline{\mathbf{w}}\|^2 = \sum_k |\overline{w_k}|^2 = \sum_k |w_k|^2$. So the inequality is Theorem 6.2.1 for $\mathbf{z}$ and $\overline{\mathbf{w}}$ (squared), and by Theorem 6.2.1(b) and Remark 6.2.2, equality holds if and only if $\mathbf{z}$ and $\overline{\mathbf{w}}$ are linearly dependent.

For $\mathbf{z} = \mathbf{w} = (1, i)$, the two vectors are equal, hence dependent, but $z_1 w_1 + z_2 w_2 = 1 + i^2 = 0$, while the right side is $2 \cdot 2 = 4$. Here $\overline{\mathbf{w}} = (1, -i)$ is not a multiple of $\mathbf{z}$. With $\mathbf{w} = (1, -i)$ instead, $\sum_k z_k w_k = 1 - i^2 = 2$ and equality $4 = 4$ holds.

Solution to Exercise 6.2.31

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Example 6.1.21 (Weighted dot products)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. The constraint says $\|\mathbf{v}\| \le 1$ for a weighted inner product. Which weights?
Hint 2. Write $x + y + z = \langle (x, y, z), \mathbf{u}\rangle$ in the inner product $x_1y_1 + 4x_2y_2 + 9x_3y_3$, for a suitable vector $\mathbf{u}$.
Hint 3. $\mathbf{u} = (1, 1/4, 1/9)$. For the maximizer, use the equality case together with $\|\mathbf{v}\| = 1$.
Sketch. In $\langle \mathbf{a}, \mathbf{b}\rangle_w = a_1b_1 + 4a_2b_2 + 9a_3b_3$, $x + y + z = \langle \mathbf{v}, \mathbf{u}\rangle_w$ with $\mathbf{u} = (1, 1/4, 1/9)$ and $\|\mathbf{u}\|_w = 7/6$. So $x + y + z \le \|\mathbf{v}\|_w \cdot 7/6 \le 7/6$, with equality at $\mathbf{v} = (6/7)\mathbf{u} = (6/7, 3/14, 2/21)$.

Let $\langle \mathbf{a}, \mathbf{b}\rangle_w = a_1b_1 + 4a_2b_2 + 9a_3b_3$, an inner product on $\mathbb{R}^3$ by Example 6.1.21, and let $\mathbf{v} = (x, y, z)$ and $\mathbf{u} = (1, 1/4, 1/9)$. Then $\langle \mathbf{v}, \mathbf{u}\rangle_w = x + y + z$, the constraint is $\|\mathbf{v}\|_w \le 1$, and

$$\|\mathbf{u}\|_w^2 = 1 + 4\cdot\frac{1}{16} + 9\cdot\frac{1}{81} = 1 + \frac{1}{4} + \frac{1}{9} = \frac{49}{36}.$$

By Theorem 6.2.1,

$$x + y + z \le \left|\langle \mathbf{v}, \mathbf{u}\rangle_w\right| \le \|\mathbf{v}\|_w\|\mathbf{u}\|_w \le \frac{7}{6}.$$

For $\mathbf{v} = (6/7)\mathbf{u} = (6/7, 3/14, 2/21)$ we have $\|\mathbf{v}\|_w = (6/7)(7/6) = 1$, so the constraint holds, and $x + y + z = (6/7)\|\mathbf{u}\|_w^2$, which is $(6/7)(49/36) = 7/6$. So the maximum is $7/6$.

It is attained only there: equality throughout requires $\|\mathbf{v}\|_w = 1$, equality in Cauchy–Schwarz, so $\mathbf{v} = t\mathbf{u}$ (as $\mathbf{u} \ne \mathbf{0}$), and $\langle \mathbf{v}, \mathbf{u}\rangle_w = t\|\mathbf{u}\|_w^2 \ge 0$, so $t \ge 0$; then $\|\mathbf{v}\|_w = t \cdot 7/6 = 1$ gives $t = 6/7$.

Solution to Exercise 6.2.32

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. The first equation says $\mathbf{v}\cdot(1, 2, 3) = 14$, and $\|(1, 2, 3)\|^2 = 14$. What does Cauchy–Schwarz say about $\mathbf{v}\cdot(1, 2, 3)$ when $\|\mathbf{v}\|^2 = 14$?
Hint 2. The data force equality in Cauchy–Schwarz. Use the equality case.
Sketch. With $\mathbf{v} = (x, y, z)$ and $\mathbf{u} = (1, 2, 3)$: $14 = \mathbf{v}\cdot\mathbf{u} \le \|\mathbf{v}\|\|\mathbf{u}\| = \sqrt{14}\sqrt{14} = 14$, so equality holds and $\mathbf{v} = t\mathbf{u}$; then $14t = 14$ gives $(1, 2, 3)$. In (b), $\mathbf{v}\cdot\mathbf{u} \le \sqrt{13\cdot 14} < 14$.

Let $\mathbf{v} = (x, y, z)$ and $\mathbf{u} = (1, 2, 3)$, so that $\|\mathbf{u}\|^2 = 14$ and the first equation is $\mathbf{v}\cdot\mathbf{u} = 14$.

(a) If $\|\mathbf{v}\|^2 = 14$, then by Theorem 6.2.1, $14 = \mathbf{v}\cdot\mathbf{u} \le \|\mathbf{v}\|\|\mathbf{u}\| = 14$, so equality holds. By Theorem 6.2.1(b), and since $\mathbf{u} \ne \mathbf{0}$, $\mathbf{v} = t\mathbf{u}$ for some $t \in \mathbb{R}$. Then $\mathbf{v}\cdot\mathbf{u} = 14t = 14$ gives $t = 1$, so $(x, y, z) = (1, 2, 3)$. Conversely, $(1, 2, 3)$ satisfies both equations, since $1 + 4 + 9 = 14$. So $(1, 2, 3)$ is the only solution.

(b) If $\|\mathbf{v}\|^2 = 13$, then $\mathbf{v}\cdot\mathbf{u} \le \sqrt{13}\sqrt{14} < 14$, so the first equation fails.

Geometrically, the plane $x + 2y + 3z = 14$ is at distance $14/\sqrt{14} = \sqrt{14}$ from the origin. It touches the sphere of radius $\sqrt{14}$ in exactly one point and misses the smaller sphere of radius $\sqrt{13}$.

Solution to Exercise 6.2.33

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.4 (Sesquilinearity), Remark 6.2.2, Notation 6.1.1 (Real and complex scalars)
Hint 1. Expand $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle$. Its real part is $\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$.
Hint 2. Then $\left|\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right| \le |\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle|$, and Cauchy–Schwarz finishes (a).
Hint 3. For (b), equality needs $\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle = 0$ and equality in Cauchy–Schwarz for $\mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}$. Show that $\mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}$ are dependent if and only if $\mathbf{x}, \mathbf{y}$ are.
Sketch. $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = \left(\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right) - 2i \operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle$, so $\left|\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right|$ is at most $|\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle|$, which Cauchy–Schwarz bounds by $\|\mathbf{x}+\mathbf{y}\|\|\mathbf{x}-\mathbf{y}\|$. Equality holds iff $\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle = 0$ and $\mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}$ are dependent, iff $\mathbf{x}, \mathbf{y}$ are dependent with a real ratio.

(a) By sesquilinearity and conjugate symmetry,

$$\begin{aligned}\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle &= \|\mathbf{x}\|^2 - \langle \mathbf{x}, \mathbf{y}\rangle + \overline{\langle \mathbf{x}, \mathbf{y}\rangle} - \|\mathbf{y}\|^2 \\ &= \left(\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right) - 2i \operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle,\end{aligned}$$

since $z - \overline{z} = 2i \operatorname{Im} z$. So $\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2$ is the real part of $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle$, and by $|\operatorname{Re} z| \le |z|$ and Theorem 6.2.1,

$$\left|\|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\right| \le |\langle \mathbf{x}+\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle| \le \|\mathbf{x}+\mathbf{y}\|\|\mathbf{x}-\mathbf{y}\|.$$

(b) Equality holds if and only if both steps are equalities. The first is an equality if and only if $\operatorname{Im}\langle \mathbf{x}, \mathbf{y}\rangle = 0$, since $|\operatorname{Re} z| = |z|$ exactly for real z. The second is an equality if and only if $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$ are linearly dependent (Remark 6.2.2). These are dependent if and only if $\mathbf{x}, \mathbf{y}$ are: a nontrivial relation $a(\mathbf{x}+\mathbf{y}) + b(\mathbf{x}-\mathbf{y}) = \mathbf{0}$ is the relation $(a+b)\mathbf{x} + (a-b)\mathbf{y} = \mathbf{0}$, which is nontrivial since $a+b = a-b = 0$ forces $a = b = 0$; conversely, $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$ becomes $((\alpha+\beta)/2)(\mathbf{x}+\mathbf{y}) + ((\alpha-\beta)/2)(\mathbf{x}-\mathbf{y}) = \mathbf{0}$, again nontrivial.

So equality holds if and only if $\mathbf{x}, \mathbf{y}$ are dependent and $\langle \mathbf{x}, \mathbf{y}\rangle$ is real. If $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$, both conditions hold, and the zero vector is the real multiple 0 of the other. If $\mathbf{y} = a\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$, then $\langle \mathbf{x}, \mathbf{y}\rangle = \overline{a}\|\mathbf{x}\|^2$ is real if and only if a is real; the case $\mathbf{x} = b\mathbf{y}$ is the same. This proves (b).

Solution to Exercise 6.2.34

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Lemma 6.2.3 (The easy half of (b))
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Equality is the same as linear dependence. Is "is dependent with" a transitive relation? Think about the zero vector.
Hint 2. If $\mathbf{y} \neq \mathbf{0}$, both $\mathbf{x}$ and $\mathbf{z}$ are multiples of $\mathbf{y}$.
Sketch. False: take $\mathbf{y} = \mathbf{0}$ and $\mathbf{x} = \mathbf{e}_1$, $\mathbf{z} = \mathbf{e}_2$ in $\mathbb{R}^2$. If $\mathbf{y} \neq \mathbf{0}$, then $\mathbf{x} = a\mathbf{y}$ and $\mathbf{z} = b\mathbf{y}$, so one of $\mathbf{x}, \mathbf{z}$ is a multiple of the other, and equality holds by Lemma 6.2.3.

The claim is false. In $\mathbb{R}^2$ with the dot product, let $\mathbf{x} = \mathbf{e}_1$, $\mathbf{y} = \mathbf{0}$, and $\mathbf{z} = \mathbf{e}_2$. Both hypotheses hold, since both sides of the inequality are 0 for each pair containing $\mathbf{0}$. But $|\mathbf{e}_1 \cdot \mathbf{e}_2| = 0 < 1 = \|\mathbf{e}_1\|\|\mathbf{e}_2\|$.

If $\mathbf{y} \neq \mathbf{0}$, the conclusion holds. By Theorem 6.2.1(b), one of $\mathbf{x}, \mathbf{y}$ is a multiple of the other, and since $\mathbf{y} \neq \mathbf{0}$, $\mathbf{x} = a\mathbf{y}$ for some a (if $\mathbf{y} = c\mathbf{x}$, then $c \neq 0$ and $\mathbf{x} = c^{-1}\mathbf{y}$). Likewise $\mathbf{z} = b\mathbf{y}$. If $a = 0$, then $\mathbf{x} = \mathbf{0} = 0\mathbf{z}$. If $a \neq 0$, then $\mathbf{y} = a^{-1}\mathbf{x}$ and $\mathbf{z} = (b/a)\mathbf{x}$. Either way one of $\mathbf{x}, \mathbf{z}$ is a multiple of the other, and equality holds by Lemma 6.2.3.

Solution to Exercise 6.2.35

Uses. Proposition 6.1.29 (Expanding a norm), Lemma 6.2.6 (Nonnegative quadratics)
Techniques. A nonnegative quadratic has nonpositive discriminant (Technique 6.2.7)
Hint 1. Expand $\|\mathbf{x} - t\mathbf{y}\|^2$ correctly with Proposition 6.1.29. Where do conjugates and real parts appear?
Hint 2. Is the right side of the displayed formula even a real number for complex t? Does "≤ 0" make sense for $\langle \mathbf{x}, \mathbf{y}\rangle^2$?
Hint 3. Try $x = 1$, $y = i$, $t = i$ in $\mathbb{C}$.
Sketch. The correct expansion is $\|\mathbf{x}\|^2 - 2\operatorname{Re}(\overline{t}\langle \mathbf{x}, \mathbf{y}\rangle) + |t|^2\|\mathbf{y}\|^2$. The displayed expression is not real in general, so it cannot be compared with 0; the discriminant criterion is about real quadratics in a real variable; and $\langle \mathbf{x}, \mathbf{y}\rangle^2$ need not be real. In $\mathbb{C}$ with $x = 1$, $y = i$, $t = i$: the formula gives -2, but $|1 - i \cdot i|^2 = 4$. The repair is the second proof of Theorem 6.2.1.

There are three errors.

First, the expansion is wrong over $\mathbb{C}$. By Proposition 6.1.29 with $a = -t$,

$$\|\mathbf{x} - t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\,\mathrm{Re}(\bar{t}\langle \mathbf{x}, \mathbf{y}\rangle) + |t|^2\|\mathbf{y}\|^2.$$

The argument's formula drops the conjugate and the real part and writes t^2 for $|t|^2$. For a numerical counterexample, take $V = \mathbb{C}$ with $\langle z, w\rangle = z\overline{w}$, $x = 1$, $y = i$, and $t = i$. Then $\langle x, y\rangle = \bar{i} = -i$, $\|x\| = \|y\| = 1$, and the formula gives $1 - 2i(-i) + i^2 = 1 - 2 - 1 = -2$, while $\|x - ty\|^2 = |1 - i \cdot i|^2 = |2|^2 = 4$.

Second, the discriminant criterion (Lemma 6.2.6) is a statement about real quadratics in a real variable. A polynomial in a complex variable t takes non-real values, and "never negative" has no meaning for it.

Third, $\langle \mathbf{x}, \mathbf{y}\rangle^2$ need not be real, so "$4\langle \mathbf{x}, \mathbf{y}\rangle^2 - 4\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \le 0$" is not a meaningful inequality, and "taking square roots" of it is not a valid step. (The claim itself is true; it is Theorem 6.2.1.)

The repair is the second proof of Theorem 6.2.1: restrict to real t, and first rotate the phase of $\mathbf{x}$ so that $\langle \mathbf{x}, \mathbf{y}\rangle$ is a nonnegative real number. Then the expansion above becomes a real quadratic in t with middle coefficient $-2|\langle \mathbf{x}, \mathbf{y}\rangle|$, and the discriminant argument applies.

Solution to Exercise 6.2.36

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.13 (Integral inner products)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. The right side contains $e^t f^2 = \left(e^{t/2} f\right)^2$. Split $f = \left(e^{t/2} f\right) \cdot e^{-t/2}$.
Hint 2. Apply Cauchy–Schwarz in $C[0,1]$ to $g = e^{t/2} f$ and $h = e^{-t/2}$. Compute $\|h\|^2$.
Sketch. $\int f = \left\langle e^{t/2} f, e^{-t/2}\right\rangle$, so $\left(\int f\right)^2 \le \left(\int e^t f^2\right)\left(\int_0^1 e^{-t}\,\mathrm{d}t\right) = \left(1 - e^{-1}\right)\int e^t f^2$. Equality iff $e^{t/2} f = ce^{-t/2}$, that is, $f(t) = ce^{-t}$.

In $C[0,1]$ with $\langle g, h\rangle = \int_0^1 gh$, let $g(t) = e^{t/2} f(t)$ and $h(t) = e^{-t/2}$, both continuous. Then $gh = f$, $g^2 = e^t f^2$, and $h^2 = e^{-t}$, so

$$\langle g, h\rangle = \int_0^1 f, \quad \|g\|^2 = \int_0^1 e^t f^2,$$

and $\|h\|^2 = \int_0^1 e^{-t}\,\mathrm{d}t = 1 - e^{-1}$. By Theorem 6.2.1, $\langle g, h\rangle^2 \le \|g\|^2\|h\|^2$, which is the inequality.

Since $h \neq \mathbf{0}$, equality holds if and only if $g = ch$ for some $c \in \mathbb{R}$, that is, $e^{t/2} f(t) = ce^{-t/2}$, or $f(t) = ce^{-t}$. Check: for $f(t) = e^{-t}$, the left side is $\left(1 - e^{-1}\right)^2$ and the right side is $\left(1 - e^{-1}\right)\int_0^1 e^{-t}\,\mathrm{d}t = \left(1 - e^{-1}\right)^2$.

Solution to Exercise 6.2.37

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Definition 1.2.17 (Evaluation)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. $p(t) = \sum_k a_k t^k$ is a dot product in $\mathbb{R}^{n+1}$. Of which two vectors?
Hint 2. For equality, choose the coefficient vector proportional to $(1, t, \ldots, t^n)$. For (b), compare $\sum_{k=0}^n t^{2k}$ with a geometric series.
Sketch. $p(t) = \mathbf{a} \cdot \boldsymbol{\tau}$ with $\mathbf{a} = (a_0, \ldots, a_n)$ and $\boldsymbol{\tau} = (1, t, \ldots, t^n)$, so Cauchy–Schwarz gives (a), with equality for $\mathbf{a} = \boldsymbol{\tau}$. For $|t| < 1$, $\sum_{k=0}^n t^{2k} \le \sum_{k \ge 0} t^{2k} = 1/\left(1 - t^2\right)$.

(a) Let $\mathbf{a} = (a_0, \ldots, a_n)$ and $\boldsymbol{\tau} = \left(1, t, t^2, \ldots, t^n\right)$ in $\mathbb{R}^{n+1}$. By Definition 1.2.17, $p(t) = \sum_{k=0}^n a_k t^k = \mathbf{a} \cdot \boldsymbol{\tau}$, and $\|\boldsymbol{\tau}\|^2 = \sum_{k=0}^n t^{2k}$. So Theorem 6.2.1 gives $|p(t)| \le \|\mathbf{a}\|\|\boldsymbol{\tau}\|$, which is the inequality. For $p = \sum_{k=0}^n t^k x^k$, the coefficient vector is $\boldsymbol{\tau}$ itself, which is nonzero (its first entry is 1), and $p(t) = \|\boldsymbol{\tau}\|^2 = \|\boldsymbol{\tau}\|\|\boldsymbol{\tau}\|$, so equality holds.

(b) For $|t| < 1$, $0 \le t^2 < 1$, so $\sum_{k=0}^n t^{2k} = \left(1 - t^{2n+2}\right)/\left(1 - t^2\right) \le 1/\left(1 - t^2\right)$. Combining with (a), $|p(t)| \le \|\mathbf{a}\|/\sqrt{1 - t^2}$, and the bound does not depend on n.

Solution to Exercise 6.2.38

Uses. Proposition 6.2.22 (The square of a sum), Theorem 6.2.1 (Cauchy–Schwarz inequality)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. (a) is Proposition 6.2.22 for $a_k = \sqrt{k}$. Recall $1 + 2 + \cdots + n = n(n+1)/2$.
Hint 2. (b) The left side is $\mathbf{a} \cdot (1, 1/2, \ldots, 1/n)$. Bound $\sum_{k=1}^n 1/k^2$ by telescoping.
Sketch. (a) $\left(\sum_k \sqrt{k}\right)^2 \le n \sum_k k = n^2(n+1)/2$. (b) Cauchy–Schwarz gives $\sum_k a_k/k \le \|\mathbf{a}\|\left(\sum_k 1/k^2\right)^{1/2}$, and the bound $1/k^2 < 1/(k-1) - 1/k$ makes $\sum_{k=2}^n 1/k^2$ telescope to less than 1.

(a) By Proposition 6.2.22 with $a_k = \sqrt{k}$,

$$\left(\sum_{k=1}^{n} \sqrt{k}\right)^2 \le n \sum_{k=1}^{n} k = n \cdot n\frac{n+1}{2},$$

and taking square roots gives the claim.

(b) Let $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{r} = (1, 1/2, \dots, 1/n)$. By Theorem 6.2.1,

$$\sum_{k=1}^{n} \frac{a_k}{k} = \mathbf{a} \cdot \mathbf{r} \le \|\mathbf{a}\|\|\mathbf{r}\| = \|\mathbf{a}\|\left(\sum_{k=1}^{n} \frac{1}{k^2}\right)^{1/2}.$$

For $k \ge 2$, $1/k^2 < 1/(k(k-1)) = 1/(k-1) - 1/k$, so the sum telescopes:

$$\sum_{k=1}^{n} \frac{1}{k^2} \le 1 + \sum_{k=2}^{n}\left(\frac{1}{k-1} - \frac{1}{k}\right) = 2 - \frac{1}{n} < 2.$$

Hence $\sum_k a_k/k \le \sqrt{2}\|\mathbf{a}\|$.

Solution to Exercise 6.2.39

Uses. Proposition 6.1.16 (Frobenius inner product), Theorem 6.2.1 (Cauchy–Schwarz inequality), Remark 6.2.2
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. The trace is an inner product with a fixed matrix in the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(B^* A)$. Which one?
Hint 2. $\langle A, I_n\rangle = \operatorname{tr} A$ and $\|I_n\|^2 = n$. For (b), the data force equality.
Sketch. $\operatorname{tr} A = \operatorname{tr}(I_n^* A) = \langle A, I_n\rangle$ and $\|I_n\|^2 = n$, so Cauchy–Schwarz gives (a), with equality iff $A = cI_n$. In (b), $|\operatorname{tr} A|^2 = 4 = 2 \cdot 2$, so $A = cI_2$, and $\operatorname{tr} A = 2$ gives $A = I_2$.

(a) In the Frobenius inner product (Proposition 6.1.16), $\langle A, I_n\rangle = \operatorname{tr}(I_n^* A) = \operatorname{tr} A$, $\|I_n\|^2 = \operatorname{tr}(I_n) = n$, and $\|A\|^2 = \operatorname{tr}(A^* A)$. By Theorem 6.2.1, $|\operatorname{tr} A|^2 \le \|A\|^2\|I_n\|^2 = n \operatorname{tr}(A^* A)$, with equality if and only if one of A, I_n is a multiple of the other, that is (as $I_n \ne 0$), $A = cI_n$ for some $c \in \mathbb{F}$.

(b) Here $\operatorname{tr}\left(A^{\mathsf{T}} A\right) = \sum_{i,j} a_{ij}^2 = 2$, so $|\operatorname{tr} A|^2 = 4 = 2 \cdot 2 = n \operatorname{tr}\left(A^{\mathsf{T}} A\right)$, which is the equality case of (a). So $A = cI_2$, and $\operatorname{tr} A = 2c = 2$ gives $c = 1$. Conversely, I_2 satisfies both conditions. So $A = I_2$ is the only such matrix.

Solution to Exercise 6.2.40

Uses. Proposition 6.1.16 (Frobenius inner product), Definition 2.4.1 (Matrix product), Theorem 6.2.1 (Cauchy–Schwarz inequality)
Hint 1. The (i, j) entry of AB is $\sum_k a_{ik} b_{kj}$, a product of row i of A with column j of B. Bound its modulus with Cauchy–Schwarz in $\mathbb{F}^n$ (watch the conjugate).
Hint 2. Square, sum over i and j, and recognize $\|A\|_F^2 = \sum_{i,k} |a_{ik}|^2$. For (b), induct on k.
Sketch. $\left|(AB)_{ij}\right|^2 \le \left(\sum_k |a_{ik}|^2\right)\left(\sum_k |b_{kj}|^2\right)$ by Cauchy–Schwarz; summing over i, j gives $\|AB\|_F^2 \le \|A\|_F^2 \|B\|_F^2$. Then $\left|\left(A^k\right)_{ij}\right| \le \left\|A^k\right\|_F \le \|A\|_F^k$. For (c), $A = B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ gives $AB = 0$.

(a) By Definition 2.4.1, $(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. Let $\mathbf{r}_i = (a_{i1}, \dots, a_{in})$ and $\mathbf{c}_j = \left(\overline{b_{1j}}, \dots, \overline{b_{nj}}\right)$ in $\mathbb{F}^n$. Then $(AB)_{ij} = \mathbf{r}_i \cdot \mathbf{c}_j$, and by Theorem 6.2.1,

$$\left|(AB)_{ij}\right|^2 \le \|\mathbf{r}_i\|^2 \|\mathbf{c}_j\|^2 = \left(\sum_k |a_{ik}|^2\right)\left(\sum_k |b_{kj}|^2\right).$$

Summing over all i and j, and using Proposition 6.1.16,

$$\begin{aligned} \|AB\|_F^2 &\le \sum_{i,j}\left(\sum_k |a_{ik}|^2\right)\left(\sum_k |b_{kj}|^2\right) \\ &= \left(\sum_{i,k} |a_{ik}|^2\right)\left(\sum_{k,j} |b_{kj}|^2\right) = \|A\|_F^2 \|B\|_F^2. \end{aligned}$$

(b) By (a) and induction on k, $\left\|A^k\right\|_F \le \|A\|_F^k$: the case $k=1$ is trivial, and $\left\|A^{k+1}\right\|_F = \left\|A^kA\right\|_F \le \left\|A^k\right\|_F\|A\|_F$. Each entry satisfies $\left|\left(A^k\right)_{ij}\right|^2 \le \sum_{i',j'}\left|\left(A^k\right)_{i'j'}\right|^2 = \left\|A^k\right\|_F^2$, so $\left|\left(A^k\right)_{ij}\right| \le \|A\|_F^k$.

(c) For $A = B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $AB = 0$, so $\|AB\|_F = 0 < 1 = \|A\|_F\|B\|_F$.

Solution to Exercise 6.2.41

Uses. Theorem 6.2.11 (Triangle inequality), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. The triangle inequality gives ≤ 1. Equality would force $(1-s)\mathbf{x}$ and $s\mathbf{y}$ onto a common ray.
Hint 2. If $(1-s)\mathbf{x} = as\mathbf{y}$ with $a \ge 0$, take norms to find a, and conclude $\mathbf{x} = \mathbf{y}$.
Hint 3. For (b), look at $\mathbf{e}_1$ and $\mathbf{e}_2$.
Sketch. By Theorem 6.2.11, $\|(1-s)\mathbf{x} + s\mathbf{y}\| \le (1-s) + s = 1$, with equality only if one of $(1-s)\mathbf{x}, s\mathbf{y}$ is a nonnegative multiple of the other. Taking norms, the multiple is the ratio of $1-s$ and s, and then $\mathbf{x} = \mathbf{y}$. In (b), $\|(\mathbf{e}_1 + \mathbf{e}_2)/2\|_1 = 1$.

(a) By Theorem 6.2.11 and Proposition 6.1.27,

$$\|(1-s)\mathbf{x} + s\mathbf{y}\| \le (1-s)\|\mathbf{x}\| + s\|\mathbf{y}\| = 1.$$

Suppose equality holds. By Theorem 6.2.11(b), one of $(1-s)\mathbf{x}$, $s\mathbf{y}$ is a nonnegative real multiple of the other. If $(1-s)\mathbf{x} = as\mathbf{y}$ with $a \ge 0$, taking norms gives $1 - s = as$, so $(1-s)\mathbf{x} = (1-s)\mathbf{y}$, and $\mathbf{x} = \mathbf{y}$ since $1 - s \neq 0$. If $s\mathbf{y} = b(1-s)\mathbf{x}$ with $b \ge 0$, the same argument gives $\mathbf{y} = \mathbf{x}$. Both contradict $\mathbf{x} \neq \mathbf{y}$, so the inequality is strict.

(b) The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are distinct and $\|\mathbf{e}_1\|_1 = \|\mathbf{e}_2\|_1 = 1$, but $\|(\mathbf{e}_1 + \mathbf{e}_2)/2\|_1 = 1/2 + 1/2 = 1$. (The whole segment from $\mathbf{e}_1$ to $\mathbf{e}_2$ is an edge of the diamond in Figure 6.1.4.) If $\|\cdot\|_1$ were the norm of an inner product, (a) would apply to it. So it is not.

Solution to Exercise 6.2.42

Uses. Theorem 6.2.11 (Triangle inequality), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For $\supseteq$, compute $\mathbf{x} - \mathbf{z}$ and $\mathbf{z} - \mathbf{y}$ for $\mathbf{z} = (1-s)\mathbf{x} + s\mathbf{y}$: both are multiples of $\mathbf{x} - \mathbf{y}$.
Hint 2. For $\subseteq$, the condition is equality in the triangle inequality for $\mathbf{u} = \mathbf{x} - \mathbf{z}$ and $\mathbf{v} = \mathbf{z} - \mathbf{y}$, whose sum is $\mathbf{x} - \mathbf{y}$.
Hint 3. If $\mathbf{v} = a\mathbf{u}$ with $a \ge 0$, then $(1+a)\mathbf{u} = \mathbf{x} - \mathbf{y}$. Solve for $\mathbf{z}$.
Sketch. For $\mathbf{z} = (1-s)\mathbf{x} + s\mathbf{y}$, $\mathbf{x} - \mathbf{z} = s(\mathbf{x} - \mathbf{y})$ and $\mathbf{z} - \mathbf{y} = (1-s)(\mathbf{x} - \mathbf{y})$, whose norms add up to $\|\mathbf{x} - \mathbf{y}\|$. Conversely, equality in the triangle inequality for $\mathbf{u} = \mathbf{x} - \mathbf{z}$, $\mathbf{v} = \mathbf{z} - \mathbf{y}$ makes one a nonnegative multiple of the other; solving $(1+a)\mathbf{u} = \mathbf{x} - \mathbf{y}$ gives $\mathbf{z} = (1-s)\mathbf{x} + s\mathbf{y}$ with $s = 1/(1+a)$, and similarly in the other case.

Let L be the set of $\mathbf{z} \in V$ satisfying the equation, and $S = \{(1-s)\mathbf{x} + s\mathbf{y} \mid s \in [0,1]\}$. We show $L = S$ by Technique 1.3.14.

$S \subseteq L$. Let $\mathbf{z} = (1-s)\mathbf{x} + s\mathbf{y}$ with $s \in [0,1]$. Then $\mathbf{x} - \mathbf{z} = s(\mathbf{x} - \mathbf{y})$ and $\mathbf{z} - \mathbf{y} = (1-s)(\mathbf{x} - \mathbf{y})$. Since $s, 1-s \ge 0$, Proposition 6.1.27 gives

$$\|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\| = (s + (1-s))\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|.$$

$L \subseteq S$. Let $\mathbf{z} \in L$, and put $\mathbf{u} = \mathbf{x} - \mathbf{z}$ and $\mathbf{v} = \mathbf{z} - \mathbf{y}$, so that $\mathbf{u} + \mathbf{v} = \mathbf{x} - \mathbf{y}$ and $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$. By Theorem 6.2.11(b), one of $\mathbf{u}, \mathbf{v}$ is a nonnegative real multiple of the other.

If $\mathbf{v} = a\mathbf{u}$ with $a \ge 0$, then $(1+a)\mathbf{u} = \mathbf{x} - \mathbf{y}$, so $\mathbf{u} = s(\mathbf{x} - \mathbf{y})$ with $s = 1/(1+a) \in (0,1]$, and $\mathbf{z} = \mathbf{x} - \mathbf{u} = (1-s)\mathbf{x} + s\mathbf{y}$. If $\mathbf{u} = b\mathbf{v}$ with $b \ge 0$, then $(1+b)\mathbf{v} = \mathbf{x} - \mathbf{y}$, so $\mathbf{v} = (1-s)(\mathbf{x} - \mathbf{y})$ with $1 - s = 1/(1+b)$, that is, $s = b/(1+b) \in [0,1)$, and $\mathbf{z} = \mathbf{y} + \mathbf{v} = (1-s)\mathbf{x} + s\mathbf{y}$. In both cases $\mathbf{z} \in S$.

Solution to Exercise 6.2.43

Uses. Remark 6.2.5 (The size of the gap), Remark 6.2.2
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Write both products as double sums over all pairs (j,k), and subtract.
Hint 2. The terms with $j = k$ cancel. Group the term for (j,k) with the term for (k,j).
Hint 3. For the equality case in (b), all the numbers $a_jb_k - a_kb_j$ vanish. If some $a_p \neq 0$, show $\mathbf{b} = \left(b_p/a_p\right)\mathbf{a}$.

Sketch. The left side is $\sum_{j,k}\left(a_j^2 b_k^2 - a_j b_j a_k b_k\right)$. Diagonal terms vanish, and the terms for (j,k) and (k,j) with $j<k$ add to $\left(a_j b_k - a_k b_j\right)^2$. A sum of squares is nonnegative and vanishes iff each term does, which says that $\mathbf{a}, \mathbf{b}$ are dependent. For $n=2$ the right side is $(a_1 b_2 - a_2 b_1)^2$, and (6.2.2) identifies the left side with the square of base times height.

(a) Expanding the products as double sums over all pairs (j,k),

$$\left(\sum_j a_j^2\right)\left(\sum_k b_k^2\right) - \left(\sum_j a_j b_j\right)\left(\sum_k a_k b_k\right)$$
$$= \sum_{j,k}\left(a_j^2 b_k^2 - a_j b_j a_k b_k\right).$$

The terms with $j=k$ are $a_j^2 b_j^2 - a_j^2 b_j^2 = 0$. For $j<k$, the terms for (j,k) and (k,j) add to

$$a_j^2 b_k^2 + a_k^2 b_j^2 - 2 a_j b_j a_k b_k = \left(a_j b_k - a_k b_j\right)^2.$$

Summing over $j<k$ gives the identity.

(b) The right side of (a) is a sum of squares, so it is nonnegative, which is the Cauchy–Schwarz inequality $(\mathbf{a}\cdot\mathbf{b})^2 \le \|\mathbf{a}\|^2\|\mathbf{b}\|^2$. By Technique 6.1.9, equality holds if and only if $a_j b_k = a_k b_j$ for all $j<k$, hence for all j,k. We show this happens if and only if $\mathbf{a}, \mathbf{b}$ are linearly dependent. If $\mathbf{a}=\mathbf{0}$, both hold. Otherwise pick p with $a_p \neq 0$. If $a_p b_k = a_k b_p$ for all k, then $\mathbf{b} = \left(b_p/a_p\right)\mathbf{a}$. Conversely, if $\mathbf{b} = c\mathbf{a}$, then $a_j b_k - a_k b_j = c\left(a_j a_k - a_k a_j\right) = 0$, and if $\mathbf{a} = c\mathbf{b}$ the same computation applies. By Remark 6.2.2, this is the equality case of Theorem 6.2.1.

(c) For $n=2$ there is one pair $j<k$, and (a) reads $\|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a}\cdot\mathbf{b})^2 = (a_1 b_2 - a_2 b_1)^2$. If $\mathbf{b}\neq\mathbf{0}$, (6.2.2) says the left side is $\|\mathbf{b}\|^2\|\mathbf{a} - c\mathbf{b}\|^2$, the square of (base) × (height) for the parallelogram on $\mathbf{a}, \mathbf{b}$ with base $\mathbf{b}$. So the area is $|a_1 b_2 - a_2 b_1|$; if $\mathbf{b}=\mathbf{0}$, both are 0. This is $\left|\det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right|$.

Solution to Exercise 6.2.44

Uses. Theorem 6.2.11 (Triangle inequality), Theorem 6.2.1 (Cauchy–Schwarz inequality), Notation 6.1.1 (Real and complex scalars)
Hint 1. The "if" direction is a computation. For "only if", first show that the sum $\mathbf{s} = \mathbf{x}_1 + \cdots + \mathbf{x}_n$ is nonzero, and let $\mathbf{u} = \mathbf{s}/\|\mathbf{s}\|$.
Hint 2. Write $\|\mathbf{s}\| = \langle \mathbf{s}, \mathbf{u}\rangle = \sum_k \langle \mathbf{x}_k, \mathbf{u}\rangle$, take real parts, and bound each term by $\|\mathbf{x}_k\|$.
Hint 3. Equality throughout forces $\langle \mathbf{x}_k, \mathbf{u}\rangle = \|\mathbf{x}_k\|\|\mathbf{u}\|$ for each k. Apply Theorem 6.2.11(b).
Sketch. If $\mathbf{x}_k = r_k\mathbf{u}$, both sides equal $\sum_k r_k$. Conversely, $\mathbf{s}\neq\mathbf{0}$ because $\|\mathbf{s}\| = \sum_k \|\mathbf{x}_k\| > 0$. With $\mathbf{u} = \mathbf{s}/\|\mathbf{s}\|$, $\|\mathbf{s}\| = \sum_k \operatorname{Re}\langle \mathbf{x}_k, \mathbf{u}\rangle \le \sum_k \|\mathbf{x}_k\| = \|\mathbf{s}\|$, so $\operatorname{Re}\langle \mathbf{x}_k, \mathbf{u}\rangle = |\langle \mathbf{x}_k, \mathbf{u}\rangle| = \|\mathbf{x}_k\|$ for each k, and Theorem 6.2.11(b) makes $\mathbf{x}_k$ a nonnegative multiple of $\mathbf{u}$.

Suppose $\mathbf{x}_k = r_k\mathbf{u}$ with $\|\mathbf{u}\| = 1$ and $r_k \ge 0$. Then $\left\|\sum_k \mathbf{x}_k\right\| = \left\|\left(\sum_k r_k\right)\mathbf{u}\right\| = \sum_k r_k = \sum_k \|\mathbf{x}_k\|$.

Conversely, suppose equality holds, and let $\mathbf{s} = \mathbf{x}_1 + \cdots + \mathbf{x}_n$. Since not all $\mathbf{x}_k$ are $\mathbf{0}$, $\|\mathbf{s}\| = \sum_k \|\mathbf{x}_k\| > 0$, so $\mathbf{s} \neq \mathbf{0}$. Let $\mathbf{u} = \mathbf{s}/\|\mathbf{s}\|$, a unit vector with $\langle \mathbf{s}, \mathbf{u}\rangle = \|\mathbf{s}\|$. By linearity in the first slot, $\operatorname{Re} z \le |z|$, and Theorem 6.2.1,

$$\|\mathbf{s}\| = \operatorname{Re}\langle \mathbf{s}, \mathbf{u}\rangle = \sum_k \operatorname{Re}\langle \mathbf{x}_k, \mathbf{u}\rangle$$
$$\le \sum_k |\langle \mathbf{x}_k, \mathbf{u}\rangle| \le \sum_k \|\mathbf{x}_k\| = \|\mathbf{s}\|.$$

So equality holds throughout. Since each term satisfies $\operatorname{Re}\langle \mathbf{x}_k, \mathbf{u}\rangle \le |\langle \mathbf{x}_k, \mathbf{u}\rangle| \le \|\mathbf{x}_k\|$, equality of the sums forces equality in every term. As in the proof of Theorem 6.2.11, $\operatorname{Re} z = |z|$ means that z is a nonnegative real number, so $\langle \mathbf{x}_k, \mathbf{u}\rangle = \|\mathbf{x}_k\| = \|\mathbf{x}_k\|\|\mathbf{u}\|$. By Theorem 6.2.11(b), one of $\mathbf{x}_k, \mathbf{u}$ is a nonnegative real multiple of the other. If $\mathbf{u} = b\mathbf{x}_k$ with $b \ge 0$, then $b > 0$ since $\mathbf{u}\neq\mathbf{0}$, and $\mathbf{x}_k = b^{-1}\mathbf{u}$. So in every case $\mathbf{x}_k = r_k\mathbf{u}$ with $r_k \ge 0$.

Solution to Exercise 6.2.45

Uses. Definition 6.2.16 (Angle), Theorem 6.2.1 (Cauchy–Schwarz inequality), (6.2.1)

Techniques. Orthogonal decomposition (Technique 6.2.4)
Hint 1. Angles do not change when vectors are scaled by positive numbers, so assume $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are unit vectors. Let $\alpha = \theta(\mathbf{x}, \mathbf{y})$ and $\beta = \theta(\mathbf{y}, \mathbf{z})$. If $\alpha + \beta \geq \pi$ there is nothing to prove.
Hint 2. Decompose $\mathbf{x} = (\cos\alpha)\mathbf{y} + \mathbf{w}$ and $\mathbf{z} = (\cos\beta)\mathbf{y} + \mathbf{w}'$ with $\mathbf{w}, \mathbf{w}'$ orthogonal to $\mathbf{y}$. Compute $\|\mathbf{w}\|$ and $\|\mathbf{w}'\|$.
Hint 3. Expand $\langle \mathbf{x}, \mathbf{z} \rangle$ and bound $\langle \mathbf{w}, \mathbf{w}' \rangle$ below with Cauchy–Schwarz. You should get $\langle \mathbf{x}, \mathbf{z} \rangle \geq \cos(\alpha + \beta)$.
Sketch. Normalize. With α, β as above, $\mathbf{x} = (\cos\alpha)\mathbf{y} + \mathbf{w}$ and $\mathbf{z} = (\cos\beta)\mathbf{y} + \mathbf{w}'$, where $\mathbf{w}, \mathbf{w}' \perp \mathbf{y}$, $\|\mathbf{w}\| = \sin\alpha$, $\|\mathbf{w}'\| = \sin\beta$. Then $\langle \mathbf{x}, \mathbf{z} \rangle = \cos\alpha\cos\beta + \langle \mathbf{w}, \mathbf{w}' \rangle$, and Cauchy–Schwarz bounds the last term below by $-\sin\alpha\sin\beta$, so $\langle \mathbf{x}, \mathbf{z} \rangle \geq \cos(\alpha + \beta)$. If $\alpha + \beta \leq \pi$, since cos is decreasing on $[0, \pi]$, $\theta(\mathbf{x}, \mathbf{z}) \leq \alpha + \beta$; otherwise the claim is trivial since $\theta \leq \pi$.

For $s, t > 0$, $\langle s\mathbf{x}, t\mathbf{y} \rangle / (\|s\mathbf{x}\|\|t\mathbf{y}\|) = \langle \mathbf{x}, \mathbf{y} \rangle / (\|\mathbf{x}\|\|\mathbf{y}\|)$, so angles are unchanged by positive scaling, and we may assume $\|\mathbf{x}\| = \|\mathbf{y}\| = \|\mathbf{z}\| = 1$. Let $\alpha = \theta(\mathbf{x}, \mathbf{y})$ and $\beta = \theta(\mathbf{y}, \mathbf{z})$, both in $[0, \pi]$. If $\alpha + \beta > \pi$, the claim holds because $\theta(\mathbf{x}, \mathbf{z}) \leq \pi$. Assume $\alpha + \beta \leq \pi$.

By Definition 6.2.16, $\langle \mathbf{x}, \mathbf{y} \rangle = \cos\alpha$ and $\langle \mathbf{z}, \mathbf{y} \rangle = \cos\beta$. As in Technique 6.2.4 (with $\|\mathbf{y}\| = 1$), let $\mathbf{w} = \mathbf{x} - (\cos\alpha)\mathbf{y}$ and $\mathbf{w}' = \mathbf{z} - (\cos\beta)\mathbf{y}$, so that $\langle \mathbf{w}, \mathbf{y} \rangle = \langle \mathbf{w}', \mathbf{y} \rangle = 0$. By (6.2.1), $\|\mathbf{w}\|^2 = 1 - \cos^2\alpha = \sin^2\alpha$, so $\|\mathbf{w}\| = \sin\alpha$, as $\sin\alpha \geq 0$ on $[0, \pi]$; similarly $\|\mathbf{w}'\| = \sin\beta$. Expanding, and using that the cross terms $\langle \mathbf{y}, \mathbf{w}' \rangle$ and $\langle \mathbf{w}, \mathbf{y} \rangle$ vanish,

$$\begin{aligned} \langle \mathbf{x}, \mathbf{z} \rangle &= \langle (\cos\alpha)\mathbf{y} + \mathbf{w}, (\cos\beta)\mathbf{y} + \mathbf{w}' \rangle \\ &= \cos\alpha\cos\beta + \langle \mathbf{w}, \mathbf{w}' \rangle. \end{aligned}$$

By Theorem 6.2.1, $\langle \mathbf{w}, \mathbf{w}' \rangle \geq -\|\mathbf{w}\|\|\mathbf{w}'\| = -\sin\alpha\sin\beta$. So

$$\begin{aligned} \cos\theta(\mathbf{x}, \mathbf{z}) = \langle \mathbf{x}, \mathbf{z} \rangle &\geq \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ &= \cos(\alpha + \beta). \end{aligned}$$

Both $\theta(\mathbf{x}, \mathbf{z})$ and $\alpha + \beta$ lie in $[0, \pi]$, where cos is strictly decreasing, so $\theta(\mathbf{x}, \mathbf{z}) \leq \alpha + \beta$.

Solution to Exercise 6.2.46

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.13 (Integral inner products)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. The constraints say $\langle f, 1 \rangle = \langle f, t \rangle = 1$. So $\langle f, g \rangle$ is known for every $g = \alpha + \beta t$. Apply Cauchy–Schwarz to f and such a g.
Hint 2. The bound is $\|f\|^2 \geq (\alpha + \beta)^2 / \|\alpha + \beta t\|^2$. The best g should itself satisfy the constraints, by the equality case.
Hint 3. Find $g = \alpha + \beta t$ with $\langle g, 1 \rangle = \langle g, t \rangle = 1$: you get $g = 6t - 2$.
Sketch. For $g = 6t - 2$: $\langle f, g \rangle = 6 - 2 = 4$ and $\|g\|^2 = 4$, so $16 \leq 4\|f\|^2$, that is, $\int f^2 \geq 4$. The function g satisfies both constraints and has $\int g^2 = 4$. Equality forces $f = cg$, and the constraints give $c = 1$. So the minimum is 4, attained only by $f(t) = 6t - 2$.

Let $g(t) = 6t - 2$. In $C[0, 1]$ with $\langle f, h \rangle = \int_0^1 fh$,

$$\langle f, g \rangle = 6\int_0^1 tf - 2\int_0^1 f = 6 - 2 = 4,$$

$$\|g\|^2 = \int_0^1 \left(36t^2 - 24t + 4\right) \mathrm{d}t = 12 - 12 + 4 = 4,$$

for every f satisfying the constraints. By Theorem 6.2.1, $16 = \langle f, g \rangle^2 \leq \|f\|^2\|g\|^2 = 4\|f\|^2$, so $\int_0^1 f^2 \geq 4$.

The function g itself satisfies the constraints: $\int_0^1 (6t - 2)\,\mathrm{d}t = 3 - 2 = 1$ and $\int_0^1 \left(6t^2 - 2t\right)\mathrm{d}t = 2 - 1 = 1$. And $\int_0^1 g^2 = 4$. So the minimum is 4. If f attains it, equality holds in Cauchy–Schwarz, so $f = cg$ (as $g \neq \mathbf{0}$), and $1 = \int_0^1 f = c\int_0^1 g = c$. So $f = g$ is the only minimizer.

Here is how to find g. The constraints determine $\langle f, \alpha + \beta t \rangle = \alpha + \beta$ for all α, β, so each $g = \alpha + \beta t \neq \mathbf{0}$ gives the bound $\|f\|^2 \geq (\alpha + \beta)^2 / \|g\|^2$. If some such g satisfies the constraints itself, then $f = g$ attains its own bound, and no f can do better. Solving $\langle \alpha + \beta t, 1 \rangle = \alpha + \beta/2 = 1$ and $\langle \alpha + \beta t, t \rangle = \alpha/2 + \beta/3 = 1$ gives $\beta = 6$ and $\alpha = -2$. (In the language of Section 6.4, the minimizer is the vector of least norm in an affine subspace, and it lies in $\mathrm{span}(1, t)$.)

Solution to Exercise 6.2.47

Uses. Corollary 6.2.26 (Engel's form)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Engel's form (Corollary 6.2.26) needs squares in the numerators. Multiply the numerator and denominator of $a/(b+c)$ by a.
Hint 2. You get a lower bound $(a+b+c)^2/(2(ab+bc+ca))$. Compare $(a+b+c)^2$ with $3(ab+bc+ca)$.
Sketch. Write $a/(b+c) = a^2/(a(b+c))$. By Corollary 6.2.26, the sum is at least $(a+b+c)^2/(2(ab+bc+ca))$. Finally $(a+b+c)^2 \geq 3(ab+bc+ca)$, because the difference is half the sum of the squares $(a-b)^2$, $(b-c)^2$, $(c-a)^2$. Equality in the last step forces $a=b=c$, and then equality holds.

Since $a, b, c > 0$, the denominators $a(b+c)$, $b(c+a)$, $c(a+b)$ are positive. By Corollary 6.2.26,

$$\begin{aligned}
&\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \\
&= \frac{a^2}{a(b+c)} + \frac{b^2}{b(c+a)} + \frac{c^2}{c(a+b)} \\
&\geq \frac{(a+b+c)^2}{2(ab+bc+ca)},
\end{aligned}$$

since the denominators add up to $2(ab+bc+ca)$. Now

$$\begin{aligned}
(a+b+c)^2 - 3(ab+bc+ca) &= a^2+b^2+c^2-ab-bc-ca \\
&= \frac{1}{2}\Big((a-b)^2+(b-c)^2+(c-a)^2\Big) \geq 0,
\end{aligned}$$

so the right side is at least $3(ab+bc+ca)/(2(ab+bc+ca)) = 3/2$.

If equality holds, the last step is an equality, so $(a-b)^2+(b-c)^2+(c-a)^2 = 0$, and $a=b=c$ by Technique 6.1.9. Conversely, for $a=b=c$ each term is $1/2$.

Solution to Exercise 6.2.48

Uses. Proposition 6.1.29 (Expanding a norm), Corollary 6.2.14 (Distances and the reverse triangle inequality), Proposition 6.1.27 (Basic properties of the norm)
Hint 1. (a) Expand the left side squared with Proposition 6.1.29; the coefficients $1/\|\mathbf{x}\|^2$ and $1/\|\mathbf{y}\|^2$ are real.
Hint 2. (b) Translating all four points by $-\mathbf{w}$ changes no distance, so you may assume $\mathbf{w} = \mathbf{0}$. Dispose of the case where one of $\mathbf{x}, \mathbf{y}, \mathbf{z}$ is $\mathbf{0}$.
Hint 3. Apply the triangle inequality to the three points $\mathbf{v}/\|\mathbf{v}\|^2$ for $\mathbf{v} = \mathbf{x}, \mathbf{y}, \mathbf{z}$, rewrite each distance with (a), and clear denominators.
Sketch. (a) Expand the square of the left side with Proposition 6.1.29 and put everything over $\|\mathbf{x}\|^2\|\mathbf{y}\|^2$; the numerator is $\|\mathbf{x}-\mathbf{y}\|^2$. (b) Translate so that $\mathbf{w} = \mathbf{0}$; the cases with a zero vector are immediate. Otherwise, let $\iota(\mathbf{v}) = \mathbf{v}/\|\mathbf{v}\|^2$. The triangle inequality for the three points $\iota(\mathbf{x}), \iota(\mathbf{y}), \iota(\mathbf{z})$ becomes, by (a), the claim divided by $\|\mathbf{x}\|\|\mathbf{y}\|\|\mathbf{z}\|$.

(a) Let $\alpha = 1/\|\mathbf{x}\|^2$ and $\beta = 1/\|\mathbf{y}\|^2$, both real. By Proposition 6.1.29 (and linearity),

$$\begin{aligned}
\|\alpha\mathbf{x} - \beta\mathbf{y}\|^2 &= \alpha^2\|\mathbf{x}\|^2 - 2\alpha\beta\,\mathrm{Re}\langle\mathbf{x},\mathbf{y}\rangle + \beta^2\|\mathbf{y}\|^2 \\
&= \frac{1}{\|\mathbf{x}\|^2} - \frac{2\,\mathrm{Re}\langle\mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2} + \frac{1}{\|\mathbf{y}\|^2} \\
&= \frac{\|\mathbf{y}\|^2 - 2\,\mathrm{Re}\langle\mathbf{x},\mathbf{y}\rangle + \|\mathbf{x}\|^2}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2} = \frac{\|\mathbf{x}-\mathbf{y}\|^2}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}.
\end{aligned}$$

Taking nonnegative square roots gives (a).

(b) Replacing $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}$ by $\mathbf{x}-\mathbf{w}, \mathbf{y}-\mathbf{w}, \mathbf{z}-\mathbf{w}, \mathbf{0}$ changes none of the six differences, so we may assume $\mathbf{w} = \mathbf{0}$. The claim is then

$$\|\mathbf{x}-\mathbf{y}\|\|\mathbf{z}\| \le \|\mathbf{x}-\mathbf{z}\|\|\mathbf{y}\| + \|\mathbf{x}\|\|\mathbf{y}-\mathbf{z}\|.$$

If $\mathbf{x}=\mathbf{0}$, both sides are $\|\mathbf{y}\|\|\mathbf{z}\|$. If $\mathbf{y}=\mathbf{0}$, both sides are $\|\mathbf{x}\|\|\mathbf{z}\|$. If $\mathbf{z}=\mathbf{0}$, the left side is 0. So assume $\mathbf{x},\mathbf{y},\mathbf{z}\neq\mathbf{0}$, and write $\iota(\mathbf{v}) = \mathbf{v}/\|\mathbf{v}\|^2$ for $\mathbf{v}\neq\mathbf{0}$. By Corollary 6.2.14(a),

$$\|\iota(\mathbf{x})-\iota(\mathbf{y})\| \le \|\iota(\mathbf{x})-\iota(\mathbf{z})\| + \|\iota(\mathbf{z})-\iota(\mathbf{y})\|,$$

and by (a) this says

$$\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|\|\mathbf{y}\|} \le \frac{\|\mathbf{x}-\mathbf{z}\|}{\|\mathbf{x}\|\|\mathbf{z}\|} + \frac{\|\mathbf{z}-\mathbf{y}\|}{\|\mathbf{z}\|\|\mathbf{y}\|}.$$

Multiplying by $\|\mathbf{x}\|\|\mathbf{y}\|\|\mathbf{z}\| > 0$ gives the claim, since $\|\mathbf{z}-\mathbf{y}\| = \|\mathbf{y}-\mathbf{z}\|$.

The map ι is *inversion* in the unit sphere, and (a) says how it distorts distances. Ptolemy's inequality is the triangle inequality, seen after inversion.

Section 6.3

Solution to Exercise 6.3.7

Uses. Definition 6.1.8 (Dot product), Definition 6.3.1 (Orthogonality), Proposition 6.1.10 (The dot product is an inner product)
Hint 1. Conjugate the entries of the second vector, multiply entrywise, and add. You get a linear equation in a.
Hint 2. The equation is $3+i-ia=0$.
Sketch. $(1,i,a)\cdot(2,1+i,i) = 2+i(1-i)-ia = 3+i-ia$, which vanishes for $a=1-3i$. Then the squared norms are 12 and 7, and the sum $(3,1+2i,1-2i)$ has squared norm $9+5+5 = 19 = 12+7$.

By Definition 6.1.8,

$$\begin{aligned}(1,i,a)\cdot(2,1+i,i) &= 1\cdot 2 + i\cdot\overline{(1+i)} + a\cdot\bar{i}\\ &= 2+i(1-i)-ia = 3+i-ia.\end{aligned}$$

This vanishes if and only if $ia = 3+i$, that is, $a = (3+i)/i = (3+i)(-i) = 1-3i$.

With $\mathbf{u}=(1,i,1-3i)$ and $\mathbf{v}=(2,1+i,i)$, Proposition 6.1.10 gives $\|\mathbf{u}\|^2 = 1+1+10 = 12$ and $\|\mathbf{v}\|^2 = 4+2+1 = 7$. The sum is $\mathbf{u}+\mathbf{v} = (3,1+2i,1-2i)$, with $\|\mathbf{u}+\mathbf{v}\|^2 = 9+5+5 = 19 = 12+7$, as Theorem 6.3.3 predicts.

Solution to Exercise 6.3.11

Uses. Definition 6.3.8 (Vectors orthogonal to a set), Proposition 6.3.9 (Properties of $A^\perp$)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For the first identity, unwind the definitions: being orthogonal to every element of $A\cup B$ means being orthogonal to every element of A and to every element of B.
Hint 2. For $A\subseteq\left(A^\perp\right)^\perp$, take $\mathbf{a}\in A$ and $\mathbf{v}\in A^\perp$, and compute $\langle\mathbf{a},\mathbf{v}\rangle$ using symmetry of orthogonality.
Hint 3. $\left(A^\perp\right)^\perp$ is always a subspace, by Proposition 6.3.9(a).
Sketch. $\mathbf{v}\perp(A\cup B)$ if and only if $\mathbf{v}\perp A$ and $\mathbf{v}\perp B$. If $\mathbf{a}\in A$ and $\mathbf{v}\in A^\perp$, then $\langle\mathbf{a},\mathbf{v}\rangle = \overline{\langle\mathbf{v},\mathbf{a}\rangle} = 0$, so $\mathbf{a}\in\left(A^\perp\right)^\perp$. For $A=\{(1,0)\}$ in $\mathbb{R}^2$, $\left(A^\perp\right)^\perp$ is a subspace, so it is not A.

A vector $\mathbf{v}$ lies in $(A\cup B)^\perp$ if and only if $\langle\mathbf{v},\mathbf{c}\rangle = 0$ for every $\mathbf{c}$ in A or in B, that is, if and only if $\mathbf{v}\in A^\perp$ and $\mathbf{v}\in B^\perp$. So $(A\cup B)^\perp = A^\perp\cap B^\perp$.

Let $\mathbf{a}\in A$. For every $\mathbf{v}\in A^\perp$, $\langle\mathbf{a},\mathbf{v}\rangle = \overline{\langle\mathbf{v},\mathbf{a}\rangle} = \bar{0} = 0$. So $\mathbf{a}$ is orthogonal to every vector of $A^\perp$, that is, $\mathbf{a}\in\left(A^\perp\right)^\perp$.

For $A=\{(1,0)\}$ in $\mathbb{R}^2$ with the dot product, $\left(A^\perp\right)^\perp$ is a subspace by Proposition 6.3.9(a), so it contains $\mathbf{0}\notin A$. In fact $A^\perp = \operatorname{span}((0,1))$ and $\left(A^\perp\right)^\perp = \operatorname{span}((1,0))$. (Section 6.4 shows that $\left(W^\perp\right)^\perp = W$ when W is a finite-dimensional subspace.)

Solution to Exercise 6.3.23

Uses. Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 6.3.22 (Parseval's identity)
Hint 1. Check the three dot products between different vectors, and the three norms. Then use Corollary 6.3.14.
Hint 2. The jth coordinate is $\mathbf{v} \cdot \mathbf{u}_j$.
Sketch. The pairwise dot products of $(1,1,0)$, $(1,-1,1)$, $(-1,1,2)$ vanish, and the scalings make them unit vectors. The coordinates are $4/\sqrt{2} = 2\sqrt{2}$, $4/\sqrt{3}$, and $2/\sqrt{6}$, and $8 + 16/3 + 2/3 = 14 = \|\mathbf{v}\|^2$.

The unscaled vectors $(1,1,0)$, $(1,-1,1)$, $(-1,1,2)$ have pairwise dot products $1-1+0=0$, $-1+1+0=0$, and $-1-1+2=0$, and squared norms 2, 3, 6. So $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is an orthonormal list of length $3 = \dim \mathbb{R}^3$, hence an orthonormal basis by Corollary 6.3.14.

By Proposition 6.3.18,

$$\begin{aligned} \mathbf{v} \cdot \mathbf{u}_1 &= \frac{4}{\sqrt{2}} = 2\sqrt{2}, \\ \mathbf{v} \cdot \mathbf{u}_2 &= \frac{3-1+2}{\sqrt{3}} = \frac{4}{\sqrt{3}}, \\ \mathbf{v} \cdot \mathbf{u}_3 &= \frac{-3+1+4}{\sqrt{6}} = \frac{2}{\sqrt{6}}, \end{aligned}$$

so $[\mathbf{v}]_{\mathcal{O}} = \left(2\sqrt{2}, 4/\sqrt{3}, 2/\sqrt{6}\right)$. By Theorem 6.3.22(b), the sum of the squares of the coordinates should be $\|\mathbf{v}\|^2 = 9+1+4 = 14$, and indeed $8 + 16/3 + 4/6 = 8 + 16/3 + 2/3 = 14$.

Solution to Exercise 6.3.37

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 6.3.34 (Characterizations of unitary matrices), Proposition 2.4.14 (Algebra of matrices)
Hint 1. Compute $(UW)^*(UW)$ with Proposition 6.3.27(b).
Hint 2. $U^{-1} = U^*$, and Proposition 6.3.34(iv). For the dot products, move U across with Proposition 6.3.27(d).
Sketch. $(UW)^*UW = W^*U^*UW = W^*W = I$. $U^{-1} = U^*$ is unitary by Proposition 6.3.34. And $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot (U^*U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$; take $\mathbf{y} = \mathbf{x}$ for the norms.

By Proposition 6.3.27(b) and associativity,

$$(UW)^*(UW) = W^*(U^*U)W = W^*I_nW = W^*W = I_n,$$

so UW is unitary. By Proposition 6.3.34, $U^{-1} = U^*$, and U^* is unitary by (iv) of the same proposition.

By Proposition 6.3.27(d) with $A = U$ and the vector $U\mathbf{y}$ in place of $\mathbf{y}$,

$$(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot (U^*U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}.$$

Taking $\mathbf{y} = \mathbf{x}$ gives $\|U\mathbf{x}\|^2 = \|\mathbf{x}\|^2$, and norms are nonnegative, so $\|U\mathbf{x}\| = \|\mathbf{x}\|$. (Section 8.4 shows that, conversely, a square matrix with $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ is unitary.)

Solution to Exercise 6.3.38

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Theorem 6.3.29 (Orthonormal columns), Definition 6.1.8 (Dot product)
Hint 1. The first column $(1,i)/\sqrt{2}$ is already a unit vector. Write the two remaining conditions on the second column.
Hint 2. Orthogonality to the first column gives one linear equation; the norm gives $|a|^2 + |b|^2 = 2$.
Sketch. Orthogonality: $(a,b) \cdot (1,i) = a - ib = 0$, so $a = ib$. Unit length: $|a|^2 + |b|^2 = 2|b|^2 = 2$. So $(a,b) = (ib, b)$ with $|b| = 1$.

By Theorem 6.3.29 the matrix is unitary if and only if its columns $\mathbf{c}_1 = (1,i)/\sqrt{2}$ and $\mathbf{c}_2 = (a,b)/\sqrt{2}$ are orthonormal. Here $\|\mathbf{c}_1\|^2 = (1+1)/2 = 1$. Next,

$$\mathbf{c}_2 \cdot \mathbf{c}_1 = \frac{1}{2}\left(a \cdot 1 + b \cdot \bar{i}\right) = \frac{1}{2}(a - ib),$$

which vanishes if and only if $a = ib$. Then $\|\mathbf{c}_2\|^2 = \left(|a|^2 + |b|^2\right)/2 = |b|^2$, which is 1 if and only if $|b| = 1$. So the matrix is unitary exactly when $a = ib$ with $|b| = 1$. For $b = -i$ we recover the first matrix of Example 6.3.35(d).

Solution to Exercise 6.3.44

Uses. Definition 6.1.8 (Dot product), Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 6.3.22 (Parseval's identity), Remark 6.3.19 (Which slot)
Hint 1. For (a), compute both squared norms and one dot product; then count.
Hint 2. For (b), the coordinates are $\mathbf{v}\cdot\mathbf{u}_1$ and $\mathbf{v}\cdot\mathbf{u}_2$. Conjugate the entries of $\mathbf{u}_j$.
Hint 3. For (c), the coefficients are the conjugates of those in (b).
Sketch. (a) $\|(1,1+i)\|^2 = \|(-1+i,1)\|^2 = 3$ and $(1,1+i)\cdot(-1+i,1) = (-1-i)+(1+i) = 0$. (b) $[\mathbf{v}]_{\mathcal{O}} = \big((5+3i)/\sqrt{3}, (-2+i)/\sqrt{3}\big)$, and $(34+5)/3 = 13 = \|\mathbf{v}\|^2$. (c) The conjugated coefficients give $((8-4i)/3, (6+i)/3) \neq \mathbf{v}$.

(a) $\|(1,1+i)\|^2 = 1+2 = 3$ and $\|(-1+i,1)\|^2 = 2+1 = 3$, so $\mathbf{u}_1$ and $\mathbf{u}_2$ are unit vectors. Also

$$\begin{aligned}(1,1+i)\cdot(-1+i,1) &= 1\cdot\overline{(-1+i)} + (1+i)\cdot 1\\ &= (-1-i)+(1+i) = 0.\end{aligned}$$

So $\mathcal{O}$ is an orthonormal list of length $2 = \dim\mathbb{C}^2$, hence an orthonormal basis by Corollary 6.3.14.

(b) By Proposition 6.3.18,

$$\begin{aligned}\langle \mathbf{v},\mathbf{u}_1\rangle &= \frac{1}{\sqrt{3}}\Big(2\cdot 1 + 3i\cdot\overline{(1+i)}\Big)\\ &= \frac{1}{\sqrt{3}}(2+3i+3) = \frac{5+3i}{\sqrt{3}},\\ \langle \mathbf{v},\mathbf{u}_2\rangle &= \frac{1}{\sqrt{3}}\Big(2\cdot\overline{(-1+i)} + 3i\cdot 1\Big)\\ &= \frac{1}{\sqrt{3}}(-2-2i+3i) = \frac{-2+i}{\sqrt{3}}.\end{aligned}$$

So $[\mathbf{v}]_{\mathcal{O}} = \big((5+3i)/\sqrt{3}, (-2+i)/\sqrt{3}\big)$. Check: $|5+3i|^2/3 + |-2+i|^2/3 = (34+5)/3 = 13 = |2|^2 + |3i|^2$, as Theorem 6.3.22(b) requires. One can also check directly that $(5+3i)(1,1+i)/3 + (-2+i)(-1+i,1)/3 = (6,9i)/3 = \mathbf{v}$, using $(5+3i)(1+i) = 2+8i$ and $(-2+i)(-1+i) = 1-3i$.

(c) The exchanged coefficients are the conjugates $\langle \mathbf{u}_1,\mathbf{v}\rangle = (5-3i)/\sqrt{3}$ and $\langle \mathbf{u}_2,\mathbf{v}\rangle = (-2-i)/\sqrt{3}$. Since $(5-3i)(1+i) = 8+2i$ and $(-2-i)(-1+i) = 3-i$,

$$\begin{aligned}&\frac{1}{3}\big((5-3i)(1,1+i) + (-2-i)(-1+i,1)\big)\\ &= \frac{1}{3}(5-3i+3-i, 8+2i-2-i) = \left(\frac{8-4i}{3}, \frac{6+i}{3}\right),\end{aligned}$$

which is not $\mathbf{v}$. Over $\mathbb{C}$ the vector belongs in the first slot (Remark 6.3.19).

Solution to Exercise 6.3.45

Uses. Example 6.3.10 (Perps are solution sets), Theorem 6.3.29 (Orthonormal columns), Definition 6.3.33 (Unitary and orthogonal matrices)
Hint 1. By Example 6.3.10, the vectors orthogonal to both are the solutions of a homogeneous system with these two rows.
Hint 2. Solve $x_1+2x_2+2x_3 = 0$, $2x_1+x_2-2x_3 = 0$, then normalize. For the matrices, the third column must be one of the vectors you found.
Sketch. Adding the equations gives $x_2 = -x_1$, and then $x_3 = x_1/2$, so the solutions are $t(2,-2,1)$, and the unit ones are $\pm\frac{1}{3}(2,-2,1)$. A 3×3 matrix is orthogonal exactly when its columns are orthonormal, so the third column must be one of these two vectors: there are exactly two such matrices.

By Example 6.3.10, $\mathbf{x}$ is orthogonal to both vectors if and only if

$$x_1+2x_2+2x_3 = 0, \quad 2x_1+x_2-2x_3 = 0.$$

Adding the equations gives $3x_1 + 3x_2 = 0$, so $x_2 = -x_1$, and then the first equation gives $-x_1 + 2x_3 = 0$, so $x_3 = x_1/2$. Conversely, every $\mathbf{x} = t(2,-2,1)$ satisfies both equations. Since $\|(2,-2,1)\| = 3$, the unit vectors among these are $\mathbf{x} = \pm\frac{1}{3}(2,-2,1)$.

By Theorem 6.3.29 and Definition 6.3.33, a real 3×3 matrix is orthogonal if and only if its columns are orthonormal. The first two given columns are orthonormal (Example 6.3.16(c)), so the matrix is orthogonal exactly when its third column is a unit vector orthogonal to both, that is, $\pm\frac{1}{3}(2,-2,1)$. So there are exactly two such matrices,

$$\frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix} \quad \text{and} \quad \frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{pmatrix}.$$

Solution to Exercise 6.3.46

Uses. Definition 6.3.1 (Orthogonality), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.29 (Expanding a norm), Proposition 6.3.9 (Properties of $A^\perp$), Non-example 6.3.4 (The converse fails over $\mathbb{C}$)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a) and (b), look for counterexamples in which a vector is repeated.
Hint 2. For (d), expand both norms with Proposition 6.1.29. Do the two hypotheses say different things?
Hint 3. For (e), combine (c) and (d) of Proposition 6.3.9.
Sketch. (a) False: $\mathbf{u} = \mathbf{w} = \mathbf{e}_1$, $\mathbf{v} = \mathbf{e}_2$. (b) False: in $\mathbb{R}^3$, $\mathbf{v}_1 = \mathbf{v}_3 = \mathbf{e}_1$, $\mathbf{v}_2 = \mathbf{e}_2$. (c) True: $\langle c_i\mathbf{v}_i, c_j\mathbf{v}_j\rangle = c_i\overline{c_j}\langle \mathbf{v}_i, \mathbf{v}_j\rangle$. (d) Both hypotheses say $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$; true over $\mathbb{R}$, false over $\mathbb{C}$ ($\mathbf{u} = 1$, $\mathbf{v} = i$). (e) True: $\mathbf{x} \in \mathcal{L}^\perp = \operatorname{span}(\mathcal{L})^\perp = V^\perp = \{\mathbf{0}\}$.

(a) False. In $\mathbb{R}^2$, take $\mathbf{u} = \mathbf{w} = \mathbf{e}_1$ and $\mathbf{v} = \mathbf{e}_2$. Then $\mathbf{u} \perp \mathbf{v}$ and $\mathbf{v} \perp \mathbf{w}$, but $\langle \mathbf{u}, \mathbf{w}\rangle = 1$.

(b) False. In $\mathbb{R}^3$, take $\mathbf{v}_1 = \mathbf{v}_3 = \mathbf{e}_1$ and $\mathbf{v}_2 = \mathbf{e}_2$. The hypotheses hold, but $\mathbf{v}_1 - \mathbf{v}_3 = \mathbf{0}$. The hypothesis says nothing about the pair $\mathbf{v}_1, \mathbf{v}_3$, so the list need not be orthogonal, and Proposition 6.3.13 does not apply. (The space is $\mathbb{R}^3$ so that the dependence is not merely forced by the dimension.)

(c) True. For $i \neq j$, Proposition 6.1.4 gives $\langle c_i\mathbf{v}_i, c_j\mathbf{v}_j\rangle = c_i\overline{c_j}\langle \mathbf{v}_i, \mathbf{v}_j\rangle = 0$.

(d) True over $\mathbb{R}$ and false over $\mathbb{C}$. By Proposition 6.1.29,

$$\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \pm 2\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2,$$

so each hypothesis says exactly that $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle = 0$; the second adds nothing to the first. Over $\mathbb{R}$ this means $\mathbf{u} \perp \mathbf{v}$ (Theorem 6.3.3(b)). Over $\mathbb{C}$, $\mathbf{u} = 1$ and $\mathbf{v} = i$ in $\mathbb{C}$ satisfy both, since $|1 \pm i|^2 = 2$, but $\langle 1, i\rangle = -i \neq 0$ (Non-example 6.3.4).

(e) True. Let $\mathcal{L} = \mathbf{v}_1, \ldots, \mathbf{v}_k$. By Proposition 6.3.9(c) and (d), $\mathbf{x} \in \mathcal{L}^\perp = \operatorname{span}(\mathcal{L})^\perp = V^\perp = \{\mathbf{0}\}$.

Solution to Exercise 6.3.47

Uses. Proposition 6.3.13 (Orthogonal lists are independent), Proposition 6.3.9 (Properties of $A^\perp$), Definition 6.1.2 (Inner product)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. In $\mathbb{F}_2$, $1 + 1 = 0$. A vector with an even number of entries 1 is orthogonal to itself.
Hint 2. Try $(1,1,0,0)$, $(0,0,1,1)$, and their sum.
Hint 3. For (c), as in Proposition 6.3.9(c), it suffices to test orthogonality to the two spanning vectors.
Sketch. (a) $(1,1,0,0)$. (b) $(1,1,0,0)$, $(0,0,1,1)$, $(1,1,1,1)$ are pairwise orthogonal and the third is the sum of the first two; the step "$a_j\mathbf{v}_j \cdot \mathbf{v}_j = 0$ forces $a_j = 0$" fails because $\mathbf{v}_j \cdot \mathbf{v}_j = 0$. (c) The conditions are $x_1 = x_2$ and $x_3 = x_4$, which describe W. So $W \cap W^\perp = W \neq \{\mathbf{0}\}$, against Proposition 6.3.9(d), whose proof used definiteness.

(a) $\mathbf{x} = (1,1,0,0)$ has $\mathbf{x} \cdot \mathbf{x} = 1 + 1 = 0$ in $\mathbb{F}_2$.

(b) Let $\mathbf{v}_1 = (1,1,0,0)$, $\mathbf{v}_2 = (0,0,1,1)$, $\mathbf{v}_3 = (1,1,1,1)$. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = 1 + 1 = 0$, and $\mathbf{v}_2 \cdot \mathbf{v}_3 = 1 + 1 = 0$. But $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = (2,2,2,2) = \mathbf{0}$, a nontrivial relation. In the first proof of Proposition 6.3.13, taking the product with $\mathbf{v}_j$ still gives $a_j(\mathbf{v}_j \cdot \mathbf{v}_j) = 0$, but now $\mathbf{v}_j \cdot \mathbf{v}_j = 0$, so nothing follows about a_j. The proof needed $\langle \mathbf{v}_j, \mathbf{v}_j\rangle \neq 0$ for $\mathbf{v}_j \neq \mathbf{0}$, which is definiteness.

(c) Since the formula is linear in each slot, the argument of Proposition 6.3.9(c) applies: $\mathbf{x}$ is orthogonal to every vector of W if and only if it is orthogonal to $(1,1,0,0)$ and $(0,0,1,1)$, that is, $x_1 + x_2 = 0$ and $x_3 + x_4 = 0$. In $\mathbb{F}_2$ these say $x_1 = x_2$ and $x_3 = x_4$, so the set is $\{(s,s,t,t) \mid s,t \in \mathbb{F}_2\} = W$. So $W \cap W^\perp = W \neq \{\mathbf{0}\}$, and the last statement of Proposition 6.3.9(d) fails; its proof used that a vector orthogonal to itself is $\mathbf{0}$. Forms over general fields are studied in Section 12.4.

Solution to Exercise 6.3.48

Uses. Proposition 6.1.16 (Frobenius inner product), Proposition 2.4.28 (Cyclicity of the trace), Definition 2.4.26 (Trace), Theorem 6.3.3 (Pythagorean theorem), (6.3.1), Proposition 6.3.27 (Properties of the conjugate transpose)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. For (a), $\langle S, K\rangle = \operatorname{tr}\left(K^{\mathsf{T}}S\right) = -\operatorname{tr}(KS)$. Compute $\operatorname{tr}(KS)$ a second way by transposing, and use cyclicity.
Hint 2. For (b), test $H = I_n$ and $K = iI_n$.
Hint 3. For the identity in (b), only $\operatorname{Re}\langle H, K\rangle$ matters ((6.3.1)). Show that $\operatorname{tr}(KH)$ is purely imaginary by comparing it with its conjugate $\operatorname{tr}\left((KH)^*\right)$.
Sketch. (a) $\operatorname{tr}(KS) = \operatorname{tr}\left((KS)^{\mathsf{T}}\right) = \operatorname{tr}\left(S^{\mathsf{T}}K^{\mathsf{T}}\right) = -\operatorname{tr}(SK) = -\operatorname{tr}(KS)$, so $\langle S, K\rangle = -\operatorname{tr}(KS) = 0$, and Pythagoras applies to $A = S + K$. (b) False: $\langle I, iI\rangle = -in \neq 0$. But $\overline{\operatorname{tr}(KH)} = \operatorname{tr}(H^*K^*) = -\operatorname{tr}(KH)$, so $\langle H, K\rangle = -\operatorname{tr}(KH)$ is purely imaginary, $\operatorname{Re}\langle H, K\rangle = 0$, and the identity holds by (6.3.1).

(a) Let $S^{\mathsf{T}} = S$ and $K^{\mathsf{T}} = -K$. A matrix and its transpose have the same diagonal, so $\operatorname{tr}\left(M^{\mathsf{T}}\right) = \operatorname{tr} M$. Hence, using cyclicity (Proposition 2.4.28),

$$\begin{aligned}\operatorname{tr}(KS) &= \operatorname{tr}\left((KS)^{\mathsf{T}}\right) = \operatorname{tr}\left(S^{\mathsf{T}}K^{\mathsf{T}}\right)\\ &= \operatorname{tr}(S(-K)) = -\operatorname{tr}(SK) = -\operatorname{tr}(KS),\end{aligned}$$

so $\operatorname{tr}(KS) = 0$, and $\langle S, K\rangle = \operatorname{tr}\left(K^{\mathsf{T}}S\right) = -\operatorname{tr}(KS) = 0$. For $A \in \mathbb{R}^{n\times n}$, $S = \left(A + A^{\mathsf{T}}\right)/2$ is symmetric, $K = \left(A - A^{\mathsf{T}}\right)/2$ is skew-symmetric, and $A = S + K$. By Theorem 6.3.3(a), $\|A\|^2 = \|S\|^2 + \|K\|^2$.

(b) False. $H = I_n$ is Hermitian and $K = iI_n$ satisfies $K^* = -iI_n = -K$, but $\langle H, K\rangle = \operatorname{tr}(K^*H) = \operatorname{tr}(-iI_n) = -in \neq 0$.

The identity still holds. Let $H^* = H$ and $K^* = -K$. The diagonal entries of M^* are the conjugates of those of M, so $\operatorname{tr}(M^*) = \overline{\operatorname{tr} M}$. By Proposition 6.3.27(b) and cyclicity,

$$\begin{aligned}\overline{\operatorname{tr}(KH)} &= \operatorname{tr}\left((KH)^*\right) = \operatorname{tr}(H^*K^*)\\ &= \operatorname{tr}(H(-K)) = -\operatorname{tr}(HK) = -\operatorname{tr}(KH).\end{aligned}$$

So $\operatorname{tr}(KH)$ is purely imaginary, and so is $\langle H, K\rangle = \operatorname{tr}(K^*H) = -\operatorname{tr}(KH)$. Thus $\operatorname{Re}\langle H, K\rangle = 0$, and by (6.3.1), $\|H + K\|^2 = \|H\|^2 + \|K\|^2$. For $A \in \mathbb{C}^{n\times n}$, $H = (A + A^*)/2$ is Hermitian and $K = (A - A^*)/2$ satisfies $K^* = -K$ (Proposition 6.3.27(a)), and $A = H + K$, which gives the identity. This is the situation of Non-example 6.3.4: Pythagoras holds without orthogonality.

Solution to Exercise 6.3.49

Uses. Example 6.3.17 (Orthogonal polynomials), Proposition 6.3.9 (Properties of $A^\perp$), Corollary 6.3.14 (Counting orthogonal vectors), Corollary 6.3.21 (Coordinates in an orthogonal basis)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. $\left\langle x^j, x^k\right\rangle = \int_{-1}^{1} t^{j+k}\,dt$ is $2/(j+k+1)$ if $j+k$ is even and 0 if it is odd. The three conditions are linear equations in a, b, c.
Hint 2. You should find $q = x^3 - (3/5)x$. For (b), $q \perp x^2 - 1/3$ follows from $q \perp 1$ and $q \perp x^2$.
Hint 3. For (c), you need $\langle p, 1\rangle$, $\langle p, x\rangle$, $\left\langle p, x^2 - 1/3\right\rangle$, $\langle p, q\rangle$ and the four squared norms.
Sketch. (a) $\langle q, 1\rangle = 2a/3 + 2c$, $\langle q, x\rangle = 2/5 + 2b/3$, $\left\langle q, x^2\right\rangle = 2a/5 + 2c/3$ force $a = c = 0$, $b = -3/5$, and $\|q\|^2 = 8/175$. (b) Four nonzero orthogonal vectors in a 4-dimensional space. (c) The coefficients are $(8/3)/2 = 4/3$, $(16/15)/(2/3) = 8/5$, $(8/45)/(8/45) = 1$, and $(8/175)/(8/175) = 1$.

Throughout, $\left\langle x^j, x^k\right\rangle = \int_{-1}^{1} t^{j+k}\,dt$, which is $2/(j+k+1)$ when $j+k$ is even and 0 when $j+k$ is odd.

(a) By linearity in the first slot,

$$\langle q,1\rangle = \frac{2}{3}a + 2c, \quad \langle q,x\rangle = \frac{2}{5} + \frac{2}{3}b, \quad \langle q,x^2\rangle = \frac{2}{5}a + \frac{2}{3}c.$$

The second equation gives $b = -3/5$. The first gives $c = -a/3$, and then the third becomes $2a/5 - 2a/9 = 8a/45 = 0$, so $a = c = 0$. Hence $q = x^3 - (3/5)x$, and

$$\begin{aligned}\|q\|^2 &= \int_{-1}^{1}\left(t^6 - \frac{6}{5}t^4 + \frac{9}{25}t^2\right)\mathrm{d}t \\ &= \frac{2}{7} - \frac{12}{25} + \frac{6}{25} = \frac{8}{175}.\end{aligned}$$

(b) By Example 6.3.17, $1, x, x^2 - 1/3$ is orthogonal. By (a), q is orthogonal to 1 and to x^2, hence to $x^2 - 1/3$ (Proposition 6.3.9(a)), and to x. So the four polynomials form an orthogonal list of nonzero vectors in the 4-dimensional space $\mathbb{R}[x]_{\le 3}$, hence an orthogonal basis (Corollary 6.3.14).

(c) We compute $\langle p, x^k\rangle = \sum_{j=0}^{3}\langle x^j, x^k\rangle$:

$$\langle p,1\rangle = 2 + \frac{2}{3} = \frac{8}{3}, \qquad \langle p,x\rangle = \frac{2}{3} + \frac{2}{5} = \frac{16}{15},$$
$$\langle p,x^2\rangle = \frac{2}{3} + \frac{2}{5} = \frac{16}{15}, \quad \langle p,x^3\rangle = \frac{2}{5} + \frac{2}{7} = \frac{24}{35}.$$

Hence $\langle p, x^2 - 1/3\rangle = 16/15 - 8/9 = 8/45$ and $\langle p, q\rangle = 24/35 - (3/5)(16/15) = 24/35 - 16/25 = 8/175$. The squared norms are 2, $2/3$, $8/45$ (Example 6.3.17), and $8/175$. By Corollary 6.3.21,

$$p = \frac{4}{3}\cdot 1 + \frac{8}{5}x + 1\cdot\left(x^2 - \frac{1}{3}\right) + 1\cdot\left(x^3 - \frac{3}{5}x\right).$$

Check: the constant term is $4/3 - 1/3 = 1$, the coefficient of x is $8/5 - 3/5 = 1$, and those of x^2 and x^3 are 1.

Solution to Exercise 6.3.50

Uses. Proposition 6.1.16 (Frobenius inner product), Proposition 2.4.28 (Cyclicity of the trace), Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.18 (Coordinates in an orthonormal basis)
Hint 1. Each σ_a is Hermitian, so $\langle\sigma_a,\sigma_b\rangle = \operatorname{tr}(\sigma_b\sigma_a)$. Compute σ_a^2 and the products $\sigma_1\sigma_2$, $\sigma_2\sigma_3$, $\sigma_3\sigma_1$.
Hint 2. $\sigma_a^2 = I_2$, and the three products are $i\sigma_3$, $i\sigma_1$, $i\sigma_2$, all of trace 0. Also $\operatorname{tr}\sigma_a = 0$ for $a = 1,2,3$.
Hint 3. For (b), the coordinate of A along $\sigma_a/\sqrt{2}$ is $\langle A, \sigma_a/\sqrt{2}\rangle = \operatorname{tr}(\sigma_a A)/\sqrt{2}$.
Sketch. (a) $\langle\sigma_a,\sigma_b\rangle = \operatorname{tr}(\sigma_b\sigma_a)$ is 2 for $a = b$ and 0 otherwise, and $\dim\mathbb{C}^{2\times 2} = 4$. (b) Proposition 6.3.18 with $\langle A,\sigma_a/\sqrt{2}\rangle = \operatorname{tr}(\sigma_a A)/\sqrt{2}$. (c) The traces are $5, 5, -i, -3$, so $A = (5\sigma_0 + 5\sigma_1 - i\sigma_2 - 3\sigma_3)/2$.

(a) Each σ_a satisfies $\sigma_a^* = \sigma_a$ (for σ_2: $\overline{-i} = i$ moves to position $(2,1)$). So $\langle\sigma_a,\sigma_b\rangle = \operatorname{tr}(\sigma_b^*\sigma_a) = \operatorname{tr}(\sigma_b\sigma_a)$. Direct multiplication gives $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I_2$ and

$$\sigma_1\sigma_2 = \begin{pmatrix} i & 0 \\ 0 & -i\end{pmatrix} = i\sigma_3,$$
$$\sigma_2\sigma_3 = \begin{pmatrix} 0 & i \\ i & 0\end{pmatrix} = i\sigma_1,$$
$$\sigma_3\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix} = i\sigma_2.$$

So $\langle\sigma_a,\sigma_a\rangle = \operatorname{tr} I_2 = 2$ for every a (including $a = 0$). For distinct $a, b \in \{1,2,3\}$, $\operatorname{tr}(\sigma_b\sigma_a) = \operatorname{tr}(\sigma_a\sigma_b)$ (Proposition 2.4.28) is the trace of i times a Pauli matrix with trace 0, so it is 0. Finally $\langle\sigma_a,\sigma_0\rangle = \operatorname{tr}\sigma_a = 0$ for $a = 1,2,3$. Hence the $\sigma_a/\sqrt{2}$ are orthonormal, and since $\dim\mathbb{C}^{2\times 2} = 4$, they form an orthonormal basis (Corollary 6.3.14).

(b) By Proposition 6.3.18, $A = \sum_a\langle A,\sigma_a/\sqrt{2}\rangle\sigma_a/\sqrt{2}$, and $\langle A,\sigma_a/\sqrt{2}\rangle = \operatorname{tr}(\sigma_a^* A)/\sqrt{2} = \operatorname{tr}(\sigma_a A)/\sqrt{2}$. Multiplying the two factors $1/\sqrt{2}$ gives the formula.

(c) We have $\operatorname{tr} A = 5$, and

$$\sigma_1 A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \quad \sigma_2 A = \begin{pmatrix} -3i & -4i \\ i & 2i \end{pmatrix},$$
$$\sigma_3 A = \begin{pmatrix} 1 & 2 \\ -3 & -4 \end{pmatrix},$$

with traces $5, -i, -3$. So $A = (5\sigma_0 + 5\sigma_1 - i\sigma_2 - 3\sigma_3)/2$. Check: $5\sigma_0 + 5\sigma_1 - i\sigma_2 - 3\sigma_3 = \begin{pmatrix} 5-3 & 5-1 \\ 5+1 & 5+3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = 2A$, since $-i\sigma_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Solution to Exercise 6.3.51

Uses. Proposition 1.6.19 (Coordinates respect linear combinations), Proposition 6.1.10 (The dot product is an inner product), Definition 6.1.2 (Inner product), Theorem 6.3.22 (Parseval's identity), Proposition 6.1.22 (Inner products on $\mathbb{F}^n$)
Hint 1. The coordinate map $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is linear and injective. Transfer each axiom from the dot product.
Hint 2. For (b), if $\mathcal{B}$ is orthonormal for some inner product, what does Theorem 6.3.22 say about it?
Hint 3. For (c), $[\mathbf{x}]_{\mathcal{B}} = (x_1 - x_2, x_2)$.
Sketch. (a) Write $C(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$. Conjugate symmetry, linearity in the first slot, and positivity pass from the dot product through the linear map C, and $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{B}} = 0$ forces $C(\mathbf{x}) = \mathbf{0}$, so $\mathbf{x} = \mathbf{0}$. Since $[\mathbf{b}_j]_{\mathcal{B}} = \mathbf{e}_j$, $\mathcal{B}$ is orthonormal. (b) By Theorem 6.3.22(a), any inner product for which $\mathcal{B}$ is orthonormal equals $[\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}}$. (c) $(x_1 - x_2)(y_1 - y_2) + x_2 y_2$, so $M = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

(a) Let $C(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$. By Proposition 1.6.19, $C : V \to \mathbb{F}^n$ is linear, and $C(\mathbf{x}) = \mathbf{0}$ only for $\mathbf{x} = 0\mathbf{b}_1 + \cdots + 0\mathbf{b}_n = \mathbf{0}$. Since the dot product is an inner product on $\mathbb{F}^n$ (Proposition 6.1.10), for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $a, b \in \mathbb{F}$:

- $\langle \mathbf{y}, \mathbf{x} \rangle_{\mathcal{B}} = C(\mathbf{y}) \cdot C(\mathbf{x}) = \overline{C(\mathbf{x}) \cdot C(\mathbf{y})} = \overline{\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{B}}}$;
- $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle_{\mathcal{B}} = (aC(\mathbf{x}) + bC(\mathbf{y})) \cdot C(\mathbf{z}) = a\langle \mathbf{x}, \mathbf{z} \rangle_{\mathcal{B}} + b\langle \mathbf{y}, \mathbf{z} \rangle_{\mathcal{B}}$;
- $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{B}} = C(\mathbf{x}) \cdot C(\mathbf{x}) \geq 0$, with equality only if $C(\mathbf{x}) = \mathbf{0}$, that is, $\mathbf{x} = \mathbf{0}$.

So $\langle \cdot, \cdot \rangle_{\mathcal{B}}$ is an inner product. Since $[\mathbf{b}_j]_{\mathcal{B}} = \mathbf{e}_j$, we get $\langle \mathbf{b}_i, \mathbf{b}_j \rangle_{\mathcal{B}} = \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, so $\mathcal{B}$ is orthonormal.

(b) Let $\langle \cdot, \cdot \rangle$ be an inner product on V for which $\mathcal{B}$ is orthonormal. Then $\mathcal{B}$ is an orthonormal basis, and by Theorem 6.3.22(a), $\langle \mathbf{x}, \mathbf{y} \rangle = [\mathbf{x}]_{\mathcal{B}} \cdot [\mathbf{y}]_{\mathcal{B}} = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y}$.

(c) Writing $\mathbf{x} = a(1,0) + c(1,1) = (a + c, c)$ gives $c = x_2$ and $a = x_1 - x_2$, so $[\mathbf{x}]_{\mathcal{B}} = (x_1 - x_2, x_2)$. Hence

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{B}} &= (x_1 - x_2)(y_1 - y_2) + x_2 y_2 \\ &= x_1 y_1 - x_1 y_2 - x_2 y_1 + 2x_2 y_2 \\ &= \mathbf{y}^{\mathsf{T}} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \mathbf{x}. \end{aligned}$$

For instance $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle_{\mathcal{B}} = -1$: the standard basis is not orthonormal for this inner product, as anticipated after Theorem 6.3.22.

Solution to Exercise 6.3.52

Uses. Definition 6.1.8 (Dot product), Corollary 6.3.14 (Counting orthogonal vectors), Theorem 6.3.24 (Bessel's inequality), Proposition 6.3.13 (Orthogonal lists are independent)
Techniques. Plug in clever test vectors (Technique 6.1.40), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. For (a), write $\mathbf{v} = (x, y)$ and expand. You need $\sum_j \cos^2 \theta_j$, $\sum_j \sin^2 \theta_j$, and $\sum_j \cos\theta_j \sin\theta_j$ for $\theta_j = 2\pi j/3$.
Hint 2. For (b), apply the hypothesis to $\mathbf{v} = \mathbf{v}_l$. The left side contains the term $|\langle \mathbf{v}_l, \mathbf{v}_l \rangle|^2 = 1$.
Hint 3. Once the list is orthonormal, the hypothesis is the equality case of Bessel's inequality for every $\mathbf{v}$.
Sketch. (a) With $\cos\theta_j = 1, -1/2, -1/2$ and $\sin\theta_j = 0, \sqrt{3}/2, -\sqrt{3}/2$, the sum is $(2/3)\big(3x^2/2 + 3y^2/2\big)$. Three nonzero vectors in $\mathbb{R}^2$ cannot be orthogonal. (b) For $\mathbf{v} = \mathbf{v}_l$: $1 = 1 + \sum_{j \neq l} |\langle \mathbf{v}_l, \mathbf{v}_j \rangle|^2$, so the cross terms vanish. Then Bessel's equality case puts every $\mathbf{v}$ in the span.

(a) For $\theta_j = 2\pi j/3$, the cosines are $1, -1/2, -1/2$ and the sines are $0, \sqrt{3}/2, -\sqrt{3}/2$. So $\sum_j \cos^2 \theta_j = 3/2$, $\sum_j \sin^2 \theta_j = 3/2$, and $\sum_j \cos\theta_j \sin\theta_j = 0 - \sqrt{3}/4 + \sqrt{3}/4 = 0$. For $\mathbf{v} = (x, y)$,

$$\sum_{j=0}^{2} \langle \mathbf{v}, \mathbf{w}_j \rangle^2 = \frac{2}{3} \sum_{j=0}^{2} \left(x \cos\theta_j + y \sin\theta_j \right)^2$$
$$= \frac{2}{3}\left(\frac{3}{2}x^2 + 0 + \frac{3}{2}y^2 \right) = \|\mathbf{v}\|^2.$$

The list is not orthonormal: each $\mathbf{w}_j$ has norm $\sqrt{2/3}$, and by Corollary 6.3.14 no three nonzero vectors in $\mathbb{R}^2$ are orthogonal.

(b) Fix l and apply the hypothesis to $\mathbf{v} = \mathbf{v}_l$. Since $\|\mathbf{v}_l\| = 1$,

$$1 = \sum_{j=1}^{k} \left| \langle \mathbf{v}_l, \mathbf{v}_j \rangle \right|^2 = 1 + \sum_{j \neq l} \left| \langle \mathbf{v}_l, \mathbf{v}_j \rangle \right|^2.$$

So the remaining sum of nonnegative terms is 0, and $\langle \mathbf{v}_l, \mathbf{v}_j \rangle = 0$ for all $j \neq l$. Thus the list is orthonormal. Now the hypothesis says that equality holds in Bessel's inequality (Theorem 6.3.24) for every $\mathbf{v} \in V$, so every $\mathbf{v}$ lies in $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. The list spans V and is independent (Proposition 6.3.13), so it is an orthonormal basis. Part (a) shows that the hypothesis "unit vectors" cannot be dropped.

Solution to Exercise 6.3.53

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 6.3.24 (Bessel's inequality), Proposition 6.2.22 (The square of a sum)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. For (a), apply Cauchy–Schwarz in $\mathbb{R}^k$ to the vectors $\left(|\langle \mathbf{x}, \mathbf{u}_j \rangle|\right)_j$ and $\left(|\langle \mathbf{y}, \mathbf{u}_j \rangle|\right)_j$, then Bessel to each norm.
Hint 2. For (b), use Proposition 6.2.22 and then Bessel. Equality requires equality in both steps.
Sketch. (a) Cauchy–Schwarz in $\mathbb{R}^k$ bounds the sum by $\left(\sum_j |\langle \mathbf{x}, \mathbf{u}_j \rangle|^2\right)^{1/2}\left(\sum_j |\langle \mathbf{y}, \mathbf{u}_j \rangle|^2\right)^{1/2}$, and Bessel bounds the factors by $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$. (b) $\left(\sum_j a_j\right)^2 \le k \sum_j a_j^2 \le k\|\mathbf{x}\|^2$ with $a_j = |\langle \mathbf{x}, \mathbf{u}_j \rangle|$. Equality: all a_j equal and $\mathbf{x} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$, that is, $\mathbf{x} = \sum_j c_j \mathbf{u}_j$ with $|c_1| = \cdots = |c_k|$.

Write $a_j = |\langle \mathbf{x}, \mathbf{u}_j \rangle|$ and $b_j = |\langle \mathbf{y}, \mathbf{u}_j \rangle|$, nonnegative reals.

(a) By Theorem 6.2.1 for the dot product on $\mathbb{R}^k$ and then Theorem 6.3.24 twice,

$$\sum_j a_j b_j \le \left(\sum_j a_j^2 \right)^{1/2} \left(\sum_j b_j^2 \right)^{1/2} \le \|\mathbf{x}\| \|\mathbf{y}\|.$$

(b) By Proposition 6.2.22 and Theorem 6.3.24,

$$\left(\sum_j a_j \right)^2 \le k \sum_j a_j^2 \le k \|\mathbf{x}\|^2,$$

and taking nonnegative square roots gives the inequality. Equality holds if and only if both steps are equalities: all a_j are equal (Proposition 6.2.22) and $\mathbf{x} \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ (Theorem 6.3.24). By Proposition 6.3.18, this says $\mathbf{x} = c_1 \mathbf{u}_1 + \cdots + c_k \mathbf{u}_k$ with $|c_1| = \cdots = |c_k|$. For example, in $\mathbb{R}^k$ with the standard basis, $\mathbf{x} = (1, -1, 1, \ldots)$ gives equality.

Solution to Exercise 6.3.54

Uses. Definition 6.3.28 (Gram matrix), Proposition 6.3.30 (Gram matrices detect independence), Example 1.5.7 (Monomials)
Techniques. Gram matrices (Technique 6.3.31)
Hint 1. Compute $\langle x^{j-1}, x^{i-1} \rangle$.
Hint 2. The monomials are linearly independent. Apply Proposition 6.3.30.
Sketch. $\langle x^{j-1}, x^{i-1} \rangle = \int_0^1 t^{i+j-2}\,dt = 1/(i+j-1)$. The monomials are independent, so H_n is invertible, and $\mathbf{c}^{\mathsf{T}} H_n \mathbf{c} = \left\| \sum_j c_j x^{j-1} \right\|^2 > 0$ because the polynomial is nonzero.

(a) The formula is an inner product on $\mathbb{R}[x]_{\leq n-1}$ (see Example 6.3.17 for definiteness; the interval does not matter). The (i,j) entry of the Gram matrix of $x^0, \ldots, x^{n-1}$ is

$$\left\langle x^{j-1}, x^{i-1} \right\rangle = \int_0^1 t^{i+j-2}\,dt = \frac{1}{i+j-1} = h_{ij}.$$

(b) The monomials $1, x, \ldots, x^{n-1}$ are linearly independent (Example 1.5.7), so H_n is invertible by Proposition 6.3.30. For $\mathbf{c} \neq \mathbf{0}$, the polynomial $p = c_1 + c_2 x + \cdots + c_n x^{n-1}$ is nonzero, and by Proposition 6.3.30(a), $\mathbf{c}^{\mathsf{T}} H_n \mathbf{c} = \|p\|^2 > 0$.

Computing $\det H_n$ directly is laborious already for $n = 4$; the Gram matrix point of view gives invertibility with no computation.

Solution to Exercise 6.3.55

Uses. Theorem 6.3.29 (Orthonormal columns), Proposition 6.3.34 (Characterizations of unitary matrices)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Check each step. Which one is a general claim about matrices that you have not seen proved?
Hint 2. Is $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ diagonal? Is it similar to a diagonal matrix? For a counterexample to the claim, try 2×2 matrices with columns $(1,1)$ and $(a,-a)$.
Sketch. The false step is "similar to a diagonal matrix implies diagonal". The claim is false: $A = \left(\begin{smallmatrix} 1 & 2 \\ 1 & -2 \end{smallmatrix}\right)$ has $A^{\mathsf{T}}A = \operatorname{diag}(2,8)$ but $AA^{\mathsf{T}} = \left(\begin{smallmatrix} 5 & -3 \\ -3 & 5 \end{smallmatrix}\right)$. With orthonormal columns ($D = I$) the argument works and gives Proposition 6.3.34.

Every step is correct except "a matrix similar to a diagonal matrix is diagonal". For instance PDP^{-1} with $D = \operatorname{diag}(1,2)$ and $P = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$, which is not diagonal. (Matrices similar to diagonal ones are the diagonalizable matrices of Section 7.4.)

The claim is false. For $A = \left(\begin{smallmatrix} 1 & 2 \\ 1 & -2 \end{smallmatrix}\right)$, the columns $(1,1)$ and $(2,-2)$ are orthogonal and nonzero, and $A^{\mathsf{T}}A = \operatorname{diag}(2,8)$. But the rows $(1,2)$ and $(1,-2)$ have dot product $1 - 4 = -3$, and indeed

$$AA^{\mathsf{T}} = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}.$$

The argument does prove something when $D = I$, that is, when the columns are orthonormal: then $AA^{\mathsf{T}} = AA^{-1} = I$, and the rows are orthonormal. This is (i) $\Rightarrow$ (iii) of Proposition 6.3.34.

Solution to Exercise 6.3.56

Uses. Definition 6.3.28 (Gram matrix), Proposition 6.1.4 (Sesquilinearity), Definition 2.4.1 (Matrix product), Definition 6.3.26 (Conjugate transpose), Definition 6.3.33 (Unitary and orthogonal matrices), Corollary 6.3.14 (Counting orthogonal vectors)
Techniques. Expand by multilinearity (Technique 5.1.7), Gram matrices (Technique 6.3.31)
Hint 1. Expand $\langle \mathbf{w}_j, \mathbf{w}_i \rangle$ in both slots. The coefficients from the second slot come out conjugated.
Hint 2. $\langle \mathbf{w}_j, \mathbf{w}_i \rangle = \sum_m \sum_l \overline{s_{mi}} g_{ml} s_{lj}$, and $\overline{s_{mi}}$ is the (i,m) entry of S^*.
Hint 3. For (b), $G = I_n$; a list is orthonormal exactly when its Gram matrix is I_n. Column j of S is $[\mathbf{w}_j]_{\mathcal{O}}$.
Sketch. (a) $\langle \mathbf{w}_j, \mathbf{w}_i \rangle = \sum_{l,m} s_{lj}\overline{s_{mi}}\langle \mathbf{v}_l, \mathbf{v}_m \rangle = \sum_{m,l} (S^*)_{im} g_{ml} s_{lj} = (S^*GS)_{ij}$. (b) With $G = I$, the $\mathbf{w}_j$ are orthonormal iff $S^*S = I$, and n orthonormal vectors form a basis. Column j of S is $[\mathbf{w}_j]_{\mathcal{O}}$, so S is the change-of-coordinates matrix from the $\mathbf{w}$-basis to $\mathcal{O}$.

(a) Expanding both slots (linearity in the first, Proposition 6.1.4 in the second),

$$\begin{aligned} \langle \mathbf{w}_j, \mathbf{w}_i \rangle &= \left\langle \sum_l s_{lj}\mathbf{v}_l, \sum_m s_{mi}\mathbf{v}_m \right\rangle = \sum_{m=1}^{n}\sum_{l=1}^{n} \overline{s_{mi}}\langle \mathbf{v}_l, \mathbf{v}_m \rangle s_{lj} \\ &= \sum_{m=1}^{n}\sum_{l=1}^{n} (S^*)_{im} g_{ml} s_{lj}, \end{aligned}$$

using $g_{ml} = \langle \mathbf{v}_l, \mathbf{v}_m \rangle$ and $(S^*)_{im} = \overline{s_{mi}}$. By Definition 2.4.1, the right side is the (i,j) entry of S^*GS.

(b) If $\mathcal{O}$ is orthonormal, its Gram matrix is I_n, so by (a) the Gram matrix of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is S^*S. A list is orthonormal exactly when its Gram matrix is the identity, so $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is orthonormal if and only if $S^*S = I_n$, that is, S is unitary. An orthonormal list of length $n = \dim V$ is an orthonormal basis (Corollary 6.3.14).

Column j of S is the coordinate vector $[\mathbf{w}_j]_{\mathcal{O}}$, so S is the change-of-coordinates matrix that converts coordinates with respect to $\mathbf{w}_1, \ldots, \mathbf{w}_n$ into coordinates with respect to $\mathcal{O}$. Hence if both lists are orthonormal bases, this matrix is unitary, and so is its inverse, the change of coordinates in the other direction (Exercise 6.3.37).

Solution to Exercise 6.3.57

Uses. Theorem 5.3.3 (The Leibniz formula), Theorem 5.4.14 (The transpose), Theorem 5.4.10 (Multiplicativity), Theorem 6.3.29 (Orthonormal columns), Definition 6.3.26 (Conjugate transpose)
Hint 1. Conjugation respects sums and products, so it passes through the Leibniz formula: $\det \overline{A} = \overline{\det A}$. Then use $\det\left(B^{\mathsf{T}}\right) = \det B$.
Hint 2. Take determinants of $U^*U = I_n$.
Hint 3. For (c), the first column of U is $u_{11}\mathbf{e}_1$. Orthogonality of the other columns to it kills the rest of the first row. Then repeat with the second column.
Sketch. (a) $\det A^* = \det \overline{A}^{\mathsf{T}} = \det \overline{A} = \overline{\det A}$, so $1 = \det(U^*U) = \overline{\det U} \det U = |\det U|^2$; a real number of modulus 1 is ± 1. (b) $\operatorname{diag}(2, 1/2)$. (c) Induct on the row: if rows $1, \ldots, i-1$ vanish off the diagonal, column i is $u_{ii}\mathbf{e}_i$, and orthogonality of column $j > i$ to it gives $u_{ij}\overline{u_{ii}} = 0$ with $|u_{ii}| = 1$, so row i vanishes off the diagonal too.

(a) By Theorem 5.3.3, $\det \overline{A}$ is a sum of signed products of entries of $\overline{A}$, which are the conjugates of the corresponding products for A; since conjugation respects sums and products, $\det \overline{A} = \overline{\det A}$. By Theorem 5.4.14, $\det(A^*) = \det\left(\overline{A}^{\mathsf{T}}\right) = \det \overline{A} = \overline{\det A}$.

If U is unitary, then by Theorem 5.4.10,

$$\begin{aligned} 1 = \det I_n = \det(U^*U) &= \det(U^*) \det U \\ &= \overline{\det U} \det U = |\det U|^2, \end{aligned}$$

so $|\det U| = 1$. If Q is real orthogonal, $\det Q$ is a real number with $|\det Q| = 1$, so $\det Q = \pm 1$.

(b) $A = \operatorname{diag}(2, 1/2)$ has $\det A = 1$, but its first column has norm 2, so $A^{\mathsf{T}}A \neq I_2$.

(c) Let U be unitary and upper triangular, with columns $\mathbf{c}_1, \ldots, \mathbf{c}_n$; by Theorem 6.3.29 they are orthonormal. We show by induction on i that $u_{ij} = 0$ for all $j > i$. Suppose this holds for the rows $1, \ldots, i-1$. Then the entries $u_{1i}, \ldots, u_{i-1,i}$ of $\mathbf{c}_i$ vanish, and so do its entries below the diagonal, so $\mathbf{c}_i = u_{ii}\mathbf{e}_i$, and $|u_{ii}| = \|\mathbf{c}_i\| = 1$. For $j > i$,

$$0 = \mathbf{c}_j \cdot \mathbf{c}_i = u_{ij}\overline{u_{ii}},$$

and $\overline{u_{ii}} \neq 0$, so $u_{ij} = 0$. This completes the induction (the case $i = 1$ needs no hypothesis). So U is diagonal, with diagonal entries of modulus 1.

Solution to Exercise 6.3.58

Uses. Theorem 6.3.29 (Orthonormal columns), Definition 6.1.8 (Dot product), Theorem 5.1.13 (The 2×2 determinant)
Hint 1. The first column (a, c) is any unit vector. Find one unit vector orthogonal to it by inspection.
Hint 2. $(-\overline{c}, \overline{a})$ works. Show that every vector orthogonal to (a, c) is a multiple of it: solve $b\overline{a} + d\overline{c} = 0$, treating $a \neq 0$ and $a = 0$ separately.
Sketch. The given matrices have orthonormal columns by a direct check. Conversely, if (a, c) is a unit vector, the solutions of $b\overline{a} + d\overline{c} = 0$ are the multiples of $(-\overline{c}, \overline{a})$, which is a unit vector; so a unit second column is $\lambda(-\overline{c}, \overline{a})$ with $|\lambda| = 1$. Then $\det U = \lambda\left(|a|^2 + |c|^2\right) = \lambda$.

By Theorem 6.3.29, U is unitary if and only if its columns are orthonormal.

Suppose U has the given form. The first column has squared norm $|a|^2 + |c|^2 = 1$, the second has squared norm $|\lambda|^2\left(|c|^2 + |a|^2\right) = 1$, and

$$(-\lambda\overline{c}, \lambda\overline{a}) \cdot (a, c) = -\lambda\overline{c}\,\overline{a} + \lambda\overline{a}\,\overline{c} = 0.$$

So U is unitary.

Conversely, let $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be unitary. Then $|a|^2 + |c|^2 = 1$, and the second column satisfies $(b,d)\cdot(a,c) = b\overline{a} + d\overline{c} = 0$. If $a \neq 0$, then $b = -d\overline{c}/\overline{a}$, so $(b,d) = (d/\overline{a})(-\overline{c},\overline{a})$. If $a = 0$, then $|c| = 1$, so $c \neq 0$ and $d = 0$, and $(b,d) = (b,0) = (-b/\overline{c})(-\overline{c},\overline{a})$ since $\overline{a} = 0$. Either way $(b,d) = \lambda(-\overline{c},\overline{a})$ for some $\lambda \in \mathbb{C}$. Since $\|(-\overline{c},\overline{a})\|^2 = |c|^2 + |a|^2 = 1$ and (b,d) is a unit vector, $|\lambda| = 1$.

Finally, by Theorem 5.1.13, $\det U = a\cdot\lambda\overline{a} - (-\lambda\overline{c})c = \lambda\left(|a|^2 + |c|^2\right) = \lambda$. The unitary matrices with determinant 1 are those with $\lambda = 1$.

Solution to Exercise 6.3.59

Uses. Definition 6.3.40 (Fourier matrix), Proposition 6.3.41 (The Fourier basis)
Hint 1. $\omega^3 = 1$, so $\omega^4 = \omega$, and $\omega^2 = \overline{\omega} = \omega^{-1}$.
Hint 2. The coefficient of $\mathbf{f}_k$ is $\langle \mathbf{x}, \mathbf{f}_k\rangle/3$ with $\langle \mathbf{x}, \mathbf{f}_k\rangle = \sum_j x_j\omega^{-jk}$.
Sketch. $\Phi_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}$ and $\Phi_3^{-1} = (1/3)\overline{\Phi_3}$. The transform of $\mathbf{x}$ is $\left(3, -\sqrt{3}i, \sqrt{3}i\right)$, so $\mathbf{x} = \mathbf{f}_0 + \left(i/\sqrt{3}\right)(\mathbf{f}_2 - \mathbf{f}_1)$, and $1 + 4 = (9+3+3)/3$.

(a) Since $\omega^3 = 1$, $\omega^4 = \omega$, and $\overline{\omega} = \omega^2$,

$$\Phi_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix},$$

$$\Phi_3^{-1} = \frac{1}{3}\overline{\Phi_3} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{pmatrix},$$

by Proposition 6.3.41(b).

(b) By Proposition 6.3.41(c), with $\omega^{-1} = \omega^2$ and $\omega^{-2} = \omega$,

$$\langle \mathbf{x}, \mathbf{f}_0\rangle = 3,$$
$$\langle \mathbf{x}, \mathbf{f}_1\rangle = 1 + 2\omega^2 = -\sqrt{3}i,$$
$$\langle \mathbf{x}, \mathbf{f}_2\rangle = 1 + 2\omega = \sqrt{3}i.$$

So $\mathbf{x} = (1/3)\left(3\mathbf{f}_0 - \sqrt{3}i\mathbf{f}_1 + \sqrt{3}i\mathbf{f}_2\right) = \mathbf{f}_0 + \left(i/\sqrt{3}\right)(\mathbf{f}_2 - \mathbf{f}_1)$. Check: $\mathbf{f}_2 - \mathbf{f}_1 = \left(0, \omega^2 - \omega, \omega - \omega^2\right) = \left(0, -\sqrt{3}i, \sqrt{3}i\right)$, so $\left(i/\sqrt{3}\right)(\mathbf{f}_2 - \mathbf{f}_1) = (0,1,-1)$, and adding $\mathbf{f}_0 = (1,1,1)$ gives $(1,2,0)$. For (d): $|1|^2 + |2|^2 + 0 = 5$ and $(9+3+3)/3 = 5$.

Solution to Exercise 6.3.60

Uses. Lemma 6.3.39 (Sums of roots of unity), Definition 5.2.29 (Permutation matrix), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.15 (Triangular matrices)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. $\left(\Phi_n^2\right)_{jl} = \sum_k \omega^{jk}\omega^{kl}$. Apply Lemma 6.3.39 with $m = j + l$.
Hint 2. For each j there is exactly one $l \in \{0,\ldots,n-1\}$ with $n \mid (j+l)$. Which?
Hint 3. For (b), take determinants in $\Phi_n^4 = n^2 I_n$.
Sketch. (a) $\left(\Phi_n^2\right)_{jl} = \sum_k \omega^{k(j+l)}$ is n if $n \mid (j+l)$ and 0 otherwise. Each row and each column of P has a single 1, at $l = \sigma(j)$ where $\sigma(j)$ is $-j$ reduced mod n, and σ is its own inverse, so $P^2 = I$ and $\Phi_n^4 = n^2P^2 = n^2 I$. (b) $(\det\Phi_n)^4 = \det\left(n^2 I_n\right) = n^{2n}$, so $|\det\Phi_n| = n^{n/2}$.

(a) By the entry formula for products and Lemma 6.3.39,

$$\left(\Phi_n^2\right)_{jl} = \sum_{k=0}^{n-1}\omega^{jk}\omega^{kl} = \sum_{k=0}^{n-1}\omega^{k(j+l)} = \begin{cases} n \text{ if } n \mid (j+l) \\ 0 \text{ otherwise,} \end{cases}$$

so $\Phi_n^2 = nP$. For $j \in \{0,\ldots,n-1\}$, let $\sigma(j) = 0$ if $j = 0$ and $\sigma(j) = n - j$ otherwise. Then $\sigma(j)$ is the only $l \in \{0,\ldots,n-1\}$ with $n \mid (j+l)$, so row j of P has its only 1 in column $\sigma(j)$. Also $\sigma(\sigma(j)) = j$, so σ is a bijection

of $\{0, \dots, n-1\}$ (its own inverse), and P is the permutation matrix of σ (Definition 5.2.29). It is symmetric, since $n \mid (j+l)$ is symmetric in j and l. Finally

$$\left(P^2\right)_{jm} = \sum_l p_{jl} p_{lm} = p_{\sigma(j)m},$$

which is 1 exactly when $m = \sigma(\sigma(j)) = j$. So $P^2 = I_n$ and $\Phi_n^4 = (nP)^2 = n^2 P^2 = n^2 I_n$.

(b) By Theorem 5.4.10 and Proposition 5.3.15, $(\det \Phi_n)^4 = \det\left(\Phi_n^4\right) = \det\left(n^2 I_n\right) = n^{2n}$. Taking moduli, $|\det \Phi_n|^4 = n^{2n}$, so $|\det \Phi_n| = n^{n/2}$. (Alternatively, $\Phi_n^* \Phi_n = n I_n$ and Exercise 6.3.57(a) give $|\det \Phi_n|^2 = n^n$.)

Solution to Exercise 6.3.61

Uses. Definition 6.3.40 (Fourier matrix), Proposition 6.3.41 (The Fourier basis), Proposition 2.4.8 (Four views of the product), Theorem 5.4.10 (Multiplicativity)
Hint 1. Write $(C\mathbf{f}_k)_j = \sum_l c_{(j-l) \bmod n} \omega^{lk}$ and substitute $m = (j-l) \bmod n$. Since $\omega^n = 1$, $\omega^{lk} = \omega^{(j-m)k}$.
Hint 2. For (b), (a) says $C\Phi_n = \Phi_n D$ column by column. Diagonal matrices commute.
Hint 3. For (c), $\mathbf{c} = (2, 1, 0)$ and $\langle \mathbf{c}, \mathbf{f}_k \rangle = 2 + \omega^{-k}$.
Sketch. (a) Reindexing by $m = (j-l) \bmod n$ gives $(C\mathbf{f}_k)_j = \omega^{jk} \sum_m c_m \omega^{-mk} = \langle \mathbf{c}, \mathbf{f}_k \rangle (\mathbf{f}_k)_j$. (b) So $C\Phi_n = \Phi_n D$, and $C = \Phi_n D \Phi_n^{-1}$; then $CC' = \Phi_n D D' \Phi_n^{-1} = C'C$, and C is invertible iff D is. (c) The values $2 + \omega^{-k}$ are 3, $2 + \omega^2$, $2 + \omega$, all nonzero, and $\det C = 3|2 + \omega|^2 = 9$.

(a) Fix j and k. As l runs over $0, \dots, n-1$, so does $m = (j-l) \bmod n$, and l is congruent to $j - m$ modulo n. Since $\omega^n = 1$, $\omega^{lk} = \omega^{(j-m)k}$. Hence

$$\begin{aligned}(C\mathbf{f}_k)_j &= \sum_{l=0}^{n-1} c_{(j-l) \bmod n} \omega^{lk} = \sum_{m=0}^{n-1} c_m \omega^{(j-m)k} \\ &= \omega^{jk} \sum_{m=0}^{n-1} c_m \omega^{-mk} = \langle \mathbf{c}, \mathbf{f}_k \rangle (\mathbf{f}_k)_j,\end{aligned}$$

using $\langle \mathbf{c}, \mathbf{f}_k \rangle = \sum_m c_m \omega^{-mk}$ (Proposition 6.3.41(c)).

(b) Column k of $C\Phi_n$ is $C\mathbf{f}_k$ (Proposition 2.4.8), and column k of $\Phi_n D$ is $\langle \mathbf{c}, \mathbf{f}_k \rangle \mathbf{f}_k$. By (a), $C\Phi_n = \Phi_n D$, and since Φ_n is invertible, $C = \Phi_n D \Phi_n^{-1}$. If C' is another circulant, with diagonal matrix D', then since diagonal matrices commute,

$$CC' = \Phi_n D D' \Phi_n^{-1} = \Phi_n D' D \Phi_n^{-1} = C'C.$$

Finally, C is invertible if and only if $D = \Phi_n^{-1} C \Phi_n$ is, and a diagonal matrix is invertible if and only if its diagonal entries are nonzero.

(c) The matrix is the circulant with $\mathbf{c} = (2, 1, 0)$. With $\omega = -1/2 + \left(\sqrt{3}/2\right)i$, the entries of D are $\langle \mathbf{c}, \mathbf{f}_k \rangle = 2 + \omega^{-k}$, namely 3, $2 + \omega^2$, and $2 + \omega$. The last two are $3/2 \mp \left(\sqrt{3}/2\right)i \neq 0$, so the matrix is invertible. By Theorem 5.4.10, $\det C = \det D = 3\left(2 + \omega^2\right)(2 + \omega) = 3|2 + \omega|^2 = 3(9/4 + 3/4) = 9$, which agrees with cofactor expansion along the first row: $2 \cdot 4 - 0 + 1 \cdot 1 = 9$.

Solution to Exercise 6.3.62

Uses. Lemma 6.3.39 (Sums of roots of unity), Definition 1.2.17 (Evaluation), de Moivre's formula (Section B.1), the binomial theorem
Hint 1. Substitute $p\left(\omega^j\right) = \sum_k a_k \omega^{jk}$, exchange the two sums, and apply Lemma 6.3.39 with $m = k - r$.
Hint 2. For (b), take $p = (1 + x)^{12}$, $n = 3$, $r = 0$. Write $1 + \omega$ and $1 + \omega^2$ in the form $\cos\theta + i \sin\theta$.
Sketch. (a) $(1/n) \sum_j \omega^{-jr} \sum_k a_k \omega^{jk} = \sum_k a_k (1/n) \sum_j \omega^{j(k-r)}$, and the inner average is 1 if $n \mid (k-r)$ and 0 otherwise. (b) With $\omega = \omega_3$, $1 + \omega = \cos(\pi/3) + i \sin(\pi/3)$ and $1 + \omega^2$ is its conjugate, so their 12th powers are 1, and the sum is $\left(2^{12} + 2\right)/3 = 1366$. (c) The same computation with exponent m gives $\left(2^m + 2\cos(m\pi/3)\right)/3$.

(a) By Definition 1.2.17, $p\left(\omega^j\right) = \sum_{k=0}^{d} a_k \omega^{jk}$. Exchanging the finite sums,

$$\frac{1}{n}\sum_{j=0}^{n-1}\omega^{-jr}p\big(\omega^j\big)=\sum_{k=0}^{d}a_k\left(\frac{1}{n}\sum_{j=0}^{n-1}\omega^{j(k-r)}\right).$$

By Lemma 6.3.39 the expression in parentheses is 1 if $n \mid (k-r)$ and 0 otherwise, which gives the claim.

(b) Take $n=3$, $r=0$, and $p=(1+x)^{12}=\sum_k\binom{12}{k}x^k$. Then the left side of (a) is the requested sum, and the right side is $\big(p(1)+p(\omega)+p\big(\omega^2\big)\big)/3$. Here $\omega=-1/2+\big(\sqrt{3}/2\big)i$, so

$$1+\omega=\frac{1}{2}+\frac{\sqrt{3}}{2}i=\cos\frac{\pi}{3}+i\sin\frac{\pi}{3},$$
$$1+\omega^2=1+\overline{\omega}=\overline{1+\omega}.$$

By de Moivre's formula, $(1+\omega)^{12}=\cos 4\pi+i\sin 4\pi=1$, and $\big(1+\omega^2\big)^{12}=\overline{(1+\omega)^{12}}=1$. So the sum is $\big(2^{12}+1+1\big)/3=4098/3=1366$.

(c) With $p=(1+x)^m$, the same computation gives $(1+\omega)^m=\cos(m\pi/3)+i\sin(m\pi/3)$ and $\big(1+\omega^2\big)^m$ its conjugate, which add up to $2\cos(m\pi/3)$. So the sum is $(2^m+2\cos(m\pi/3))/3$.

Solution to Exercise 6.3.63

Uses. Theorem 6.3.29 (Orthonormal columns), Theorem 2.4.34 (Block multiplication)
Techniques. Block decomposition (Technique 2.4.35)
Hint 1. Multiplying a row of H by -1, or permuting the rows, keeps the columns orthogonal. Use this to make the first column all 1s.
Hint 2. A ±1 column orthogonal to the all-1s column has as many $+1$s as -1s. Arrange the second column as $n/2$ entries $+1$ followed by $n/2$ entries -1.
Hint 3. Count the $+1$s of the third column in each half, and use its orthogonality to the first two columns.
Sketch. (a) Normalize the first column to all 1s and sort the rows so that column 2 is $+1$ in the first half and -1 in the second. If column 3 has a entries $+1$ in the first half and b in the second, orthogonality to column 1 gives $a+b=n/2$ and to column 2 gives $2a-2b=0$, so $a=n/4$. (b) Block multiplication gives $\begin{pmatrix}H & H\\ H & -H\end{pmatrix}^{\mathsf{T}}\begin{pmatrix}H & H\\ H & -H\end{pmatrix}=\operatorname{diag}\big(2H^{\mathsf{T}}H, 2H^{\mathsf{T}}H\big)=2nI_{2n}$; start from (1).

(a) Let H be a Hadamard matrix of order $n\ge 3$, with columns $\mathbf{h}_1,\dots,\mathbf{h}_n$, so $\mathbf{h}_k\cdot\mathbf{h}_l=\sum_j h_{jk}h_{jl}$. Multiplying row j by -1 multiplies both h_{jk} and h_{jl} by -1, so it leaves each term $h_{jk}h_{jl}$ unchanged, and permuting the rows only reorders the sum. So both operations produce Hadamard matrices. Multiply each row whose first entry is -1 by -1; now $\mathbf{h}_1=(1,\dots,1)$.

Every other column satisfies $0=\mathbf{h}_k\cdot\mathbf{h}_1=\sum_j h_{jk}$, so it has equally many entries $+1$ and -1; in particular n is even. Permute the rows so that $\mathbf{h}_2$ has $+1$ in rows $1,\dots,n/2$ and -1 in rows $n/2+1,\dots,n$. Let a and b be the numbers of entries $+1$ of $\mathbf{h}_3$ in the first and in the second half. From $\mathbf{h}_3\cdot\mathbf{h}_1=0$, $\mathbf{h}_3$ has $n/2$ entries $+1$ in total, so $a+b=n/2$. From $\mathbf{h}_3\cdot\mathbf{h}_2=0$,

$$0=\left(a-\left(\frac{n}{2}-a\right)\right)-\left(b-\left(\frac{n}{2}-b\right)\right)=2a-2b.$$

So $a=b=n/4$, and since a is an integer, $4\mid n$.

(b) Let $K=\begin{pmatrix}H & H\\ H & -H\end{pmatrix}$. Its entries are ±1, and by Theorem 2.4.34,

$$K^{\mathsf{T}}K=\begin{pmatrix}H^{\mathsf{T}} & H^{\mathsf{T}}\\ H^{\mathsf{T}} & -H^{\mathsf{T}}\end{pmatrix}\begin{pmatrix}H & H\\ H & -H\end{pmatrix}$$
$$=\begin{pmatrix}2H^{\mathsf{T}}H & 0\\ 0 & 2H^{\mathsf{T}}H\end{pmatrix}=\begin{pmatrix}2nI_n & 0\\ 0 & 2nI_n\end{pmatrix}=2nI_{2n}.$$

So K is a Hadamard matrix of order $2n$ (Theorem 6.3.29). Starting from the Hadamard matrix (1) of order 1 and applying this construction k times gives one of order 2^k. The first step gives $\Phi_2=\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$. Whether Hadamard matrices exist for every order divisible by 4 is a famous open problem.

Solution to Exercise 6.3.64

Uses. Lemma 6.3.39 (Sums of roots of unity)
Hint 1. Write $|g|^2 = g\overline{g} = \sum_j \sum_k \omega^{j^2-k^2}$. The exponent $j^2 - k^2$ suggests substituting $j = k + d$.
Hint 2. Since $\omega^n = 1$, ω^{j^2} depends only on j modulo n, so for fixed k you may let d run over $0, \ldots, n-1$ with $j \equiv k + d$. Then $j^2 - k^2 \equiv d^2 + 2dk$.
Hint 3. The sum over k of ω^{2dk} is given by Lemma 6.3.39. For which d is it nonzero when n is odd?
Sketch. $|g|^2 = \sum_k \sum_d \omega^{(k+d)^2-k^2} = \sum_d \omega^{d^2} \sum_k \omega^{2dk}$. The inner sum is n if $n \mid 2d$ and 0 otherwise. For odd n, $n \mid 2d$ only for $d = 0$, so $|g|^2 = n$. For $n = 4$ the values $d = 0, 2$ contribute, giving 8; for $n = 6$, $d = 0, 3$ contribute $6\left(1 + \omega^9\right) = 0$.

(a) Since $\omega^n = 1$, ω^m depends only on m modulo n; as $(j+n)^2 = j^2 + n(2j+n)$, the term ω^{j^2} depends only on j modulo n. Since $\overline{\omega^{k^2}} = \omega^{-k^2}$,

$$|g|^2 = g\overline{g} = \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} \omega^{j^2-k^2}.$$

Fix k. As d runs over $0, \ldots, n-1$, the residue of $k + d$ modulo n runs over all residues once, so we may replace j by $k + d$ in the inner sum:

$$\sum_{j=0}^{n-1} \omega^{j^2-k^2} = \sum_{d=0}^{n-1} \omega^{(k+d)^2-k^2} = \sum_{d=0}^{n-1} \omega^{d^2+2dk}.$$

Summing over k and exchanging the sums,

$$|g|^2 = \sum_{d=0}^{n-1} \omega^{d^2} \sum_{k=0}^{n-1} \omega^{k(2d)}.$$

By Lemma 6.3.39 the inner sum is n if $n \mid 2d$ and 0 otherwise. Since n is odd, $n \mid 2d$ implies $n \mid d$, which for $0 \le d \le n-1$ means $d = 0$. So $|g|^2 = \omega^0 \cdot n = n$.

(b) The computation up to the last step did not use that n is odd, so in general $|g|^2 = n \sum \omega^{d^2}$ over the $d \in \{0, \ldots, n-1\}$ with $n \mid 2d$. For $n = 4$ (where $\omega = i$) these are $d = 0, 2$, and $|g|^2 = 4\left(1 + i^4\right) = 8$; directly, $g = 1 + i + i^4 + i^9 = 2 + 2i$. For $n = 6$ they are $d = 0, 3$, and $|g|^2 = 6\left(1 + \omega^9\right) = 6\left(1 + \omega^3\right) = 0$, since $\omega^3 = -1$ for $n = 6$. The oddness of n was used only to conclude that $n \mid 2d$ forces $d = 0$.

Section 6.4

Solution to Exercise 6.4.13

Uses. Definition 6.4.1 (Projection onto the span of an orthonormal list), Definition 6.1.8 (Dot product), Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$)
Hint 1. $\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v} \cdot \mathbf{u}$ conjugates the entries of $\mathbf{u}$.
Hint 2. You should get $\langle \mathbf{v}, \mathbf{u} \rangle = -i/\sqrt{2}$, while $\langle \mathbf{u}, \mathbf{v} \rangle$ is its conjugate.
Sketch. (a) $\langle \mathbf{v}, \mathbf{u} \rangle = \left(1 \cdot \bar{i}\right)/\sqrt{2} = -i/\sqrt{2}$, so $\mathrm{Pr}_{\mathbf{u}}(\mathbf{v}) = (-i/2)(1, i, 0) = (-i/2, 1/2, 0)$, and $\mathbf{v} - \mathrm{Pr}_{\mathbf{u}}(\mathbf{v}) = (i/2, 1/2, 1)$ has dot product $i/2 + (1/2)(-i) = 0$ with $(1, i, 0)$. (b) $\langle \mathbf{u}, \mathbf{v} \rangle \mathbf{u} = (i/2, -1/2, 0)$, and $\mathbf{v}$ minus it has dot product $-2i \neq 0$ with $(1, i, 0)$.

(a) By Definition 6.1.8, $\langle \mathbf{v}, \mathbf{u} \rangle = \left(0 \cdot 1 + 1 \cdot \bar{i} + 1 \cdot 0\right)/\sqrt{2} = -i/\sqrt{2}$. So

$$\begin{aligned} \mathrm{Pr}_{\mathbf{u}}(\mathbf{v}) &= -\frac{i}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(1, i, 0) \\ &= -\frac{i}{2}(1, i, 0) = \left(-\frac{i}{2}, \frac{1}{2}, 0\right), \end{aligned}$$

and $\mathbf{v} - \mathrm{Pr}_{\mathbf{u}}(\mathbf{v}) = (i/2, 1/2, 1)$. Its dot product with $(1, i, 0)$ is $(i/2) \cdot 1 + (1/2) \cdot \bar{i} + 0 = i/2 - i/2 = 0$, so it is orthogonal to $\mathbf{u}$, as Proposition 6.4.4(b) says.

(b) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} = i/\sqrt{2}$, so the student's vector is $(i/2)(1, i, 0) = (i/2, -1/2, 0)$, and $\mathbf{v}$ minus it is $(-i/2, 3/2, 1)$. Its dot product with $(1, i, 0)$ is $-i/2 + (3/2)(-i) = -2i \neq 0$. Over $\mathbb{C}$ the vector being projected belongs in the first slot.

Solution to Exercise 6.4.21

Uses. Algorithm 6.4.14 (Gram–Schmidt), Definition 6.1.8 (Dot product), Corollary 6.4.10 (Projecting with an orthogonal list)
Hint 1. Work with residuals: $\mathbf{f}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{f}_1\rangle / \|\mathbf{f}_1\|^2 \, \mathbf{f}_1$ with $\mathbf{f}_1 = \mathbf{v}_1$.
Hint 2. $\langle \mathbf{v}_2, \mathbf{f}_1\rangle = 1 + \bar{i} = 1 - i$ and $\|\mathbf{f}_1\|^2 = 2$.
Sketch. $\mathbf{f}_2 = (1,1,1) - ((1-i)/2)(1,i,0) = ((1+i)/2, (1-i)/2, 1)$, of squared norm 2. So $\mathbf{u}_1 = (1,i,0)/\sqrt{2}$ and $\mathbf{u}_2 = ((1+i)/2, (1-i)/2, 1)/\sqrt{2}$. With the conjugated coefficient $(1+i)/2$, the residual has dot product $-2i \neq 0$ with $\mathbf{v}_1$.

We use Remark 6.4.16. First $\mathbf{f}_1 = (1,i,0)$ with $\|\mathbf{f}_1\|^2 = 2$. Next $\langle \mathbf{v}_2, \mathbf{f}_1\rangle = 1 \cdot 1 + 1 \cdot \bar{i} + 0 = 1 - i$, so

$$\begin{aligned}\mathbf{f}_2 &= (1,1,1) - \frac{1-i}{2}(1,i,0) \\ &= \left(1 - \frac{1-i}{2}, 1 - \frac{i+1}{2}, 1\right) \\ &= \left(\frac{1+i}{2}, \frac{1-i}{2}, 1\right).\end{aligned}$$

Check: $\mathbf{f}_2 \cdot \mathbf{f}_1 = (1+i)/2 + ((1-i)/2)\bar{i} = (1+i)/2 + (-i-1)/2 = 0$. Also $\|\mathbf{f}_2\|^2 = 1/2 + 1/2 + 1 = 2$. So

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,i,0), \quad \mathbf{u}_2 = \frac{1}{\sqrt{2}}\left(\frac{1+i}{2}, \frac{1-i}{2}, 1\right).$$

With the slots exchanged, the coefficient becomes $\overline{1-i}/2 = (1+i)/2$, and the "residual" is $(1,1,1) - ((1+i)/2)(1,i,0) = ((1-i)/2, (3-i)/2, 1)$. Its dot product with $\mathbf{v}_1$ is $(1-i)/2 + ((3-i)/2)(-i) = (1 - i - 3i - 1)/2 = -2i \neq 0$, so the output would not be orthogonal.

Solution to Exercise 6.4.28

Uses. Theorem 6.4.26 (The double complement), Corollary 6.4.10 (Projecting with an orthogonal list), Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.3.9 (Properties of $A^\perp$)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. $W = \operatorname{span}(\mathbf{j})^\perp$ for $\mathbf{j} = (1,1,1,1)$. Apply Theorem 6.4.26.
Hint 2. $\operatorname{Pr}_{W^\perp}$ is the projection onto a line. For the closest vector, guess $\mathbf{e}_1 - \operatorname{Pr}_{W^\perp}(\mathbf{e}_1)$ and check the two conditions of Technique 6.4.7 for the subspace W.
Sketch. $W = \operatorname{span}(\mathbf{j})^\perp$, so $W^\perp = \operatorname{span}(\mathbf{j})$ by Theorem 6.4.26, and $\operatorname{Pr}_{W^\perp}(\mathbf{x}) = ((x_1 + x_2 + x_3 + x_4)/4)\mathbf{j}$. The vector $\mathbf{p} = \mathbf{e}_1 - \mathbf{j}/4 = (3/4, -1/4, -1/4, -1/4)$ lies in W, and $\mathbf{e}_1 - \mathbf{p} = \mathbf{j}/4 \in W^\perp$, so $\mathbf{p} = \operatorname{Pr}_W(\mathbf{e}_1)$.

Let $\mathbf{j} = (1,1,1,1)$. Then $W = \{\mathbf{j}\}^\perp = \operatorname{span}(\mathbf{j})^\perp$ by Proposition 6.3.9(c), and $\operatorname{span}(\mathbf{j})$ is finite-dimensional, so Theorem 6.4.26 gives $W^\perp = \left(\operatorname{span}(\mathbf{j})^\perp\right)^\perp = \operatorname{span}(\mathbf{j})$. By Corollary 6.4.10 with the single vector $\mathbf{j}$, $\|\mathbf{j}\|^2 = 4$,

$$\operatorname{Pr}_{W^\perp}(\mathbf{x}) = \frac{x_1 + x_2 + x_3 + x_4}{4}\mathbf{j}.$$

In particular $\operatorname{Pr}_{W^\perp}(\mathbf{e}_1) = \mathbf{j}/4$. Let

$$\mathbf{p} = \mathbf{e}_1 - \frac{1}{4}\mathbf{j} = \left(\frac{3}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}\right).$$

Its coordinates add up to 0, so $\mathbf{p} \in W$, and $\mathbf{e}_1 - \mathbf{p} = \mathbf{j}/4 \in W^\perp$. By Theorem 6.4.6 (applied with an orthonormal basis of W), $\mathbf{p} = \operatorname{Pr}_W(\mathbf{e}_1)$. By Theorem 6.4.33 below, this is the vector of W closest to $\mathbf{e}_1$, at distance $\|\mathbf{j}/4\| = 1/2$. (Proposition 6.4.29(d) below explains the shortcut: $\operatorname{Pr}_W = \mathrm{id} - \operatorname{Pr}_{W^\perp}$.)

Solution to Exercise 6.4.36

Uses. Example 6.4.17 (Gram–Schmidt in $\mathbb{R}^3$), Theorem 6.4.33 (Best approximation), Example 6.4.35 (Distance to a hyperplane)
Hint 1. Example 6.4.17 already contains an orthonormal basis of W.
Hint 2. $\mathbf{u}_1 = (1,1,0)/\sqrt{2}$ and $\mathbf{u}_2 = (1,-1,2)/\sqrt{6}$. For the check, a normal vector of W is $\mathbf{u}_3$.
Sketch. With $\mathbf{u}_1, \mathbf{u}_2$ from Example 6.4.17, $\langle \mathbf{v}, \mathbf{u}_1\rangle = 0$ and $\langle \mathbf{v}, \mathbf{u}_2\rangle = 6/\sqrt{6}$, so $\operatorname{Pr}_W(\mathbf{v}) = (1,-1,2)$ and $\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) = (-1,1,1)$, of length $\sqrt{3}$. With $\mathbf{n} = (-1,1,1)$, $|\langle \mathbf{v}, \mathbf{n}\rangle| / \|\mathbf{n}\| = 3/\sqrt{3} = \sqrt{3}$.

By Example 6.4.17, $\mathbf{u}_1 = (1,1,0)/\sqrt{2}$, $\mathbf{u}_2 = (1,-1,2)/\sqrt{6}$ is an orthonormal basis of $\operatorname{span}(\mathbf{v}_1,\mathbf{v}_2) = W$. Here $\langle \mathbf{v}, \mathbf{u}_1\rangle = 0$ and $\langle \mathbf{v}, \mathbf{u}_2\rangle = 6/\sqrt{6}$, so by Theorem 6.4.33 the closest point is

$$\operatorname{Pr}_W(\mathbf{v}) = 0 \cdot \mathbf{u}_1 + \frac{6}{\sqrt{6}} \cdot \frac{1}{\sqrt{6}}(1,-1,2) = (1,-1,2).$$

The distance is $\|\mathbf{v} - \operatorname{Pr}_W(\mathbf{v})\| = \|(-1,1,1)\| = \sqrt{3}$. Check: $(-1,1,1)$ is orthogonal to $(1,1,0)$ and $(1,0,1)$, so $W = \{\mathbf{n}\}^\perp$ with $\mathbf{n} = (-1,1,1)$ (both are planes, and one contains the other), and Example 6.4.35 gives $|\langle \mathbf{v}, \mathbf{n}\rangle|/\|\mathbf{n}\| = 3/\sqrt{3} = \sqrt{3}$.

Solution to Exercise 6.4.40

Uses. Algorithm 6.4.14 (Gram–Schmidt), Remark 6.4.16 (Postponing the normalization), Corollary 6.4.10 (Projecting with an orthogonal list), Theorem 6.4.33 (Best approximation), Corollary 6.4.25 (Dimension of the orthogonal complement)
Hint 1. Run the algorithm with residuals only (Remark 6.4.16). The coefficients are $\langle \mathbf{v}_j, \mathbf{f}_i\rangle/\|\mathbf{f}_i\|^2$.
Hint 2. The residuals are $(1,1,1,1)$, $(1,-1,-1,1)$, $(0,1,-1,0)$. For (b), use Corollary 6.4.10 with these.
Hint 3. For (c), $\dim W^\perp = 1$, and $\mathbf{b} - \operatorname{Pr}_W(\mathbf{b})$ is a nonzero vector of $W^\perp$.
Sketch. $\mathbf{f}_1 = (1,1,1,1)$, $\mathbf{f}_2 = \mathbf{v}_2 - \mathbf{f}_1 = (1,-1,-1,1)$, $\mathbf{f}_3 = \mathbf{v}_3 - \mathbf{f}_1/2 + \mathbf{f}_2/2 = (0,1,-1,0)$. Then $\operatorname{Pr}_W(\mathbf{b}) = (10/4)\mathbf{f}_1 + 0\mathbf{f}_2 - (1/2)\mathbf{f}_3 = (5/2,2,3,5/2)$, and $\mathbf{b} - \operatorname{Pr}_W(\mathbf{b}) = (3/2)(-1,0,0,1)$, of length $3/\sqrt{2}$. $W^\perp = \operatorname{span}((1,0,0,-1))$.

(a) By Remark 6.4.16 we may work with residuals. First $\mathbf{f}_1 = (1,1,1,1)$, with $\|\mathbf{f}_1\|^2 = 4$. Next $\langle \mathbf{v}_2, \mathbf{f}_1\rangle = 4$, so

$$\mathbf{f}_2 = (2,0,0,2) - (1,1,1,1) = (1,-1,-1,1),$$

with $\|\mathbf{f}_2\|^2 = 4$. Then $\langle \mathbf{v}_3, \mathbf{f}_1\rangle = 2$ and $\langle \mathbf{v}_3, \mathbf{f}_2\rangle = -2$, so

$$\begin{aligned}\mathbf{f}_3 &= (0,2,0,0) - \frac{1}{2}(1,1,1,1) + \frac{1}{2}(1,-1,-1,1)\\ &= (0,1,-1,0),\end{aligned}$$

with $\|\mathbf{f}_3\|^2 = 2$. No residual is $\mathbf{0}$, so $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is independent, and

$$\mathbf{u}_1 = \frac{1}{2}(1,1,1,1), \quad \mathbf{u}_2 = \frac{1}{2}(1,-1,-1,1),$$

$$\mathbf{u}_3 = \frac{1}{\sqrt{2}}(0,1,-1,0)$$

is an orthonormal basis of W (Theorem 6.4.15).

(b) We have $\langle \mathbf{b}, \mathbf{f}_1\rangle = 10$, $\langle \mathbf{b}, \mathbf{f}_2\rangle = 1 - 2 - 3 + 4 = 0$, and $\langle \mathbf{b}, \mathbf{f}_3\rangle = -1$. By Corollary 6.4.10,

$$\begin{aligned}\operatorname{Pr}_W(\mathbf{b}) &= \frac{10}{4}(1,1,1,1) + 0 - \frac{1}{2}(0,1,-1,0)\\ &= \left(\frac{5}{2}, 2, 3, \frac{5}{2}\right).\end{aligned}$$

So $\mathbf{b} - \operatorname{Pr}_W(\mathbf{b}) = (-3/2,0,0,3/2)$, which is indeed orthogonal to $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$, and by Theorem 6.4.33, $\operatorname{dist}(\mathbf{b}, W) = \sqrt{9/4 + 9/4} = 3/\sqrt{2}$.

(c) By Corollary 6.4.25, $\dim W^\perp = 4 - 3 = 1$. The vector $(1,0,0,-1)$ is orthogonal to $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$, which span W, so it lies in $W^\perp$ (Proposition 6.3.9(c)). Hence $W^\perp = \operatorname{span}((1,0,0,-1))$. As it must, $\mathbf{b} - \operatorname{Pr}_W(\mathbf{b}) = -(3/2)(1,0,0,-1)$ lies in $W^\perp$.

Solution to Exercise 6.4.41

Uses. Theorem 6.4.31 (The matrix of an orthogonal projection), Corollary 6.4.10 (Projecting with an orthogonal list), Proposition 6.4.29 (The orthogonal projection as a linear map), Corollary 6.4.25 (Dimension of the orthogonal complement)
Hint 1. Gram–Schmidt with residuals gives $(1,0,1)$ and $(1,2,-1)$. Use $\sum_i \mathbf{f}_i\mathbf{f}_i^{\mathsf{T}}/\|\mathbf{f}_i\|^2$.
Hint 2. For (b), find $\mathbf{n} \neq \mathbf{0}$ orthogonal to both spanning vectors; then $P = I_3 - \mathbf{n}\mathbf{n}^{\mathsf{T}}/\|\mathbf{n}\|^2$ by Proposition 6.4.29(d).

Sketch. Both ways give $P = (1/3)\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$: in (a) as $(1/2)\mathbf{f}_1\mathbf{f}_1^{\mathsf{T}} + (1/6)\mathbf{f}_2\mathbf{f}_2^{\mathsf{T}}$ with $\mathbf{f}_1 = (1,0,1)$, $\mathbf{f}_2 = (1,2,-1)$, and in (b) as $I_3 - (1/3)\mathbf{n}\mathbf{n}^{\mathsf{T}}$ with $\mathbf{n} = (-1,1,1)$. The trace is $6/3 = 2 = \dim W$.

(a) Gram–Schmidt with residuals: $\mathbf{f}_1 = (1,0,1)$, and $(1,1,0) - (1/2)(1,0,1) = (1/2, 1, -1/2)$, which we replace by $\mathbf{f}_2 = (1,2,-1)$. By Theorem 6.4.31 and Corollary 6.4.10,

$$P = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}.$$

(b) The vector $\mathbf{n} = (-1,1,1)$ is orthogonal to $(1,0,1)$ and $(1,1,0)$, so it lies in $W^{\perp}$, which is a line by Corollary 6.4.25. So $W^{\perp} = \operatorname{span}(\mathbf{n})$ and, by Proposition 6.4.29(d),

$$P = I_3 - \frac{\mathbf{n}\mathbf{n}^{\mathsf{T}}}{3} = I_3 - \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$
$$= \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}.$$

The matrix is symmetric, and $\operatorname{tr} P = (2+2+2)/3 = 2$. For $P^2 = P$, by Theorem 6.4.31 there is nothing to check, but directly: the first row of $3P$ times the columns of $3P$ gives $(4+1+1, 2+2-1, 2-1+2) = (6,3,3)$, which is 3 times the first row of $3P$, and similarly for the other rows; so $(3P)^2 = 3(3P)$, that is, $P^2 = P$.

Solution to Exercise 6.4.42

Uses. Corollary 6.4.10 (Projecting with an orthogonal list), Theorem 6.4.6 (The projection is characterized by orthogonality), Algorithm 6.4.14 (Gram–Schmidt)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Which hypothesis of Corollary 6.4.10 does the list $\mathbf{a}, \mathbf{b}$ fail?
Hint 2. Check whether $\mathbf{v} - (5/2, 3/2, 0)$ is orthogonal to $\mathbf{a}$. To fix the computation, orthogonalize $\mathbf{a}, \mathbf{b}$ first.
Sketch. The formula needs an orthogonal list, and $\langle \mathbf{a}, \mathbf{b} \rangle = 1 \neq 0$. The student's $\mathbf{v} - \mathbf{p} = (-3/2, 1/2, 3)$ has inner product $-3/2$ with $\mathbf{a}$. Gram–Schmidt replaces $\mathbf{b}$ by $(0,1,0)$, and then $\operatorname{Pr}_W(\mathbf{v}) = (1,0,0) + (0,2,0) = (1,2,0)$.

The formula of Corollary 6.4.10 holds for orthogonal lists, and $\langle \mathbf{a}, \mathbf{b} \rangle = 1 \neq 0$. The test of Theorem 6.4.6 detects the error: for $\mathbf{p} = (5/2, 3/2, 0)$, the vector $\mathbf{v} - \mathbf{p} = (-3/2, 1/2, 3)$ has $\langle \mathbf{v} - \mathbf{p}, \mathbf{a} \rangle = -3/2 \neq 0$, so $\mathbf{v} - \mathbf{p} \notin W^{\perp}$.

Gram–Schmidt turns $\mathbf{a}, \mathbf{b}$ into the orthogonal list $\mathbf{a}$, $\mathbf{b} - \langle \mathbf{b}, \mathbf{a} \rangle \mathbf{a} = (0,1,0)$ with the same span, so W is the x_1x_2-plane. Now Corollary 6.4.10 applies:

$$\operatorname{Pr}_W(\mathbf{v}) = 1 \cdot (1,0,0) + 2 \cdot (0,1,0) = (1,2,0),$$

and $\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) = (0,0,3)$ is orthogonal to W. The student's vector even lies in W; it is just the wrong vector of W.

Solution to Exercise 6.4.43

Uses. Algorithm 6.4.14 (Gram–Schmidt), Theorem 6.4.15 (Gram–Schmidt is correct), Remark 6.4.16 (Postponing the normalization)
Hint 1. Use residuals. $\mathbf{f}_1 = (1,0,1,0)$ and $\mathbf{f}_2$ is a multiple of $(1,2,-1,0)$.
Hint 2. The third residual is $\mathbf{0}$. Its formula expresses $\mathbf{v}_3$ through $\mathbf{f}_1$ and $\mathbf{f}_2$; rewrite those in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Sketch. $\mathbf{f}_1 = (1,0,1,0)$; $\mathbf{f}_2 = \mathbf{v}_2 - \mathbf{f}_1/2$, rescaled to $\mathbf{g} = (1,2,-1,0) = 2\mathbf{v}_2 - \mathbf{v}_1$; $\mathbf{v}_3 - (-1/2)\mathbf{f}_1 - (3/6)\mathbf{g} = \mathbf{0}$, so $\mathbf{v}_3$ is skipped and $\mathbf{v}_3 = -\mathbf{v}_1/2 + \mathbf{g}/2 = \mathbf{v}_2 - \mathbf{v}_1$; $\mathbf{v}_4$ is already orthogonal to the others. The dimension is 3.

With residuals (Remark 6.4.16): $\mathbf{f}_1 = (1,0,1,0)$, $\|\mathbf{f}_1\|^2 = 2$. Then $\langle \mathbf{v}_2, \mathbf{f}_1 \rangle = 1$, and

$$\mathbf{v}_2 - \frac{1}{2}\mathbf{f}_1 = \left(\frac{1}{2}, 1, -\frac{1}{2}, 0\right),$$

which we replace by $\mathbf{g} = (1,2,-1,0) = 2\mathbf{v}_2 - \mathbf{v}_1$, with $\|\mathbf{g}\|^2 = 6$. In the third pass, $\langle \mathbf{v}_3, \mathbf{f}_1 \rangle = -1$ and $\langle \mathbf{v}_3, \mathbf{g} \rangle = 2 + 1 = 3$, so

$$\begin{aligned} &\mathbf{v}_3 + \frac{1}{2}(1,0,1,0) - \frac{3}{6}(1,2,-1,0) \\ &= \left(0 + \frac{1}{2} - \frac{1}{2}, 1 - 1, -1 + \frac{1}{2} + \frac{1}{2}, 0\right) = \mathbf{0}. \end{aligned}$$

So $\mathbf{v}_3$ is skipped, and by Theorem 6.4.15(b) it lies in $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. Explicitly, $\mathbf{v}_3 = -(1/2)\mathbf{v}_1 + (1/2)(2\mathbf{v}_2 - \mathbf{v}_1) = \mathbf{v}_2 - \mathbf{v}_1$, which is easily checked. Finally $\mathbf{v}_4$ is orthogonal to $\mathbf{f}_1$ and $\mathbf{g}$, so its residual is $\mathbf{v}_4$ itself. Three vectors are kept, so the output

$$\frac{1}{\sqrt{2}}(1,0,1,0), \quad \frac{1}{\sqrt{6}}(1,2,-1,0), \quad (0,0,0,1)$$

is an orthonormal basis of $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_4)$ (Theorem 6.4.15(c)), whose dimension is therefore 3.

Solution to Exercise 6.4.44

Uses. Theorem 6.4.33 (Best approximation), Example 6.4.19 (Legendre polynomials), Corollary 6.4.10 (Projecting with an orthogonal list)
Hint 1. Use the orthogonal basis $1, t, t^2 - 1/3$ of Example 6.4.19, with squared norms $2, 2/3, 8/45$. Each new degree adds one term to the projection.
Hint 2. $\langle |t|, 1 \rangle = 1$, $\langle |t|, t \rangle = 0$ (odd integrand), and $\langle |t|, t^2 \rangle = 2\int_0^1 t^3\,dt = 1/2$.
Hint 3. The squared distance is $\|f\|^2$ minus the squared norm of the projection, and each term contributes $\langle f, \mathbf{f}_j \rangle^2 / \|\mathbf{f}_j\|^2$.
Sketch. The coefficients are $1/2, 0$, and $(1/6)/(8/45) = 15/16$. So the best constant and the best linear function are both $1/2$, with squared distance $2/3 - 1/2 = 1/6$, and the best quadratic is $1/2 + (15/16)\left(t^2 - 1/3\right) = \left(3 + 15t^2\right)/16$, with squared distance $1/6 - (1/36)(45/8) = 1/96$.

Let $f(t) = |t|$. The polynomial functions of degree at most n form a finite-dimensional subspace W_n, and by Theorem 6.4.33 the closest one to f is $\operatorname{Pr}_{W_n}(f)$. By Example 6.4.19, 1; $1, t$; and $1, t, t^2 - 1/3$ are orthogonal bases of W_0, W_1, W_2, with squared norms 2, $2/3$, $8/45$. We need

$$\langle f, 1 \rangle = 2\int_0^1 t\,dt = 1, \quad \langle f, t \rangle = 0,$$

$$\left\langle f, t^2 - \frac{1}{3} \right\rangle = 2\int_0^1 t^3\,dt - \frac{1}{3} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6},$$

where $\langle f, t \rangle = 0$ because $|t|t$ is odd. By Corollary 6.4.10,

$$\begin{aligned} \operatorname{Pr}_{W_0}(f) &= \operatorname{Pr}_{W_1}(f) = \frac{1}{2}, \\ \operatorname{Pr}_{W_2}(f) &= \frac{1}{2} + \frac{1/6}{8/45}\left(t^2 - \frac{1}{3}\right) \\ &= \frac{1}{2} + \frac{15}{16}\left(t^2 - \frac{1}{3}\right) = \frac{3 + 15t^2}{16}. \end{aligned}$$

For the distances, $\|f\|^2 = \int_{-1}^1 t^2\,dt = 2/3$, and by Theorem 6.4.33 (applied with the normalized basis, whose squared inner products with f are $\langle f, \mathbf{f}_j \rangle^2 / \|\mathbf{f}_j\|^2$),

$$\operatorname{dist}(f, W_0)^2 = \operatorname{dist}(f, W_1)^2 = \frac{2}{3} - \frac{1^2}{2} = \frac{1}{6},$$

$$\operatorname{dist}(f, W_2)^2 = \frac{1}{6} - \frac{(1/6)^2}{8/45} = \frac{1}{6} - \frac{5}{32} = \frac{1}{96}.$$

The best linear approximation is no better than the best constant, because f is even and t is odd. Allowing t^2 cuts the squared error by a factor of 16, although $|t|$ is not differentiable at 0 and has no Taylor polynomial there.

Solution to Exercise 6.4.45

Uses. Theorem 6.4.33 (Best approximation), Example 6.4.19 (Legendre polynomials), Exercise 6.3.49
Hint 1. The integral is $\|t^3 - p\|^2$ for $p = a + bt + ct^2$ in the subspace W of polynomial functions of degree at most 2. Which p minimizes it?
Hint 2. $p = \operatorname{Pr}_W(t^3)$, and $t^3 - \operatorname{Pr}_W(t^3)$ is the fourth Gram–Schmidt residual of Example 6.4.19.
Sketch. By Theorem 6.4.33 the minimizer is $p = \operatorname{Pr}_W(t^3)$, and $t^3 - \operatorname{Pr}_W(t^3)$ is the residual $t^3 - (3/5)t$ from Gram–Schmidt on $1, t, t^2, t^3$. So $a = c = 0$, $b = 3/5$, and the minimum is $\|t^3 - (3/5)t\|^2 = 8/175$ (Exercise 6.3.49).

Let W be the space of polynomial functions of degree at most 2 in $C[-1, 1]$ with $\langle f, g\rangle = \int_{-1}^{1} fg$. The integral is $\|t^3 - p\|^2$ with $p = a + bt + ct^2 \in W$, and as (a, b, c) ranges over $\mathbb{R}^3$, p ranges over all of W. By Theorem 6.4.33 the minimum is attained exactly at $p = \operatorname{Pr}_W(t^3)$.

By Example 6.4.19, $W = \operatorname{span}(1, t, t^2)$, and the fourth Gram–Schmidt residual is $t^3 - \operatorname{Pr}_W(t^3) = t^3 - (3/5)t$. So $\operatorname{Pr}_W(t^3) = (3/5)t$: the minimum is attained exactly at $a = 0$, $b = 3/5$, $c = 0$, and its value is

$$\begin{aligned}\left\|t^3 - \frac{3}{5}t\right\|^2 &= \int_{-1}^{1}\left(t^6 - \frac{6}{5}t^4 + \frac{9}{25}t^2\right)\mathrm{d}t \\ &= \frac{2}{7} - \frac{12}{25} + \frac{6}{25} = \frac{8}{175},\end{aligned}$$

as computed in Exercise 6.3.49. (Calculus would set three partial derivatives to 0 and solve a 3×3 system; the orthogonal basis does that work once and for all.)

Solution to Exercise 6.4.46

Uses. Algorithm 6.4.14 (Gram–Schmidt), Remark 6.4.16 (Postponing the normalization), Theorem 6.4.33 (Best approximation)
Hint 1. $\langle x^i, x^j\rangle = 1/(i + j + 1)$. Work with residuals.
Hint 2. The residuals are 1, $x - 1/2$, $x^2 - x + 1/6$, with squared norms 1, $1/12$, $1/180$.
Hint 3. For (b), the minimizer is $a + bx = \operatorname{Pr}(x^2)$ onto $\operatorname{span}(1, x)$, and the minimum is the squared norm of the third residual.
Sketch. $f_1 = 1$; $f_2 = x - \langle x, 1\rangle = x - 1/2$; $f_3 = x^2 - 1/3 - (1/12)/(1/12)(x - 1/2) = x^2 - x + 1/6$. Normalizing with $\|f_2\|^2 = 1/12$ and $\|f_3\|^2 = 1/180$ gives the list. (b) The minimum is $\|f_3\|^2 = 1/180$, at $a + bt = t - 1/6$.

(a) Here $\langle x^i, x^j\rangle = \int_0^1 t^{i+j}\,\mathrm{d}t = 1/(i + j + 1)$. With residuals: $f_1 = 1$, $\|f_1\|^2 = 1$, and $f_2 = x - \langle x, 1\rangle \cdot 1 = x - 1/2$, with

$$\|f_2\|^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 \mathrm{d}t = \frac{1}{12}.$$

Next, $\langle x^2, 1\rangle = 1/3$ and $\langle x^2, x - 1/2\rangle = 1/4 - 1/6 = 1/12$, so

$$f_3 = x^2 - \frac{1}{3} - \frac{1/12}{1/12}\left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}.$$

Expanding, $\|f_3\|^2 = \int_0^1 \left(t^4 - 2t^3 + (4/3)t^2 - t/3 + 1/36\right)\mathrm{d}t = 1/5 - 1/2 + 4/9 - 1/6 + 1/36 = 1/180$. Normalizing,

$$u_1 = 1,$$

$$u_2 = \sqrt{12}\left(x - \frac{1}{2}\right) = \sqrt{3}(2x - 1),$$
$$u_3 = \sqrt{180}\left(x^2 - x + \frac{1}{6}\right) = \sqrt{5}\left(6x^2 - 6x + 1\right).$$

The first two are the output of the lecture's example.

(b) As (a, b) ranges over $\mathbb{R}^2$, $a + bx$ ranges over $W = \operatorname{span}(1, x) = \operatorname{span}(f_1, f_2)$. By Theorem 6.4.33, the minimum is $\operatorname{dist}\left(x^2, W\right)^2 = \left\|x^2 - \operatorname{Pr}_W\left(x^2\right)\right\|^2$, and $x^2 - \operatorname{Pr}_W\left(x^2\right)$ is the residual f_3. So the minimum is $1/180$, attained at $a + bx = x^2 - f_3 = x - 1/6$, that is, $a = -1/6$, $b = 1$.

Solution to Exercise 6.4.47

Uses. Definition 6.4.22 (Orthogonal complement), Proposition 1.4.24 (The sum of two subspaces), Theorem 6.4.26 (The double complement), Proposition 6.3.9 (Properties of $A^\perp$)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For (a), unwind the definitions. One containment uses $U, W \subseteq U + W$; the other expands $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle$.
Hint 2. For (b), apply (a) to the subspaces $U^\perp$ and $W^\perp$, then take $\perp$ of both sides.
Hint 3. Theorem 6.4.26 says $\left(X^\perp\right)^\perp = X$ for every subspace X of the finite-dimensional space V.
Sketch. (a) $U, W \subseteq U + W$ gives $(U + W)^\perp \subseteq U^\perp \cap W^\perp$; conversely $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle = 0$. (b) By (a) and Theorem 6.4.26, $\left(U^\perp + W^\perp\right)^\perp = U^{\perp\perp} \cap W^{\perp\perp} = U \cap W$. Taking $\perp$ and using Theorem 6.4.26 again, $U^\perp + W^\perp = (U \cap W)^\perp$.

(a) Since $U \subseteq U + W$ and $W \subseteq U + W$ (Proposition 1.4.24), Proposition 6.3.9(b) gives $(U + W)^\perp \subseteq U^\perp$ and $(U + W)^\perp \subseteq W^\perp$. Conversely, let $\mathbf{v} \in U^\perp \cap W^\perp$. Every vector of $U + W$ is $\mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$, and

$$\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle = 0 + 0 = 0.$$

So $\mathbf{v} \in (U + W)^\perp$.

(b) Every subspace X of the finite-dimensional space V is finite-dimensional, so $\left(X^\perp\right)^\perp = X$ by Theorem 6.4.26. Apply (a) to the subspaces $U^\perp$ and $W^\perp$:

$$\left(U^\perp + W^\perp\right)^\perp = \left(U^\perp\right)^\perp \cap \left(W^\perp\right)^\perp = U \cap W.$$

Taking orthogonal complements of both sides and applying Theorem 6.4.26 to $X = U^\perp + W^\perp$,

$$U^\perp + W^\perp = \left(\left(U^\perp + W^\perp\right)^\perp\right)^\perp = (U \cap W)^\perp,$$

as desired.

Solution to Exercise 6.4.48

Uses. Theorem 6.4.24 (Orthogonal decomposition), Proposition 6.4.4 (Properties of $\operatorname{Pr}_{\mathcal{O}}$), Remark 6.4.12 (Why "orthogonal"), Definition 4.4.4 (Direct sum)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), (c), and (e), look in $\mathbb{R}^2$, using lines that are neither equal nor perpendicular.
Hint 2. For (b), $W \subseteq U^\perp$ is clear. For the other containment, write $\mathbf{v} \in U^\perp$ as $\mathbf{u} + \mathbf{w}$ and take the inner product with $\mathbf{u}$.
Hint 3. For (d), use Proposition 6.4.4(e).
Sketch. (a) False: $\mathbb{R}^2 = \operatorname{span}(\mathbf{e}_1) \oplus \operatorname{span}((1, 1))$. (b) True: if $\mathbf{v} = \mathbf{u} + \mathbf{w} \in U^\perp$, then $0 = \langle \mathbf{v}, \mathbf{u} \rangle = \|\mathbf{u}\|^2$. (c) False: for $U = \operatorname{span}(\mathbf{e}_1)$, $W = \operatorname{span}((1, 1))$, $\operatorname{Pr}_U \operatorname{Pr}_W \mathbf{e}_1 = (1/2, 0)$ but $\operatorname{Pr}_W \operatorname{Pr}_U \mathbf{e}_1 = (1/2, 1/2)$. (d) True, by the equality case of Bessel. (e) False: $P(x, y) = (x + y, 0)$ has $P^2 = P$ and $\|P(1, 1)\| = 2 > \sqrt{2}$.

(a) False. In $\mathbb{R}^2$, let $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \operatorname{span}((1, 1))$. Then $U \cap W = \{\mathbf{0}\}$ and $U + W = \mathbb{R}^2$, so $V = U \oplus W$, but $U^\perp = \operatorname{span}(\mathbf{e}_2) \neq W$. Every subspace has many complements, and $U^\perp$ is only one of them.

(b) True. By hypothesis $W \subseteq U^\perp$. Conversely, let $\mathbf{v} \in U^\perp$, and write $\mathbf{v} = \mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Then

$$0 = \langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = \|\mathbf{u}\|^2,$$

so $\mathbf{u} = \mathbf{0}$ and $\mathbf{v} = \mathbf{w} \in W$.

(c) False. With U and W as in (a), $\operatorname{Pr}_W(\mathbf{e}_1) = (1/2)(1, 1)$, so

$$\operatorname{Pr}_U \operatorname{Pr}_W(\mathbf{e}_1) = \operatorname{Pr}_U\left(\frac{1}{2}, \frac{1}{2}\right) = \left(\frac{1}{2}, 0\right),$$
$$\operatorname{Pr}_W \operatorname{Pr}_U(\mathbf{e}_1) = \operatorname{Pr}_W(\mathbf{e}_1) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

(Exercise 6.4.58 shows when the two products agree.)

(d) True. By Proposition 6.4.4(e) applied to an orthonormal basis of W, $\|\operatorname{Pr}_W(\mathbf{v})\| = \|\mathbf{v}\|$ holds only if $\mathbf{v} \in W$.

(e) False. $P(x, y) = (x + y, 0)$ is linear with $P^2(x, y) = P(x + y, 0) = (x + y, 0) = P(x, y)$, but $\|P(1, 1)\| = \|(2, 0)\| = 2 > \sqrt{2} = \|(1, 1)\|$. Like the map of Remark 6.4.12, P projects onto the x-axis along a direction that is not perpendicular to it (here $(1, -1)$), and such projections can stretch vectors. Exercise 6.4.57 shows that this is the only way an idempotent can fail to shrink norms.

Solution to Exercise 6.4.49

Uses. Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 1.4.24 (The sum of two subspaces), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. For (a), check the two conditions of Technique 6.4.7 for $\mathbf{p} = \operatorname{Pr}_U(\mathbf{v}) + \operatorname{Pr}_W(\mathbf{v})$ and the subspace $U + W$. ($U + W$ is finite-dimensional: concatenate orthonormal bases.)
Hint 2. For $\operatorname{Pr}_U \operatorname{Pr}_W = 0$: $\operatorname{Pr}_W(\mathbf{v}) \in W \subseteq U^\perp = \ker \operatorname{Pr}_U$.
Hint 3. For (b), compute $\langle \mathbf{w}, \mathbf{u} \rangle = \langle \operatorname{Pr}_W \mathbf{w}, \operatorname{Pr}_U \mathbf{u} \rangle$ and move Pr_U across with Proposition 6.4.29(c).
Sketch. (a) $U + W$ is spanned by the concatenation of orthonormal bases of U and W, which is orthonormal. For $\mathbf{p} = \operatorname{Pr}_U(\mathbf{v}) + \operatorname{Pr}_W(\mathbf{v}) \in U + W$, $\langle \mathbf{v} - \mathbf{p}, \mathbf{u} \rangle = \langle \mathbf{v} - \operatorname{Pr}_U(\mathbf{v}), \mathbf{u} \rangle - \langle \operatorname{Pr}_W(\mathbf{v}), \mathbf{u} \rangle = 0$ for $\mathbf{u} \in U$, and similarly for W; so $\mathbf{p} = \operatorname{Pr}_{U+W}(\mathbf{v})$. Also $\operatorname{Pr}_W(\mathbf{v}) \in W \subseteq U^\perp = \ker \operatorname{Pr}_U$. (b) $\langle \mathbf{w}, \mathbf{u} \rangle = \langle \operatorname{Pr}_W \mathbf{w}, \operatorname{Pr}_U \mathbf{u} \rangle = \langle \operatorname{Pr}_U \operatorname{Pr}_W \mathbf{w}, \mathbf{u} \rangle = 0$.

(a) Let $\mathbf{u}_1, \ldots, \mathbf{u}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_l$ be orthonormal bases of U and W (Theorem 6.4.20). By hypothesis $\langle \mathbf{u}_i, \mathbf{w}_j \rangle = 0$, so the concatenated list is orthonormal, and it spans $U + W$. So $U + W$ is finite-dimensional and Pr_{U+W} is defined. (In fact the formula of Definition 6.4.11 for this basis already gives the claim; we give the argument of Technique 6.4.7 instead.)

Let $\mathbf{v} \in V$ and $\mathbf{p} = \operatorname{Pr}_U(\mathbf{v}) + \operatorname{Pr}_W(\mathbf{v})$, which lies in $U + W$. For $\mathbf{u} \in U$, we have $\operatorname{Pr}_W(\mathbf{v}) \in W$, so $\langle \operatorname{Pr}_W(\mathbf{v}), \mathbf{u} \rangle = 0$ by hypothesis, and

$$\begin{aligned} \langle \mathbf{v} - \mathbf{p}, \mathbf{u} \rangle &= \langle \mathbf{v} - \operatorname{Pr}_U(\mathbf{v}), \mathbf{u} \rangle - \langle \operatorname{Pr}_W(\mathbf{v}), \mathbf{u} \rangle \\ &= 0 - 0 = 0. \end{aligned}$$

In the same way $\langle \mathbf{v} - \mathbf{p}, \mathbf{w} \rangle = 0$ for $\mathbf{w} \in W$. So $\mathbf{v} - \mathbf{p}$ is orthogonal to every $\mathbf{u} + \mathbf{w}$, that is, to $U + W$, and $\mathbf{p} = \operatorname{Pr}_{U+W}(\mathbf{v})$ by Theorem 6.4.6.

The hypothesis says $W \subseteq U^\perp$, and $U^\perp = \ker \operatorname{Pr}_U$ by Proposition 6.4.29(a). Since $\operatorname{Pr}_W(\mathbf{v}) \in W$, we get $\operatorname{Pr}_U \operatorname{Pr}_W(\mathbf{v}) = \mathbf{0}$ for every $\mathbf{v}$.

(b) Let $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Then $\mathbf{u} = \operatorname{Pr}_U(\mathbf{u})$ and $\mathbf{w} = \operatorname{Pr}_W(\mathbf{w})$, and by Proposition 6.4.29(c) for Pr_U,

$$\begin{aligned} \langle \mathbf{w}, \mathbf{u} \rangle &= \langle \operatorname{Pr}_W(\mathbf{w}), \operatorname{Pr}_U(\mathbf{u}) \rangle = \langle \operatorname{Pr}_U(\operatorname{Pr}_W(\mathbf{w})), \mathbf{u} \rangle \\ &= \langle \mathbf{0}, \mathbf{u} \rangle = 0. \end{aligned}$$

Solution to Exercise 6.4.50

Uses. Theorem 6.4.6 (The projection is characterized by orthogonality), Exercise 6.4.49
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Every step but one is an instance of something proved in this section. Which statement about $\operatorname{Pr}_W(\mathbf{v})$ has no justification?
Hint 2. Test the claim with $U = W$.
Sketch. The step "$\operatorname{Pr}_W(\mathbf{v}) \perp U$" is unjustified: $\operatorname{Pr}_W(\mathbf{v})$ lies in W, and W need not be orthogonal to U. The claim is false: for $U = W = \operatorname{span}(\mathbf{e}_1)$ in $\mathbb{R}^2$, $\operatorname{Pr}_U + \operatorname{Pr}_W = 2 \operatorname{Pr}_U$, which sends $\mathbf{e}_1$ to $2\mathbf{e}_1$, while $\operatorname{Pr}_{U+W}(\mathbf{e}_1) = \mathbf{e}_1$. It is true when $U \perp W$ (Exercise 6.4.49).

The error is the assertion $\mathrm{Pr}_W(\mathbf{v}) \perp U$. All we know is that $\mathrm{Pr}_W(\mathbf{v}) \in W$, and a vector of W is orthogonal to U only if, for instance, every vector of W is orthogonal to every vector of U.

The claim is false. Let $U = W = \mathrm{span}(\mathbf{e}_1)$ in $\mathbb{R}^2$. Then $U + W = U$, and

$$(\mathrm{Pr}_U + \mathrm{Pr}_W)(\mathbf{e}_1) = 2\mathbf{e}_1 \neq \mathbf{e}_1 = \mathrm{Pr}_{U+W}(\mathbf{e}_1).$$

The argument becomes correct under the extra hypothesis that every vector of U is orthogonal to every vector of W, and then it proves Exercise 6.4.49(a).

Solution to Exercise 6.4.51

Uses. Theorem 6.4.31 (The matrix of an orthogonal projection), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.4.33 (Best approximation), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. By Theorem 6.4.31, $P = QQ^*$ with $Q^*Q = I_k$. Use cyclicity of the trace.
Hint 2. For (b), $p_{ii} = (P\mathbf{e}_i)\cdot\mathbf{e}_i$. Use Proposition 6.4.29(c).
Hint 3. For (c), $\mathrm{dist}(\mathbf{e}_i, W)^2 = 1 - \|\mathrm{Pr}_W(\mathbf{e}_i)\|^2$, and the n numbers $\|\mathrm{Pr}_W(\mathbf{e}_i)\|^2$ add up to k.
Sketch. (a) $\mathrm{rank}\, P = \dim \mathrm{im}\, \mathrm{Pr}_W = k$, and $\mathrm{tr}(QQ^*) = \mathrm{tr}(Q^*Q) = \mathrm{tr}\, I_k = k$. (b) $p_{ii} = \langle P\mathbf{e}_i, \mathbf{e}_i\rangle = \langle P\mathbf{e}_i, P\mathbf{e}_i\rangle$; sum over i. (c) $\mathrm{dist}(\mathbf{e}_i, W)^2 = 1 - p_{ii}$, and the average of the p_{ii} is k/n, so some $p_{ii} \geq k/n$.

(a) By Proposition 6.4.29(a), the image of Pr_W is W, so $\mathrm{rank}\, P = \dim W = k$. By Theorem 6.4.20, W has an orthonormal basis, and by Theorem 6.4.31, $P = QQ^*$ with $Q \in \mathbb{F}^{n\times k}$ and $Q^*Q = I_k$. By Proposition 2.4.28, $\mathrm{tr}\, P = \mathrm{tr}(QQ^*) = \mathrm{tr}(Q^*Q) = \mathrm{tr}\, I_k = k$.

(b) The ith entry of $P\mathbf{e}_i$ is p_{ii}, so $p_{ii} = (P\mathbf{e}_i)\cdot\mathbf{e}_i = \langle P\mathbf{e}_i, \mathbf{e}_i\rangle$. By Proposition 6.4.29(c), this equals $\langle P\mathbf{e}_i, P\mathbf{e}_i\rangle = \|P\mathbf{e}_i\|^2$. Summing over i and using (a), $\sum_i \|\mathrm{Pr}_W(\mathbf{e}_i)\|^2 = \mathrm{tr}\, P = k$.

(c) By Theorem 6.4.33, $\mathrm{dist}(\mathbf{e}_i, W)^2 = \|\mathbf{e}_i\|^2 - \|\mathrm{Pr}_W(\mathbf{e}_i)\|^2 = 1 - p_{ii}$. The n numbers p_{ii} add up to k, so the largest of them is at least k/n. For that i, $\mathrm{dist}(\mathbf{e}_i, W)^2 \leq 1 - k/n$. (For the plane of Example 6.4.32, every p_{ii} equals $2/3$, so the bound is attained by all three.)

Solution to Exercise 6.4.52

Uses. Theorem 6.4.24 (Orthogonal decomposition), Theorem 6.2.1 (Cauchy–Schwarz inequality), Definition 6.4.34 (Distance to a subspace), Exercise 6.4.45
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. For $\mathbf{x} \in W^\perp$, $\langle \mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v} - \mathrm{Pr}_W(\mathbf{v}), \mathbf{x}\rangle$. Why?
Hint 2. Apply Cauchy–Schwarz, and find when equality holds.
Hint 3. For (b), W is the space of polynomial functions of degree at most 2; the distance from t^3 to it was found in Exercise 6.4.45.
Sketch. For a unit $\mathbf{x} \in W^\perp$, $\langle \mathrm{Pr}_W(\mathbf{v}), \mathbf{x}\rangle = 0$, so $|\langle \mathbf{v}, \mathbf{x}\rangle| = |\langle \mathbf{v} - \mathrm{Pr}_W(\mathbf{v}), \mathbf{x}\rangle| \leq \|\mathbf{v} - \mathrm{Pr}_W(\mathbf{v})\|$ by Cauchy–Schwarz. Equality holds for $\mathbf{x} = (\mathbf{v} - \mathrm{Pr}_W(\mathbf{v}))/\mathrm{dist}(\mathbf{v}, W)$, which is a unit vector in $W^\perp$. (b) The maximum is $\mathrm{dist}(t^3, W) = \sqrt{8/175}$, at $g = (t^3 - 3t/5)/\sqrt{8/175}$.

(a) Let $\mathbf{r} = \mathbf{v} - \mathrm{Pr}_W(\mathbf{v})$, so $\mathbf{r} \in W^\perp$ and $\|\mathbf{r}\| = \mathrm{dist}(\mathbf{v}, W) > 0$ because $\mathbf{v} \notin W$. For a unit vector $\mathbf{x} \in W^\perp$, $\langle \mathrm{Pr}_W(\mathbf{v}), \mathbf{x}\rangle = 0$ since $\mathrm{Pr}_W(\mathbf{v}) \in W$, so by Theorem 6.2.1

$$|\langle \mathbf{v}, \mathbf{x}\rangle| = |\langle \mathbf{r}, \mathbf{x}\rangle| \leq \|\mathbf{r}\|\|\mathbf{x}\| = \mathrm{dist}(\mathbf{v}, W).$$

For $\mathbf{x} = \mathbf{r}/\|\mathbf{r}\|$, which is a unit vector in $W^\perp$, we get $\langle \mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{r}, \mathbf{r}\rangle/\|\mathbf{r}\| = \|\mathbf{r}\|$. So the maximum exists and equals $\mathrm{dist}(\mathbf{v}, W)$. By the equality case of Cauchy–Schwarz, the maximizers are the unit vectors $\mathbf{x} = \lambda\mathbf{r}/\|\mathbf{r}\|$ with $|\lambda| = 1$.

(b) Let W be the space of polynomial functions of degree at most 2. A function orthogonal to $1, t, t^2$ is orthogonal to W (Proposition 6.3.9(c)), so we are maximizing $\langle t^3, g\rangle$ over unit vectors $g \in W^\perp$. By (a) and Exercise 6.4.45, the maximum of $|\langle t^3, g\rangle|$ is $\mathrm{dist}(t^3, W) = \sqrt{8/175}$, attained at $g = \pm q/\|q\|$ with $q = t^3 - (3/5)t$. Taking the plus sign makes $\langle t^3, g\rangle$ positive, so the largest value of $\int_{-1}^{1} t^3 g$ is $\sqrt{8/175} \approx 0.214$.

So the distance from $\mathbf{v}$ to W can be computed in two dual ways: as a minimum over W, or as a maximum over unit vectors of $W^\perp$.

Solution to Exercise 6.4.53

Uses. Theorem 6.4.33 (Best approximation), Proposition 3.1.19 (Solution sets are translates of the kernel), Corollary 6.4.10 (Projecting with an orthogonal list)
Hint 1. $\|\mathbf{y} - (\mathbf{x}_0 + \mathbf{u})\| = \|(\mathbf{y} - \mathbf{x}_0) - \mathbf{u}\|$. So you are approximating $\mathbf{y} - \mathbf{x}_0$ by vectors of U.
Hint 2. For (b), the solution set is $\mathbf{x}_0 + \ker A$ (Proposition 3.1.19), and the closest point to $\mathbf{y} = \mathbf{0}$ is wanted.
Sketch. (a) By Theorem 6.4.33 applied to $\mathbf{y} - \mathbf{x}_0$, $\|(\mathbf{y} - \mathbf{x}_0) - \mathbf{u}\|$ is minimized exactly at $\mathbf{u} = \operatorname{Pr}_U(\mathbf{y} - \mathbf{x}_0)$. (b) With $\mathbf{x}_0 = (2, 0, 2)$ and $\ker A = \operatorname{span}((1, -1, 1))$, the answer is $\mathbf{x}_0 - \operatorname{Pr}_{\ker A}(\mathbf{x}_0) = (2, 0, 2) - (4/3)(1, -1, 1) = (2/3, 4/3, 2/3)$, which is orthogonal to $(1, -1, 1)$.

(a) For $\mathbf{u} \in U$, $\|\mathbf{y} - (\mathbf{x}_0 + \mathbf{u})\| = \|(\mathbf{y} - \mathbf{x}_0) - \mathbf{u}\|$. By Theorem 6.4.33, applied to the vector $\mathbf{y} - \mathbf{x}_0$ and the subspace U, this is smallest exactly when $\mathbf{u} = \operatorname{Pr}_U(\mathbf{y} - \mathbf{x}_0)$. So $\mathbf{x}_0 + \operatorname{Pr}_U(\mathbf{y} - \mathbf{x}_0)$ is the unique closest vector.

(b) Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. One solution is $\mathbf{x}_0 = (2, 0, 2)$, and $\ker A = \operatorname{span}((1, -1, 1))$: the equations say $x_1 = -x_2 = x_3$. By Proposition 3.1.19 the solution set is $\mathbf{x}_0 + \ker A$. The solution of smallest norm is the one closest to $\mathbf{y} = \mathbf{0}$, which by (a) is

$$\begin{aligned} \mathbf{x}_0 + \operatorname{Pr}_{\ker A}(-\mathbf{x}_0) &= \mathbf{x}_0 - \frac{\langle \mathbf{x}_0, (1, -1, 1) \rangle}{3}(1, -1, 1) \\ &= (2, 0, 2) - \frac{4}{3}(1, -1, 1) \\ &= \left(\frac{2}{3}, \frac{4}{3}, \frac{2}{3} \right), \end{aligned}$$

using Corollary 6.4.10. Check: $2/3 + 4/3 = 2$ and $4/3 + 2/3 = 2$; the norm is $\sqrt{24/9} = \sqrt{8/3}$; and $(2/3, 4/3, 2/3) \cdot (1, -1, 1) = 0$. In general the least-norm solution is $\mathbf{x}_0 - \operatorname{Pr}_{\ker A}(\mathbf{x}_0)$, the component of any solution in $(\ker A)^\perp$, which for real A is the span of the rows (Example 6.4.23(b), Theorem 6.4.26). It is the only solution in $(\ker A)^\perp$.

Solution to Exercise 6.4.54

Uses. Example 6.3.2 (Orthogonal vectors), Proposition 6.1.27 (Basic properties of the norm), Theorem 6.3.3 (Pythagorean theorem), Definition 6.4.22 (Orthogonal complement)
Hint 1. $O \subseteq E^\perp$ because the product of an even and an odd function is odd. For $E^\perp \subseteq O$, write $f \in E^\perp$ as $f_e + f_o$ and pair f with f_e.
Hint 2. For (b), $f - f_e = f_o$ is odd, hence orthogonal to E. Run the Pythagoras argument of Theorem 6.4.33 by hand.
Sketch. (a) Odd times even is odd, with integral 0, so $O \subseteq E^\perp$. If $f \in E^\perp$, then $0 = \langle f, f_e \rangle = \|f_e\|^2 + \langle f_o, f_e \rangle = \|f_e\|^2$, so $f = f_o \in O$. The same with E and O exchanged gives $O^\perp = E$. (b) For $g \in E$, $f - g = f_o + (f_e - g)$ with $f_o \perp f_e - g$, so $\|f - g\|^2 = \|f_o\|^2 + \|f_e - g\|^2 \geq \|f_o\|^2$, with equality only for $g = f_e$.

Every $f \in C[-1, 1]$ is $f = f_e + f_o$ with $f_e(t) = (f(t) + f(-t))/2$ even and $f_o(t) = (f(t) - f(-t))/2$ odd, both continuous.

(a) If g is even and h is odd, then gh is odd, so $\langle g, h \rangle = 0$ (Example 6.3.2(c)). This gives $O \subseteq E^\perp$ and $E \subseteq O^\perp$. Conversely, let $f \in E^\perp$. Since $f_e \in E$,

$$0 = \langle f, f_e \rangle = \langle f_e, f_e \rangle + \langle f_o, f_e \rangle = \|f_e\|^2,$$

so $f_e = \mathbf{0}$ (Proposition 6.1.27) and $f = f_o \in O$. Similarly, if $f \in O^\perp$, then $0 = \langle f, f_o \rangle = \|f_o\|^2$, so $f = f_e \in E$.

(b) Let $g \in E$. Then $f - g = f_o + (f_e - g)$, where $f_o \in O = E^\perp$ and $f_e - g \in E$. By Theorem 6.3.3,

$$\|f - g\|^2 = \|f_o\|^2 + \|f_e - g\|^2 \geq \|f_o\|^2 = \|f - f_e\|^2,$$

with equality if and only if $g = f_e$. So f_e is the unique even function closest to f. The proof of Theorem 6.4.33 only needed the decomposition $f = f_e + f_o$ into a part in E and a part orthogonal to E. For finite-dimen-

sional subspaces that decomposition always exists (Theorem 6.4.24); here it exists for a reason specific to E, and Exercise 6.4.63 shows that it need not exist in general.

Solution to Exercise 6.4.55

Uses. Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 6.4.15 (Gram–Schmidt is correct), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. $\mathbf{u}'_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$, so expand it with Proposition 6.3.18(a). Which coefficients vanish?
Hint 2. For $i < j$, $\mathbf{u}_i \in \operatorname{span}(\mathbf{u}'_1, \ldots, \mathbf{u}'_{j-1})$, which is orthogonal to $\mathbf{u}'_j$.
Hint 3. For (b), compare an arbitrary such list with the Gram–Schmidt output using (a), and use $\langle \mathbf{v}_j, \lambda \mathbf{u}_j \rangle = \overline{\lambda} \langle \mathbf{v}_j, \mathbf{u}_j \rangle$.
Sketch. (a) $\mathbf{u}'_j = \sum_{i \le j} \langle \mathbf{u}'_j, \mathbf{u}_i \rangle \mathbf{u}_i$, and for $i < j$, $\mathbf{u}_i \in \operatorname{span}(\mathbf{u}'_1, \ldots, \mathbf{u}'_{j-1})$, which is orthogonal to $\mathbf{u}'_j$. So $\mathbf{u}'_j = \lambda_j \mathbf{u}_j$, and $|\lambda_j| = \|\mathbf{u}'_j\| = 1$. (b) By Theorem 6.4.15(d) the output satisfies the conditions. For another such list, (a) gives $\mathbf{u}'_j = \lambda_j \mathbf{u}_j$, and $\langle \mathbf{v}_j, \mathbf{u}'_j \rangle = \overline{\lambda_j} \langle \mathbf{v}_j, \mathbf{u}_j \rangle$ with both inner products positive, so $\overline{\lambda_j} > 0$ and $|\lambda_j| = 1$ force $\lambda_j = 1$.

(a) Fix j. Since $\mathbf{u}'_j \in \operatorname{span}(\mathbf{u}'_1, \ldots, \mathbf{u}'_j) = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$, Proposition 6.3.18(a) gives

$$\mathbf{u}'_j = \sum_{i=1}^{j} \langle \mathbf{u}'_j, \mathbf{u}_i \rangle \mathbf{u}_i.$$

For $i < j$, $\mathbf{u}_i \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}) = \operatorname{span}(\mathbf{u}'_1, \ldots, \mathbf{u}'_{j-1})$, and $\mathbf{u}'_j$ is orthogonal to each of $\mathbf{u}'_1, \ldots, \mathbf{u}'_{j-1}$, hence to their span (Proposition 6.3.9(c)). So $\langle \mathbf{u}'_j, \mathbf{u}_i \rangle = 0$ for $i < j$, and $\mathbf{u}'_j = \lambda_j \mathbf{u}_j$ with $\lambda_j = \langle \mathbf{u}'_j, \mathbf{u}_j \rangle$. Then $1 = \|\mathbf{u}'_j\| = |\lambda_j| \|\mathbf{u}_j\| = |\lambda_j|$ by Proposition 6.1.27(b).

(b) By Theorem 6.4.15(d), the output $\mathbf{u}_1, \ldots, \mathbf{u}_m$ of Gram–Schmidt has both properties. Let $\mathbf{u}'_1, \ldots, \mathbf{u}'_m$ be another orthonormal list with both properties. The two lists have the same spans $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j)$, so by (a), $\mathbf{u}'_j = \lambda_j \mathbf{u}_j$ with $|\lambda_j| = 1$. By Proposition 6.1.4,

$$\langle \mathbf{v}_j, \mathbf{u}'_j \rangle = \overline{\lambda_j} \langle \mathbf{v}_j, \mathbf{u}_j \rangle.$$

Both inner products are positive reals, so $\overline{\lambda_j}$ is a positive real number of modulus 1, that is, $\lambda_j = 1$. Hence $\mathbf{u}'_j = \mathbf{u}_j$ for every j.

Solution to Exercise 6.4.56

Uses. Theorem 6.4.33 (Best approximation), Example 6.4.19 (Legendre polynomials), Exercise 6.3.49
Hint 1. The polynomials of degree n with leading coefficient 1 are exactly the $x^n - q$ with $q \in \mathbb{R}[x]_{\le n-1}$.
Hint 2. Minimizing $\|x^n - q\|$ over $q \in W$ is a best-approximation problem.
Sketch. Such polynomials are $x^n - q$ with $q \in W$, and $\|x^n - q\|$ is minimized exactly at $q = \operatorname{Pr}_W(x^n)$. By Theorem 6.4.15(d), W is the span of the first n Gram–Schmidt outputs, so $x^n - \operatorname{Pr}_W(x^n)$ is the next residual. For $n = 2, 3$ these are $x^2 - 1/3$ and $x^3 - (3/5)x$, with squared norms $8/45$ and $8/175$.

A polynomial of degree n with leading coefficient 1 is $x^n - q$ with $q \in W = \mathbb{R}[x]_{\le n-1}$, and every such difference is one. The space W is finite-dimensional, so by Theorem 6.4.33, $\|x^n - q\|$ is smallest for exactly one $q \in W$, namely $q = \operatorname{Pr}_W(x^n)$. When Gram–Schmidt is run on $1, x, x^2, \ldots$, the first n outputs span $\operatorname{span}(1, \ldots, x^{n-1}) = W$ (Theorem 6.4.15(d)), so the next residual is exactly $x^n - \operatorname{Pr}_W(x^n)$.

By Example 6.4.19, for $n = 2$ the minimizer is $x^2 - 1/3$, with $\|x^2 - 1/3\|^2 = 8/45$ (Example 6.3.17), and for $n = 3$ it is $x^3 - (3/5)x$, with squared norm $8/175$ (Exercise 6.3.49). So among all monic quadratics, $x^2 - 1/3$ is the smallest on $[-1, 1]$ in the mean-square sense.

Solution to Exercise 6.4.57

Uses. Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 6.4.4 (Properties of $\operatorname{Pr}_{\mathcal{O}}$), Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.1.29 (Expanding a norm)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7), Orthogonal decomposition (Technique 6.2.4), Proving TFAE (Technique 2.7.1)
Hint 1. (i) $\Rightarrow$ (iii) is Bessel. For (ii) $\Rightarrow$ (i), check the two conditions of Technique 6.4.7 for $P\mathbf{v}$; note $\mathbf{v} - P\mathbf{v} \in \ker P$.
Hint 2. For (iii) $\Rightarrow$ (ii), take $\mathbf{u} \in \operatorname{im} P$ and $\mathbf{w} \in \ker P$. Then $P\mathbf{u} = \mathbf{u}$, so $P(\mathbf{u} + \lambda \mathbf{w}) = \mathbf{u}$ for every scalar λ, and (iii) gives $\|\mathbf{u}\| \le \|\mathbf{u} + \lambda \mathbf{w}\|$.
Hint 3. Choose λ to make $\mathbf{u} + \lambda \mathbf{w}$ the part of $\mathbf{u}$ orthogonal to $\mathbf{w}$, as in Technique 6.2.4.

Sketch. (i) ⇒ (iii): Proposition 6.4.4(e). (iii) ⇒ (ii): for $\mathbf{u} \in \operatorname{im} P$, $\mathbf{w} \in \ker P$, $\mathbf{w} \neq \mathbf{0}$, and $\lambda = -\langle \mathbf{u}, \mathbf{w}\rangle / \|\mathbf{w}\|^2$, we get $\|\mathbf{u}\|^2 = \|P(\mathbf{u} + \lambda \mathbf{w})\|^2 \leq \|\mathbf{u} + \lambda \mathbf{w}\|^2 = \|\mathbf{u}\|^2 - |\langle \mathbf{u}, \mathbf{w}\rangle|^2 / \|\mathbf{w}\|^2$, so $\langle \mathbf{u}, \mathbf{w}\rangle = 0$. (ii) ⇒ (i): $P\mathbf{v} \in \operatorname{im} P$ and $\mathbf{v} - P\mathbf{v} \in \ker P \subseteq (\operatorname{im} P)^\perp$, so $P\mathbf{v} = \operatorname{Pr}_{\operatorname{im} P}(\mathbf{v})$.

We prove (i) ⇒ (iii) ⇒ (ii) ⇒ (i).

(i) ⇒ (iii). This is Proposition 6.4.4(e), applied to an orthonormal basis of $\operatorname{im} P$.

(iii) ⇒ (ii). Let $\mathbf{u} \in \operatorname{im} P$ and $\mathbf{w} \in \ker P$; we may assume $\mathbf{w} \neq \mathbf{0}$. Write $\mathbf{u} = P\mathbf{x}$; then $P\mathbf{u} = P^2\mathbf{x} = P\mathbf{x} = \mathbf{u}$. So for every scalar λ, $P(\mathbf{u} + \lambda \mathbf{w}) = \mathbf{u} + \lambda \mathbf{0} = \mathbf{u}$, and (iii) gives $\|\mathbf{u}\| \leq \|\mathbf{u} + \lambda \mathbf{w}\|$. Take $\lambda = -\langle \mathbf{u}, \mathbf{w}\rangle / \|\mathbf{w}\|^2$. By Proposition 6.1.29,

$$\begin{aligned} \|\mathbf{u} + \lambda \mathbf{w}\|^2 &= \|\mathbf{u}\|^2 + 2\operatorname{Re}\left(\overline{\lambda}\langle \mathbf{u}, \mathbf{w}\rangle\right) + |\lambda|^2 \|\mathbf{w}\|^2 \\ &= \|\mathbf{u}\|^2 - 2\frac{|\langle \mathbf{u}, \mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2} + \frac{|\langle \mathbf{u}, \mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2} \\ &= \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2}. \end{aligned}$$

Combined with $\|\mathbf{u}\|^2 \leq \|\mathbf{u} + \lambda \mathbf{w}\|^2$, this gives $|\langle \mathbf{u}, \mathbf{w}\rangle|^2 \leq 0$, so $\langle \mathbf{u}, \mathbf{w}\rangle = 0$.

(ii) ⇒ (i). Let $W = \operatorname{im} P$, which is finite-dimensional. For $\mathbf{v} \in V$, $P\mathbf{v} \in W$, and $P(\mathbf{v} - P\mathbf{v}) = P\mathbf{v} - P^2\mathbf{v} = \mathbf{0}$, so $\mathbf{v} - P\mathbf{v} \in \ker P$. By (ii), $\mathbf{v} - P\mathbf{v} \in W^\perp$. By Theorem 6.4.6, $P\mathbf{v} = \operatorname{Pr}_W(\mathbf{v})$.

So an idempotent operator fails to be an orthogonal projection exactly when it stretches some vector, as the map in Exercise 6.4.48(e) does.

Solution to Exercise 6.4.58

Uses. Proposition 6.4.30 (Recognizing orthogonal projections), Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 6.1.37 (Inner products distinguish vectors)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Both directions go through Proposition 6.4.30. Compute $\langle PQ\mathbf{v}, \mathbf{x}\rangle$ by moving P and then Q to the other slot.
Hint 2. You get $\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, QP\mathbf{x}\rangle$. Compare this with the symmetry condition for PQ.
Hint 3. For the image: $\operatorname{im} PQ \subseteq U$ always, and $PQ = QP$ puts it inside W too. Conversely, $PQ\mathbf{z} = \mathbf{z}$ for $\mathbf{z} \in U \cap W$.
Sketch. By Proposition 6.4.29(c), $\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle Q\mathbf{v}, P\mathbf{x}\rangle = \langle \mathbf{v}, QP\mathbf{x}\rangle$. If $PQ = QP$, then $(PQ)^2 = P^2Q^2 = PQ$ and $\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, PQ\mathbf{x}\rangle$, so PQ is an orthogonal projection onto $\operatorname{im} PQ = U \cap W$. Conversely, if PQ is an orthogonal projection, its symmetry gives $\langle \mathbf{v}, PQ\mathbf{x}\rangle = \langle \mathbf{v}, QP\mathbf{x}\rangle$ for all $\mathbf{v}$, so $PQ = QP$.

For all $\mathbf{v}, \mathbf{x} \in V$, Proposition 6.4.29(c), applied first to P and then to Q, gives

$$\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle Q\mathbf{v}, P\mathbf{x}\rangle = \langle \mathbf{v}, QP\mathbf{x}\rangle. \tag{C.1.8}$$

Suppose $PQ = QP$. Then $(PQ)^2 = PQPQ = PPQQ = P^2Q^2 = PQ$ by Proposition 6.4.29(b), and by (C.1.8), $\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, QP\mathbf{x}\rangle = \langle \mathbf{v}, PQ\mathbf{x}\rangle$. By Proposition 6.4.30, $PQ = \operatorname{Pr}_{\operatorname{im} PQ}$. Its image lies in $\operatorname{im} P = U$, and since $PQ = QP$ also in $\operatorname{im} Q = W$. Conversely, if $\mathbf{z} \in U \cap W$, then $PQ\mathbf{z} = P\mathbf{z} = \mathbf{z}$, so $\mathbf{z} \in \operatorname{im} PQ$. Hence $\operatorname{im} PQ = U \cap W$ and $PQ = \operatorname{Pr}_{U \cap W}$.

Conversely, suppose PQ is an orthogonal projection. By Proposition 6.4.30, $\langle PQ\mathbf{v}, \mathbf{x}\rangle = \langle \mathbf{v}, PQ\mathbf{x}\rangle$ for all $\mathbf{v}, \mathbf{x}$, and comparing with (C.1.8) gives $\langle \mathbf{v}, PQ\mathbf{x} - QP\mathbf{x}\rangle = 0$ for all $\mathbf{v}$. Taking $\mathbf{v} = PQ\mathbf{x} - QP\mathbf{x}$ shows that this vector has norm 0, so $PQ\mathbf{x} = QP\mathbf{x}$ for every $\mathbf{x}$.

Exercise 6.4.48(c) gives two subspaces for which PQ is not an orthogonal projection. The commuting case includes $U \subseteq W$, where $PQ = QP = P$, and $U \perp W$, where $PQ = QP = 0$ (Exercise 6.4.49).

Solution to Exercise 6.4.59

Uses. Theorem 6.4.15 (Gram–Schmidt is correct), Theorem 6.3.22 (Parseval's identity), Proposition 5.3.15 (Triangular matrices), Theorem 5.4.10 (Multiplicativity), Exercise 6.3.57, Technique 6.4.5 (Projections shrink norms), Theorem 6.3.29

(Orthonormal columns), Theorem 5.4.7 (Invertibility and the determinant), Proposition 6.4.29 (The orthogonal projection as a linear map)
Techniques. Gram matrices (Technique 6.3.31)
Hint 1. For $i > j$, $\mathbf{v}_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$ is orthogonal to $\mathbf{u}_i$. For the product, $(R^*R)_{ij} = \sum_l \overline{r_{li}} r_{lj}$; recognize Parseval's identity in $\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m)$.
Hint 2. For (b), take determinants: $\det R^* = \overline{\det R}$. Gram–Schmidt on $\mathbf{v}_1, \ldots, \mathbf{v}_j$ produces the same first j residuals.
Hint 3. For (c), $\mathbf{f}_j = \mathbf{v}_j - \operatorname{Pr}(\mathbf{v}_j)$, and projections shrink norms.
Sketch. (a) $r_{ij} = 0$ for $i > j$ since $\mathbf{v}_j \perp \mathbf{u}_i$, and $r_{jj} = \|\mathbf{f}_j\|$ by Theorem 6.4.15(d). By Parseval in the span of the $\mathbf{u}$'s, $(R^*R)_{ij} = \sum_l \langle \mathbf{v}_j, \mathbf{u}_l \rangle \overline{\langle \mathbf{v}_i, \mathbf{u}_l \rangle} = \langle \mathbf{v}_j, \mathbf{v}_i \rangle = g_{ij}$. (b) $\det G = |\det R|^2 = \prod_j \|\mathbf{f}_j\|^2$; the ratio of consecutive Gram determinants is $\|\mathbf{f}_m\|^2$, the squared distance. (c) $\|\mathbf{f}_j\| \le \|\mathbf{v}_j\|$ with equality iff $\mathbf{v}_j \perp \mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$. For A with independent columns, $G = A^*A$ and $\det G = |\det A|^2$.

(a) Let $U = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$. For $i > j$, $\mathbf{v}_j \in \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$ (Theorem 6.4.15(d)), which is orthogonal to $\mathbf{u}_i$, so $r_{ij} = 0$: R is upper triangular. Also $r_{jj} = \langle \mathbf{v}_j, \mathbf{u}_j \rangle = \|\mathbf{f}_j\|$ by Theorem 6.4.15(d). Note that $r_{li} = \langle \mathbf{v}_i, \mathbf{u}_l \rangle$ for all l and i. Since $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is an orthonormal basis of the inner product space U, which contains every $\mathbf{v}_i$, Theorem 6.3.22(a) gives

$$(R^*R)_{ij} = \sum_{l=1}^{m} \overline{r_{li}} r_{lj} = \sum_{l=1}^{m} \langle \mathbf{v}_j, \mathbf{u}_l \rangle \overline{\langle \mathbf{v}_i, \mathbf{u}_l \rangle} = \langle \mathbf{v}_j, \mathbf{v}_i \rangle = g_{ij}.$$

(b) By Theorem 5.4.10 and Exercise 6.3.57(a), $\det G = \det(R^*) \det R = \overline{\det R} \det R = |\det R|^2$, and by Proposition 5.3.15, $\det R = r_{11} \cdots r_{mm} = \|\mathbf{f}_1\| \cdots \|\mathbf{f}_m\|$. So $\det G = \prod_j \|\mathbf{f}_j\|^2$. Gram–Schmidt processes the list in order, so running it on $\mathbf{v}_1, \ldots, \mathbf{v}_j$ produces the residuals $\mathbf{f}_1, \ldots, \mathbf{f}_j$, and $\det G_j = \prod_{i \le j} \|\mathbf{f}_i\|^2$. Hence $\det G_m / \det G_{m-1} = \|\mathbf{f}_m\|^2$. Finally $\mathbf{f}_m = \mathbf{v}_m - \operatorname{Pr}_X(\mathbf{v}_m)$ with $X = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_{m-1}) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{m-1})$, so $\|\mathbf{f}_m\| = \operatorname{dist}(\mathbf{v}_m, X)$ by Theorem 6.4.33.

(c) Write $X_j = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$, so $\mathbf{f}_j = \mathbf{v}_j - \operatorname{Pr}_{X_j}(\mathbf{v}_j)$. By Technique 6.4.5, $\|\mathbf{f}_j\| \le \|\mathbf{v}_j\|$, with equality if and only if $\operatorname{Pr}_{X_j}(\mathbf{v}_j) = \mathbf{0}$, that is (Proposition 6.4.29(a)), if and only if $\mathbf{v}_j \perp X_j$. Multiplying these inequalities between nonnegative numbers, $\det G \le \prod_j \|\mathbf{v}_j\|^2$, with equality if and only if $\mathbf{v}_j \perp \mathbf{v}_i$ for all $i < j$, that is, if and only if the list is orthogonal.

For $A \in \mathbb{F}^{n \times n}$: if its columns are dependent, then A is not invertible and $\det A = 0$ (Theorem 5.4.7), and there is nothing to prove. Otherwise, by Theorem 6.3.29, A^*A is the Gram matrix G of the columns, so $|\det A|^2 = \overline{\det A} \det A = \det(A^*A) = \det G \le \prod_j \|\mathbf{a}_j\|^2$. Geometrically: the volume of a box is at most the product of its side lengths, with equality only for a rectangular box.

Solution to Exercise 6.4.60

Uses. Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 6.1.27 (Basic properties of the norm)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Projections shrink norms (Technique 6.4.5)
Hint 1. Write $P_i = \operatorname{Pr}_{W_i}$. For any $\mathbf{v}$, compute $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \sum_i P_i \mathbf{v} \rangle$ and rewrite each term with Proposition 6.4.29(c).
Hint 2. You get $\|\mathbf{v}\|^2 = \sum_i \|P_i \mathbf{v}\|^2$. Now take $\mathbf{v} \in W_1$.
Sketch. With $P_i = \operatorname{Pr}_{W_i}$, $\|\mathbf{v}\|^2 = \sum_i \langle P_i \mathbf{v}, \mathbf{v} \rangle = \sum_i \|P_i \mathbf{v}\|^2$. For $\mathbf{v} \in W_i$, the ith term is $\|\mathbf{v}\|^2$, so $P_j \mathbf{v} = \mathbf{0}$ for $j \ne i$, that is, $\mathbf{v} \in \ker P_j = W_j^\perp$. Existence of the decomposition: $\mathbf{v} = \sum_i P_i \mathbf{v}$. Uniqueness: apply P_i to $\sum_j \mathbf{w}_j$, which kills $\mathbf{w}_j$ for $j \ne i$ and fixes $\mathbf{w}_i$.

Write $P_i = \operatorname{Pr}_{W_i}$. For every $\mathbf{v} \in V$, by the hypothesis and Proposition 6.4.29(c),

$$\begin{aligned} \|\mathbf{v}\|^2 &= \left\langle \sum_{i=1}^{m} P_i \mathbf{v}, \mathbf{v} \right\rangle = \sum_{i=1}^{m} \langle P_i \mathbf{v}, \mathbf{v} \rangle \\ &= \sum_{i=1}^{m} \langle P_i \mathbf{v}, P_i \mathbf{v} \rangle = \sum_{i=1}^{m} \|P_i \mathbf{v}\|^2. \end{aligned}$$

Now let $\mathbf{v} \in W_i$. Then $P_i \mathbf{v} = \mathbf{v}$, so the identity reads $\|\mathbf{v}\|^2 = \|\mathbf{v}\|^2 + \sum_{j \ne i} \|P_j \mathbf{v}\|^2$. A sum of nonnegative terms is 0 only if each term is, so $P_j \mathbf{v} = \mathbf{0}$ for every $j \ne i$, that is, $\mathbf{v} \in \ker P_j = W_j^\perp$ (Proposition 6.4.29(a)). So every vector of W_i is orthogonal to every vector of W_j.

For $\mathbf{v} \in V$, the hypothesis gives $\mathbf{v} = P_1\mathbf{v} + \cdots + P_m\mathbf{v}$ with $P_i\mathbf{v} \in W_i$. If also $\mathbf{v} = \mathbf{w}_1 + \cdots + \mathbf{w}_m$ with $\mathbf{w}_j \in W_j$, apply P_i: for $j \neq i$, $\mathbf{w}_j \in W_j \subseteq W_i^\perp = \ker P_i$, and $P_i\mathbf{w}_i = \mathbf{w}_i$. So $P_i\mathbf{v} = \mathbf{w}_i$, and the decomposition is unique.

For example, if $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is an orthonormal basis of V, then $\mathrm{Pr}_{\mathrm{span}(\mathbf{u}_1)} + \cdots + \mathrm{Pr}_{\mathrm{span}(\mathbf{u}_n)} = \mathrm{id}_V$ by Proposition 6.3.18(b). The exercise says that every decomposition of id_V into orthogonal projections looks like a grouping of such a basis.

Solution to Exercise 6.4.61

Uses. Lemma 6.1.14 (Nonnegative functions with zero integral), Proposition 7.1.22 (Counting roots), the intermediate value theorem
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Let $t_1 < \cdots < t_r$ be the points of $(-1, 1)$ where p changes sign. Build a polynomial s of degree r such that ps does not change sign on $[-1, 1]$.
Hint 2. Take $s(t) = (t - t_1)\cdots(t - t_r)$. If $r < n$, then $\int_{-1}^{1} ps = 0$ by hypothesis. Why is that impossible?
Hint 3. A continuous function that is ≥ 0 on $[-1, 1]$ and has integral 0 vanishes on $[-1, 1]$ (Lemma 6.1.14).
Sketch. p has finitely many roots, so $(-1, 1)$ splits into finitely many open intervals on which p has constant sign. Let $t_1 < \cdots < t_r$ be the roots at which the sign changes, and $s = \prod_i (t - t_i)$. Then ps has constant sign on $[-1, 1]$ and is not identically zero there, so $\int ps \neq 0$. If $r < n$, $\deg s < n$ and the hypothesis gives $\int ps = 0$, a contradiction. So p has at least n distinct roots in $(-1, 1)$, and at most n roots in all.

The polynomial p is nonzero, so it has at most n roots (Proposition 7.1.22). Let $a_1 < \cdots < a_N$ be its roots in $(-1, 1)$, and put $a_0 = -1$, $a_{N+1} = 1$. On each open interval (a_k, a_{k+1}), p has no root, so by the intermediate value theorem it has constant sign there. Say that p *changes sign* at a_k $(1 \leq k \leq N)$ if its signs on (a_{k-1}, a_k) and (a_k, a_{k+1}) differ, and let $t_1 < \cdots < t_r$ be the roots at which p changes sign. Let

$$s(t) = (t - t_1)(t - t_2)\cdots(t - t_r)$$

($s = 1$ if $r = 0$). Each factor $t - t_i$ changes sign at t_i and nowhere else, so s has constant sign on each interval (a_k, a_{k+1}) and changes sign exactly at $t_1, \ldots, t_r$. Hence the product ps has constant sign on each (a_k, a_{k+1}), and at each a_k either both factors change sign or neither does, so ps has the same sign on all of these intervals. Replacing s by $-s$ if necessary, $ps \geq 0$ on $[-1, 1]$ by continuity.

Suppose $r < n$. Then $\deg s = r < n$, so $\int_{-1}^{1} p(t)s(t)\,\mathrm{d}t = 0$ by hypothesis. Since ps is continuous and nonnegative on $[-1, 1]$, Lemma 6.1.14 gives $ps = 0$ on $[-1, 1]$. But ps is a nonzero polynomial, with finitely many roots, a contradiction. So $r \geq n$: p has at least n distinct roots in $(-1, 1)$. Since it has at most n roots, these are all of its roots.

The argument used only that p is continuous, has finitely many zeros in $[-1, 1]$, and is orthogonal to all polynomials of degree less than n. So it applies to the error $e^t - p(t)$ of Example 6.4.37 with $n = 3$: that error has at most three zeros (see the remark after Example 6.4.37), and therefore changes sign at least three times in $(-1, 1)$.

Solution to Exercise 6.4.62

Uses. Corollary 6.4.25 (Dimension of the orthogonal complement), Corollary 6.4.10 (Projecting with an orthogonal list), Proposition 6.4.4 (Properties of $\mathrm{Pr}_{\mathcal{O}}$)
Techniques. Induction on the length of a list (Technique 1.6.9), Orthogonal decomposition (Technique 6.2.4)
Hint 1. Induct on the dimension, proving the statement for every n-dimensional subspace of some $\mathbb{R}^N$. Project the first $m - 1$ vectors onto the hyperplane $\mathbf{v}_m^\perp$.
Hint 2. Let $\mathbf{w}_i = \mathbf{v}_i - c_i\mathbf{v}_m$ with $c_i = \langle \mathbf{v}_i, \mathbf{v}_m \rangle / \|\mathbf{v}_m\|^2 < 0$. Compute $\langle \mathbf{w}_i, \mathbf{w}_j \rangle$ for $i \neq j$.
Hint 3. $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - c_i c_j \|\mathbf{v}_m\|^2 < 0$, since $c_i c_j > 0$. For the example, try $\mathbf{e}_i - t(1, \ldots, 1)$ and $-(1, \ldots, 1)$.
Sketch. Prove, by induction on n, that pairwise negative inner products allow at most $n + 1$ vectors in any n-dimensional subspace V of $\mathbb{R}^N$. For $n = 0$ only $\mathbf{0}$ exists and $\langle \mathbf{0}, \mathbf{0} \rangle = 0$, so $m \leq 1$. For $n \geq 1$ and $m \geq 2$, all $\mathbf{v}_i \neq \mathbf{0}$; project $\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}$ onto $U = \mathbf{v}_m^\perp \cap V$, of dimension $n - 1$. The projections $\mathbf{w}_i = \mathbf{v}_i - c_i\mathbf{v}_m$ have $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - c_i c_j \|\mathbf{v}_m\|^2 < 0$, so $m - 1 \leq n$. Example: $\mathbf{e}_i - (1/2n)\mathbf{j}$ for $i \leq n$, and $-\mathbf{j}$, with $\mathbf{j} = (1, \ldots, 1)$.

We prove, by induction on $n \geq 0$, the following statement: if V is an n-dimensional subspace of $\mathbb{R}^N$ (for some N) and $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ satisfy $\langle \mathbf{v}_i, \mathbf{v}_j \rangle < 0$ for all $i \neq j$, then $m \leq n + 1$. The case $V = \mathbb{R}^n$ is the claim.

If $n = 0$, then $V = \{\mathbf{0}\}$ and $\langle \mathbf{0}, \mathbf{0} \rangle = 0$, so there is no pair $i \neq j$, and $m \leq 1$.

Let $n \geq 1$, and assume the statement for $(n-1)$-dimensional subspaces. If $m \leq 2$ there is nothing to prove, so let $m \geq 3$. Every $\mathbf{v}_i$ is nonzero, since $\langle \mathbf{v}_i, \mathbf{v}_j \rangle \neq 0$ for some $j \neq i$. Let $U = \{\mathbf{x} \in V \mid \langle \mathbf{x}, \mathbf{v}_m \rangle = 0\}$, the orthogonal complement of $\operatorname{span}(\mathbf{v}_m)$ inside the inner product space V, so $\dim U = n - 1$ by Corollary 6.4.25. For $i < m$ let

$$c_i = \frac{\langle \mathbf{v}_i, \mathbf{v}_m \rangle}{\|\mathbf{v}_m\|^2} < 0, \quad \mathbf{w}_i = \mathbf{v}_i - c_i \mathbf{v}_m.$$

By Corollary 6.4.10 and Proposition 6.4.4(b), $\mathbf{w}_i = \mathbf{v}_i - \operatorname{Pr}_{\operatorname{span}(\mathbf{v}_m)}(\mathbf{v}_i) \in U$. For $i \neq j$, both less than m,

$$\begin{aligned} \langle \mathbf{w}_i, \mathbf{w}_j \rangle &= \langle \mathbf{v}_i, \mathbf{v}_j \rangle - c_j \langle \mathbf{v}_i, \mathbf{v}_m \rangle - c_i \langle \mathbf{v}_m, \mathbf{v}_j \rangle + c_i c_j \|\mathbf{v}_m\|^2 \\ &= \langle \mathbf{v}_i, \mathbf{v}_j \rangle - c_i c_j \|\mathbf{v}_m\|^2 - c_i c_j \|\mathbf{v}_m\|^2 + c_i c_j \|\mathbf{v}_m\|^2 \\ &= \langle \mathbf{v}_i, \mathbf{v}_j \rangle - c_i c_j \|\mathbf{v}_m\|^2 < 0, \end{aligned}$$

because $\langle \mathbf{v}_i, \mathbf{v}_j \rangle < 0$ and $c_i c_j > 0$. So $\mathbf{w}_1, \ldots, \mathbf{w}_{m-1}$ are vectors in the $(n-1)$-dimensional space U with pairwise negative inner products, and the induction hypothesis gives $m - 1 \leq n$, that is, $m \leq n + 1$.

For an example with $m = n + 1$, let $\mathbf{j} = (1, \ldots, 1) \in \mathbb{R}^n$, $\mathbf{v}_i = \mathbf{e}_i - (1/2n)\mathbf{j}$ for $i = 1, \ldots, n$, and $\mathbf{v}_{n+1} = -\mathbf{j}$. For $i \neq k \leq n$,

$$\langle \mathbf{v}_i, \mathbf{v}_k \rangle = 0 - \frac{1}{2n} - \frac{1}{2n} + \frac{n}{4n^2} = -\frac{3}{4n} < 0,$$

$$\langle \mathbf{v}_i, \mathbf{v}_{n+1} \rangle = -1 + \frac{n}{2n} = -\frac{1}{2} < 0.$$

The proof is an induction on dimension in which projecting onto a hyperplane preserves the hypothesis. The triangularization theorems of Section 7.5 are proved by the same pattern (Technique 7.5.12).

Solution to Exercise 6.4.63

Uses. Definition 1.2.13 (Polynomials), Definition 1.2.17 (Evaluation), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 6.3.9 (Properties of $A^{\perp}$), Theorem 6.4.24 (Orthogonal decomposition)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. The sum defining $\langle p, q \rangle$ is finite. W is the set of polynomials killed by evaluation at 1, and $1 \notin W$.
Hint 2. For $q \in W^{\perp}$, test against $x^k - x^{k+1} \in W$. What does $\langle q, x^k - x^{k+1} \rangle = 0$ say about the coefficients of q?
Hint 3. For (c), try $w_n = 1 - (1 + x + \cdots + x^n)/(n+1)$.
Sketch. (a) W is closed under the operations since evaluation respects them, and $1 \notin W$. If $q \perp W$, then $0 = \langle q, x^k - x^{k+1} \rangle = q_k - q_{k+1}$ for all k, so all coefficients of q are equal, and since almost all are 0, $q = 0$. (b) $W + W^{\perp} = W$ and $(W^{\perp})^{\perp} = V$. (c) $w_n \in W$ and $\|1 - w_n\|^2 = 1/(n+1) \to 0$, but $\|1 - w\| > 0$ for every $w \in W$ since $1 \notin W$.

(a) For each p only finitely many p_k are nonzero, so the sum is finite. It is symmetric and linear in each slot, and $\langle p, p \rangle = \sum_k p_k^2 > 0$ for $p \neq 0$, so it is an inner product (it is the dot product of coefficient sequences). By Proposition 1.2.18, $(ap + bq)(1) = ap(1) + bq(1)$, so W is a subspace, and $1 \notin W$, so $W \neq \mathbb{R}[x]$.

Let $q \in W^{\perp}$. For each $k \geq 0$, the polynomial $x^k - x^{k+1}$ vanishes at 1, so it lies in W, and

$$0 = \langle q, x^k - x^{k+1} \rangle = q_k - q_{k+1}.$$

So $q_0 = q_1 = q_2 = \cdots$. Since $q_k = 0$ for all large k, every coefficient is 0, and $q = 0$. Hence $W^{\perp} = \{\mathbf{0}\}$.

(b) $W + W^{\perp} = W \neq \mathbb{R}[x]$, so $\mathbb{R}[x]$ is not $W \oplus W^{\perp}$. And $(W^{\perp})^{\perp} = \{\mathbf{0}\}^{\perp} = \mathbb{R}[x] \neq W$ (Proposition 6.3.9(d)). Both conclusions of Theorem 6.4.24 and Theorem 6.4.26 fail. The hypothesis that breaks is finite-dimensionality: W contains the independent list $1 - x, x - x^2, x^2 - x^3, \ldots$, of every length, so it has no (finite) orthonormal basis, and Pr_W cannot be defined.

(c) Let $w_n = 1 - (1 + x + \cdots + x^n)/(n+1)$. Then $w_n(1) = 1 - 1 = 0$, so $w_n \in W$, and

$$\|1 - w_n\|^2 = \left\| \frac{1 + x + \cdots + x^n}{n+1} \right\|^2 = \frac{n+1}{(n+1)^2} = \frac{1}{n+1}.$$

So $\inf_{w \in W} \|1 - w\| = 0$. But $1 \notin W$, so $\|1 - w\| > 0$ for every $w \in W$. If some $w \in W$ were closest to 1, then $\|1 - w\| \le \|1 - w_n\| = (n+1)^{-1/2}$ for every n, forcing $\|1 - w\| = 0$, which is impossible. In the spaces of Section 15.3, which are *complete*, closest points do exist for closed subspaces, and this kind of example disappears.

Solution to Exercise 6.4.64

Uses. Theorem 6.4.38 (QR factorization), Theorem 6.4.31 (The matrix of an orthogonal projection), Algorithm 6.4.14 (Gram–Schmidt)
Hint 1. Gram–Schmidt on the columns: $\|\mathbf{a}_1\| = 3$, and $\langle \mathbf{a}_2, \mathbf{u}_1 \rangle$ is an integer.
Hint 2. $\mathbf{u}_1 = (1,2,2)/3$ and $\mathbf{u}_2 = (2,1,-2)/3$. For (b), $\mathrm{Pr}_{\mathrm{im}\,A} = QQ^{\mathsf{T}}$.
Sketch. $\mathbf{u}_1 = (1,2,2)/3$, $r_{11} = 3$; $r_{12} = \langle \mathbf{a}_2, \mathbf{u}_1 \rangle = 3$, $\mathbf{f}_2 = (3,3,0) - (1,2,2) = (2,1,-2)$, $r_{22} = 3$. So $R = \left(\begin{smallmatrix} 3 & 3 \\ 0 & 3 \end{smallmatrix}\right)$. Then $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{e}_1) = (1/9)(1,2,2) + (2/9)(2,1,-2) = (5,4,-2)/9$, and $A^{\mathsf{T}}A = \left(\begin{smallmatrix} 9 & 9 \\ 9 & 18 \end{smallmatrix}\right) = R^{\mathsf{T}}R$.

(a) Let $\mathbf{a}_1 = (1,2,2)$ and $\mathbf{a}_2 = (3,3,0)$. Then $\|\mathbf{a}_1\| = 3$, so $\mathbf{u}_1 = (1,2,2)/3$ and $r_{11} = 3$. Next $r_{12} = \langle \mathbf{a}_2, \mathbf{u}_1 \rangle = (3+6+0)/3 = 3$, and $\mathbf{f}_2 = \mathbf{a}_2 - 3\mathbf{u}_1 = (2,1,-2)$, with $r_{22} = \|\mathbf{f}_2\| = 3$ and $\mathbf{u}_2 = (2,1,-2)/3$. By Theorem 6.4.38,

$$A = QR, \quad Q = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}, \quad R = \begin{pmatrix} 3 & 3 \\ 0 & 3 \end{pmatrix}.$$

(b) The columns of Q are an orthonormal basis of $\mathrm{im}\,A$, so by Theorem 6.4.31,

$$\begin{aligned} \mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{e}_1) &= \langle \mathbf{e}_1, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{e}_1, \mathbf{u}_2 \rangle \mathbf{u}_2 \\ &= \frac{1}{9}(1,2,2) + \frac{2}{9}(2,1,-2) = \frac{1}{9}(5,4,-2). \end{aligned}$$

(c) $A^{\mathsf{T}}A = \left(\begin{smallmatrix} 1+4+4 & 3+6+0 \\ 3+6+0 & 9+9+0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 9 & 9 \\ 9 & 18 \end{smallmatrix}\right)$ and $R^{\mathsf{T}}R = \left(\begin{smallmatrix} 3 & 0 \\ 3 & 3 \end{smallmatrix}\right)\left(\begin{smallmatrix} 3 & 3 \\ 0 & 3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 9 & 9 \\ 9 & 18 \end{smallmatrix}\right)$. In general $A^*A = R^*Q^*QR = R^*R$, which is Exercise 6.4.59(a) for the columns of A.

Section 6.5

Solution to Exercise 6.5.8

Uses. Theorem 6.5.5 (The normal equations), Definition 6.1.8 (Dot product), Definition 6.3.26 (Conjugate transpose)
Hint 1. Here $A^*A = \mathbf{a}^*\mathbf{a} = \|\mathbf{a}\|^2$ and $A^*\mathbf{b} = \mathbf{a}^*\mathbf{b} = \langle \mathbf{b}, \mathbf{a} \rangle$, with the entries of $\mathbf{a}$ conjugated.
Hint 2. You should get $\hat{z} = (1-i)/3$. For (b), $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + i^2 + 1 = 1$.
Sketch. (a) $\mathbf{a}^*\mathbf{a} = 3$ and $\mathbf{a}^*\mathbf{b} = 1 - i$, so $\hat{z} = (1-i)/3$, and the residual $((2+i)/3, (2-i)/3, (-1+i)/3)$ has inner product 0 with $\mathbf{a}$. (b) $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1$ and $\mathbf{a}^{\mathsf{T}}\mathbf{b} = 1 + i$, so $z = 1 + i$, whose residual $(-i, 2-i, -1-i)$ has inner product $-2 - 4i \neq 0$ with $\mathbf{a}$; by Theorem 6.5.5 it is not a least-squares solution.

(a) With one column, the normal equations read $(\mathbf{a}^*\mathbf{a})z = \mathbf{a}^*\mathbf{b}$. Here $\mathbf{a}^* = (1, -i, 1)$, so

$$\mathbf{a}^*\mathbf{a} = 1 + (-i)(i) + 1 = 3, \quad \mathbf{a}^*\mathbf{b} = 1 + (-i) \cdot 1 + 0 = 1 - i,$$

and $\hat{z} = (1-i)/3$ by Theorem 6.5.5. The residual is

$$\begin{aligned} \mathbf{b} - \hat{z}\mathbf{a} &= \left(1 - \frac{1-i}{3}, 1 - \frac{i+1}{3}, -\frac{1-i}{3}\right) \\ &= \left(\frac{2+i}{3}, \frac{2-i}{3}, \frac{-1+i}{3}\right), \end{aligned}$$

using $i(1-i) = i + 1$. Its inner product with $\mathbf{a}$ is

$$\frac{1}{3}\big((2+i)\cdot 1+(2-i)\cdot \bar{i}+(-1+i)\cdot 1\big)$$
$$=\frac{1}{3}((2+i)+(-1-2i)+(-1+i))=0.$$

(b) $\mathbf{a}^{\mathsf{T}}\mathbf{a}=1+i^2+1=1$ and $\mathbf{a}^{\mathsf{T}}\mathbf{b}=1+i$, so the student gets $z=1+i$. Its residual is $\mathbf{b}-(1+i)\mathbf{a}=(1-(1+i),1-(i-1),-(1+i))=(-i,2-i,-1-i)$, and its inner product with $\mathbf{a}$ is $-i+(2-i)(-i)+(-1-i)=-i-2i-1-1-i=-2-4i\neq 0$. By Theorem 6.5.5, $1+i$ is not a least-squares solution. Over $\mathbb{C}$ the transpose does not compute inner products; the conjugate transpose does.

Solution to Exercise 6.5.15

Uses. Proposition 6.5.13 (Projection onto a column space), Corollary 6.5.11 (The least-squares solution for independent columns), Proposition 6.5.2 (Least squares is projection)
Hint 1. With one column, $A^{\mathsf{T}}A=\mathbf{a}^{\mathsf{T}}\mathbf{a}=9$.
Hint 2. $P=\mathbf{a}\mathbf{a}^{\mathsf{T}}/9$, and $\hat{x}=\mathbf{a}^{\mathsf{T}}\mathbf{b}/9$.
Sketch. $P=(1/9)\begin{pmatrix}1&2&2\\2&4&4\\2&4&4\end{pmatrix}$, $\hat{x}=9/9=1$, $P\mathbf{b}=(1,2,2)$, and the residual $(2,-2,1)$ has length 3.

Here $A=\mathbf{a}$ is a single nonzero column, so $A^{\mathsf{T}}A=\mathbf{a}\cdot\mathbf{a}=9$ and, by Proposition 6.5.13,

$$P=\frac{1}{9}\mathbf{a}\mathbf{a}^{\mathsf{T}}=\frac{1}{9}\begin{pmatrix}1&2&2\\2&4&4\\2&4&4\end{pmatrix}.$$

By Corollary 6.5.11, $\hat{x}=\left(\mathbf{a}^{\mathsf{T}}\mathbf{a}\right)^{-1}\mathbf{a}^{\mathsf{T}}\mathbf{b}=(3+0+6)/9=1$, and indeed $P\mathbf{b}=\hat{x}\mathbf{a}=(1,2,2)$. The residual is $\mathbf{b}-\mathbf{a}=(2,-2,1)$, which is orthogonal to $\mathbf{a}$, and the least-squares error is $\|(2,-2,1)\|=3$.

Solution to Exercise 6.5.20

Uses. Corollary 6.5.11 (The least-squares solution for independent columns), Definition 6.5.1 (Least-squares solution)
Hint 1. The design matrix has the single column $\mathbf{x}$. What are $\mathbf{x}^{\mathsf{T}}\mathbf{x}$ and $\mathbf{x}^{\mathsf{T}}\mathbf{y}$?
Hint 2. For (b), $\sum_i I_i^2=7.5$.
Sketch. (a) $A=\mathbf{x}$ has an independent column since $\mathbf{x}\neq\mathbf{0}$, so $\hat{c}=\left(\mathbf{x}^{\mathsf{T}}\mathbf{x}\right)^{-1}\mathbf{x}^{\mathsf{T}}\mathbf{y}=(\sum x_iy_i)/\left(\sum x_i^2\right)$. (b) $\sum I_iV_i=0.5+2.1+4.35+8.2=15.15$ and $\sum I_i^2=7.5$, so $R\approx 2.02$ ohms.

(a) The sum $\sum_i (y_i-cx_i)^2$ is $\|\mathbf{y}-c\mathbf{x}\|^2$, so we want the least-squares solution of $Ac=\mathbf{y}$ for the $m\times 1$ matrix A with column $\mathbf{x}$. Since $\mathbf{x}\neq\mathbf{0}$, this column is independent, and by Corollary 6.5.11

$$\hat{c}=\left(\mathbf{x}^{\mathsf{T}}\mathbf{x}\right)^{-1}\mathbf{x}^{\mathsf{T}}\mathbf{y}=\frac{\sum_i x_iy_i}{\sum_i x_i^2}.$$

(b) Here $\sum_i I_iV_i=0.5\cdot 1.0+1.0\cdot 2.1+1.5\cdot 2.9+2.0\cdot 4.1=0.5+2.1+4.35+8.2=15.15$ and $\sum_i I_i^2=0.25+1+2.25+4=7.5$. So $R\approx 15.15/7.5=2.02$ ohms. The model has no intercept, because a resistor carries no voltage at zero current; fitting $V=c_0+c_1I$ instead would let the data choose a nonzero intercept that the physics rules out.

Solution to Exercise 6.5.30

Uses. Theorem 6.5.5 (The normal equations), Proposition 6.5.9 ($\operatorname{rank}(A^*A)=\operatorname{rank}A$)
Hint 1. Form $A^{\mathsf{T}}A$ and $A^{\mathsf{T}}\mathbf{b}$ for $A=\begin{pmatrix}1&1\\1&-1\\2&1\end{pmatrix}$ and $\mathbf{b}=(2,0,5)$.
Hint 2. $A^{\mathsf{T}}A=\begin{pmatrix}6&2\\2&3\end{pmatrix}$ and $A^{\mathsf{T}}\mathbf{b}=(12,7)$.
Sketch. The normal equations $6x_1+2x_2=12$, $2x_1+3x_2=7$ give $\hat{\mathbf{x}}=(11/7,9/7)$. The residual is $(-6,-2,4)/7$, orthogonal to $(1,1,2)$ and $(1,-1,1)$, and the error is $\sqrt{56}/7=\sqrt{8/7}$.

Let $A=\begin{pmatrix}1&1\\1&-1\\2&1\end{pmatrix}$ and $\mathbf{b}=(2,0,5)$. Then

$$A^{\mathsf{T}}A = \begin{pmatrix} 1+1+4 & 1-1+2 \\ 1-1+2 & 1+1+1 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix},$$
$$A^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 2+0+10 \\ 2-0+5 \end{pmatrix} = \begin{pmatrix} 12 \\ 7 \end{pmatrix}.$$

The columns of A are independent, so by Theorem 6.5.5 and Proposition 6.5.9 the unique least-squares solution solves $6x_1 + 2x_2 = 12$, $2x_1 + 3x_2 = 7$. Subtracting 3 times the second equation from the first gives $-7x_2 = -9$, so $\hat{x}_2 = 9/7$ and $\hat{x}_1 = (7 - 27/7)/2 = 11/7$. Then $A\hat{\mathbf{x}} = (20/7, 2/7, 31/7)$, and the residual is

$$\mathbf{b} - A\hat{\mathbf{x}} = \frac{1}{7}(14 - 20, 0 - 2, 35 - 31) = \frac{1}{7}(-6, -2, 4).$$

Its dot products with the columns are $(-6 - 2 + 8)/7 = 0$ and $(-6 + 2 + 4)/7 = 0$. The least-squares error is $\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{36 + 4 + 16}/7 = \sqrt{56}/7 = \sqrt{8/7}$.

Solution to Exercise 6.5.31

Uses. Theorem 6.5.5 (The normal equations), Definition 6.3.26 (Conjugate transpose), Theorem 6.3.29 (Orthonormal columns)
Hint 1. The entries of A^*A are $\mathbf{a}_i^*\mathbf{a}_j$; remember to conjugate the entries of $\mathbf{a}_i$.
Hint 2. $A^*A = \left(\begin{smallmatrix} 3 & 1 \\ 1 & 3 \end{smallmatrix}\right)$ and $A^*\mathbf{b} = (3, 1 - 2i)$.
Sketch. The normal equations $3x_1 + x_2 = 3$, $x_1 + 3x_2 = 1 - 2i$ give $\hat{\mathbf{x}} = (1 + i/4, -3i/4)$. The residual is $((1 - i)/4, (-1 - i)/4, i/2)$, orthogonal to both columns, with squared norm $1/2$.

With $\mathbf{a}_1 = (1, 1, 1)$ and $\mathbf{a}_2 = (i, -i, 1)$,

$$\mathbf{a}_1^*\mathbf{a}_1 = 3, \quad \mathbf{a}_1^*\mathbf{a}_2 = i - i + 1 = 1,$$
$$\mathbf{a}_2^*\mathbf{a}_2 = 1 + 1 + 1 = 3,$$

and $\mathbf{a}_2^*\mathbf{a}_1 = \overline{\mathbf{a}_1^*\mathbf{a}_2} = 1$. So $A^*A = \left(\begin{smallmatrix} 3 & 1 \\ 1 & 3 \end{smallmatrix}\right)$ (Theorem 6.3.29). Also $A^*\mathbf{b} = \left(2 + 0 + 1, \bar{i} \cdot 2 + \overline{-i} \cdot 0 + 1\right) = (3, 1 - 2i)$. The normal equations are

$$3x_1 + x_2 = 3, \quad x_1 + 3x_2 = 1 - 2i.$$

Subtracting 3 times the second from the first gives $-8x_2 = 6i$, so $\hat{x}_2 = -3i/4$, and then $\hat{x}_1 = 1 - 3\hat{x}_2 + (-2i) = 1 + 9i/4 - 2i = 1 + i/4$. By Theorem 6.5.5 this is the least-squares solution, and it is unique since A^*A is invertible. Now

$$A\hat{\mathbf{x}} = (\hat{x}_1 + i\hat{x}_2, \hat{x}_1 - i\hat{x}_2, \hat{x}_1 + \hat{x}_2)$$
$$= \left(\frac{7}{4} + \frac{i}{4}, \frac{1}{4} + \frac{i}{4}, 1 - \frac{i}{2}\right),$$

so the residual is $\mathbf{b} - A\hat{\mathbf{x}} = ((1 - i)/4, (-1 - i)/4, i/2)$. Its inner product with $\mathbf{a}_1$ is $(1 - i - 1 - i)/4 + i/2 = 0$, and with $\mathbf{a}_2$ it is $((1 - i)/4)\bar{i} + ((-1 - i)/4)\overline{-i} + i/2 = (-i - 1)/4 + (-i + 1)/4 + i/2 = 0$. The squared error is $2/16 + 2/16 + 1/4 = 1/2$.

Solution to Exercise 6.5.32

Uses. Proposition 6.5.17 (The least-squares line)
Hint 1. Center the data: $m_x = 3$ and $m_y = 16.6$.
Hint 2. $\tilde{\mathbf{x}} = (-2, -1, 0, 1, 2)$ and $\tilde{\mathbf{y}} = (4.4, 2.4, -0.6, -1.6, -4.6)$.
Sketch. $\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = -22$ and $\|\tilde{\mathbf{x}}\|^2 = 10$, so $c_1 = -2.2$ and $c_0 = 16.6 + 6.6 = 23.2$. Residuals $0, 0.2, -0.6, 0.6, -0.2$. Prediction at 6: 10.0. At 11: -1.0, a negative price.

Here $m_x = 3$ and $m_y = 83/5 = 16.6$, so $\tilde{\mathbf{x}} = (-2, -1, 0, 1, 2)$ and $\tilde{\mathbf{y}} = (4.4, 2.4, -0.6, -1.6, -4.6)$. Then

$$\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = -8.8 - 2.4 + 0 - 1.6 - 9.2 = -22, \quad \|\tilde{\mathbf{x}}\|^2 = 10,$$

so by Proposition 6.5.17 $c_1 = -2.2$ and $c_0 = 16.6 + 2.2 \cdot 3 = 23.2$. The line is $y = 23.2 - 2.2x$: the car loses about 2.2 thousand dollars a year. Its values at $x = 1, \dots, 5$ are $21.0, 18.8, 16.6, 14.4, 12.2$, so the residuals are $0, 0.2, -0.6, 0.6, -0.2$, which add up to 0. At age 6 it predicts $23.2 - 13.2 = 10.0$. At age 11 it predicts $23.2 - 24.2 = -1.0$, a negative price. The fit describes the data on the range where it was measured; nothing in least squares says that the linear model stays valid outside that range, and here it clearly does not.

Solution to Exercise 6.5.33

Uses. Proposition 6.5.17 (The least-squares line), Proposition 6.5.21 (Polynomial fits), Theorem 6.5.5 (The normal equations), Proposition 6.5.2 (Least squares is projection), Theorem 6.4.33 (Best approximation)

Hint 1. For (b), $A^{\mathsf{T}}A$ has entries $\sum_i x_i^{j+k}$; here $A^{\mathsf{T}}A = \begin{pmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{pmatrix}$ and $A^{\mathsf{T}}\mathbf{y} = (8, 7, 19)$.

Hint 2. For (c), compare the column spaces of the two design matrices.

Sketch. (a) $y = 1.7 + 0.6x$ with residuals $0.9, -0.7, -1.3, 1.1$ and squared error 4.2. (b) $y = 0.7 - 0.4x + x^2$ with residuals $-0.1, 0.3, -0.3, 0.1$ and squared error 0.2. (c) $\operatorname{span}(\mathbf{1}, \mathbf{x}) \subseteq \operatorname{span}(\mathbf{1}, \mathbf{x}, \mathbf{x}^2)$, and the distance from $\mathbf{y}$ to a larger subspace is no larger; equivalently, every line is a parabola with $c_2 = 0$.

(a) For the line, $A^{\mathsf{T}}A = \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix}$ and $A^{\mathsf{T}}\mathbf{y} = (2 + 1 + 1 + 4, -2 + 0 + 1 + 8) = (8, 7)$. The normal equations $4c_0 + 2c_1 = 8$, $2c_0 + 6c_1 = 7$ give $c_1 = 3/5$ and $c_0 = 17/10$, so the line is $y = 1.7 + 0.6x$. Its values at $-1, 0, 1, 2$ are $1.1, 1.7, 2.3, 2.9$, the residuals are $0.9, -0.7, -1.3, 1.1$, and the squared error is $0.81 + 0.49 + 1.69 + 1.21 = 4.2$.

(b) For the parabola, row i of A is $\left(1, x_i, x_i^2\right)$, and with $\sum x_i = 2$, $\sum x_i^2 = 6$, $\sum x_i^3 = 8$, $\sum x_i^4 = 18$, $\sum y_i = 8$, $\sum x_i y_i = 7$, $\sum x_i^2 y_i = 19$, the normal equations are

$$\begin{aligned} 4c_0 + 2c_1 + 6c_2 &= 8, \\ 2c_0 + 6c_1 + 8c_2 &= 7, \\ 6c_0 + 8c_1 + 18c_2 &= 19. \end{aligned}$$

Row reduction gives $(c_0, c_1, c_2) = (7/10, -2/5, 1)$; check: $2.8 - 0.8 + 6 = 8$, $1.4 - 2.4 + 8 = 7$, $4.2 - 3.2 + 18 = 19$. The four x_i are distinct, so this is the unique least-squares parabola (Proposition 6.5.21). Its values are $2.1, 0.7, 1.3, 3.9$, the residuals are $-0.1, 0.3, -0.3, 0.1$, and the squared error is 0.2.

(c) Let A_1 and A_2 be the two design matrices. The columns $\mathbf{1}, \mathbf{x}$ of A_1 are also columns of A_2, so $\operatorname{im} A_1 \subseteq \operatorname{im} A_2$. By Proposition 6.5.2(d) the two errors are $\operatorname{dist}(\mathbf{y}, \operatorname{im} A_1)$ and $\operatorname{dist}(\mathbf{y}, \operatorname{im} A_2)$, and the distance from $\mathbf{y}$ to a larger set is no larger, since Theorem 6.4.33 minimizes over all its points. Concretely, the least-squares line is the parabola with $c_2 = 0$, so the best parabola does at least as well. (Exercise 6.5.39 says exactly when the error drops.)

Solution to Exercise 6.5.34

Uses. Theorem 6.5.5 (The normal equations), Lemma 6.5.4 (The orthogonal complement of a column space)

Hint 1. By Theorem 6.5.5, $(1, 1)$ is a least-squares solution exactly when $\mathbf{b} - A(1, 1) \in (\operatorname{im} A)^{\perp}$.

Hint 2. $A(1, 1) = (1, 2, 3)$, and $(\operatorname{im} A)^{\perp} = \ker A^{\mathsf{T}}$ (Lemma 6.5.4).

Sketch. $(\operatorname{im} A)^{\perp} = \ker A^{\mathsf{T}} = \operatorname{span}((1, -2, 1))$, so the answer is the line $\mathbf{b} = (1, 2, 3) + t(1, -2, 1)$, $t \in \mathbb{R}$, through $A(1, 1)$ perpendicular to the plane $\operatorname{im} A$. The data sets are $(0, 1 + t), (1, 2 - 2t), (2, 3 + t)$.

By Theorem 6.5.5, $(1, 1)$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\mathbf{b} - A(1, 1) \in (\operatorname{im} A)^{\perp}$. Here $A(1, 1) = (1, 2, 3)$, and by Lemma 6.5.4, $(\operatorname{im} A)^{\perp} = \ker A^{\mathsf{T}}$, the solutions of $z_1 + z_2 + z_3 = 0$ and $z_2 + 2z_3 = 0$. These give $z_2 = -2z_3$ and $z_1 = z_3$, so $(\operatorname{im} A)^{\perp} = \operatorname{span}((1, -2, 1))$. Hence the answer is

$$\left\{\mathbf{b} \in \mathbb{R}^3 \;\middle|\; \mathbf{b} = (1, 2, 3) + t(1, -2, 1) \text{ for some } t \in \mathbb{R}\right\},$$

the line through the point $A(1, 1)$ of the plane $\operatorname{im} A$ perpendicular to that plane. The columns of A are independent, so for each such $\mathbf{b}$ the vector $(1, 1)$ is the only least-squares solution, and the least-squares error is $|t|\sqrt{6}$, which is 0 only for $\mathbf{b} = (1, 2, 3)$. In terms of data, the least-squares line of $(0, y_1), (1, y_2), (2, y_3)$ is $y = 1 + x$ exactly when $(y_1, y_2, y_3) = (1 + t, 2 - 2t, 3 + t)$ for some t: the points are pushed off the line alternately up, down, up, in the pattern $(1, -2, 1)$.

Solution to Exercise 6.5.35

Uses. Theorem 6.5.5 (The normal equations), Definition 6.3.33 (Unitary and orthogonal matrices), Proposition 6.3.27 (Properties of the conjugate transpose), Corollary 6.5.11 (The least-squares solution for independent columns)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), (b), and (d), compare normal equations.
Hint 2. For (c), multiplying one equation by 2 makes its squared error count four times as much. Try a single column $A = (1, 1)$.
Sketch. (a) True: the normal equations are linear in $\mathbf{b}$. (b) True: $(UA)^*(UA) = A^*A$ and $(UA)^*U\mathbf{b} = A^*\mathbf{b}$. (c) False: $A = (1, 1)$, $\mathbf{b} = (0, 2)$, $G = \operatorname{diag}(1, 2)$ give $\hat{x} = 1$ and $\hat{x} = 8/5$. (d) True: $\mathbf{y} \mapsto AS\mathbf{y}$ has the same values as $\mathbf{x} \mapsto A\mathbf{x}$, reached at $\mathbf{x} = S\mathbf{y}$.

(a) True. By Theorem 6.5.5, $A^*A\hat{\mathbf{x}} = A^*\mathbf{b}$ and $A^*A\hat{\mathbf{x}}' = A^*\mathbf{b}'$. Adding, $A^*A(\hat{\mathbf{x}} + \hat{\mathbf{x}}') = A^*(\mathbf{b} + \mathbf{b}')$, so $\hat{\mathbf{x}} + \hat{\mathbf{x}}'$ is a least-squares solution of $A\mathbf{x} = \mathbf{b} + \mathbf{b}'$.

(b) True. Since $U^*U = I_m$ (Definition 6.3.33) and $(UA)^* = A^*U^*$ (Proposition 6.3.27(b)),

$$(UA)^*(UA) = A^*U^*UA = A^*A,$$
$$(UA)^*(U\mathbf{b}) = A^*U^*U\mathbf{b} = A^*\mathbf{b}.$$

So the two problems have the same normal equations, and by Theorem 6.5.5 the same least-squares solutions. (Geometrically, U preserves lengths, so $\|U\mathbf{b} - UA\mathbf{x}\| = \|\mathbf{b} - A\mathbf{x}\|$.)

(c) False. Let $A = \begin{pmatrix}1\\1\end{pmatrix}$, $\mathbf{b} = (0, 2)$, and $G = \operatorname{diag}(1, 2)$. By Corollary 6.5.11, the least-squares solution of $Ax = \mathbf{b}$ is $\left(A^{\mathsf{T}}\mathbf{b}\right)/\left(A^{\mathsf{T}}A\right) = 2/2 = 1$. But $GA = \begin{pmatrix}1\\2\end{pmatrix}$ and $G\mathbf{b} = (0, 4)$, and the least-squares solution of $GAx = G\mathbf{b}$ is $8/5$. Multiplying an equation by 2 multiplies its squared error by 4, which changes the balance between the equations; this is weighted least squares (Aside (Weighted least squares)).

(d) True. As $\mathbf{y}$ runs over $\mathbb{F}^n$, so does $\mathbf{x} = S\mathbf{y}$, since S is invertible, and $(AS)\mathbf{y} = A(S\mathbf{y})$. So $\|\mathbf{b} - AS\mathbf{y}\|$ takes the same values as $\|\mathbf{b} - A\mathbf{x}\|$, and $\mathbf{y}$ minimizes the first if and only if $S\mathbf{y}$ minimizes the second. With $\mathbf{y} = S^{-1}\hat{\mathbf{x}}$, this is the claim.

Solution to Exercise 6.5.36

Uses. Proposition 2.5.17 (Algebra of inverses), Proposition 6.5.13 (Projection onto a column space), Corollary 3.5.25 (The invertible matrix theorem)
Hint 1. The rule $(BC)^{-1} = C^{-1}B^{-1}$ requires B and C to be invertible. Are A and A^*?
Hint 2. Try $A = (1, 2, 2)$, a single column. When is $\operatorname{im} A = \mathbb{F}^m$?
Sketch. $(A^*A)^{-1} = A^{-1}(A^*)^{-1}$ needs A invertible, hence square; for $m > n$, A has no inverse. For $A = (1, 2, 2)$, the projection matrix is $\mathbf{a}\mathbf{a}^{\mathsf{T}}/9 \neq I_3$. The conclusion holds exactly when $m = n$, that is, when A is invertible, and then the computation is valid.

The step $(A^*A)^{-1} = A^{-1}(A^*)^{-1}$ applies Proposition 2.5.17 to the factors A^* and A, which requires both to be invertible. An invertible matrix is square, so for $m > n$ the matrix A^{-1} does not exist, even though A^*A is invertible. For instance, for the single column $\mathbf{a} = (1, 2, 2)$ the projection matrix is $\mathbf{a}\mathbf{a}^{\mathsf{T}}/9$ (Exercise 6.5.15), which has rank 1 and is not I_3. In general $\operatorname{Pr}_{\operatorname{im} A}$ has rank $\dim \operatorname{im} A = n$, so $\operatorname{Pr}_{\operatorname{im} A} = I_m$ forces $n = m$.

Conversely, if $m = n$, then A is a square matrix with independent columns, hence invertible (Corollary 3.5.25), and so is A^*. Then the student's computation is valid, $\operatorname{Pr}_{\operatorname{im} A} = I_m$, and indeed $\operatorname{im} A = \mathbb{F}^m$.

Solution to Exercise 6.5.37

Uses. Theorem 6.5.5 (The normal equations), Exercise 6.5.20
Hint 1. By Theorem 6.5.5(ii), the residual is orthogonal to every vector of $\operatorname{im} A$.
Hint 2. For (b), two points suffice, for example $(1, 1)$ and $(2, 1)$.
Sketch. (a) The residual $\mathbf{r}$ is orthogonal to $\mathbf{1} \in \operatorname{im} A$, and $\mathbf{r} \cdot \mathbf{1} = \sum_i r_i$. (b) For $(1, 1), (2, 1)$, $c = 3/5$ and the residuals $2/5, -1/5$ add up to $1/5$.

(a) Let $\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$. By Theorem 6.5.5, $\mathbf{r} \perp \operatorname{im} A$, and $\mathbf{1} \in \operatorname{im} A$, so $0 = \mathbf{r} \cdot \mathbf{1} = \sum_i r_i$. Hence $\sum_i (A\hat{\mathbf{x}})_i = \sum_i b_i$, and dividing by m, the two vectors have the same mean. For a polynomial fit, $\mathbf{1}$ is the first column of the design matrix.

(b) For the data $(1,1),(2,1)$, Exercise 6.5.20 gives $c=(1+2)/(1+4)=3/5$. The residuals are $1-3/5=2/5$ and $1-6/5=-1/5$, which add up to $1/5\neq 0$. Here $\operatorname{im} A=\operatorname{span}((1,2))$ does not contain $\mathbf{1}$.

Solution to Exercise 6.5.38

Uses. Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.5.9 ($\operatorname{rank}(A^*A)=\operatorname{rank} A$)
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. For (a), cancel $(A^*A)(A^*A)^{-1}$ in the middle of P^2. For P^*, note that A^*A is Hermitian.
Hint 2. For (b), $\operatorname{im} P\subseteq\operatorname{im} A$ because $P=A(\cdots)$; and $A=PA$ gives the reverse.
Hint 3. For (c), move the factor A from the front to the back with Proposition 2.4.28.
Sketch. (a) $P^2=A(A^*A)^{-1}(A^*A)(A^*A)^{-1}A^*=P$, and $P^*=A\big((A^*A)^{-1}\big)^*A^*=P$ since $(A^*A)^*=A^*A$. (b) $PA=A$; $\operatorname{im} P\subseteq\operatorname{im} A=\operatorname{im}(PA)\subseteq\operatorname{im} P$. (c) $\operatorname{tr}\big(A(A^*A)^{-1}A^*\big)=\operatorname{tr}\big((A^*A)^{-1}A^*A\big)=\operatorname{tr} I_n=n$. (d) $P=(1/10)\begin{pmatrix}7&4&1&-2\\4&3&2&1\\1&2&3&4\\-2&1&4&7\end{pmatrix}$.

The matrix A^*A is invertible by Proposition 6.5.9(c), and Hermitian, since $(A^*A)^*=A^*A^{**}=A^*A$ by Proposition 6.3.27(a) and (b).

(a) $P^2=A(A^*A)^{-1}(A^*A)(A^*A)^{-1}A^*=A(A^*A)^{-1}A^*=P$. By Proposition 6.3.27, $P^*=(A^*)^*\big((A^*A)^{-1}\big)^*A^*=A\big((A^*A)^*\big)^{-1}A^*=A(A^*A)^{-1}A^*=P$.

(b) $PA=A(A^*A)^{-1}(A^*A)=A$. Every $P\mathbf{y}=A\big((A^*A)^{-1}A^*\mathbf{y}\big)$ lies in $\operatorname{im} A$, and every $A\mathbf{x}=P(A\mathbf{x})$ lies in $\operatorname{im} P$. So $\operatorname{im} P=\operatorname{im} A$.

(c) By Proposition 2.4.28, $\operatorname{tr} P=\operatorname{tr}\big(A\cdot(A^*A)^{-1}A^*\big)=\operatorname{tr}\big((A^*A)^{-1}A^*\cdot A\big)=\operatorname{tr} I_n=n$.

(d) Here $A^{\mathsf T}A=\begin{pmatrix}4&10\\10&30\end{pmatrix}$, with determinant 20, so $\big(A^{\mathsf T}A\big)^{-1}=(1/20)\begin{pmatrix}30&-10\\-10&4\end{pmatrix}$. Row i of A is $(1,x_i)$, so

$$p_{ij}=\frac{1}{20}\big(30-10(x_i+x_j)+4x_ix_j\big),$$

so

$$P=\frac{1}{10}\begin{pmatrix}7&4&1&-2\\4&3&2&1\\1&2&3&4\\-2&1&4&7\end{pmatrix}.$$

Each row adds up to $10/10=1$, so $P\mathbf{1}=\mathbf{1}$; the rows times $(1,2,3,4)$ give $10,20,30,40$, divided by 10, so $P\mathbf{x}=\mathbf{x}$; and $\operatorname{tr} P=(7+3+3+7)/10=2$. The diagonal entries p_{ii} are largest at the ends: a data point far from m_x pulls harder on the fitted line.

Solution to Exercise 6.5.39

Uses. Theorem 6.5.5 (The normal equations), Proposition 6.5.2 (Least squares is projection), Proposition 6.1.29 (Expanding a norm), Lemma 6.5.4 (The orthogonal complement of a column space)
Techniques. Perturb a minimizer (Technique 6.5.6)
Hint 1. If $\hat{\mathbf{x}}$ is a least-squares solution for A, try $\mathbf{x}'=(\hat{\mathbf{x}},t)$ for a scalar t. What is its residual?
Hint 2. Its residual is $\mathbf{r}-t\mathbf{a}$. Expand $\|\mathbf{r}-t\mathbf{a}\|^2$ and choose $t=\langle\mathbf{r},\mathbf{a}\rangle/\|\mathbf{a}\|^2$.
Hint 3. If $\langle\mathbf{r},\mathbf{a}\rangle=0$, check that $(\hat{\mathbf{x}},0)$ satisfies condition (ii) of Theorem 6.5.5 for A'.
Sketch. For $\mathbf{x}'=(\hat{\mathbf{x}},t)$, the residual is $\mathbf{r}-t\mathbf{a}$, so $E'^2\le\|\mathbf{r}-t\mathbf{a}\|^2$ for every t; $t=0$ gives $E'\le E$. If $\langle\mathbf{r},\mathbf{a}\rangle\neq0$, then $t=\langle\mathbf{r},\mathbf{a}\rangle/\|\mathbf{a}\|^2$ gives $E'^2\le E^2-|\langle\mathbf{r},\mathbf{a}\rangle|^2/\|\mathbf{a}\|^2<E^2$. If $\langle\mathbf{r},\mathbf{a}\rangle=0$, then $\mathbf{r}$ is orthogonal to every column of A', so $(\hat{\mathbf{x}},0)$ is a least-squares solution for A' with residual $\mathbf{r}$, and $E'=E$.

Let $\hat{\mathbf{x}}$ be a least-squares solution of $A\mathbf{x}=\mathbf{b}$, so $\mathbf{r}=\mathbf{b}-A\hat{\mathbf{x}}$ and $E=\|\mathbf{r}\|$. (By Proposition 6.5.2(a), $\mathbf{r}=\mathbf{b}-\operatorname{Pr}_{\operatorname{im} A}(\mathbf{b})$ does not depend on the choice of $\hat{\mathbf{x}}$.) For $t\in\mathbb{F}$, the vector $\mathbf{x}'=(\hat{\mathbf{x}},t)$ has $A'\mathbf{x}'=A\hat{\mathbf{x}}+t\mathbf{a}$, so its residual is $\mathbf{r}-t\mathbf{a}$, and

$$E'^2\le\|\mathbf{r}-t\mathbf{a}\|^2\quad\text{for every } t\in\mathbb{F}.$$

Taking $t=0$ gives $E'\le E$.

Suppose $\langle \mathbf{r}, \mathbf{a}\rangle \neq 0$. Then $\mathbf{a} \neq \mathbf{0}$; let $t = \langle \mathbf{r}, \mathbf{a}\rangle / \|\mathbf{a}\|^2$. By Proposition 6.1.29,

$$\begin{aligned}\|\mathbf{r} - t\mathbf{a}\|^2 &= \|\mathbf{r}\|^2 - 2\,\mathrm{Re}(\bar{t}\langle \mathbf{r}, \mathbf{a}\rangle) + |t|^2\|\mathbf{a}\|^2 \\ &= \|\mathbf{r}\|^2 - 2\frac{|\langle \mathbf{r}, \mathbf{a}\rangle|^2}{\|\mathbf{a}\|^2} + \frac{|\langle \mathbf{r}, \mathbf{a}\rangle|^2}{\|\mathbf{a}\|^2} = E^2 - \frac{|\langle \mathbf{r}, \mathbf{a}\rangle|^2}{\|\mathbf{a}\|^2},\end{aligned}$$

which is less than E^2. So $E' < E$.

Suppose $\langle \mathbf{r}, \mathbf{a}\rangle = 0$. By Theorem 6.5.5(ii), $\mathbf{r}$ is orthogonal to the columns of A, and now also to $\mathbf{a}$, so to every column of A'. The vector $(\hat{\mathbf{x}}, 0)$ has residual $\mathbf{b} - A\hat{\mathbf{x}} = \mathbf{r}$ for A', so by Theorem 6.5.5 it is a least-squares solution of $A'\mathbf{x}' = \mathbf{b}$, and $E' = \|\mathbf{r}\| = E$.

In words: adding a column to the design matrix never makes the fit worse, and it makes it strictly better exactly when the new column is correlated with what the old model leaves unexplained. This is also why the error alone cannot decide between models: a model with more columns always wins.

Solution to Exercise 6.5.40

Uses. Proposition 6.5.17 (The least-squares line), Proposition 6.1.29 (Expanding a norm), Example 6.2.19 (Correlation is a cosine)
Hint 1. Write $\mathbf{y} = m_y\mathbf{1} + \tilde{\mathbf{y}}$ and $A\mathbf{c} = c_0\mathbf{1} + c_1\mathbf{x}$, and use $c_0 = m_y - c_1 m_x$.
Hint 2. Expand $\|\tilde{\mathbf{y}} - c_1\tilde{\mathbf{x}}\|^2$ and substitute $c_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{x}}\|^2$.
Sketch. (a) $\mathbf{y} - A\mathbf{c} = m_y\mathbf{1} + \tilde{\mathbf{y}} - \big(m_y - c_1 m_x\big)\mathbf{1} - c_1\mathbf{x} = \tilde{\mathbf{y}} - c_1\tilde{\mathbf{x}}$. (b) $\|\tilde{\mathbf{y}} - c_1\tilde{\mathbf{x}}\|^2 = \|\tilde{\mathbf{y}}\|^2 - (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})^2/\|\tilde{\mathbf{x}}\|^2 = \|\tilde{\mathbf{y}}\|^2\big(1 - r^2\big)$. (c) $\|\tilde{\mathbf{y}}\|^2 = 9.5$ and the squared error is 0.091, so $r^2 = 1 - 0.091/9.5 \approx 0.990$.

(a) By Proposition 6.5.17, $c_0 = m_y - c_1 m_x$, so

$$\begin{aligned}\mathbf{y} - A\mathbf{c} &= \big(m_y\mathbf{1} + \tilde{\mathbf{y}}\big) - \big(m_y - c_1 m_x\big)\mathbf{1} - c_1\mathbf{x} \\ &= \tilde{\mathbf{y}} - c_1(\mathbf{x} - m_x\mathbf{1}) = \tilde{\mathbf{y}} - c_1\tilde{\mathbf{x}}.\end{aligned}$$

(b) By Proposition 6.1.29 (all vectors are real) and $c_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{x}}\|^2$,

$$\begin{aligned}\|\tilde{\mathbf{y}} - c_1\tilde{\mathbf{x}}\|^2 &= \|\tilde{\mathbf{y}}\|^2 - 2c_1(\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}) + c_1^2\|\tilde{\mathbf{x}}\|^2 \\ &= \|\tilde{\mathbf{y}}\|^2 - \frac{(\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})^2}{\|\tilde{\mathbf{x}}\|^2} = \|\tilde{\mathbf{y}}\|^2\left(1 - \frac{(\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})^2}{\|\tilde{\mathbf{x}}\|^2\|\tilde{\mathbf{y}}\|^2}\right),\end{aligned}$$

and the last fraction is r^2. Since the left side is nonnegative, this reproves $r^2 \le 1$, with equality exactly when every residual is 0, that is, when the data lie on a line.

(c) For the spring data, $\|\tilde{\mathbf{y}}\|^2 = 1.9^2 + 1.1^2 + 0.2^2 + 0.8^2 + 2.0^2 = 9.5$, and the squared error is 0.091 (Example 6.5.18). So $1 - r^2 = 0.091/9.5$, and $r^2 \approx 0.990$: the line accounts for 99% of the variation in the lengths.

Solution to Exercise 6.5.41

Uses. Proposition 6.5.17 (The least-squares line), Example 6.2.19 (Correlation is a cosine), Theorem 6.2.1 (Cauchy–Schwarz inequality)
Hint 1. By Proposition 6.5.17 with the roles exchanged, $d_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{y}}\|^2$.
Hint 2. If $r \neq 0$, the second line is $y = m_y + (1/d_1)(x - m_x)$. When is $1/d_1 = c_1$?
Sketch. (a) $c_1 d_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})^2/\big(\|\tilde{\mathbf{x}}\|^2\|\tilde{\mathbf{y}}\|^2\big) = r^2$. (b) Both lines pass through the means. If $r = 0$, they are $y = m_y$ and $x = m_x$. Otherwise they agree iff $c_1 = 1/d_1$, iff $r^2 = 1$, iff the points are collinear (Example 6.2.19). (c) $c_1 = d_1 = 4/5$: $y = 3/10 + (4/5)x$ and $y = -3/8 + (5/4)x$.

(a) By Proposition 6.5.17, $c_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{x}}\|^2$, and, with the roles of x and y exchanged, $d_1 = (\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})/\|\tilde{\mathbf{y}}\|^2$. So

$$c_1 d_1 = \frac{(\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}})^2}{\|\tilde{\mathbf{x}}\|^2\|\tilde{\mathbf{y}}\|^2} = r^2.$$

(b) By Proposition 6.5.17, $m_y = c_0 + c_1 m_x$ and $m_x = d_0 + d_1 m_y$, so both lines pass through $\left(m_x, m_y\right)$. If $r = 0$, then $c_1 = d_1 = 0$, and the lines are $y = m_y$ and $x = m_x$, which are different. If $r \neq 0$, then $d_1 \neq 0$ and the second line is $y = m_y + (1/d_1)(x - m_x)$. Two lines through the same point coincide exactly when their slopes agree, that is, when $c_1 = 1/d_1$, which by (a) means $r^2 = 1$. By Example 6.2.19, $r = \pm 1$ exactly when all data points lie on one line.

(c) Here $m_x = m_y = 3/2$, $\tilde{\mathbf{x}} = (-3, -1, 1, 3)/2$, and $\tilde{\mathbf{y}} = (-3, 1, -1, 3)/2$. So $\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = (9 - 1 - 1 + 9)/4 = 4$ and $\|\tilde{\mathbf{x}}\|^2 = \|\tilde{\mathbf{y}}\|^2 = 5$. Hence $c_1 = d_1 = 4/5$ and $r^2 = 16/25$. The line of y on x is $y = 3/2 + (4/5)(x - 3/2) = 3/10 + (4/5)x$, and the line of x on y is $y = 3/2 + (5/4)(x - 3/2) = -3/8 + (5/4)x$.

Since $r^2 \leq 1$, (a) gives $|c_1 d_1| \leq 1$. So when $r \neq 0$, $|c_1| \leq 1/|d_1|$, and when $r = 0$ the line of x on y is vertical: either way, the line of y on x is never steeper than the line of x on y. Predictions of y from x are pulled toward the mean m_y, which is where the word "regression" comes from.

Solution to Exercise 6.5.42

Uses. Proposition 6.5.2 (Least squares is projection), Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 2.6.18 (The column space is spanned by the columns)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Let A and A' be the design matrices for the old and new data. Compare $\operatorname{im} A$ and $\operatorname{im} A'$.
Hint 2. $\operatorname{im} A = \operatorname{im} A'$. The fitted values of the new data are $\Pr_{\operatorname{im} A}(\gamma \mathbf{y} + \delta \mathbf{1})$; use linearity of the projection and $\Pr_{\operatorname{im} A}(\mathbf{1}) = \mathbf{1}$.
Sketch. $\mathbf{x}' = \alpha \mathbf{x} + \beta \mathbf{1}$ and $\mathbf{x} = (\mathbf{x}' - \beta \mathbf{1})/\alpha$, so $\operatorname{im} A' = \operatorname{span}(\mathbf{1}, \mathbf{x}') = \operatorname{span}(\mathbf{1}, \mathbf{x}) = \operatorname{im} A$. By linearity, $\Pr(\mathbf{y}') = \gamma \Pr(\mathbf{y}) + \delta \mathbf{1} = \gamma(c_0 \mathbf{1} + c_1 \mathbf{x}) + \delta \mathbf{1}$; substituting $\mathbf{x} = (\mathbf{x}' - \beta \mathbf{1})/\alpha$ writes this as $A' \mathbf{c}'$ with the coefficients of the claimed line.

Let A and A' have columns $\mathbf{1}, \mathbf{x}$ and $\mathbf{1}, \mathbf{x}'$. Since $\mathbf{x}' = \alpha \mathbf{x} + \beta \mathbf{1}$ and $\mathbf{x} = (\mathbf{x}' - \beta \mathbf{1})/\alpha$, each pair of columns lies in the span of the other, so $W = \operatorname{im} A = \operatorname{im} A'$ (Proposition 2.6.18). The x_i' are not all equal, so the new least-squares line is unique, and by Proposition 6.5.2(a) its fitted values are $\Pr_W(\mathbf{y}')$. By linearity of $\Pr_W$ (Proposition 6.4.29(a)), $\Pr_W(\mathbf{1}) = \mathbf{1}$, and $\Pr_W(\mathbf{y}) = c_0 \mathbf{1} + c_1 \mathbf{x}$,

$$\begin{aligned} \Pr_W(\mathbf{y}') &= \gamma \Pr_W(\mathbf{y}) + \delta \Pr_W(\mathbf{1}) = \gamma(c_0 \mathbf{1} + c_1 \mathbf{x}) + \delta \mathbf{1} \\ &= \left(\gamma c_0 + \delta - \frac{\gamma c_1 \beta}{\alpha}\right) \mathbf{1} + \frac{\gamma c_1}{\alpha} \mathbf{x}'. \end{aligned}$$

So the new line is $y' = c_0' + c_1' x'$ with $c_1' = \gamma c_1/\alpha$ and $c_0' = \gamma c_0 + \delta - \gamma c_1 \beta/\alpha$, which is $y' = \gamma(c_0 + c_1(x' - \beta)/\alpha) + \delta$, as claimed. Measuring in other units, or from another origin, therefore gives the same fitted line. That fails for a nonlinear change such as $x' = x^2$, which changes the column space.

Solution to Exercise 6.5.43

Uses. Proposition 6.5.23 (Normal equations for a spanning list), Theorem 6.4.33 (Best approximation), Proposition 6.4.4 (Properties of $\Pr_{\mathcal{O}}$)
Techniques. Gram matrices (Technique 6.3.31)
Hint 1. The Gram matrix of $1, t, t^2$ has entries $\langle t^j, t^i \rangle = 1/(i + j + 1)$, and $\langle \sqrt{t}, t^i \rangle = 1/(i + 3/2) = 2/(2i + 3)$.
Hint 2. Clear denominators: the equations become $6c_0 + 3c_1 + 2c_2 = 4$, $6c_0 + 4c_1 + 3c_2 = 24/5$, $20c_0 + 15c_1 + 12c_2 = 120/7$. For the error, $\|f - p\|^2 = \|f\|^2 - \langle p, f \rangle$ because $f - p \perp p$.
Sketch. $G = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}$, $\mathbf{d} = (2/3, 2/5, 2/7)$, solution $p = 6/35 + (48/35)t - (4/7)t^2$. Then $\|f - p\|^2 = 1/2 - 612/1225 = 1/2450$. The best line is $4/15 + (4/5)t$, with squared error $1/450$.

Let W be the polynomials of degree at most 2. By Proposition 6.5.23 with the list $1, t, t^2$, the coefficients of $p = \Pr_W(f) = c_0 + c_1 t + c_2 t^2$ solve $G\mathbf{c} = \mathbf{d}$, where $g_{ij} = \int_0^1 t^{i+j} \, \mathrm{d}t = 1/(i + j + 1)$ (indices from 0) and $d_i = \int_0^1 t^{i+1/2} \, \mathrm{d}t = 2/(2i + 3)$:

$$\begin{aligned} c_0 + \frac{c_1}{2} + \frac{c_2}{3} &= \frac{2}{3}, \\ \frac{c_0}{2} + \frac{c_1}{3} + \frac{c_2}{4} &= \frac{2}{5}, \end{aligned}$$

$$\frac{c_0}{3}+\frac{c_1}{4}+\frac{c_2}{5}=\frac{2}{7}.$$

Multiplying by 6, 12, and 60 gives $6c_0+3c_1+2c_2=4$, $6c_0+4c_1+3c_2=24/5$, and $20c_0+15c_1+12c_2=120/7$. The second minus the first is $c_1+c_2=4/5$. The third minus $10/3$ times the first is $5c_1+(16/3)c_2=120/7-40/3=80/21$, that is, $15c_1+16c_2=80/7$. Substituting $c_1=4/5-c_2$ gives $12+c_2=80/7$, so $c_2=-4/7$, $c_1=48/35$, and $c_0=(4-144/35+8/7)/6=6/35$. So

$$p(t)=\frac{6}{35}+\frac{48}{35}t-\frac{4}{7}t^2.$$

Since $f-p\perp W$ and $p\in W$ (Proposition 6.4.4), $\|f-p\|^2=\langle f-p,f\rangle=\|f\|^2-\langle p,f\rangle$, and $\langle p,f\rangle=\sum_i c_i d_i=(6/35)(2/3)+(48/35)(2/5)-(4/7)(2/7)=612/1225$. With $\|f\|^2=\int_0^1 t\,dt=1/2$,

$$\|f-p\|^2=\frac{1}{2}-\frac{612}{1225}=\frac{1}{2450}.$$

For the best line, drop c_2 and the third equation: $c_0+c_1/2=2/3$ and $c_0/2+c_1/3=2/5$ give $4/15+(4/5)t$, with squared error $1/2-(8/45+8/25)=1/450$. Allowing t^2 reduces the squared error by a factor of 5.4, and changes the constant and linear coefficients too, because $1,t,t^2$ are not orthogonal.

Solution to Exercise 6.5.44

Uses. Theorem 6.5.5 (The normal equations), Proposition 6.5.9 ($\operatorname{rank}(A^*A)=\operatorname{rank}A$), Remark 6.5.10 (The conjugate is essential)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Look for a nonzero column $\mathbf{a}\in\mathbb{F}_2^m$ with $\mathbf{a}^{\mathsf{T}}\mathbf{a}=0$.
Hint 2. $\mathbf{a}=(1,1)$ works. For (a), choose $\mathbf{b}$ with $\mathbf{a}^{\mathsf{T}}\mathbf{b}=1$.
Sketch. (a) False: $A=\begin{pmatrix}1\\1\end{pmatrix}$ and $\mathbf{b}=(1,0)$ give $0x=1$. (b) False: the same A has an independent column and $A^{\mathsf{T}}A=0$. (c) True: over $\mathbb{R}$, $A^{\mathsf{T}}=A^*$, so these are Theorem 6.5.5 and Proposition 6.5.9(c).

(a) False. Let $A=\begin{pmatrix}1\\1\end{pmatrix}\in\mathbb{F}_2^{2\times1}$ and $\mathbf{b}=(1,0)$. Then $A^{\mathsf{T}}A=1+1=0$ and $A^{\mathsf{T}}\mathbf{b}=1$, so the system reads $0x=1$, which has no solution.

(b) False. The same A has a single nonzero, hence independent, column, but $A^{\mathsf{T}}A=0$ is not invertible.

(c) True. Over $\mathbb{R}$, $A^{\mathsf{T}}=A^*$. The system $A^{\mathsf{T}}A\mathbf{x}=A^{\mathsf{T}}\mathbf{b}$ is the normal equations, which always have a solution by Theorem 6.5.5, and $A^{\mathsf{T}}A$ is invertible when the columns are independent by Proposition 6.5.9(c). Both proofs used $\mathbf{v}^*\mathbf{v}=\|\mathbf{v}\|^2>0$ for $\mathbf{v}\neq\mathbf{0}$, which fails over $\mathbb{F}_2$ for $\mathbf{v}=(1,1)$ (Remark 6.5.10).

Solution to Exercise 6.5.45

Uses. Lemma 6.5.4 (The orthogonal complement of a column space), Proposition 6.5.9 ($\operatorname{rank}(A^*A)=\operatorname{rank}A$), Proposition 6.3.27 (Properties of the conjugate transpose), Theorem 6.4.26 (The double complement)
Hint 1. Apply Lemma 6.5.4 to $M=A^*A$ and to $M=A^*$. What are $(A^*A)^*$ and $(A^*)^*$?
Hint 2. Both complements equal $\ker A$, by Proposition 6.5.9(a). Now take complements again with Theorem 6.4.26.
Sketch. $(\operatorname{im}A^*A)^\perp=\ker(A^*A)^*=\ker(A^*A)=\ker A$ and $(\operatorname{im}A^*)^\perp=\ker A^{**}=\ker A$. Taking complements, $\operatorname{im}(A^*A)=(\ker A)^\perp=\operatorname{im}A^*$ by Theorem 6.4.26. Then $A^*\mathbf{b}\in\operatorname{im}A^*=\operatorname{im}(A^*A)$.

By Proposition 6.3.27, $(A^*A)^*=A^*A$ and $(A^*)^*=A$. By Lemma 6.5.4, applied to A^*A and to A^*, and by Proposition 6.5.9(a),

$$(\operatorname{im}(A^*A))^\perp=\ker(A^*A)=\ker A,$$
$$(\operatorname{im}A^*)^\perp=\ker A.$$

Subspaces of $\mathbb{F}^n$ are finite-dimensional, so Theorem 6.4.26 gives

$$\operatorname{im}(A^*A)=\left((\operatorname{im}(A^*A))^\perp\right)^\perp=(\ker A)^\perp$$
$$=\left((\operatorname{im}A^*)^\perp\right)^\perp=\operatorname{im}A^*.$$

Hence $A^*\mathbf{b} \in \operatorname{im} A^* = \operatorname{im}(A^*A)$, that is, $A^*A\mathbf{x} = A^*\mathbf{b}$ for some $\mathbf{x}$.

Solution to Exercise 6.5.45

Uses. Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$), Theorem 4.2.7 (The image of T_A), Theorem 4.3.8 (Row rank equals column rank), Proposition 3.5.2 (First properties of rank), Proposition 4.1.21 (Dimensions of subspaces)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. One containment is clear: $A^*A\mathbf{x} = A^*(A\mathbf{x})$. Compare dimensions.
Hint 2. $\dim\operatorname{im}(A^*A) = \operatorname{rank} A$ by Proposition 6.5.9(b). For $\dim\operatorname{im} A^* = \operatorname{rank} A^*$, write $A^* = \overline{A}^{\mathsf{T}}$: transposing does not change the rank, and neither does conjugating the entries.
Sketch. $\operatorname{im}(A^*A) \subseteq \operatorname{im} A^*$. Both have dimension rank A: the first by Proposition 6.5.9(b), the second because $\operatorname{rank} A^* = \operatorname{rank} \overline{A}^{\mathsf{T}} = \operatorname{rank} \overline{A} = \operatorname{rank} A$. Equal dimensions and containment give equality.

Since $A^*A\mathbf{x} = A^*(A\mathbf{x})$, $\operatorname{im}(A^*A) \subseteq \operatorname{im} A^*$. By Theorem 4.2.7 and Proposition 6.5.9(b), $\dim\operatorname{im}(A^*A) = \operatorname{rank}(A^*A) = \operatorname{rank} A$. For A^*, write $A^* = \overline{A}^{\mathsf{T}}$. By Theorem 4.3.8, $\operatorname{rank} A^* = \operatorname{rank} \overline{A}$. Conjugating the coefficients of a relation shows that $\overline{\mathbf{a}}_j \in \operatorname{span}(\overline{\mathbf{a}}_1, \ldots, \overline{\mathbf{a}}_{j-1})$ if and only if $\mathbf{a}_j \in \operatorname{span}(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1})$, so by Proposition 3.5.2(d), $\operatorname{rank} \overline{A} = \operatorname{rank} A$. Hence $\dim\operatorname{im} A^* = \operatorname{rank} A = \dim\operatorname{im}(A^*A)$, and Proposition 4.1.21 gives $\operatorname{im}(A^*A) = \operatorname{im} A^*$. As before, $A^*\mathbf{b} \in \operatorname{im}(A^*A)$.

Solution to Exercise 6.5.46

Uses. Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 3.1.19 (Solution sets are translates of the kernel), Theorem 6.3.3 (Pythagorean theorem)
Hint 1. The columns of A^* are the conjugates of the rows of A. Apply Proposition 6.5.9(c) to A^*.
Hint 2. For $\mathbf{z} \in \ker A$, $\langle A^*\mathbf{y}, \mathbf{z}\rangle = \langle \mathbf{y}, A\mathbf{z}\rangle$ by Proposition 6.3.27(d).
Hint 3. Every solution is $\mathbf{x}_0 + \mathbf{z}$ with $\mathbf{z} \in \ker A$. Use Pythagoras.
Sketch. (a) The columns of A^* are independent, so $(A^*)^*A^* = AA^*$ is invertible. (b) $A\mathbf{x}_0 = AA^*(AA^*)^{-1}\mathbf{b} = \mathbf{b}$, and $\mathbf{x}_0 = A^*\mathbf{y}$ is orthogonal to $\ker A$ since $\langle A^*\mathbf{y}, \mathbf{z}\rangle = \langle \mathbf{y}, A\mathbf{z}\rangle = 0$. (c) Solutions are $\mathbf{x}_0 + \mathbf{z}$, $\mathbf{z} \in \ker A$, with $\|\mathbf{x}_0 + \mathbf{z}\|^2 = \|\mathbf{x}_0\|^2 + \|\mathbf{z}\|^2$. (d) $AA^{\mathsf{T}} = \operatorname{diag}(3, 2)$, so $\mathbf{x}_0 = A^{\mathsf{T}}(1, 1/2) = (3/2, 1, 1/2)$.

(a) Row i of A, read as a vector, has conjugate equal to column i of A^* (Definition 6.3.26). A relation $\sum_i c_i \overline{\mathbf{r}_i} = \mathbf{0}$ among the conjugated rows gives, after conjugating, the relation $\sum_i \overline{c_i}\mathbf{r}_i = \mathbf{0}$ among the rows, so the columns of A^* are independent. By Proposition 6.5.9(c) applied to $A^* \in \mathbb{F}^{n\times m}$, the matrix $(A^*)^*A^* = AA^*$ is invertible.

(b) $A\mathbf{x}_0 = (AA^*)(AA^*)^{-1}\mathbf{b} = \mathbf{b}$. Let $\mathbf{y} = (AA^*)^{-1}\mathbf{b}$, so $\mathbf{x}_0 = A^*\mathbf{y}$. For $\mathbf{z} \in \ker A$, Proposition 6.3.27(d) and conjugate symmetry give $\langle \mathbf{x}_0, \mathbf{z}\rangle = \langle A^*\mathbf{y}, \mathbf{z}\rangle = \overline{\langle \mathbf{z}, A^*\mathbf{y}\rangle} = \overline{\langle A\mathbf{z}, \mathbf{y}\rangle} = 0$. So $\mathbf{x}_0 \in (\ker A)^{\perp}$.

(c) By Proposition 3.1.19, the solutions are the vectors $\mathbf{x}_0 + \mathbf{z}$ with $\mathbf{z} \in \ker A$. Since $\mathbf{x}_0 \perp \mathbf{z}$, Theorem 6.3.3 gives $\|\mathbf{x}_0 + \mathbf{z}\|^2 = \|\mathbf{x}_0\|^2 + \|\mathbf{z}\|^2 \geq \|\mathbf{x}_0\|^2$, with equality if and only if $\mathbf{z} = \mathbf{0}$.

(d) Here $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$, whose rows are independent, and $AA^{\mathsf{T}} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$. So $\left(AA^{\mathsf{T}}\right)^{-1}(3, 1) = (1, 1/2)$, and

$$\mathbf{x}_0 = A^{\mathsf{T}}\begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1 \\ 1/2 \end{pmatrix}.$$

Check: $3/2 + 1 + 1/2 = 3$ and $3/2 - 1/2 = 1$, and $\mathbf{x}_0$ is orthogonal to $(1, -2, 1)$, which spans $\ker A$. In the language of Section 6.5.7, $\mathbf{x}_0 = A^+\mathbf{b}$, and $A^+ = A^*(AA^*)^{-1}$ for A with independent rows.

Solution to Exercise 6.5.47

Uses. Proposition 6.5.17 (The least-squares line)
Hint 1. Show that $\sum_{i<j}(u_j - u_i)(v_j - v_i) = m\sum_k (u_k - m_u)(v_k - m_v)$ for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$. Sum over all ordered pairs (i, j) first.
Hint 2. Apply this with $(\mathbf{u}, \mathbf{v}) = (\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}, \mathbf{x})$, and note that $(x_j - x_i)(y_j - y_i) = (x_j - x_i)^2 s_{ij}$.
Sketch. Summing over all ordered pairs, $\sum_{i,j}(u_j - u_i)(v_j - v_i) = 2m\sum_k u_k v_k - 2\left(\sum_k u_k\right)\left(\sum_k v_k\right) = 2m\sum_k (u_k - m_u)(v_k - m_v)$, and the sum over $i < j$ is half of it. So the numerator and denominator in the claim are $m\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}}$ and $m\|\tilde{\mathbf{x}}\|^2$, whose ratio is c_1. A weighted average with positive weights lies between the smallest and largest values.

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ with means m_u, m_v. Summing over all ordered pairs (i, j), including $i = j$, whose terms are 0,

$$\begin{aligned}&\sum_{i,j}(u_j-u_i)(v_j-v_i)\\&=\sum_{i,j}(u_jv_j+u_iv_i-u_jv_i-u_iv_j)\\&=2m\sum_k u_kv_k-2\left(\sum_k u_k\right)\left(\sum_k v_k\right).\end{aligned}$$

Since $\sum_k u_k = mm_u$ and $\sum_k v_k = mm_v$, the right side is $2m\big(\sum_k u_kv_k - mm_um_v\big) = 2m\sum_k(u_k-m_u)(v_k-m_v)$, by the identity in the proof of Proposition 6.5.17. The terms for (i,j) and (j,i) are equal, so

$$\sum_{i<j}(u_j-u_i)(v_j-v_i) = m\sum_k(u_k-m_u)(v_k-m_v).$$

With $(\mathbf{u},\mathbf{v}) = (\mathbf{x},\mathbf{y})$ and $(\mathbf{x},\mathbf{x})$, this gives $\sum_{i<j}(x_j-x_i)(y_j-y_i) = m\,\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}$ and $\sum_{i<j}(x_j-x_i)^2 = m\|\tilde{\mathbf{x}}\|^2 > 0$. In the first sum, a pair with $x_i = x_j$ contributes 0, and a pair with $x_i \neq x_j$ contributes $(x_j-x_i)^2 s_{ij}$. Dividing, and using Proposition 6.5.17,

$$\frac{\sum_{i<j}(x_j-x_i)^2 s_{ij}}{\sum_{i<j}(x_j-x_i)^2} = \frac{\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}}{\|\tilde{\mathbf{x}}\|^2} = c_1.$$

The weights $w_{ij} = (x_j-x_i)^2$ of the remaining pairs are positive. If $s_{\min}$ and $s_{\max}$ are the smallest and largest s_{ij}, then $s_{\min}\sum w_{ij} \le \sum w_{ij}s_{ij} \le s_{\max}\sum w_{ij}$, so $s_{\min} \le c_1 \le s_{\max}$. Pairs of points far apart in x get the most weight, which matches the intuition that they determine a slope most reliably.

Solution to Exercise 6.5.48

Uses. Theorem 6.5.5 (The normal equations), Proposition 6.3.27 (Properties of the conjugate transpose), Aside (Weighted least squares) (Weighted least squares)
Hint 1. $(W\mathbf{v})_i = \sqrt{w_i}v_i$. So the weighted sum is an ordinary least-squares problem for the matrix WA and the vector $W\mathbf{b}$.
Hint 2. For (b), $A^{\mathsf{T}}DA = \begin{pmatrix}13/4 & 15/4\\ 15/4 & 29/4\end{pmatrix}$ and $A^{\mathsf{T}}D\mathbf{y} = (8,14)$.
Sketch. (a) $\|W(\mathbf{b}-A\mathbf{x})\|^2 = \sum_i w_i|b_i-(A\mathbf{x})_i|^2$, so the minimizers are the least-squares solutions of $(WA)\mathbf{x} = W\mathbf{b}$, whose normal equations are $A^*W^*WA\mathbf{x} = A^*W^*W\mathbf{b}$, and $W^*W = D$. (b) Ordinary: $y = 1/5 + (11/5)x$. Weighted: $y = 11/19 + (31/19)x \approx 0.58 + 1.63x$.

(a) The ith entry of $W(\mathbf{b}-A\mathbf{x})$ is $\sqrt{w_i}(b_i-(A\mathbf{x})_i)$, so $\|W\mathbf{b}-WA\mathbf{x}\|^2 = \sum_i w_i|b_i-(A\mathbf{x})_i|^2$. Hence the minimizers are the least-squares solutions of $(WA)\mathbf{x} = W\mathbf{b}$, which by Theorem 6.5.5 are the solutions of $(WA)^*(WA)\mathbf{x} = (WA)^*W\mathbf{b}$. Since W is real and diagonal, $W^* = W$ and $W^*W = W^2 = D$, and by Proposition 6.3.27(b) the equations read $A^*DA\mathbf{x} = A^*D\mathbf{b}$. This is a second derivation of the weighted normal equations of Aside (Weighted least squares).

(b) Ordinary least squares: $A^{\mathsf{T}}A = \begin{pmatrix}4 & 6\\ 6 & 14\end{pmatrix}$ and $A^{\mathsf{T}}\mathbf{y} = (14, 0+2+6+24) = (14,32)$. The determinant is 20, and $c_0 = (14\cdot 14 - 6\cdot 32)/20 = 1/5$, $c_1 = (4\cdot 32 - 6\cdot 14)/20 = 11/5$. So the line is $y = 0.2 + 2.2x$.

Weighted, with $D = \operatorname{diag}(1,1,1,1/4)$:

$$A^{\mathsf{T}}DA = \begin{pmatrix}3+1/4 & 3+3/4\\ 3+3/4 & 5+9/4\end{pmatrix} = \frac{1}{4}\begin{pmatrix}13 & 15\\ 15 & 29\end{pmatrix},$$

$$A^{\mathsf{T}}D\mathbf{y} = \begin{pmatrix}6+2\\ 8+6\end{pmatrix} = \begin{pmatrix}8\\ 14\end{pmatrix}.$$

Multiplying by 4, the system is $13c_0 + 15c_1 = 32$, $15c_0 + 29c_1 = 56$, with determinant $377 - 225 = 152$. So $c_0 = (32\cdot 29 - 15\cdot 56)/152 = 88/152 = 11/19$ and $c_1 = (13\cdot 56 - 15\cdot 32)/152 = 248/152 = 31/19$. The weighted line is $y \approx 0.58 + 1.63x$. The first three points lie on $y = 1 + x$, and giving the unreliable fourth point less weight pulls the slope from 2.2 toward 1.

Solution to Exercise 6.5.49

Uses. Proposition 6.5.27 (The least-squares solution of minimal norm), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.4.31 (The matrix of an orthogonal projection), Proposition 6.5.2 (Least squares is projection)
Hint 1. $A^+(A\mathbf{x})$ is the least-squares solution of $A\mathbf{z} = A\mathbf{x}$ of smallest norm. This system is consistent; what are its solutions?
Hint 2. For (b), use $AA^+ = \mathrm{Pr}_{\mathrm{im}\,A}$, (a), and the fact that $A^+\mathbf{b} \in (\ker A)^\perp$. Matrices of orthogonal projections are Hermitian.
Sketch. (a) The solutions of $A\mathbf{z} = A\mathbf{x}$ are $\mathbf{x} + \ker A$, so $A^+A\mathbf{x} = \mathbf{x} - \mathrm{Pr}_{\ker A}(\mathbf{x}) = \mathrm{Pr}_{(\ker A)^\perp}(\mathbf{x})$ by Proposition 6.5.27(a) and Proposition 6.4.29(d). (b) $AA^+A = \mathrm{Pr}_{\mathrm{im}\,A}\,A = A$; $A^+AA^+ = \mathrm{Pr}_{(\ker A)^\perp}\,A^+ = A^+$; orthogonal projection matrices are Hermitian. (c) $A^+ = (1/4)\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$, with $AA^+ = A^+A = (1/2)\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$.

(a) Let $\mathbf{x} \in \mathbb{F}^n$. Then $A^+(A\mathbf{x})$ is the least-squares solution of smallest norm of $A\mathbf{z} = A\mathbf{x}$. This system is consistent, so its least-squares solutions are its solutions, $\mathbf{x} + \ker A$ (Proposition 6.5.2(b)). By Proposition 6.5.27(a) with $\hat{\mathbf{x}} = \mathbf{x}$, $A^+A\mathbf{x} = \mathbf{x} - \mathrm{Pr}_{\ker A}(\mathbf{x})$, and since $\mathbb{F}^n$ is finite-dimensional this is $\mathrm{Pr}_{(\ker A)^\perp}(\mathbf{x})$ by Proposition 6.4.29(d).

(b) By Proposition 6.5.2(a), $AA^+ = \mathrm{Pr}_{\mathrm{im}\,A}$, which fixes every column of A; so $AA^+A = A$. By Proposition 6.5.27(a), every $A^+\mathbf{b}$ lies in $(\ker A)^\perp$, which $\mathrm{Pr}_{(\ker A)^\perp}$ fixes; with (a), $A^+AA^+ = A^+$. Finally AA^+ and A^+A are matrices of orthogonal projections, which are Hermitian by Theorem 6.4.31.

(c) Here $\mathrm{im}\,A = \mathrm{span}((1,1))$, $\ker A = \mathrm{span}((1,-1))$, and $(\ker A)^\perp = \mathrm{span}((1,1))$. For $\mathbf{b} \in \mathbb{F}^2$, $\mathrm{Pr}_{\mathrm{im}\,A}(\mathbf{b}) = ((b_1+b_2)/2)(1,1)$, so the least-squares solutions satisfy $x_1 + x_2 = (b_1+b_2)/2$, and the one in $(\ker A)^\perp$ is $s(1,1)$ with $2s = (b_1+b_2)/2$. So $A^+\mathbf{b} = ((b_1+b_2)/4)(1,1)$, that is, $A^+ = (1/4)\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$. Then $AA^+ = A^+A = (1/2)\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$, which is Hermitian, $AA^+A = (1/2)\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix} = A$, and $A^+AA^+ = (1/8)\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix} = A^+$.

Solution to Exercise 6.5.50

Uses. Example 6.4.35 (Distance to a hyperplane), Exercise 6.5.20, Theorem 6.3.22 (Parseval's identity), Theorem 6.2.1 (Cauchy–Schwarz inequality), Exercise 6.5.41
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. For (a), $\ell = \mathbf{q} + \{\mathbf{n}\}^\perp$ for any $\mathbf{q} \in \ell$; translate Example 6.4.35. Then $\sum_i (\langle \mathbf{p}_i, \mathbf{n}\rangle - c)^2$ is a least-squares problem for a constant.
Hint 2. For (b), $\mathbf{u}, \mathbf{n}$ is an orthonormal basis of $\mathbb{R}^2$, so $\|\tilde{\mathbf{p}}\|^2 = \langle \tilde{\mathbf{p}}, \mathbf{u}\rangle^2 + \langle \tilde{\mathbf{p}}, \mathbf{n}\rangle^2$.
Hint 3. For (c), expand $(\tilde{x}_i \cos\theta + \tilde{y}_i \sin\theta)^2$ with the double-angle formulas, and bound $\mathbf{w} \cdot (\cos 2\theta, \sin 2\theta)$ by Cauchy–Schwarz.
Sketch. (a) With $\mathbf{q} \in \ell$, $\mathrm{dist}(\mathbf{p}, \ell) = \mathrm{dist}\big(\mathbf{p} - \mathbf{q}, \{\mathbf{n}\}^\perp\big) = |\langle \mathbf{p} - \mathbf{q}, \mathbf{n}\rangle| = |\langle \mathbf{p}, \mathbf{n}\rangle - c|$; minimizing $\sum_i (\langle \mathbf{p}_i, \mathbf{n}\rangle - c)^2$ over c gives the mean $c = \langle \mathbf{g}, \mathbf{n}\rangle$. (b) Parseval in the basis $\mathbf{u}, \mathbf{n}$. (c) The double-angle formulas give the expression, and Cauchy–Schwarz bounds $\mathbf{w} \cdot (\cos 2\theta, \sin 2\theta) \le \|\mathbf{w}\|$ with equality iff $(\cos 2\theta, \sin 2\theta) = \mathbf{w}/\|\mathbf{w}\|$. (d) $\mathbf{g} = (3/2, 3/2)$, $\mathbf{w} = (0, 4)$, so $\theta = \pi/4$ and the line is $y = x$, whose slope 1 lies between $4/5$ and $5/4$.

(a) Pick $\mathbf{q} \in \ell$, so $\langle \mathbf{q}, \mathbf{n}\rangle = c$ and $\ell = \mathbf{q} + H$ with $H = \{\mathbf{n}\}^\perp$. Translating by $-\mathbf{q}$ preserves distances, so by Example 6.4.35, $\mathrm{dist}(\mathbf{p}, \ell) = \mathrm{dist}(\mathbf{p} - \mathbf{q}, H) = |\langle \mathbf{p} - \mathbf{q}, \mathbf{n}\rangle| = |\langle \mathbf{p}, \mathbf{n}\rangle - c|$. For fixed $\mathbf{n}$, $D(\ell) = \sum_i (t_i - c)^2$ with $t_i = \langle \mathbf{p}_i, \mathbf{n}\rangle$. This is the least-squares problem $c\mathbf{1} \approx \mathbf{t}$, whose solution (Exercise 6.5.20 with all $x_i = 1$) is the mean $c = (1/m)\sum_i t_i = \langle \mathbf{g}, \mathbf{n}\rangle$, and it is unique. So D is smallest when $\mathbf{g} \in \ell$.

(b) Let $\mathbf{n}$ be a unit vector orthogonal to $\mathbf{u}$, so that $\mathbf{u}, \mathbf{n}$ is an orthonormal basis of $\mathbb{R}^2$ and $\ell = \{\mathbf{p} \mid \langle \mathbf{p}, \mathbf{n}\rangle = \langle \mathbf{g}, \mathbf{n}\rangle\}$. By (a), $\mathrm{dist}(\mathbf{p}_i, \ell)^2 = \langle \tilde{\mathbf{p}}_i, \mathbf{n}\rangle^2$, and by Parseval (Theorem 6.3.22), $\|\tilde{\mathbf{p}}_i\|^2 = \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2 + \langle \tilde{\mathbf{p}}_i, \mathbf{n}\rangle^2$. Summing over i gives the formula.

(c) We have $\langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2 = \tilde{x}_i^2 \cos^2\theta + 2\tilde{x}_i\tilde{y}_i \cos\theta \sin\theta + \tilde{y}_i^2 \sin^2\theta$. Summing, and using $\cos^2\theta = (1 + \cos 2\theta)/2$, $\sin^2\theta = (1 - \cos 2\theta)/2$, and $2\sin\theta\cos\theta = \sin 2\theta$,

$$\sum_i \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2 = \frac{S_{xx} + S_{yy}}{2} + \frac{S_{xx} - S_{yy}}{2}\cos 2\theta + S_{xy}\sin 2\theta$$
$$= \frac{S_{xx} + S_{yy}}{2} + \mathbf{w} \cdot (\cos 2\theta, \sin 2\theta).$$

By Cauchy–Schwarz (Theorem 6.2.1), $\mathbf{w} \cdot (\cos 2\theta, \sin 2\theta) \le \|\mathbf{w}\|$, with equality if and only if the unit vector $(\cos 2\theta, \sin 2\theta)$ is a positive multiple of $\mathbf{w}$, that is, equals $\mathbf{w}/\|\mathbf{w}\|$. By (a), a minimizing line passes through $\mathbf{g}$, and by (b), among those, D is smallest exactly when $\sum_i \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2$ is largest. The condition determines 2θ modulo 2π, hence $\mathbf{u}$ up to sign, hence the line.

(d) Here $\mathbf{g} = (3/2, 3/2)$, $S_{xx} = S_{yy} = 5$, and $S_{xy} = 4$ (Exercise 6.5.41(c)), so $\mathbf{w} = (0, 4)$, $2\theta = \pi/2$, and $\mathbf{u} = (1, 1)/\sqrt{2}$. The best line is $y = x$, with $D = 10 - (5 + 4) = 1$. Its slope 1 lies between the slope $4/5$ of the regression of y on x and the slope $5/4$ of the regression of x on y, as it treats the two coordinates symmetrically. The direction $\mathbf{u}$ maximizing $\sum_i \langle \tilde{\mathbf{p}}_i, \mathbf{u}\rangle^2$ is the first principal component of the points (Section 16.2).

Section 6.6

Solution to Exercise 6.6.6

Uses. Theorem 6.6.3 (Riesz representation theorem), Definition 6.1.8 (Dot product), Example 6.6.5 (Representing evaluation at 0)
Hint 1. In (a), $\mathbf{z} \cdot \mathbf{u} = \sum_j z_j \overline{u_j}$, so each u_j is the conjugate of a coefficient.
Hint 2. In (b), use the orthonormal basis $1, \sqrt{3}(2x-1)$, or write $q = c + dx$ and impose the two conditions $\langle 1, q\rangle = 1$ and $\langle x, q\rangle = 1$.
Sketch. (a) $\mathbf{u} = (2, i, 1-i)$, the conjugated coefficients. (b) $q = 1 + \sqrt{3}\cdot\sqrt{3}(2x-1) = 6x - 2$; check: $\int_0^1 (6t-2)\,dt = 1$ and $\int_0^1 t(6t-2)\,dt = 1$.

(a) For $\mathbf{u} = (2, i, 1-i)$,

$$\mathbf{z} \cdot \mathbf{u} = z_1\overline{2} + z_2\overline{i} + z_3\overline{1-i} = 2z_1 - iz_2 + (1+i)z_3,$$

as required. By Theorem 6.6.3 (or by comparing coefficients at $\mathbf{z} = \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$) it is the only such vector.

(b) With $e_1 = 1$ and $e_2 = \sqrt{3}(2x-1)$, the functional $\varphi(p) = p(1)$ has $\varphi(e_1) = 1$ and $\varphi(e_2) = \sqrt{3}$, so Theorem 6.6.3 gives

$$q = 1 + \sqrt{3}\cdot\sqrt{3}(2x-1) = 6x - 2.$$

Check: $\int_0^1 (6t-2)\,dt = 3 - 2 = 1 = \varphi(1)$ and $\int_0^1 t(6t-2)\,dt = 2 - 1 = 1 = \varphi(x)$, so $\langle p, q\rangle = p(1)$ for all $p \in \mathbb{R}[x]_{\le 1}$ by linearity. As expected from the symmetry $t \mapsto 1 - t$ of $[0,1]$, $q(x)$ is $4 - 6(1-x)$, the polynomial of Example 6.6.5 with x replaced by $1 - x$.

Solution to Exercise 6.6.17

Uses. Proposition 6.1.16 (Frobenius inner product), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.3.27 (Properties of the conjugate transpose), Definition 6.6.8 (Adjoint)
Techniques. Find the adjoint by isolating the input (Technique 6.6.11)
Hint 1. For (a), use $\operatorname{tr}(BC) = \operatorname{tr}(CB)$ to move A next to Y^*, and recognize AY^* as $(YA^*)^*$.
Hint 2. For (b), use the entrywise formula $\langle X, Y\rangle = \sum_{i,j} x_{ij}\overline{y_{ij}}$ and rename the indices.
Sketch. (a) $\operatorname{tr}(Y^*XA) = \operatorname{tr}(AY^*X) = \operatorname{tr}\big((YA^*)^*X\big)$, so $R_A^* = R_{A^*}$. (b) $\langle X^{\mathsf{T}}, Y\rangle = \sum_{i,j} x_{ji}\overline{y_{ij}} = \langle X, Y^{\mathsf{T}}\rangle$, so the transpose map is its own adjoint. (c) $(iX)^* = -iX^*$, so the map is not linear over $\mathbb{C}$, and Definition 6.6.8 applies only to linear maps.

(a) By Proposition 2.4.28 and Proposition 6.3.27,

$$\langle XA, Y\rangle = \operatorname{tr}(Y^*XA) = \operatorname{tr}(AY^*X) = \operatorname{tr}\big((YA^*)^*X\big) = \langle X, YA^*\rangle,$$

since $(YA^*)^* = (A^*)^*Y^* = AY^*$. So $R_A^*(Y) = YA^*$, that is, $R_A^* = R_{A^*}$.

(b) By Proposition 6.1.16, $\langle X, Y\rangle = \sum_{i,j} x_{ij}\overline{y_{ij}}$. The (i,j) entry of X^{T} is x_{ji}, so

$$\langle X^{\mathsf{T}}, Y\rangle = \sum_{i,j} x_{ji}\overline{y_{ij}} = \sum_{i,j} x_{ij}\overline{y_{ji}} = \langle X, Y^{\mathsf{T}}\rangle,$$

after renaming i and j. So the transpose map is its own adjoint.

(c) $(iX)^* = \bar{i}X^* = -iX^*$ (Proposition 6.3.27(a)), so $X \mapsto X^*$ is not linear over $\mathbb{C}$ (unless $n = 0$). Adjoints are defined for linear maps only. (Over $\mathbb{R}$ the map is the transpose, covered by (b).)

Solution to Exercise 6.6.19

Uses. Proposition 6.6.18 (Properties of the adjoint)
Techniques. Induction on the length of a list (Technique 1.6.9)
Hint 1. First show $\left(T^j\right)^* = (T^*)^j$ by induction on j, using (d) and (e) of Proposition 6.6.18.
Hint 2. Then expand the left side with (a) and (b).
Sketch. $\left(T^0\right)^* = \mathrm{id}^* = \mathrm{id}$, and $\left(T^{j+1}\right)^* = \left(TT^j\right)^* = \left(T^j\right)^* T^* = (T^*)^j T^* = (T^*)^{j+1}$. Then (a) and (b) give $\left(\sum_j c_j T^j\right)^* = \sum_j \overline{c_j}\left(T^j\right)^* = \sum_j \overline{c_j}(T^*)^j$.

All adjoints exist since V is finite-dimensional (Theorem 6.6.15). We show $\left(T^j\right)^* = (T^*)^j$ by induction on $j \geq 0$. For $j = 0$, $T^0 = \mathrm{id}$ and $\mathrm{id}^* = \mathrm{id}$ by Proposition 6.6.18(e). If $\left(T^j\right)^* = (T^*)^j$, then by Proposition 6.6.18(d),

$$\left(T^{j+1}\right)^* = \left(TT^j\right)^* = \left(T^j\right)^* T^* = (T^*)^j T^* = (T^*)^{j+1}.$$

By (a) and (b) of Proposition 6.6.18, applied term by term,

$$\left(\sum_{j=0}^{k} c_j T^j\right)^* = \sum_{j=0}^{k} \overline{c_j}\left(T^j\right)^* = \sum_{j=0}^{k} \overline{c_j}(T^*)^j,$$

as desired.

Solution to Exercise 6.6.25

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Corollary 6.6.22 (T and T^* have the same rank), Proposition 2.6.18 (The column space is spanned by the columns), Proposition 6.3.9 (Properties of $A^\perp$)
Hint 1. The third row is the sum of the first two, so rank $A = 2$, and each kernel is a line.
Hint 2. Solve $A\mathbf{x} = \mathbf{0}$ and $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$, then check orthogonality against the rows and the columns of A.
Sketch. $\ker A = \mathrm{span}((1,-1,1))$ and $\mathrm{im}\, A^{\mathsf{T}} = \mathrm{span}((1,2,1),(0,1,1))$; the kernel vector is orthogonal to both rows. $\mathrm{im}\, A = \mathrm{span}((1,0,1),(2,1,3))$ and $\ker A^{\mathsf{T}} = \mathrm{span}((1,1,-1))$, orthogonal to every column. Dimensions $1+2=3$ in both cases.

Row 3 is row 1 plus row 2, and rows 1 and 2 are not proportional, so $\operatorname{rank} A = 2$ and $\mathrm{im}\, A^{\mathsf{T}} = \mathrm{span}((1,2,1),(0,1,1))$ has dimension 2. Solving $x_1 + 2x_2 + x_3 = 0$ and $x_2 + x_3 = 0$ gives $x_2 = -x_3$ and $x_1 = x_3$, so $\ker A = \mathrm{span}((1,-1,1))$. Indeed $(1,-1,1)\cdot(1,2,1) = 0$ and $(1,-1,1)\cdot(0,1,1) = 0$, so $\ker A \perp \mathrm{im}\, A^{\mathsf{T}}$ (Proposition 6.3.9(c)), and the dimensions add up to 3, as Theorem 6.6.21 predicts.

The columns are $(1,0,1)$, $(2,1,3)$, $(1,1,2)$, and the third is the second minus the first, so $\mathrm{im}\, A = \mathrm{span}((1,0,1),(2,1,3))$ (Proposition 2.6.18). The equations of $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$ are $y_1 + y_3 = 0$, $2y_1 + y_2 + 3y_3 = 0$, $y_1 + y_2 + 2y_3 = 0$; the first gives $y_3 = -y_1$, and then the second gives $y_2 = y_1$, which also satisfies the third. So $\ker A^{\mathsf{T}} = \mathrm{span}((1,1,-1))$, which is orthogonal to $(1,0,1)$ and $(2,1,3)$. Again the dimensions are 2 and 1.

Solution to Exercise 6.6.31

Uses. Example 6.6.10 (The adjoint of a matrix), Proposition 6.6.9 (Adjoints are unique and linear), Proposition 6.1.4 (Sesquilinearity), Theorem 6.6.15 (Existence of adjoints)
Techniques. Find the adjoint by isolating the input (Technique 6.6.11)
Hint 1. For (a), use Example 6.6.10 with $A = \begin{pmatrix} 1 & i \\ 1-i & 0 \end{pmatrix}$.
Hint 2. For (b), expand $\langle T\mathbf{z}, \mathbf{w}\rangle$ and collect the terms with z_1 and with z_2. The coefficient of z_2 must be written as 2 times a conjugate, to match the weight 2.
Sketch. (a) $A^* = \begin{pmatrix} 1 & 1+i \\ -i & 0 \end{pmatrix}$, so $T^*(\mathbf{w}) = (w_1 + (1+i)w_2, -iw_1)$. (b) $\langle T\mathbf{z}, \mathbf{w}\rangle = z_1\overline{w_1 + 2(1+i)w_2} + 2z_2\overline{-iw_1/2}$, so $T^*(\mathbf{w}) = (w_1 + (2+2i)w_2, -iw_1/2)$. (c) No; the standard basis is not orthonormal for the weighted inner product.

(a) The matrix of T is $A = \begin{pmatrix} 1 & i \\ 1-i & 0 \end{pmatrix}$, so by Example 6.6.10, T^* is multiplication by $A^* = \begin{pmatrix} 1 & 1+i \\ -i & 0 \end{pmatrix}$:

$$T^*(w_1, w_2) = (w_1 + (1+i)w_2, -iw_1).$$

(b) For $\mathbf{z}, \mathbf{w} \in \mathbb{C}^2$,

$$\begin{aligned} \langle T\mathbf{z}, \mathbf{w}\rangle &= (z_1 + iz_2)\overline{w_1} + 2(1-i)z_1\overline{w_2} \\ &= z_1(\overline{w_1} + 2(1-i)\overline{w_2}) + 2z_2 \cdot \frac{i\overline{w_1}}{2} \end{aligned}$$

$$= z_1\overline{w_1 + 2(1+i)w_2} + 2z_2\overline{-iw_1/2},$$

which is $\langle \mathbf{z}, \mathbf{u}\rangle$ for $\mathbf{u} = (w_1 + (2+2i)w_2, -iw_1/2)$. By Proposition 6.6.9,

$$T^*(w_1, w_2) = (w_1 + (2+2i)w_2, -(i/2)w_1).$$

(c) The matrix of this T^* in the standard basis is $\begin{pmatrix} 1 & 2+2i \\ -i/2 & 0 \end{pmatrix} \neq A^*$. There is no contradiction: Theorem 6.6.15(b) needs orthonormal bases, and $\mathbf{e}_1, \mathbf{e}_2$ is not orthonormal for the weighted inner product, since $\langle \mathbf{e}_2, \mathbf{e}_2\rangle = 2$. The adjoint depends on the inner products, not only on the map.

Solution to Exercise 6.6.32

Uses. Theorem 6.6.3 (Riesz representation theorem), Example 6.4.19 (Legendre polynomials), Corollary 6.4.10 (Projecting with an orthogonal list)
Hint 1. Use the orthogonal basis $1, x, x^2 - 1/3$ of Example 6.4.19, with squared norms $2, 2/3, 8/45$. For an orthogonal basis f_j, the formula of Theorem 6.6.3 becomes $q = \sum_j \varphi(f_j)/\|f_j\|^2 f_j$ over $\mathbb{R}$.
Hint 2. (a) $\varphi(1) = 0$, $\varphi(x) = 1$, $\varphi(x^2 - 1/3) = 0$. (b) $\varphi(1) = 1$, $\varphi(x) = 1/2$, $\varphi(x^2 - 1/3) = 0$.
Sketch. (a) $q_1 = (1/(2/3))x = 3x/2$. (b) $q_2 = 1/2 + (1/2)/(2/3)\,x = 1/2 + 3x/4$. Check both on $1, x, x^2$.

Let $f_1 = 1$, $f_2 = x$, $f_3 = x^2 - 1/3$, an orthogonal basis with $\|f_1\|^2 = 2$, $\|f_2\|^2 = 2/3$, $\|f_3\|^2 = 8/45$ (Example 6.4.19). Normalizing, $e_j = f_j/\|f_j\|$, the formula of Theorem 6.6.3 gives, over $\mathbb{R}$,

$$q = \sum_j \varphi(e_j)e_j = \sum_j \frac{\varphi(f_j)}{\|f_j\|^2} f_j.$$

(a) For $\varphi(p) = p'(0)$: $\varphi(1) = 0$, $\varphi(x) = 1$, and $\varphi(x^2 - 1/3) = 0$. So $q_1 = (3/2)x$. Check: for $p = a + bx + cx^2$, $\langle p, q_1\rangle = (3/2)\left(a\int_{-1}^{1} t\,dt + b\int_{-1}^{1} t^2\,dt + c\int_{-1}^{1} t^3\,dt\right) = (3/2)(2b/3) = b = p'(0)$.

(b) For $\varphi(p) = \int_0^1 p$: $\varphi(1) = 1$, $\varphi(x) = 1/2$, and $\varphi(x^2 - 1/3) = 1/3 - 1/3 = 0$. So

$$q_2 = \frac{1}{2} + \frac{1/2}{2/3}x = \frac{1}{2} + \frac{3}{4}x.$$

Check: $\langle 1, q_2\rangle = 1 = \int_0^1 1$, $\langle x, q_2\rangle = (3/4)(2/3) = 1/2 = \int_0^1 t\,dt$, and $\langle x^2, q_2\rangle = (1/2)(2/3) = 1/3 = \int_0^1 t^2\,dt$.

Solution to Exercise 6.6.33

Uses. Definition 6.1.8 (Dot product), Proposition 6.6.9 (Adjoints are unique and linear), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Definition 6.4.11 (Orthogonal projection)
Techniques. Find the adjoint by isolating the input (Technique 6.6.11)
Hint 1. Write $\langle R\mathbf{x}, \mathbf{y}\rangle = \sum_{j=1}^{n-1} x_j\overline{y_{j+1}}$ and read off the vector in the second slot.
Hint 2. R^* shifts the other way. For (b), C sends $\mathbf{e}_j$ to $\mathbf{e}_{j+1}$ and $\mathbf{e}_n$ to $\mathbf{e}_1$.
Sketch. (a) $R^*(\mathbf{y}) = (y_2, \ldots, y_n, 0)$. Then $R^*R\mathbf{x} = (x_1, \ldots, x_{n-1}, 0)$ projects onto $\operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1})$ and $RR^*\mathbf{y} = (0, y_2, \ldots, y_n)$ projects onto $\operatorname{span}(\mathbf{e}_2, \ldots, \mathbf{e}_n)$. (b) $C^*(\mathbf{y}) = (y_2, \ldots, y_n, y_1)$, which undoes C.

(a) For $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$,

$$\langle R\mathbf{x}, \mathbf{y}\rangle = \sum_{j=1}^{n-1} x_j\overline{y_{j+1}} = \langle \mathbf{x}, (y_2, \ldots, y_n, 0)\rangle,$$

so $R^*(y_1, \ldots, y_n) = (y_2, \ldots, y_n, 0)$, the shift to the left. Then

$$R^*R\mathbf{x} = R^*(0, x_1, \ldots, x_{n-1}) = (x_1, \ldots, x_{n-1}, 0), \quad RR^*\mathbf{y} = R(y_2, \ldots, y_n, 0) = (0, y_2, \ldots, y_n).$$

The first keeps the coordinates along $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ and kills the last, so it is Pr_W for $W = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_{n-1})$ (Definition 6.4.11, with the orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}$ of W). Likewise $RR^* = \operatorname{Pr}_{\operatorname{span}(\mathbf{e}_2, \ldots, \mathbf{e}_n)}$. Note $R^*R \neq RR^*$: the shift is not normal (Section 8.1).

(b) For $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$,

$$\langle C\mathbf{x}, \mathbf{y}\rangle = x_n\overline{y_1} + \sum_{j=1}^{n-1} x_j\overline{y_{j+1}} = \langle \mathbf{x}, (y_2, ..., y_n, y_1)\rangle,$$

so $C^*(y_1, ..., y_n) = (y_2, ..., y_n, y_1)$. Then $C(C^*\mathbf{y}) = C(y_2, ..., y_n, y_1) = (y_1, y_2, ..., y_n)$ and $C^*(C\mathbf{x}) = C^*(x_n, x_1, ..., x_{n-1}) = (x_1, ..., x_n)$. So C^* is a two-sided inverse of C.

Solution to Exercise 6.6.34

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Definition 6.3.26 (Conjugate transpose), Definition 6.1.8 (Dot product), Definition 6.6.26 (Self-adjoint operator)
Hint 1. The second row is i times the first, so A has rank 1 and every subspace in question is a line.
Hint 2. $A^* = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix}$. Remember to conjugate the second vector in each dot product.
Sketch. (a) $\ker A = \operatorname{span}((-i, 1))$, $\operatorname{im} A = \operatorname{span}((1, i))$, $\ker A^* = \operatorname{span}((i, 1))$, $\operatorname{im} A^* = \operatorname{span}((1, -i))$. Then $(i, 1) \cdot (1, i) = 0$ and $(-i, 1) \cdot (1, -i) = 0$, and each pair consists of two lines in $\mathbb{C}^2$, so they are orthogonal complements. (b) $(-i, 1) \cdot (1, i) = -2i$. (c) No, $A^* \neq A$.

(a) The second row of A is i times the first, and the second column is i times the first, so $\operatorname{rank} A = 1$. The kernel is given by $x_1 + ix_2 = 0$, so $\ker A = \operatorname{span}((-i, 1))$, and $\operatorname{im} A = \operatorname{span}((1, i))$. Next, $A^* = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix}$, again of rank 1: $\ker A^*$ is given by $x_1 - ix_2 = 0$, so $\ker A^* = \operatorname{span}((i, 1))$, and $\operatorname{im} A^* = \operatorname{span}((1, -i))$. Now

$$(i, 1) \cdot (1, i) = i + \bar{i} = 0, \quad (-i, 1) \cdot (1, -i) = -i + \overline{-i} = 0.$$

So $\ker A^* \perp \operatorname{im} A$ and $\ker A \perp \operatorname{im} A^*$. In each case both are lines in $\mathbb{C}^2$, and the perp of a line in $\mathbb{C}^2$ is a line (Corollary 6.4.25), so $\ker A^* = (\operatorname{im} A)^\perp$ and $\ker A = (\operatorname{im} A^*)^\perp$, which are (a) and (b) of Theorem 6.6.21. Taking perps once more gives (c) and (d).

(b) $\operatorname{im} A^{\mathsf{T}} = \operatorname{im} A = \operatorname{span}((1, i))$, and $(-i, 1) \cdot (1, i) = -i + \bar{i} = -2i \neq 0$.

(c) No: $A^* = \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix} \neq A$. A complex symmetric matrix need not be Hermitian, and (b) shows what goes wrong: the orthogonal complement of the kernel is spanned by the conjugated rows, not by the rows.

Solution to Exercise 6.6.35

Uses. Proposition 6.6.18 (Properties of the adjoint), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents)
Hint 1. Take the adjoint of both sides of $T^*T = T$.
Hint 2. You should find that T is self-adjoint. Then $T^2 = T^*T$.
Sketch. $T^* = (T^*T)^* = T^*T^{**} = T^*T = T$, so T is self-adjoint, and $T^2 = T^*T = T$. By Corollary 6.6.29, $T = \operatorname{Pr}_{\operatorname{im} T}$.

By (c) and (d) of Proposition 6.6.18,

$$T^* = (T^*T)^* = T^*(T^*)^* = T^*T = T.$$

So T is self-adjoint, and then $T^2 = T^*T = T$. By Corollary 6.6.29, T is the orthogonal projection onto $\operatorname{im} T$.

Solution to Exercise 6.6.36

Uses. Proposition 6.6.18 (Properties of the adjoint), Corollary 6.6.28 (Self-adjointness in an orthonormal basis)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Compute each adjoint with Proposition 6.6.18 and compare with the original operator.
Hint 2. For (b), $(ST)^* = TS$. Look for two symmetric 2×2 matrices that do not commute.
Sketch. (a) True. (b) False in general: $(ST)^* = TS$, and $S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ give $ST = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; ST is self-adjoint exactly when $ST = TS$. (c) True: $(STS)^* = STS$. (d) False unless $T = 0$: $(iT)^* = -iT$. (e) True: $\left(T^{-1}\right)^* = (T^*)^{-1} = T^{-1}$.

(a) True: $(S + T)^* = S^* + T^* = S + T$ by Proposition 6.6.18(a).

(b) False in general. By Proposition 6.6.18(d), $(ST)^* = T^*S^* = TS$, so ST is self-adjoint exactly when $ST = TS$. On $\mathbb{R}^2$ with the dot product, $S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ are symmetric, but $ST = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is not, so $T_S T_T$ is not self-adjoint (Corollary 6.6.28).

(c) True: $(STS)^* = S^*T^*S^* = STS$, applying Proposition 6.6.18(d) twice.

(d) False unless $T = 0$. By Proposition 6.6.18(b), $(iT)^* = \bar{i}T^* = -iT$, which equals iT only if $2iT = 0$, that is, $T = 0$. So the self-adjoint operators on a complex space are closed under addition and real scalar multiples, but they do not form a complex subspace of $\mathrm{L}(V)$.

(e) True: by Proposition 6.6.18(f), $\left(T^{-1}\right)^* = (T^*)^{-1} = T^{-1}$.

Solution to Exercise 6.6.37

Uses. Proposition 6.6.18 (Properties of the adjoint), Definition 6.6.26 (Self-adjoint operator), Example 6.6.10 (The adjoint of a matrix)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. Imitate $x = (z + \bar{z})/2$ and $y = (z - \bar{z})/(2i)$.
Hint 2. For uniqueness, if $A + iB = A' + iB'$, then $C = A - A'$ is self-adjoint and also i times a self-adjoint operator. Compute C^* in two ways.
Sketch. (a) $A = (T + T^*)/2$ and $B = (T - T^*)/(2i)$ are self-adjoint with $A + iB = T$. If also $T = A' + iB'$, then $C = A - A' = i(B' - B)$ satisfies $C^* = C$ and $C^* = -i(B' - B) = -C$, so $C = 0$. (b) $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & -i \\ i & 1 \end{smallmatrix}\right)$.

(a) Let $A = (T + T^*)/2$ and $B = (T - T^*)/(2i)$. By Proposition 6.6.18, $A^* = (T^* + T)/2 = A$, and since $\overline{1/(2i)} = -1/(2i)$,

$$B^* = -\frac{1}{2i}(T^* - T) = \frac{1}{2i}(T - T^*) = B.$$

Also $A + iB = (T + T^*)/2 + (T - T^*)/2 = T$. For uniqueness, suppose $T = A' + iB'$ with A', B' self-adjoint, and let $C = A - A' = i(B' - B)$. Then C is self-adjoint, as a difference of self-adjoint operators, and also $C^* = \bar{i}(B' - B)^* = -i(B' - B) = -C$. So $C = -C$, hence $C = 0$, $A = A'$, and $B' - B = -iC = 0$.

(b) By Example 6.6.10 we may compute with matrices. Here $M^* = \left(\begin{smallmatrix} 1 & 0 \\ 2 & -i \end{smallmatrix}\right)$, so

$$A = \frac{1}{2}(M + M^*) = \frac{1}{2}\begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

$$B = \frac{1}{2i}(M - M^*) = \frac{1}{2i}\begin{pmatrix} 0 & 2 \\ -2 & 2i \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 1 \end{pmatrix}.$$

Both are Hermitian, and $A + iB = \left(\begin{smallmatrix} 1 & 1+1 \\ 1-1 & i \end{smallmatrix}\right) = M$.

Solution to Exercise 6.6.38

Uses. Definition 6.6.8 (Adjoint), Definition 6.4.22 (Orthogonal complement), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.4.26 (The double complement)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Take $\mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$, and move T^* across in $\langle T^*\mathbf{w}, \mathbf{u}\rangle$.
Hint 2. For the converse, apply the first direction to T^* and the subspace $U^\perp$.
Sketch. If $T(U) \subseteq U$, then for $\mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$, $\langle T^*\mathbf{w}, \mathbf{u}\rangle = \langle \mathbf{w}, T\mathbf{u}\rangle = 0$, so $T^*\mathbf{w} \in U^\perp$. Conversely apply this to T^* and $U^\perp$, using $(T^*)^* = T$ and $\left(U^\perp\right)^\perp = U$.

Suppose T maps U into U. Let $\mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$. Then $T\mathbf{u} \in U$, so

$$\langle T^*\mathbf{w}, \mathbf{u}\rangle = \langle \mathbf{w}, T\mathbf{u}\rangle = 0.$$

As $\mathbf{u} \in U$ was arbitrary, $T^*\mathbf{w} \in U^\perp$. Conversely, suppose T^* maps $U^\perp$ into $U^\perp$. Applying what we just proved to the operator T^* and the subspace $U^\perp$, the operator $(T^*)^* = T$ (Proposition 6.6.18(c)) maps $\left(U^\perp\right)^\perp$ into itself, and $\left(U^\perp\right)^\perp = U$ by Theorem 6.4.26. If T is self-adjoint and maps U into U, then $T^* = T$ maps $U^\perp$ into $U^\perp$. This is the step that lets the spectral theorem (Section 8.2) split V into perpendicular pieces one at a time.

Solution to Exercise 6.6.39

Uses. Proposition 6.6.9 (Adjoints are unique and linear), Theorem 6.6.15 (Existence of adjoints), Proposition 6.6.18 (Properties of the adjoint)
Hint 1. Integrate by parts: $\int_0^{2\pi} f'g = [fg]_0^{2\pi} - \int_0^{2\pi} fg'$. What is $[fg]_0^{2\pi}$ for these functions?

Hint 2. For (c), normalizing the basis divides $\cos kt$ and $\sin kt$ by the same number, so the 2×2 blocks of the matrix do not change. Then use Theorem 6.6.15(b).
Sketch. (a) D sends each spanning function into V. Every $f \in V$ satisfies $f(2\pi) = f(0)$, so integration by parts gives $\langle Df, g\rangle = -\langle f, Dg\rangle$, that is, $D^* = -D$. (b) $\langle Df, f\rangle = -\langle f, Df\rangle = -\langle Df, f\rangle$. (c) The matrix is block diagonal with blocks (0), $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$; it stays the same in the orthonormal basis obtained by normalizing, and it is antisymmetric.

(a) $D1 = 0$, $D\cos kt = -k\sin kt$, and $D\sin kt = k\cos kt$, so D maps the spanning functions, and hence all of V, into V. Every $f \in V$ is 2π-periodic, so $f(2\pi) = f(0)$. For $f, g \in V$, integration by parts gives

$$\langle Df, g\rangle = \int_0^{2\pi} f'(t)g(t)\,dt = f(2\pi)g(2\pi) - f(0)g(0) - \int_0^{2\pi} f(t)g'(t)\,dt = \langle f, -Dg\rangle.$$

By Proposition 6.6.9, $D^* = -D$.

(b) For $f \in V$, $\langle Df, f\rangle = \langle f, -Df\rangle = -\langle f, Df\rangle = -\langle Df, f\rangle$, by symmetry of a real inner product. So $\langle Df, f\rangle = 0$. Since $D\sin t = \cos t \neq 0$, $D \neq 0$, so $D^* = -D \neq D$, and D is not self-adjoint.

(c) From the formulas in (a), the matrix of D in the basis $1, \cos t, \sin t, \cos 2t, \sin 2t$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & -2 & 0 \end{pmatrix}.$$

The basis is orthogonal, with $\|1\| = \sqrt{2\pi}$ and $\|\cos kt\| = \|\sin kt\| = \sqrt{\pi}$. Dividing each basis function by its norm rescales $\cos kt$ and $\sin kt$ by the same factor, so in the resulting orthonormal basis the matrix is unchanged. It is antisymmetric, so by Theorem 6.6.15(b) the matrix of D^* is its transpose, which is its negative: $D^* = -D$ again.

Solution to Exercise 6.6.40

Uses. Exercise 6.1.51, Corollary 6.1.38 (Linear maps are determined by inner products), Definition 6.6.8 (Adjoint), Theorem 6.3.29 (Orthonormal columns), Example 6.6.10 (The adjoint of a matrix)
Techniques. Proving TFAE (Technique 2.7.1), Move an operator across the inner product (Technique 6.6.20)
Hint 1. (i) $\Rightarrow$ (ii) is polarization, Exercise 6.1.51.
Hint 2. For (ii) $\Rightarrow$ (iii), move T out of the second slot: $\langle T\mathbf{v}, T\mathbf{w}\rangle = \langle T^*T\mathbf{v}, \mathbf{w}\rangle$. Then use Corollary 6.1.38.
Sketch. (i) $\Rightarrow$ (ii) by polarization. (ii) $\Rightarrow$ (iii): $\langle T^*T\mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, T\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w}$, so $T^*T = \mathrm{id}_V$. (iii) $\Rightarrow$ (i): $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|^2$. For T_A, (iii) is $A^*A = I_n$: the columns of A are orthonormal.

We prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i). The first implication is Exercise 6.1.51. For the second, the conjugated form of the defining identity, $\langle \mathbf{x}, T\mathbf{w}\rangle = \langle T^*\mathbf{x}, \mathbf{w}\rangle$, with $\mathbf{x} = T\mathbf{v}$ gives

$$\langle T^*T\mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, T\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathrm{id}_V\mathbf{v}, \mathbf{w}\rangle$$

for all $\mathbf{v}, \mathbf{w} \in V$, so $T^*T = \mathrm{id}_V$ by Corollary 6.1.38. For the third, $\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|^2$.

For $T = T_A$, $T^*T = T_{A^*A}$ (Example 6.6.10), so (iii) says $A^*A = I_n$, that is, the columns of A are orthonormal (Theorem 6.3.29). Such maps are studied in Section 8.4.

Solution to Exercise 6.6.41

Uses. Definition 2.5.12 (Invertible), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.6.18 (Properties of the adjoint), Proposition 2.5.4 (Left inverses give uniqueness), Theorem 4.2.11 (Rank–nullity)
Hint 1. $T^*T = \mathrm{id}_V$ says that T^* is a left inverse. Try $T : \mathbb{R} \to \mathbb{R}^2$, $Tx = (x, 0)$.
Hint 2. For (b), check that $P = TT^*$ satisfies $P^2 = P = P^*$, and compare $\mathrm{im}\,P$ with $\mathrm{im}\,T$.
Hint 3. For (c), T is injective. What does rank–nullity say when $\dim V = \dim W$?
Sketch. (a) A left inverse need not be a right inverse; for $Tx = (x, 0)$, $T^*(y_1, y_2) = y_1$, $T^*T = \mathrm{id}_{\mathbb{R}}$, but $TT^*(y_1, y_2) = (y_1, 0)$. (b) $(TT^*)^2 = T(T^*T)T^* = TT^*$ and $(TT^*)^* = TT^*$, so TT^* is the orthogonal projection onto its image, which is $\mathrm{im}\,T$ because $T = (TT^*)T$. (c) T is injective, so if $\dim V = \dim W$ it is surjective, and $TT^* = \mathrm{Pr}_W = \mathrm{id}_W$.

(a) The equation $T^*T = \mathrm{id}_V$ only says that T^* is a left inverse of T; an inverse must also be a right inverse (Definition 2.5.12). Let $T : \mathbb{R} \to \mathbb{R}^2$, $Tx = (x, 0)$, with dot products. Then $\langle Tx, \mathbf{y}\rangle = xy_1 = \langle x, y_1\rangle$, so $T^*(y_1, y_2) = y_1$ and $T^*Tx = x$. But $TT^*(y_1, y_2) = (y_1, 0)$, which is not $\mathrm{id}_{\mathbb{R}^2}$.

(b) Let $P = TT^*$. By Proposition 6.6.18,

$$P^2 = T(T^*T)T^* = TT^* = P, \quad P^* = (T^*)^*T^* = TT^* = P.$$

By Corollary 6.6.29, $P = \mathrm{Pr}_{\mathrm{im}\, P}$. Clearly $\mathrm{im}\, P \subseteq \mathrm{im}\, T$, and $T = T(T^*T) = PT$ shows $\mathrm{im}\, T \subseteq \mathrm{im}\, P$. So $TT^* = \mathrm{Pr}_{\mathrm{im}\, T}$.

(c) T has a left inverse, so it is injective (Proposition 2.5.4). By Theorem 4.2.11, $\mathrm{rank}\, T = \dim V = \dim W$, so $\mathrm{im}\, T = W$, and by (b), $TT^* = \mathrm{Pr}_W = \mathrm{id}_W$.

Solution to Exercise 6.6.42

Uses. Definition 6.6.8 (Adjoint), Theorem 6.2.1 (Cauchy–Schwarz inequality), Notation 6.1.1 (Real and complex scalars), Proposition 6.6.18 (Properties of the adjoint)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. For (a), move T across and apply Cauchy–Schwarz to $\langle \mathbf{v}, T^*\mathbf{w}\rangle$.
Hint 2. For (b), write $\|T^*\mathbf{w}\|^2 = \langle TT^*\mathbf{w}, \mathbf{w}\rangle$ and bound it by Cauchy–Schwarz and the hypothesis on T.
Sketch. (a) $\mathrm{Re}\langle T\mathbf{v}, \mathbf{w}\rangle = \mathrm{Re}\langle \mathbf{v}, T^*\mathbf{w}\rangle \le |\langle \mathbf{v}, T^*\mathbf{w}\rangle| \le \|T^*\mathbf{w}\|$, with equality at $\mathbf{v} = T^*\mathbf{w}/\|T^*\mathbf{w}\|$. (b) $\|T^*\mathbf{w}\|^2 = \langle TT^*\mathbf{w}, \mathbf{w}\rangle \le \|TT^*\mathbf{w}\|\|\mathbf{w}\| \le c\|T^*\mathbf{w}\|\|\mathbf{w}\|$; divide by $\|T^*\mathbf{w}\|$. The converse is the same argument for T^*, whose adjoint is T.

(a) For a unit vector $\mathbf{v}$, by $|\mathrm{Re}\, z| \le |z|$ (Notation 6.1.1) and Cauchy–Schwarz (Theorem 6.2.1),

$$\mathrm{Re}\langle T\mathbf{v}, \mathbf{w}\rangle = \mathrm{Re}\langle \mathbf{v}, T^*\mathbf{w}\rangle \le |\langle \mathbf{v}, T^*\mathbf{w}\rangle| \le \|\mathbf{v}\|\|T^*\mathbf{w}\| = \|T^*\mathbf{w}\|.$$

If $T^*\mathbf{w} \neq \mathbf{0}$ and $\mathbf{v} = T^*\mathbf{w}/\|T^*\mathbf{w}\|$, then $\langle \mathbf{v}, T^*\mathbf{w}\rangle = \|T^*\mathbf{w}\|^2/\|T^*\mathbf{w}\| = \|T^*\mathbf{w}\|$, a real number, so equality holds.

(b) Suppose $\|T\mathbf{v}\| \le c\|\mathbf{v}\|$ for all $\mathbf{v}$, and let $\mathbf{w} \in W$. By the defining identity with $\mathbf{v} = T^*\mathbf{w}$, Cauchy–Schwarz, and the hypothesis,

$$\|T^*\mathbf{w}\|^2 = \langle TT^*\mathbf{w}, \mathbf{w}\rangle \le \|TT^*\mathbf{w}\|\|\mathbf{w}\| \le c\|T^*\mathbf{w}\|\|\mathbf{w}\|.$$

If $T^*\mathbf{w} \neq \mathbf{0}$, dividing by $\|T^*\mathbf{w}\|$ gives $\|T^*\mathbf{w}\| \le c\|\mathbf{w}\|$; if $T^*\mathbf{w} = \mathbf{0}$ this is clear. Conversely, T^* has adjoint T (Proposition 6.6.18(c)), so the same argument with the roles of T and T^* exchanged proves the other implication.

Solution to Exercise 6.6.43

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Definition 6.3.8 (Vectors orthogonal to a set), Theorem 4.2.11 (Rank–nullity), Example 6.6.10 (The adjoint of a matrix)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Every vector of $\mathrm{im}\, T$ is orthogonal to $\mathrm{im}\, T^*$ exactly when $\mathrm{im}\, T \subseteq (\mathrm{im}\, T^*)^\perp$. Which subspace is $(\mathrm{im}\, T^*)^\perp$?
Hint 2. For (b), $\mathrm{im}\, T \subseteq \ker T$; apply rank–nullity.
Sketch. (a) The condition says $\mathrm{im}\, T \subseteq (\mathrm{im}\, T^*)^\perp = \ker T$ (Theorem 6.6.21(b)), that is, $T(T\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$. (b) $\mathrm{rank}\, T \le \dim \ker T = \dim V - \mathrm{rank}\, T$. (c) T_N with N sending $\mathbf{e}_3 \mapsto \mathbf{e}_1$ and killing $\mathbf{e}_1, \mathbf{e}_2$: $\mathrm{im}\, T_N = \mathrm{span}(\mathbf{e}_1)$ and $\mathrm{im}\, T_N^* = \mathrm{span}(\mathbf{e}_3)$.

(a) Every vector of $\mathrm{im}\, T$ is orthogonal to every vector of $\mathrm{im}\, T^*$ if and only if $\mathrm{im}\, T \subseteq (\mathrm{im}\, T^*)^\perp$, and $(\mathrm{im}\, T^*)^\perp = \ker T$ by Theorem 6.6.21(b). Now $\mathrm{im}\, T \subseteq \ker T$ means $T(T\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$, that is, $T^2 = 0$.

(b) If $T^2 = 0$, then $\mathrm{im}\, T \subseteq \ker T$, so by Theorem 4.2.11, $\mathrm{rank}\, T \le \dim \ker T = \dim V - \mathrm{rank}\, T$, and $2\,\mathrm{rank}\, T \le \dim V$.

(c) Let $N = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, so that $N\mathbf{e}_3 = \mathbf{e}_1$ and $N\mathbf{e}_1 = N\mathbf{e}_2 = \mathbf{0}$. Then $N^2\mathbf{e}_3 = N\mathbf{e}_1 = \mathbf{0}$, so $N^2 = 0$, while $N \neq 0$. Here $\mathrm{im}\, T_N = \mathrm{span}(\mathbf{e}_1)$, and $T_N^* = T_{N^\mathsf{T}}$ (Example 6.6.10) has $N^\mathsf{T}\mathbf{e}_1 = \mathbf{e}_3$, so $\mathrm{im}\, T_N^* = \mathrm{span}(\mathbf{e}_3)$, which is orthogonal to $\mathbf{e}_1$. Also $\mathrm{rank}\, N = 1 \le 3/2$.

Solution to Exercise 6.6.44

Uses. Proposition 6.6.23 (T^*T and T), Corollary 6.6.22 (T and T^* have the same rank), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.6.21 (Kernels and images of adjoints), Definition 6.6.8 (Adjoint)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. For (a), apply Proposition 6.6.23 to T and to T^*.
Hint 2. For (b), $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle$ and $\|T^*\mathbf{v}\|^2 = \langle TT^*\mathbf{v}, \mathbf{v}\rangle$. For the images, use Theorem 6.6.21(d) and (c).
Sketch. (a) $\operatorname{rank}(T^*T) = \operatorname{rank} T$, and applied to T^* (whose adjoint is T), $\operatorname{rank}(TT^*) = \operatorname{rank} T^* = \operatorname{rank} T$. (b) $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \|T^*\mathbf{v}\|^2$, so the kernels agree, and $\operatorname{im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{im} T^*$. (c) $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

(a) By Proposition 6.6.23(b), $\operatorname{rank}(T^*T) = \operatorname{rank} T$. Applying the same result to T^*, whose adjoint is T (Proposition 6.6.18(c)), gives $\operatorname{rank}(TT^*) = \operatorname{rank} T^*$, which is $\operatorname{rank} T$ by Corollary 6.6.22.

(b) By the defining identity of T^*, and of $(T^*)^* = T$,

$$\|T\mathbf{v}\|^2 = \langle T\mathbf{v}, T\mathbf{v}\rangle = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \langle T^*\mathbf{v}, T^*\mathbf{v}\rangle = \|T^*\mathbf{v}\|^2.$$

So $T\mathbf{v} = \mathbf{0}$ if and only if $T^*\mathbf{v} = \mathbf{0}$: $\ker T = \ker T^*$. By (d) and (c) of Theorem 6.6.21, $\operatorname{im} T = (\ker T^*)^\perp = (\ker T)^\perp = \operatorname{im} T^*$.

(c) For $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $\ker N = \operatorname{span}(\mathbf{e}_1)$, while $\ker N^{\mathsf{T}} = \operatorname{span}(\mathbf{e}_2)$. Operators with $T^*T = TT^*$ are called normal; they are the subject of Section 8.1.

Solution to Exercise 6.6.45

Uses. Definition 6.6.26 (Self-adjoint operator), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.6.18 (Properties of the adjoint)
Hint 1. Substitute $s = 1 - t$ in $\int_0^1 p(1-t)q(t)\,dt$. For the second inner product, test $p = x$ and $q = 1$.
Hint 2. $S^2 = \mathrm{id}$. Check $P^2 = P = P^*$ for $P = (\mathrm{id} + S)/2$, and identify $\operatorname{im} P$.
Sketch. (a) The substitution $s = 1 - t$ maps $[0,1]$ to itself, giving $\langle Sp, q\rangle = \langle p, Sq\rangle$. On $[0,2]$, $p = x$, $q = 1$ give $\langle Sp, q\rangle = 0$ but $\langle p, Sq\rangle = 2$. (b) $S^2 = \mathrm{id}$, so $P^2 = P$, and $P^* = P$. $\operatorname{im} P = \{p \mid Sp = p\} = E$, with basis $1, x^2 - x$.

(a) For the first inner product, substituting $s = 1 - t$,

$$\langle Sp, q\rangle = \int_0^1 p(1-t)q(t)\,dt = \int_0^1 p(s)q(1-s)\,ds = \langle p, Sq\rangle.$$

For the second, take $p = x$ and $q = 1$: $\langle Sp, q\rangle = \int_0^2 (1-t)\,dt = 2 - 2 = 0$, while $\langle p, Sq\rangle = \int_0^2 t\,dt = 2$. So S is not self-adjoint for $\int_0^2$; the substitution maps $[0,2]$ to $[-1,1]$, not to itself.

(b) $S^2p(x) = p(1-(1-x)) = p(x)$, so $S^2 = \mathrm{id}$. Let $P = (\mathrm{id} + S)/2$. Then $P^2 = \left(\mathrm{id} + 2S + S^2\right)/4 = (\mathrm{id} + S)/2 = P$, and $P^* = (\mathrm{id} + S^*)/2 = P$ by (a) and Proposition 6.6.18. By Corollary 6.6.29, $P = \mathrm{Pr}_{\operatorname{im} P}$. If $Sp = p$, then $Pp = p$, so $E \subseteq \operatorname{im} P$; conversely $SP = \left(S + S^2\right)/2 = P$, so S fixes every vector of $\operatorname{im} P$. Hence $\operatorname{im} P = E$. For $p = a + bx + cx^2$,

$$p(1-x) = (a + b + c) - (b + 2c)x + cx^2,$$

which equals $p(x)$ if and only if $b = -c$. So $E = \left\{a + c\left(x^2 - x\right)\right\}$, with basis $1, x^2 - x$, and $\dim E = 2$.

Solution to Exercise 6.6.46

Uses. Proposition 6.1.4 (Sesquilinearity), Definition 6.3.28 (Gram matrix), Proposition 6.3.30 (Gram matrices detect independence), Theorem 2.3.9 (The matrix of a map computes the map), Proposition 6.3.27 (Properties of the conjugate transpose), Theorem 6.3.22 (Parseval's identity), Example 6.6.16 (The adjoint of differentiation)
Techniques. Translate between maps and matrices (Technique 2.3.27), Gram matrices (Technique 6.3.31)
Hint 1. For (a), expand $\left\langle \sum_j x_j \mathbf{b}_j, \sum_i y_i \mathbf{b}_i \right\rangle$ by sesquilinearity, as in Proposition 6.1.22(a).
Hint 2. For (b), with $A = [T]_{\mathcal{B}}$ and $C = [T^*]_{\mathcal{B}}$, (a) turns the defining identity into $\mathbf{y}^*GA\mathbf{x} = \mathbf{y}^*C^*G\mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$. Conclude $GA = C^*G$; is G invertible?
Sketch. (a) $\langle \mathbf{v}, \mathbf{w}\rangle = \sum_{i,j} \overline{y_i} x_j \langle \mathbf{b}_j, \mathbf{b}_i\rangle = \mathbf{y}^*G\mathbf{x}$. (b) The defining identity reads $\mathbf{y}^*GA\mathbf{x} = \mathbf{y}^*C^*G\mathbf{x}$ for all $\mathbf{x}, \mathbf{y}$; taking standard basis vectors gives $GA = C^*G$. G is invertible (Gram matrix of an independent list) and Hermitian, so $C = G^{-1}A^*G$. (c) $G^{-1}\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)G = \left(\begin{smallmatrix} -6 & -3 \\ 12 & 6 \end{smallmatrix}\right)$. (d) The formula $\langle \mathbf{v}, \mathbf{w}\rangle = [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}}$ is Parseval's identity, valid only for orthonormal $\mathcal{B}$; the claim holds whenever G is a scalar multiple of I_n.

(a) Write $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ and $\mathbf{y} = [\mathbf{w}]_{\mathcal{B}}$. By linearity in the first slot and Proposition 6.1.4,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \left\langle \sum_j x_j \mathbf{b}_j, \sum_i y_i \mathbf{b}_i \right\rangle = \sum_{i,j} \overline{y_i} \langle \mathbf{b}_j, \mathbf{b}_i \rangle x_j = \sum_i \overline{y_i} \left(\sum_j g_{ij} x_j \right) = \mathbf{y}^* G \mathbf{x}.$$

(b) Let $A = [T]_{\mathcal{B}}$ and $C = [T^*]_{\mathcal{B}}$. For $\mathbf{v}, \mathbf{w} \in V$ with coordinates $\mathbf{x}, \mathbf{y}$, Theorem 2.3.9 gives $[T\mathbf{v}]_{\mathcal{B}} = A\mathbf{x}$ and $[T^*\mathbf{w}]_{\mathcal{B}} = C\mathbf{y}$, so by (a) and Proposition 6.3.27,

$$\langle T\mathbf{v}, \mathbf{w} \rangle = \mathbf{y}^* G A \mathbf{x}, \quad \langle \mathbf{v}, T^*\mathbf{w} \rangle = (C\mathbf{y})^* G \mathbf{x} = \mathbf{y}^* C^* G \mathbf{x}.$$

These agree for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, since coordinates range over all of $\mathbb{F}^n$. Taking $\mathbf{x} = \mathbf{e}_j$ and $\mathbf{y} = \mathbf{e}_i$ shows that GA and C^*G have the same (i, j) entry, so $GA = C^*G$. The Gram matrix of the independent list $\mathcal{B}$ is invertible (Proposition 6.3.30), and $G^* = G$, so $C^* = GAG^{-1}$ and

$$C = \left(GAG^{-1}\right)^* = \left(G^{-1}\right)^* A^* G^* = G^{-1} A^* G,$$

using Proposition 6.3.27 and $\left(G^{-1}\right)^* = (G^*)^{-1} = G^{-1}$.

(c) For $\mathcal{B} = 1, x$ with $\langle p, q \rangle = \int_0^1 pq$, $G = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}$, whose determinant is $1/12$, so $G^{-1} = \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix}$. With $[D]_{\mathcal{B}}^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$,

$$G^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} G = \begin{pmatrix} -6 & 0 \\ 12 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} = \begin{pmatrix} -6 & -3 \\ 12 & 6 \end{pmatrix},$$

in agreement with Example 6.6.16.

(d) The identity $\langle \mathbf{v}, \mathbf{w} \rangle = [\mathbf{v}]_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}}$ is Parseval's identity (Theorem 6.3.22), which needs $\mathcal{B}$ orthonormal. By (a) it holds exactly when $G = I_n$. If $G = cI_n$ for some $c > 0$, that is, the basis vectors are orthogonal and of equal length, then (b) gives $[T^*]_{\mathcal{B}} = A^*$ as well. For other bases the claim can fail, as in (c).

Solution to Exercise 6.6.47

Uses. Definition 4.4.4 (Direct sum), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.6.21 (Kernels and images of adjoints), Theorem 6.4.26 (The double complement), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents)
Hint 1. Any $Q \in \mathrm{L}(V)$ with $Q^2 = Q$ satisfies $V = \operatorname{im} Q \oplus \ker Q$ and is the projection onto $\operatorname{im} Q$ along $\ker Q$: split $\mathbf{v} = Q\mathbf{v} + (\mathbf{v} - Q\mathbf{v})$.
Hint 2. Apply this to $Q = P^*$, using $(P^*)^2 = \left(P^2\right)^*$ and Theorem 6.6.21 for $\operatorname{im} P^*$ and $\ker P^*$.
Sketch. (a) is immediate from uniqueness of the decomposition. (b) An idempotent Q is the projection onto $\operatorname{im} Q$ along $\ker Q$, and $V = \operatorname{im} Q \oplus \ker Q$. Now $(P^*)^2 = \left(P^2\right)^* = P^*$, $\operatorname{im} P^* = (\ker P)^\perp = X^\perp$, and $\ker P^* = (\operatorname{im} P)^\perp = U^\perp$. (c) P is orthogonal iff $P = P^*$ (Corollary 6.6.29), iff they have the same image and kernel, iff $U = X^\perp$, iff $X = U^\perp$.

(a) For $\mathbf{u} \in U$, the decomposition of $\mathbf{u}$ is $\mathbf{u} + \mathbf{0}$, so $P\mathbf{u} = \mathbf{u}$. Hence $P(P\mathbf{v}) = P\mathbf{v}$ for every $\mathbf{v}$, since $P\mathbf{v} \in U$, and $\operatorname{im} P = U$. Finally $P(\mathbf{u} + \mathbf{x}) = \mathbf{0}$ exactly when $\mathbf{u} = \mathbf{0}$, so $\ker P = X$.

(b) First a general fact. Let $Q \in \mathrm{L}(V)$ with $Q^2 = Q$. Every $\mathbf{v}$ is $Q\mathbf{v} + (\mathbf{v} - Q\mathbf{v})$, with $Q\mathbf{v} \in \operatorname{im} Q$ and $Q(\mathbf{v} - Q\mathbf{v}) = Q\mathbf{v} - Q^2\mathbf{v} = \mathbf{0}$. If $\mathbf{y} \in \operatorname{im} Q \cap \ker Q$, say $\mathbf{y} = Q\mathbf{z}$, then $\mathbf{0} = Q\mathbf{y} = Q^2\mathbf{z} = Q\mathbf{z} = \mathbf{y}$. So $V = \operatorname{im} Q \oplus \ker Q$, and Q sends $\mathbf{v} = Q\mathbf{v} + (\mathbf{v} - Q\mathbf{v})$ to its component $Q\mathbf{v}$ in $\operatorname{im} Q$: Q is the projection onto $\operatorname{im} Q$ along $\ker Q$.

Apply this to $Q = P^*$. By Proposition 6.6.18(d), $(P^*)^2 = \left(P^2\right)^* = P^*$. By Theorem 6.6.21 and (a), $\operatorname{im} P^* = (\ker P)^\perp = X^\perp$ and $\ker P^* = (\operatorname{im} P)^\perp = U^\perp$. So $V = X^\perp \oplus U^\perp$, and P^* is the projection onto $X^\perp$ along $U^\perp$.

(c) Since $P^2 = P$, Corollary 6.6.29 says that P is an orthogonal projection if and only if $P = P^*$. A projection along a complement is determined by its image and kernel (by the general fact in (b)), so $P = P^*$ if and only if $U = X^\perp$ and $X = U^\perp$. These two conditions are equivalent to each other by Theorem 6.4.26: if $X = U^\perp$, then $X^\perp = \left(U^\perp\right)^\perp = U$, and conversely. So P is an orthogonal projection if and only if $X = U^\perp$.

Solution to Exercise 6.6.48

Uses. Proposition 6.1.22 (Inner products on $\mathbb{F}^n$), Definition 6.6.26 (Self-adjoint operator)
Techniques. Undetermined coefficients (Technique 3.1.9)
Hint 1. For $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}} M\mathbf{x}$, show that T_A is self-adjoint if and only if MA is symmetric.
Hint 2. Write $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. For (a), solve for the symmetry condition and then make M positive definite. For (b), the symmetry condition forces $c = -a$; why is that impossible?
Sketch. $\langle A\mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}} MA\mathbf{x}$ and $\langle \mathbf{x}, A\mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}} A^{\mathsf{T}} M\mathbf{x}$, so self-adjointness means $MA = A^{\mathsf{T}} M = (MA)^{\mathsf{T}}$. (a) $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ gives $MA = \begin{pmatrix} a & a+2b \\ b & b+2c \end{pmatrix}$, symmetric iff $b = -a$; take $M = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. (b) $MJ = \begin{pmatrix} b & -a \\ c & -b \end{pmatrix}$ is symmetric iff $c = -a$, but $a = \mathbf{e}_1^{\mathsf{T}} M\mathbf{e}_1 > 0$ and $c = \mathbf{e}_2^{\mathsf{T}} M\mathbf{e}_2 > 0$.

By Proposition 6.1.22, the inner products on $\mathbb{R}^2$ are exactly $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}} M\mathbf{x}$ with $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ symmetric and $\mathbf{x}^{\mathsf{T}} M\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$. For such an inner product and any $A \in \mathbb{R}^{2\times 2}$,

$$\langle A\mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^{\mathsf{T}}(MA)\mathbf{x}, \quad \langle \mathbf{x}, A\mathbf{y}\rangle = (A\mathbf{y})^{\mathsf{T}} M\mathbf{x} = \mathbf{y}^{\mathsf{T}}\left(A^{\mathsf{T}} M\right)\mathbf{x}.$$

These agree for all $\mathbf{x}, \mathbf{y}$ if and only if $MA = A^{\mathsf{T}} M$ (test on standard basis vectors), that is, $MA = (MA)^{\mathsf{T}}$, since $(MA)^{\mathsf{T}} = A^{\mathsf{T}} M^{\mathsf{T}} = A^{\mathsf{T}} M$.

(a) $MA = \begin{pmatrix} a & a+2b \\ b & b+2c \end{pmatrix}$ is symmetric if and only if $a + 2b = b$, that is, $b = -a$. Take $a = 1, b = -1, c = 2$. Then $M = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ is symmetric and $\mathbf{x}^{\mathsf{T}} M\mathbf{x} = x_1^2 - 2x_1x_2 + 2x_2^2 = (x_1 - x_2)^2 + x_2^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$. So T_A is self-adjoint for

$$\langle \mathbf{x}, \mathbf{y}\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2.$$

For instance, $A\mathbf{e}_1 = \mathbf{e}_1$ and $A(1,1) = 2(1,1)$, and these two vectors are orthogonal for this inner product: $\langle \mathbf{e}_1, (1,1)\rangle = 1 - 1 = 0$. That is what Section 8.2 leads one to expect.

(b) $MJ = \begin{pmatrix} b & -a \\ c & -b \end{pmatrix}$ is symmetric if and only if $c = -a$. But $a = \mathbf{e}_1^{\mathsf{T}} M\mathbf{e}_1 > 0$ and $c = \mathbf{e}_2^{\mathsf{T}} M\mathbf{e}_2 > 0$, so $c \neq -a$. Hence T_J is self-adjoint for no inner product on $\mathbb{R}^2$. (By Section 8.2, a self-adjoint operator on $\mathbb{R}^2$ maps some line into itself, and a rotation by $\pi/2$ maps no line into itself.)

Solution to Exercise 6.6.49

Uses. Theorem 6.6.3 (Riesz representation theorem), Theorem 6.2.1 (Cauchy–Schwarz inequality)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. For (a), apply Cauchy–Schwarz to $p(0) = \langle p, q_n\rangle$.
Hint 2. For (c), if q worked for all polynomials, compare q with q_n for $n \geq \deg q$.
Sketch. (a) $|p(0)| = |\langle p, q_n\rangle| \leq \|p\|\|q_n\|$. (b) $p(0) = 1$ and $\|p\|^2 = \int_0^1 (1-t)^{2n}\,dt = 1/(2n+1)$. (c) Such a q would lie in P_n for $n \geq \deg q$ and represent evaluation there, so $q = q_n$ by uniqueness, and then $\|q\| \geq \sqrt{2n+1}$ for all large n.

(a) By Cauchy–Schwarz, $|p(0)| = |\langle p, q_n\rangle| \leq \|p\|\|q_n\|$. Divide by $\|p\| > 0$.

(b) For $p(t) = (1-t)^n$, $p(0) = 1$ and $\|p\|^2 = \int_0^1 (1-t)^{2n}\,dt = 1/(2n+1)$. By (a), $\|q_n\| \geq 1/\|p\| = \sqrt{2n+1}$. For $n = 1$ this gives $\|q_1\| \geq \sqrt{3}$; in fact $\|q_1\| = 2$ (Example 6.6.5).

(c) Suppose q is a polynomial with $\langle p, q\rangle = p(0)$ for all polynomials p, and let $n \geq \deg q$. Then $q \in P_n$ and $\langle p, q\rangle = p(0)$ for all $p \in P_n$, so $q = q_n$ by the uniqueness in Theorem 6.6.3. By (b), $\|q\| = \|q_n\| \geq \sqrt{2n+1}$ for every $n \geq \deg q$, which is impossible for a fixed q.

Solution to Exercise 6.6.50

Uses. Proposition 6.1.16 (Frobenius inner product), Corollary 6.1.19 ($A^*A = 0$ forces $A = 0$), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.3.27 (Properties of the conjugate transpose)
Techniques. The trace of A^*A (Technique 6.1.18)
Hint 1. It suffices to show that $B = A - A^*$ satisfies $\operatorname{tr}(B^*B) = 0$.
Hint 2. Expand $\operatorname{tr}((A^* - A)(A - A^*))$ into four traces, and express each through $\operatorname{tr}(A^*A)$ using $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, $\operatorname{tr}(X^*) = \overline{\operatorname{tr} X}$, and the hypothesis.
Sketch. With $B = A - A^*$, $\operatorname{tr}(B^*B) = \operatorname{tr}(A^*A) - \operatorname{tr}\left((A^*)^2\right) - \operatorname{tr}\left(A^2\right) + \operatorname{tr}(AA^*)$. Here $\operatorname{tr}\left(A^2\right) = \operatorname{tr}(A^*A)$, which is real, so $\operatorname{tr}\left((A^*)^2\right) = \overline{\operatorname{tr}\left(A^2\right)} = \operatorname{tr}(A^*A)$, and $\operatorname{tr}(AA^*) = \operatorname{tr}(A^*A)$. So $\operatorname{tr}(B^*B) = 0$, and $B = 0$.

Let $B = A - A^*$, so that $B^* = A^* - A$ (Proposition 6.3.27(a)). Then

$$\operatorname{tr}(B^*B) = \operatorname{tr}(A^*A) - \operatorname{tr}\big((A^*)^2\big) - \operatorname{tr}\big(A^2\big) + \operatorname{tr}(AA^*).$$

We express each term through $t = \operatorname{tr}(A^*A) = \sum_{i,j} |a_{ij}|^2$, a real number (Proposition 6.1.16). By hypothesis $\operatorname{tr}\big(A^2\big) = t$. The diagonal entries of X^* are the conjugates of those of X, so $\operatorname{tr}(X^*) = \overline{\operatorname{tr} X}$, and with Proposition 6.3.27(b), $\operatorname{tr}\big((A^*)^2\big) = \operatorname{tr}\big((A^2)^*\big) = \overline{\operatorname{tr}\big(A^2\big)} = \bar{t} = t$. By Proposition 2.4.28, $\operatorname{tr}(AA^*) = \operatorname{tr}(A^*A) = t$. Hence $\operatorname{tr}(B^*B) = t - t - t + t = 0$. By Proposition 6.1.16, $\operatorname{tr}(B^*B) = \sum_{i,j} |b_{ij}|^2$, so every entry of B is 0 (as in Corollary 6.1.19). That is, $A^* = A$.

Solution to Exercise 6.6.51

Uses. Exercise 6.6.42, Proposition 6.1.29 (Expanding a norm), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.6.21 (Kernels and images of adjoints), Theorem 6.4.24 (Orthogonal decomposition), Theorem 6.2.11 (Triangle inequality)
Techniques. Show a vector is zero via its norm (Technique 6.1.36), Move an operator across the inner product (Technique 6.6.20)
Hint 1. By Exercise 6.6.42(b), $\|T^*\mathbf{w}\| \le \|\mathbf{w}\|$ as well. If $T\mathbf{v} = \mathbf{v}$, expand $\|T^*\mathbf{v} - \mathbf{v}\|^2$ and compute $\langle T^*\mathbf{v}, \mathbf{v}\rangle$ by moving T^* across.
Hint 2. For (b), (a) says $\ker(T - \mathrm{id}) = \ker(T^* - \mathrm{id}) = \ker(T - \mathrm{id})^*$. Apply Theorem 6.6.21 to $T - \mathrm{id}$.
Hint 3. For (c), write $\mathbf{v} = \mathbf{u} + (T - \mathrm{id})\mathbf{z}$ by (b). The averages fix $\mathbf{u}$, and the sum over the second part telescopes.
Sketch. (a) If $T\mathbf{v} = \mathbf{v}$, then $\langle T^*\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, T\mathbf{v}\rangle = \|\mathbf{v}\|^2$, so $\|T^*\mathbf{v} - \mathbf{v}\|^2 = \|T^*\mathbf{v}\|^2 - \|\mathbf{v}\|^2 \le 0$. The converse is the same for T^*. (b) $\ker(T - \mathrm{id}) = \ker(T - \mathrm{id})^* = (\operatorname{im}(T - \mathrm{id}))^\perp$. (c) For $\mathbf{v} = \mathbf{u} + (T - \mathrm{id})\mathbf{z}$ with $T\mathbf{u} = \mathbf{u}$: $A_N\mathbf{u} = \mathbf{u}$, and $A_N(T - \mathrm{id})\mathbf{z} = \big(T^N\mathbf{z} - \mathbf{z}\big)/N$ has norm at most $2\|\mathbf{z}\|/N$.

(a) By Exercise 6.6.42(b) with $c = 1$, $\|T^*\mathbf{w}\| \le \|\mathbf{w}\|$ for all $\mathbf{w}$. Suppose $T\mathbf{v} = \mathbf{v}$. Moving T^* across, $\langle T^*\mathbf{v}, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, T^*\mathbf{v}\rangle} = \overline{\langle T\mathbf{v}, \mathbf{v}\rangle} = \overline{\langle \mathbf{v}, \mathbf{v}\rangle} = \|\mathbf{v}\|^2$, which is real. By Proposition 6.1.29,

$$\|T^*\mathbf{v} - \mathbf{v}\|^2 = \|T^*\mathbf{v}\|^2 - 2\operatorname{Re}\langle T^*\mathbf{v}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 = \|T^*\mathbf{v}\|^2 - \|\mathbf{v}\|^2 \le 0,$$

so $T^*\mathbf{v} = \mathbf{v}$. Conversely, T^* also satisfies $\|T^*\mathbf{w}\| \le \|\mathbf{w}\|$, and its adjoint is T (Proposition 6.6.18(c)), so the same argument applied to T^* shows that $T^*\mathbf{v} = \mathbf{v}$ implies $T\mathbf{v} = \mathbf{v}$.

(b) By Proposition 6.6.18, $(T - \mathrm{id})^* = T^* - \mathrm{id}$, so (a) says $\ker(T - \mathrm{id}) = \ker(T - \mathrm{id})^*$. By Theorem 6.6.21(a) applied to $T - \mathrm{id}$, this is $(\operatorname{im}(T - \mathrm{id}))^\perp$. Since $\operatorname{im}(T - \mathrm{id})$ is finite-dimensional, Theorem 6.4.24 gives $V = \operatorname{im}(T - \mathrm{id}) \oplus (\operatorname{im}(T - \mathrm{id}))^\perp = \operatorname{im}(T - \mathrm{id}) \oplus \ker(T - \mathrm{id})$, an orthogonal decomposition.

(c) Let $\mathbf{v} \in V$. By (b), $\mathbf{v} = \mathbf{u} + (T - \mathrm{id})\mathbf{z}$ with $\mathbf{u} \in \ker(T - \mathrm{id})$ and $\mathbf{z} \in V$, and $\mathbf{u} = \operatorname{Pr}_{\ker(T - \mathrm{id})}(\mathbf{v})$ because the second piece lies in $\operatorname{im}(T - \mathrm{id}) = (\ker(T - \mathrm{id}))^\perp$. Since $T\mathbf{u} = \mathbf{u}$, $A_N\mathbf{u} = \mathbf{u}$. The other piece telescopes:

$$A_N(T - \mathrm{id})\mathbf{z} = \frac{1}{N}\sum_{k=0}^{N-1}\big(T^{k+1}\mathbf{z} - T^k\mathbf{z}\big) = \frac{1}{N}\big(T^N\mathbf{z} - \mathbf{z}\big).$$

By the triangle inequality and $\big\|T^N\mathbf{z}\big\| \le \|\mathbf{z}\|$, this has norm at most $2\|\mathbf{z}\|/N$, which tends to 0. Hence $\|A_N\mathbf{v} - \mathbf{u}\| \le 2\|\mathbf{z}\|/N \to 0$. This is the finite-dimensional case of von Neumann's mean ergodic theorem.

Section 6.7

Solution to Exercise 6.7.5

Uses. Corollary 6.4.10 (Projecting with an orthogonal list), Example 6.4.35 (Distance to a hyperplane), Theorem 6.5.5 (The normal equations), Theorem 6.4.33 (Best approximation)
Techniques. Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. In (a), compute the inner product of the two spanning vectors. In (b), the complement of the hyperplane is a line. In (c), there are only two unknowns.
Hint 2. (a) Use the projection formula for an orthogonal list. (b) $\operatorname{dist}(\mathbf{v}, H) = |\langle \mathbf{v}, \mathbf{n}\rangle|/\|\mathbf{n}\|$. (c) Solve $A^{\mathsf{T}}A\mathbf{x} = A^{\mathsf{T}}\mathbf{b}$, a 2×2 system.
Sketch. (a) The spanning vectors are orthogonal, so $\operatorname{Pr}_W(\mathbf{v}) = (9/4)\mathbf{w}_1 + (5/4)\mathbf{w}_2 = (7/2, 1, 7/2, 1)$ and the distance is $\|(-1/2, 0, 1/2, 0)\| = 1/\sqrt{2}$. (b) With $\mathbf{n} = (1, -1, 1, -1)$, the distance is $|-2|/2 = 1$. (c) $A^{\mathsf{T}}A = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}$ and $A^{\mathsf{T}}\mathbf{b} = (5, 11)$ give $\hat{\mathbf{x}} = (2/3, 1/2)$.

(a) The spanning vectors $\mathbf{w}_1 = (1,1,1,1)$ and $\mathbf{w}_2 = (1,-1,1,-1)$ are orthogonal, so no orthogonalization is needed and Corollary 6.4.10 applies directly. Here $\langle \mathbf{v}, \mathbf{w}_1 \rangle = 9$, $\langle \mathbf{v}, \mathbf{w}_2 \rangle = 5$, and $\|\mathbf{w}_1\|^2 = \|\mathbf{w}_2\|^2 = 4$, so

$$\operatorname{Pr}_W(\mathbf{v}) = \frac{9}{4}\mathbf{w}_1 + \frac{5}{4}\mathbf{w}_2 = (7/2, 1, 7/2, 1).$$

Then $\mathbf{v} - \operatorname{Pr}_W(\mathbf{v}) = (-1/2, 0, 1/2, 0)$, and the distance is $1/\sqrt{2}$ by Theorem 6.4.33. As a check, $\|\mathbf{v}\|^2 - \|\operatorname{Pr}_W(\mathbf{v})\|^2 = 27 - 53/2 = 1/2$.

(b) The hyperplane is $H = \{\mathbf{n}\}^\perp$ with $\mathbf{n} = (1,-1,1,-1)$, and its complement is the line through $\mathbf{n}$, so it is quickest to project onto that line. By Example 6.4.35, $\operatorname{dist}(\mathbf{v}, H) = |\langle \mathbf{v}, \mathbf{n} \rangle| / \|\mathbf{n}\| = |1 - 2 + 3 - 4|/2 = 1$.

(c) The columns $(1,1,1)$ and $(1,2,3)$ of the coefficient matrix A are not orthogonal, and there are only two unknowns, so we solve the normal equations (Theorem 6.5.5):

$$A^{\mathsf{T}}A = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}, \quad A^{\mathsf{T}}\mathbf{b} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}, \quad \hat{\mathbf{x}} = \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}\begin{pmatrix} 5 \\ 11 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 1/2 \end{pmatrix}.$$

The residual $\mathbf{b} - A\hat{\mathbf{x}} = (-1/6, 1/3, -1/6)$ is orthogonal to both columns, as it must be.

Solution to Exercise 6.7.6

Uses. Proposition 6.1.29 (Expanding a norm)
Techniques. Perturb a minimizer (Technique 6.5.6), Rotate the phase (Technique 6.2.8)
Hint 1. Expand $\|\mathbf{x} + t\mathbf{y}\|^2 - \|\mathbf{x}\|^2$ as a function of t.
Hint 2. For the converse, try $t = -s\langle \mathbf{x}, \mathbf{y} \rangle$ with a small real $s > 0$. For (b), look at $V = \mathbb{C}$ with $\mathbf{x} = 1$ and $\mathbf{y} = i$.
Sketch. $\|\mathbf{x} + t\mathbf{y}\|^2 - \|\mathbf{x}\|^2 = 2\operatorname{Re}(\bar{t}\langle \mathbf{x}, \mathbf{y} \rangle) + |t|^2\|\mathbf{y}\|^2$. If $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ this is ≥ 0. If $c = \langle \mathbf{x}, \mathbf{y} \rangle \neq 0$, then $t = -sc$ gives $s|c|^2\big(s\|\mathbf{y}\|^2 - 2\big)$, which is negative for small $s > 0$. (b) In $\mathbb{C}$, $|1 + ti|^2 = 1 + t^2 \geq 1$ for real t, but $\langle 1, i \rangle = -i$.

(a) By Proposition 6.1.29, for every $t \in \mathbb{F}$,

$$\|\mathbf{x} + t\mathbf{y}\|^2 - \|\mathbf{x}\|^2 = 2\operatorname{Re}(\bar{t}\langle \mathbf{x}, \mathbf{y} \rangle) + |t|^2\|\mathbf{y}\|^2.$$

If $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, the right side is $|t|^2\|\mathbf{y}\|^2 \geq 0$, so the inequality holds for every t.

Conversely, suppose $c = \langle \mathbf{x}, \mathbf{y} \rangle \neq 0$, and take $t = -sc$ with $s > 0$ real. Then $\bar{t}c = -s|c|^2$, and the right side becomes

$$-2s|c|^2 + s^2|c|^2\|\mathbf{y}\|^2 = s|c|^2\big(s\|\mathbf{y}\|^2 - 2\big).$$

For $0 < s < 2/\big(\|\mathbf{y}\|^2 + 1\big)$ this is negative, so $\|\mathbf{x} + t\mathbf{y}\| < \|\mathbf{x}\|$ for that t. This proves the contrapositive. The choice of t is Technique 6.5.6 in the direction $-c\mathbf{y}$: the factor c rotates the phase so that the linear term is real and negative.

(b) Take $V = \mathbb{C}$ with $\langle z, w \rangle = z\overline{w}$, $\mathbf{x} = 1$, and $\mathbf{y} = i$. For real t, $\|1 + ti\|^2 = 1 + t^2 \geq 1$, but $\langle 1, i \rangle = \bar{i} = -i \neq 0$. Real t only see $\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle$, and the imaginary part needs $t = is$.

The condition in (a) mentions only the norm, so it makes sense in any normed space, where it is used as a definition of orthogonality (*Birkhoff–James orthogonality*). In an inner product space it is ordinary orthogonality. Exercise 6.7.13 puts it to work.

Solution to Exercise 6.7.7

Uses. Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 2.2.1 (Linear maps are determined by a basis)
Techniques. Plug in clever test vectors (Technique 6.1.40), Take an inner product to isolate a coefficient (Technique 6.3.12)
Hint 1. Choose an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. What does the hypothesis say about the coordinates of $T\mathbf{e}_i$?
Hint 2. Each $T\mathbf{e}_i$ is a multiple $c_i\mathbf{e}_i$. To compare c_i with c_j, find two orthogonal vectors built from $\mathbf{e}_i$ and $\mathbf{e}_j$ other than $\mathbf{e}_i, \mathbf{e}_j$ themselves.
Sketch. In an orthonormal basis, $\langle T\mathbf{e}_i, \mathbf{e}_j \rangle = 0$ for $j \neq i$, so $T\mathbf{e}_i = c_i\mathbf{e}_i$. Since $\mathbf{e}_i + \mathbf{e}_j \perp \mathbf{e}_i - \mathbf{e}_j$, $0 = \big\langle T\big(\mathbf{e}_i + \mathbf{e}_j\big), \mathbf{e}_i - \mathbf{e}_j \big\rangle = c_i - c_j$. So all c_i equal some c, and $T = c\mathrm{id}_V$ on a basis.

If $V = \{\mathbf{0}\}$, there is nothing to prove. Otherwise choose an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V (Theorem 6.4.20). For $j \neq i$, $\mathbf{e}_i \perp \mathbf{e}_j$, so $\langle T\mathbf{e}_i, \mathbf{e}_j\rangle = 0$. By Proposition 6.3.18, $T\mathbf{e}_i = \sum_j \langle T\mathbf{e}_i, \mathbf{e}_j\rangle \mathbf{e}_j = c_i \mathbf{e}_i$, where $c_i = \langle T\mathbf{e}_i, \mathbf{e}_i\rangle$.

Now let $i \neq j$. The vectors $\mathbf{e}_i + \mathbf{e}_j$ and $\mathbf{e}_i - \mathbf{e}_j$ are orthogonal, since $\langle \mathbf{e}_i + \mathbf{e}_j, \mathbf{e}_i - \mathbf{e}_j\rangle = 1 - 0 + 0 - 1 = 0$. By the hypothesis,

$$0 = \langle T(\mathbf{e}_i + \mathbf{e}_j), \mathbf{e}_i - \mathbf{e}_j\rangle = \langle c_i\mathbf{e}_i + c_j\mathbf{e}_j, \mathbf{e}_i - \mathbf{e}_j\rangle = c_i - c_j.$$

So $c_1 = \cdots = c_n = c$, and T agrees with $c\mathrm{id}_V$ on a basis. By Theorem 2.2.1, $T = c\mathrm{id}_V$.

A basis-free version: for fixed $\mathbf{u}$, the hypothesis says $T\mathbf{u} \perp \operatorname{span}(\mathbf{u})^\perp$, so $T\mathbf{u} \in \left(\operatorname{span}(\mathbf{u})^\perp\right)^\perp = \operatorname{span}(\mathbf{u})$ by Theorem 6.4.26. A linear map that sends every vector to a multiple of itself is a multiple of the identity, by comparing $T(\mathbf{u} + \mathbf{v})$ with $T\mathbf{u} + T\mathbf{v}$ for independent $\mathbf{u}, \mathbf{v}$.

Solution to Exercise 6.7.8

Uses. Definition 6.6.8 (Adjoint), Proposition 6.6.18 (Properties of the adjoint), Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v}, \mathbf{v}\rangle$ determine T), Corollary 6.1.38 (Linear maps are determined by inner products)
Techniques. Move an operator across the inner product (Technique 6.6.20), Polarization (Technique 6.1.33), Plug in clever test vectors (Technique 6.1.40)
Hint 1. Write $\|T\mathbf{v}\|^2$ and $\|T^*\mathbf{v}\|^2$ as $\langle S\mathbf{v}, \mathbf{v}\rangle$ for suitable operators S.
Hint 2. For the converse, $S = T^*T - TT^*$ is self-adjoint with $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for all $\mathbf{v}$. Over $\mathbb{C}$ cite Proposition 6.1.41. Over $\mathbb{R}$, expand $\langle S(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}\rangle$.
Sketch. $\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2 = \langle (T^*T - TT^*)\mathbf{v}, \mathbf{v}\rangle$. So normality gives equal norms. Conversely $S = T^*T - TT^*$ is self-adjoint with zero diagonal values; expanding $\langle S(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}\rangle = 0$ gives $2\operatorname{Re}\langle S\mathbf{x}, \mathbf{y}\rangle = 0$, and replacing $\mathbf{y}$ by $i\mathbf{y}$ over $\mathbb{C}$ kills the imaginary part. So $S = 0$.

For every $\mathbf{v}$, moving the operators across the inner product (Definition 6.6.8, with $(T^*)^* = T$ by Proposition 6.6.18),

$$\|T\mathbf{v}\|^2 - \|T^*\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle - \langle TT^*\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, \mathbf{v}\rangle,$$

where $S = T^*T - TT^*$. If T is normal, then $S = 0$ and the norms agree.

Conversely, suppose $\langle S\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$. By Proposition 6.6.18, $S^* = T^*T - TT^* = S$. For $\mathbf{x}, \mathbf{y} \in V$,

$$0 = \langle S(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}\rangle = \langle S\mathbf{x}, \mathbf{y}\rangle + \langle S\mathbf{y}, \mathbf{x}\rangle = \langle S\mathbf{x}, \mathbf{y}\rangle + \overline{\langle S\mathbf{x}, \mathbf{y}\rangle},$$

using $\langle S\mathbf{y}, \mathbf{x}\rangle = \langle \mathbf{y}, S\mathbf{x}\rangle = \overline{\langle S\mathbf{x}, \mathbf{y}\rangle}$. So $\operatorname{Re}\langle S\mathbf{x}, \mathbf{y}\rangle = 0$. Over $\mathbb{R}$ this says $\langle S\mathbf{x}, \mathbf{y}\rangle = 0$. Over $\mathbb{C}$, apply it to $i\mathbf{y}$ in place of $\mathbf{y}$: $\operatorname{Re}\langle S\mathbf{x}, i\mathbf{y}\rangle = \operatorname{Re}(-i\langle S\mathbf{x}, \mathbf{y}\rangle) = \operatorname{Im}\langle S\mathbf{x}, \mathbf{y}\rangle = 0$. Either way $\langle S\mathbf{x}, \mathbf{y}\rangle = 0$ for all $\mathbf{x}, \mathbf{y}$, so $S = 0$ by Corollary 6.1.38. (Over $\mathbb{C}$ one can instead cite Proposition 6.1.41, which does not need S to be self-adjoint.)

Over $\mathbb{R}$, self-adjointness was used to replace $\langle S\mathbf{y}, \mathbf{x}\rangle$ by $\langle S\mathbf{x}, \mathbf{y}\rangle$. Without it the argument fails, as the rotation of Example 6.1.39 shows: it has zero diagonal values and is not 0.

Solution to Exercise 6.7.9

Uses. Proposition 6.6.18 (Properties of the adjoint), Proposition 6.1.27 (Basic properties of the norm), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents)
Techniques. Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Let $S = T^2 - T$. Is S self-adjoint? Compute S^2 using $T^3 = T^2$.
Hint 2. $S^2 = 0$ and $S^* = S$. Now $\|S\mathbf{v}\|^2 = \langle S^*S\mathbf{v}, \mathbf{v}\rangle$. For (b), try a nilpotent matrix.
Sketch. $S = T^2 - T$ is self-adjoint, and $T^4 = TT^3 = T^3 = T^2$, so $S^2 = T^4 - 2T^3 + T^2 = 0$. Then $\|S\mathbf{v}\|^2 = \langle S^2\mathbf{v}, \mathbf{v}\rangle = 0$, so $S = 0$, and $T^2 = T = T^*$ is an orthogonal projection. (b) $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ has $N^3 = N^2 = 0 \neq N$.

(a) Let $S = T^2 - T$. By Proposition 6.6.18, $S^* = (T^*)^2 - T^* = S$. Since $T^4 = TT^3 = TT^2 = T^3 = T^2$,

$$S^2 = T^4 - 2T^3 + T^2 = T^2 - 2T^2 + T^2 = 0.$$

For every $\mathbf{v}$, $\|S\mathbf{v}\|^2 = \langle S\mathbf{v}, S\mathbf{v}\rangle = \langle S^*S\mathbf{v}, \mathbf{v}\rangle = \langle S^2\mathbf{v}, \mathbf{v}\rangle = 0$, so $S\mathbf{v} = \mathbf{0}$ (Proposition 6.1.27). Hence $T^2 = T$. Together with $T^* = T$, this makes T an orthogonal projection by Corollary 6.6.29.

(b) Let $N \in \mathbb{R}^{2\times 2}$ have $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_2 = \mathbf{e}_1$, that is, $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $N^2 = 0$, so $N^3 = N^2$, but $N^2 = 0 \neq N$.

The proof used self-adjointness only to make $S^*S = S^2$, so that the algebraic fact $S^2 = 0$ became the positive fact $\|S\mathbf{v}\|^2 = 0$. For a general T, $S^2 = 0$ says nothing about S, as N shows.

Solution to Exercise 6.7.10

Uses. Theorem 6.3.22 (Parseval's identity), Definition 6.6.8 (Adjoint), Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 6.1.16 (Frobenius inner product)
Techniques. Move an operator across the inner product (Technique 6.6.20), The trace of A^*A (Technique 6.1.18)
Hint 1. Apply Theorem 6.3.22(b) to $T\mathbf{e}_j$ in the basis $\mathbf{f}_1, \dots, \mathbf{f}_m$, and to $T^*\mathbf{f}_i$ in the basis $\mathbf{e}_1, \dots, \mathbf{e}_n$.
Hint 2. The middle sum in (a) is symmetric in the two bases. Which of the outer sums does not mention the basis of V? For $\operatorname{tr}(A^*A)$, recall Proposition 6.1.16.
Sketch. By Parseval in W, $\|T\mathbf{e}_j\|^2 = \sum_i |\langle T\mathbf{e}_j, \mathbf{f}_i\rangle|^2$; by Parseval in V, $\|T^*\mathbf{f}_i\|^2 = \sum_j |\langle T^*\mathbf{f}_i, \mathbf{e}_j\rangle|^2 = \sum_j |\langle T\mathbf{e}_j, \mathbf{f}_i\rangle|^2$. The right side of (a) does not involve the $\mathbf{e}_j$, so the left side does not depend on them. The entries of A are $\langle T\mathbf{e}_j, \mathbf{f}_i\rangle$, so the middle sum is $\operatorname{tr}(A^*A)$.

(a) By Theorem 6.3.22(b) in W, $\|T\mathbf{e}_j\|^2 = \sum_{i=1}^m |\langle T\mathbf{e}_j, \mathbf{f}_i\rangle|^2$; summing over j gives the first equality. By Theorem 6.3.22(b) in V and the definition of the adjoint,

$$\|T^*\mathbf{f}_i\|^2 = \sum_{j=1}^n |\langle T^*\mathbf{f}_i, \mathbf{e}_j\rangle|^2 = \sum_{j=1}^n |\langle \mathbf{f}_i, T\mathbf{e}_j\rangle|^2 = \sum_{j=1}^n |\langle T\mathbf{e}_j, \mathbf{f}_i\rangle|^2,$$

using $|\overline{z}| = |z|$. Summing over i gives the second equality.

(b) Fix the basis $\mathbf{f}_1, \dots, \mathbf{f}_m$ of W. By (a), $\sum_j \|T\mathbf{e}_j\|^2 = \sum_i \|T^*\mathbf{f}_i\|^2$ for every orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ of V, and the right side does not involve that basis. So the left side is the same for all of them. By Proposition 6.3.18, the (i,j) entry of A is the ith coordinate of $T\mathbf{e}_j$, namely $\langle T\mathbf{e}_j, \mathbf{f}_i\rangle$, so the middle sum of (a) is $\sum_{i,j} |a_{ij}|^2 = \operatorname{tr}(A^*A)$ (Proposition 6.1.16). Since the left side of (a) does not depend on the bases either, neither does $\operatorname{tr}(A^*A)$.

The number $\sum_j \|T\mathbf{e}_j\|^2$ is the squared *Hilbert–Schmidt norm* of T; for $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ with dot products it is the squared Frobenius norm of the matrix. The proof never chose a convenient basis: it compared two arbitrary ones through a double sum that is symmetric in them.

Solution to Exercise 6.7.11

Uses. Theorem 6.4.20 (Orthonormal bases exist), Theorem 6.3.22 (Parseval's identity), Proposition 6.3.30 (Gram matrices detect independence), Theorem 5.4.10 (Multiplicativity), Exercise 6.3.57, Theorem 5.4.7 (Invertibility and the determinant)
Techniques. Pass to orthonormal coordinates (Technique 6.7.2), Gram matrices (Technique 6.3.31)
Hint 1. If the list is dependent, Proposition 6.3.30 already says that G is singular.
Hint 2. If it is independent, $W = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ has dimension k. Let $B \in \mathbb{F}^{k\times k}$ have columns $[\mathbf{v}_j]_{\mathcal{O}}$ for an orthonormal basis $\mathcal{O}$ of W, and show $G = B^*B$.
Sketch. If the list is dependent, G is singular (Proposition 6.3.30) and $\det G = 0$. If it is independent, choose an orthonormal basis $\mathcal{O}$ of its span, which has k vectors, and let B have columns $[\mathbf{v}_j]_{\mathcal{O}}$. Parseval gives $G = B^*B$, and B is invertible, so $\det G = \overline{\det B} \det B = |\det B|^2 > 0$.

If the list is dependent, then G is not invertible by Proposition 6.3.30, so $\det G = 0$ (Theorem 5.4.7).

Suppose the list is independent. Then it is a basis of $W = \operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$, so $\dim W = k$, and W has an orthonormal basis $\mathcal{O}$ with k vectors (Theorem 6.4.20). Let $B \in \mathbb{F}^{k\times k}$ be the matrix with columns $[\mathbf{v}_1]_{\mathcal{O}}, \dots, [\mathbf{v}_k]_{\mathcal{O}}$. By Theorem 6.3.22, the (i,j) entry of B^*B is

$$[\mathbf{v}_i]_{\mathcal{O}}^* [\mathbf{v}_j]_{\mathcal{O}} = \langle \mathbf{v}_j, \mathbf{v}_i\rangle = g_{ij},$$

so $G = B^*B$. The columns of B are the coordinate vectors of a basis, so they are independent and B is invertible, with $\det B \neq 0$. By Theorem 5.4.10 and Exercise 6.3.57(a),

$$\det G = \det(B^*)\det B = \overline{\det B}\det B = |\det B|^2 > 0.$$

So $\det G \geq 0$ always, with equality exactly for dependent lists. This is Technique 6.7.2: the list becomes the columns of a square matrix, and the Gram determinant becomes a squared modulus. The Gram–Schmidt route of Exercise 6.4.59 gives the more precise formula $\det G = \|\mathbf{f}_1\|^2 \cdots \|\mathbf{f}_k\|^2$, and with it the geometric meaning, base times height, repeated.

Solution to Exercise 6.7.12

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm)
Techniques. A sum of nonnegative terms (Technique 6.1.9), Orthogonal decomposition (Technique 6.2.4)
Hint 1. For the lower bound, expand $\|\mathbf{v}_1 + \cdots + \mathbf{v}_n\|^2 \geq 0$.
Hint 2. For the construction, work in $\mathbb{R}^n$ with $\mathbf{u} = (1/\sqrt{n})(1, \ldots, 1)$ and $\mathbf{f}_i = \mathbf{e}_i - (1/\sqrt{n})\mathbf{u}$, which are orthogonal to $\mathbf{u}$. Try $\mathbf{v}_i = \alpha\mathbf{f}_i + \beta\mathbf{u}$.
Sketch. Necessity: $x \leq 1$ by Cauchy–Schwarz, and $0 \leq \|\sum \mathbf{v}_i\|^2 = n + n(n-1)x$. Sufficiency: with $\mathbf{f}_i = \mathbf{e}_i - (1/n)(1, \ldots, 1)$, $\langle \mathbf{f}_i, \mathbf{f}_j \rangle = \delta_{ij} - 1/n$ and $\mathbf{f}_i \perp \mathbf{u}$. Then $\mathbf{v}_i = \alpha\mathbf{f}_i + \beta\mathbf{u}$ has $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \alpha^2(\delta_{ij} - 1/n) + \beta^2$, and $\alpha^2 = 1 - x$, $\beta^2 = (1 + (n-1)x)/n$ work exactly when $-1/(n-1) \leq x \leq 1$.

Necessity. By Theorem 6.2.1, $x = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \leq \|\mathbf{v}_1\|\|\mathbf{v}_2\| = 1$. By Proposition 6.1.29 (with all $n(n-1)$ cross terms equal to x),

$$0 \leq \|\mathbf{v}_1 + \cdots + \mathbf{v}_n\|^2 = \sum_{i,j} \langle \mathbf{v}_j, \mathbf{v}_i \rangle = n + n(n-1)x,$$

so $x \geq -1/(n-1)$.

Sufficiency. Let $-1/(n-1) \leq x \leq 1$, and work in $\mathbb{R}^n$ with the dot product. Let $\mathbf{1} = (1, \ldots, 1)$, $\mathbf{u} = \mathbf{1}/\sqrt{n}$, and $\mathbf{f}_i = \mathbf{e}_i - (1/n)\mathbf{1}$, the part of $\mathbf{e}_i$ orthogonal to $\mathbf{u}$ (Technique 6.2.4). Then $\langle \mathbf{f}_i, \mathbf{u} \rangle = (1-1)/\sqrt{n} = 0$ and

$$\langle \mathbf{f}_i, \mathbf{f}_j \rangle = \delta_{ij} - \frac{1}{n} - \frac{1}{n} + \frac{n}{n^2} = \delta_{ij} - \frac{1}{n}.$$

For real α, β, let $\mathbf{v}_i = \alpha\mathbf{f}_i + \beta\mathbf{u}$. Since $\mathbf{f}_i \perp \mathbf{u}$ and $\|\mathbf{u}\| = 1$,

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \alpha^2\left(\delta_{ij} - \frac{1}{n}\right) + \beta^2.$$

We need $\alpha^2(1 - 1/n) + \beta^2 = 1$ and $-\alpha^2/n + \beta^2 = x$. Subtracting, $\alpha^2 = 1 - x$, and then $\beta^2 = x + (1-x)/n = (1 + (n-1)x)/n$. Both right sides are nonnegative exactly when $-1/(n-1) \leq x \leq 1$, so we can take $\alpha = \sqrt{1-x}$ and $\beta = \sqrt{(1 + (n-1)x)/n}$.

Endpoints. At $x = 1$, equality in Cauchy–Schwarz for unit vectors with $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 1$ forces $\mathbf{v}_i = \mathbf{v}_j$: all the vectors coincide. At $x = -1/(n-1)$, the expansion above gives $\|\mathbf{v}_1 + \cdots + \mathbf{v}_n\| = 0$, so the vectors sum to $\mathbf{0}$; in the construction $\beta = 0$, and the vectors $\mathbf{v}_i = \sqrt{n/(n-1)}\mathbf{f}_i$ point from the center of the regular simplex with vertices $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to its vertices. For $n = 3$ these are three unit vectors in a plane at angles $2\pi/3$. Every list attaining the lower endpoint has the same Gram matrix as this one, so by Problem 6.9 it is this configuration up to an isometry of spans.

The Gram matrix of the list is $(1-x)I_n + xJ$, where J has every entry 1. The exercise therefore says for which x this matrix is a Gram matrix; Problem 6.17 answers that question for every Hermitian matrix.

Solution to Exercise 6.7.13

Uses. Theorem 6.3.3 (Pythagorean theorem), Theorem 6.4.6 (The projection is characterized by orthogonality), Exercise 6.7.6
Techniques. Recognize a projection by its two conditions (Technique 6.4.7), Perturb a minimizer (Technique 6.5.6)
Hint 1. Which hypothesis does the argument never use? What does the Pythagorean theorem require of the two pieces?
Hint 2. For a correct proof, take $\mathbf{u} \in \operatorname{im} P$ and $\mathbf{w} \in \ker P$, and compute $P(\mathbf{u} + t\mathbf{w})$. Then use Exercise 6.7.6.

Sketch. The Pythagorean theorem needs $P\mathbf{v} \perp \mathbf{v} - P\mathbf{v}$, which is what has to be proved; the argument never uses $\|P\mathbf{v}\| \le \|\mathbf{v}\|$, and the oblique projection $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ satisfies everything else. Correct proof: for $\mathbf{u} \in \operatorname{im} P$, $\mathbf{w} \in \ker P$, $P(\mathbf{u} + t\mathbf{w}) = \mathbf{u}$, so $\|\mathbf{u}\| \le \|\mathbf{u} + t\mathbf{w}\|$ for all t, and $\mathbf{u} \perp \mathbf{w}$ by Exercise 6.7.6. Then $\mathbf{v} - P\mathbf{v} \perp \operatorname{im} P$ and Theorem 6.4.6 finishes.

The error. The Pythagorean theorem (Theorem 6.3.3) applies to orthogonal vectors; its converse, which the argument uses in the form "the norms add up, so the pieces are orthogonal", needs the equation $\|\mathbf{v}\|^2 = \|P\mathbf{v}\|^2 + \|\mathbf{v} - P\mathbf{v}\|^2$, which nobody has shown. The argument assumes the conclusion. A sign of this is that the hypothesis $\|P\mathbf{v}\| \le \|\mathbf{v}\|$ is never used. Every step except the Pythagorean one holds for the oblique projection $P = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ on $\mathbb{R}^2$, which has $P^2 = P$, $\ker P = \operatorname{span}((1,-1))$, and $\operatorname{im} P = \operatorname{span}(\mathbf{e}_1)$, and for which the pieces of $\mathbf{v} = (0,1)$ are $(1,0)$ and $(-1,1)$, not orthogonal. (Consistently, this P violates the norm hypothesis: $P(1,1) = (2,0)$ is longer than $(1,1)$.) Over $\mathbb{C}$ there is a second gap: even if the norms did add up, the converse of Pythagoras fails there (Theorem 6.3.3(b) is stated for $\mathbb{F} = \mathbb{R}$ only), so orthogonality would still not follow.

A correct proof. Let $\mathbf{u} \in \operatorname{im} P$ and $\mathbf{w} \in \ker P$. Since $\mathbf{u} = P\mathbf{x}$ for some $\mathbf{x}$, $P\mathbf{u} = P^2\mathbf{x} = \mathbf{u}$, so for every $t \in \mathbb{F}$, $P(\mathbf{u} + t\mathbf{w}) = \mathbf{u} + tP\mathbf{w} = \mathbf{u}$. The norm hypothesis applied to $\mathbf{u} + t\mathbf{w}$ gives

$$\|\mathbf{u}\| = \|P(\mathbf{u} + t\mathbf{w})\| \le \|\mathbf{u} + t\mathbf{w}\| \quad \text{for all } t \in \mathbb{F},$$

so $\langle \mathbf{u}, \mathbf{w} \rangle = 0$ by Exercise 6.7.6. Thus every vector of $\ker P$ is orthogonal to $\operatorname{im} P$. Now for $\mathbf{v} \in V$, $P\mathbf{v} \in \operatorname{im} P$ and $\mathbf{v} - P\mathbf{v} \in \ker P \subseteq (\operatorname{im} P)^{\perp}$, so $P\mathbf{v} = \operatorname{Pr}_{\operatorname{im} P}(\mathbf{v})$ by Theorem 6.4.6.

Solution to Exercise 6.7.14

Uses. Definition 6.6.8 (Adjoint), Theorem 6.2.1 (Cauchy–Schwarz inequality), Exercise 6.7.8
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Which step replaces one operator by another? When is that legitimate?
Hint 2. For a correct proof, start from $\|T^*\mathbf{v}\|^2 = \langle TT^*\mathbf{v}, \mathbf{v} \rangle$ and apply Cauchy–Schwarz, then the hypothesis to the vector $T^*\mathbf{v}$.
Sketch. The step $TT^* = T^*T$ assumes that T is normal; the argument would prove $\|T^*\mathbf{v}\| = \|T\mathbf{v}\|$ for every T, which fails for $T = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ at $\mathbf{v} = \mathbf{e}_2$. Correct proof: $\|T^*\mathbf{v}\|^2 = \langle T(T^*\mathbf{v}), \mathbf{v} \rangle \le \|TT^*\mathbf{v}\|\|\mathbf{v}\| \le \|T^*\mathbf{v}\|\|\mathbf{v}\|$, and divide by $\|T^*\mathbf{v}\|$ when it is nonzero.

The error. The middle equality replaces TT^* by T^*T. These are equal only when T is normal, and the argument, which never uses the hypothesis on T, would prove $\|T^*\mathbf{v}\| = \|T\mathbf{v}\|$ for every operator. By Exercise 6.7.8 that holds only for normal T. For $T = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ on $\mathbb{R}^2$ with the dot product, $T^* = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, and at $\mathbf{v} = \mathbf{e}_2$: $\|T^*\mathbf{e}_2\|^2 = 0$ while $\|T\mathbf{e}_2\|^2 = \|\mathbf{e}_1\|^2 = 1$.

A correct proof. Let $\mathbf{v} \in V$. Moving T across the inner product, then using Cauchy–Schwarz (Theorem 6.2.1) and the hypothesis applied to $T^*\mathbf{v}$,

$$\|T^*\mathbf{v}\|^2 = \langle TT^*\mathbf{v}, \mathbf{v} \rangle \le \|T(T^*\mathbf{v})\|\|\mathbf{v}\| \le \|T^*\mathbf{v}\|\|\mathbf{v}\|.$$

If $T^*\mathbf{v} = \mathbf{0}$ the claim is clear; otherwise divide by $\|T^*\mathbf{v}\| > 0$. So $\|T^*\mathbf{v}\| \le \|\mathbf{v}\|$.

In terms of the operator norm of Section 9.2, the claim says $\|T^*\| \le \|T\|$, and applying it to T^* gives equality.

Solution to Exercise 6.7.15

Uses. Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v}, \mathbf{v} \rangle$ determine T), Example 6.1.39 (A rotation), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.1.32 (Polarization identities), Proposition 6.6.30 (Self-adjoint means real values)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Plug in clever test vectors (Technique 6.1.40)
Hint 1. Each of the four holds over exactly one of $\mathbb{R}$ and $\mathbb{C}$. Test the complex failures on $V = \mathbb{C}$ with $\mathbf{x} = 1$ and $\mathbf{y} = i$.
Hint 2. For the real failures, use the rotation $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.
Sketch. (a) True over $\mathbb{C}$ (Proposition 6.1.41), false over $\mathbb{R}$ (rotation by $\pi/2$). (b) True over $\mathbb{R}$, false over $\mathbb{C}$: $|1+i|^2 = 2 = 1 + 1$ but $\langle 1, i \rangle = -i$. (c) True over $\mathbb{R}$; over $\mathbb{C}$ the right side is $\operatorname{Re}\langle \mathbf{x}, \mathbf{y} \rangle$, so it fails for $1, i$. (d) True over $\mathbb{C}$ (Proposition 6.6.30); over $\mathbb{R}$ the hypothesis holds for every T, and the rotation is not self-adjoint.

(a) Over $\mathbb{C}$ this is Proposition 6.1.41. Over $\mathbb{R}$ it is false: the rotation $T\mathbf{v} = J\mathbf{v}$ with $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v} \in \mathbb{R}^2$ (Example 6.1.39), and $T \neq 0$.

(b) Over $\mathbb{R}$ this is Theorem 6.3.3(b). Over $\mathbb{C}$ it is false: in $V = \mathbb{C}$ with $\langle z, w\rangle = z\overline{w}$, $|1+i|^2 = 2 = |1|^2 + |i|^2$, but $\langle 1, i\rangle = -i \neq 0$.

(c) Over $\mathbb{R}$ this is Proposition 6.1.32(a). Over $\mathbb{C}$, Proposition 6.1.29 shows that the right side equals $\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$, so the formula fails whenever $\langle \mathbf{x}, \mathbf{y}\rangle$ is not real: for $\mathbf{x} = 1$ and $\mathbf{y} = i$ in $\mathbb{C}$ the right side is $(2-2)/4 = 0$ while $\langle 1, i\rangle = -i$.

(d) Over $\mathbb{C}$ this is Proposition 6.6.30. Over $\mathbb{R}$, every $\langle T\mathbf{v}, \mathbf{v}\rangle$ is real, so the hypothesis holds for every operator, and the rotation T of (a) is not self-adjoint: $\langle T\mathbf{e}_1, \mathbf{e}_2\rangle = 1$ while $\langle \mathbf{e}_1, T\mathbf{e}_2\rangle = -1$.

Statements (a) and (d) hold over $\mathbb{C}$ because a complex inner product can separate real and imaginary parts with the test vector $\mathbf{x} + i\mathbf{y}$; statements (b) and (c) hold over $\mathbb{R}$ because there the cross term $2\operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$ is all of $2\langle \mathbf{x}, \mathbf{y}\rangle$.

Problems for Chapter 6

Solution to Problem 6.1

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Proposition 6.1.29 (Expanding a norm), Lemma 6.2.3 (The easy half of (b)), Theorem 6.2.11 (Triangle inequality)
Techniques. Orthogonal decomposition (Technique 6.2.4), A nonnegative quadratic has nonpositive discriminant (Technique 6.2.7)
Hint 1. For (b), if $\mathbf{y} \neq \mathbf{0}$, write $\mathbf{x} = c\mathbf{y} + \mathbf{w}$ with $\mathbf{w} \perp \mathbf{y}$ and expand $\|\mathbf{x}\|^2$.
Hint 2. For (c), write $a = \|\mathbf{x}\|$, $b = \|\mathbf{y}\|$, $r = \operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$, and express the squares of both sides through a, b, r.
Hint 3. After clearing denominators, the difference of the two sides of (c) factors as $(a-b)^2(ab+r)$. Why is $ab + r \geq 0$?
Sketch. (b) With $c = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$ and $\mathbf{w} = \mathbf{x} - c\mathbf{y} \perp \mathbf{y}$, $\|\mathbf{x}\|^2 = \|\mathbf{w}\|^2 + |\langle \mathbf{x}, \mathbf{y}\rangle|^2/\|\mathbf{y}\|^2$. (c) The left side squared is $2(ab-r)/(ab)$ and the right side squared is $4\big((a-b)^2 + 2(ab-r)\big)/(a+b)^2$. After clearing denominators, the inequality between the squares becomes $(ab-r)(a+b)^2 \leq 2ab\big((a-b)^2 + 2(ab-r)\big)$, and the right side minus the left is $(a-b)^2(ab+r) \geq 0$, since $-r \leq |\langle \mathbf{x}, \mathbf{y}\rangle| \leq ab$ by Cauchy–Schwarz. Equality iff $a = b$ or $r = -ab$, and the latter means $-\mathbf{y}$ is a positive multiple of $\mathbf{x}$.

(a) For all vectors $\mathbf{x}, \mathbf{y}$ in an inner product space, $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$, with equality if and only if one of $\mathbf{x}, \mathbf{y}$ is a scalar multiple of the other (Theorem 6.2.1).

(b) If $\mathbf{y} = \mathbf{0}$, both sides are 0 and $\mathbf{y} = 0\mathbf{x}$. Let $\mathbf{y} \neq \mathbf{0}$, $c = \langle \mathbf{x}, \mathbf{y}\rangle / \|\mathbf{y}\|^2$, and $\mathbf{w} = \mathbf{x} - c\mathbf{y}$. Then $\langle \mathbf{w}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle - c\|\mathbf{y}\|^2 = 0$, so by Proposition 6.1.29 the cross term in $\|\mathbf{w} + c\mathbf{y}\|^2$ vanishes:

$$\|\mathbf{x}\|^2 = \|\mathbf{w}\|^2 + |c|^2\|\mathbf{y}\|^2 = \|\mathbf{w}\|^2 + \frac{|\langle \mathbf{x}, \mathbf{y}\rangle|^2}{\|\mathbf{y}\|^2}.$$

Dropping $\|\mathbf{w}\|^2 \geq 0$ and multiplying by $\|\mathbf{y}\|^2$ gives the inequality. Equality holds exactly when $\mathbf{w} = \mathbf{0}$, that is, $\mathbf{x} = c\mathbf{y}$; the converse is Lemma 6.2.3.

(c) Let $a = \|\mathbf{x}\| > 0$, $b = \|\mathbf{y}\| > 0$, and $r = \operatorname{Re}\langle \mathbf{x}, \mathbf{y}\rangle$. By Proposition 6.1.29 with the real scalars $1/a$ and $1/b$,

$$\left\|\frac{\mathbf{x}}{a} - \frac{\mathbf{y}}{b}\right\|^2 = 1 - \frac{2r}{ab} + 1 = \frac{2(ab-r)}{ab},$$

and $\|\mathbf{x} - \mathbf{y}\|^2 = a^2 - 2r + b^2 = (a-b)^2 + 2(ab-r)$. Both sides of the claimed inequality are nonnegative, so it is equivalent to the inequality between their squares, and after multiplying by $ab(a+b)^2/2 > 0$, to

$$(ab-r)(a+b)^2 \leq 2ab\big((a-b)^2 + 2(ab-r)\big).$$

Since $(a+b)^2 = (a-b)^2 + 4ab$, the left side is $(ab-r)(a-b)^2 + 4ab(ab-r)$, and the right side minus the left side is

$$2ab(a-b)^2-(ab-r)(a-b)^2=(a-b)^2(ab+r).$$

By Cauchy–Schwarz, $-r\le|\langle\mathbf{x},\mathbf{y}\rangle|\le ab$, so $ab+r\ge 0$, which proves the inequality.

Equality holds if and only if $(a-b)^2(ab+r)=0$, that is, $a=b$ or $r=-ab$. If $r=-ab$, then $\operatorname{Re}\langle\mathbf{x},-\mathbf{y}\rangle=\|\mathbf{x}\|\|-\mathbf{y}\|$. Since $\operatorname{Re}z\le|z|$ and Cauchy–Schwarz bounds $|\langle\mathbf{x},-\mathbf{y}\rangle|$ by the same number, $\langle\mathbf{x},-\mathbf{y}\rangle=\|\mathbf{x}\|\|-\mathbf{y}\|$, and by Theorem 6.2.11(b) one of $\mathbf{x},-\mathbf{y}$ is a nonnegative real multiple of the other. Both are nonzero, so $\mathbf{y}=-s\mathbf{x}$ with $s>0$. Conversely, if $\mathbf{y}=-s\mathbf{x}$ with $s>0$, then $r=-sa^2=-ab$.

The inequality says that sending nonzero vectors to the unit sphere along rays moves them apart by at most $2/(\|\mathbf{x}\|+\|\mathbf{y}\|)$ times their distance. Its proof used only $\operatorname{Re}\langle\mathbf{x},\mathbf{y}\rangle\ge-\|\mathbf{x}\|\|\mathbf{y}\|$. In a general normed space the same inequality holds with 4 in place of 2 (Dunkl and Williams, 1964).

Solution to Problem 6.2

> *Uses.* Theorem 6.4.24 (Orthogonal decomposition), Theorem 6.4.33 (Best approximation), Proposition 6.3.9 (Properties of $A^\perp$), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.1.30 (Parallelogram law), Aside (Norms that do not come from inner products) (Norms that do not come from inner products)
> *Techniques.* Projections shrink norms (Technique 6.4.5), Recognize a projection by its two conditions (Technique 6.4.7)
> *Hint 1.* For existence in (b), check $\langle\mathbf{v}-\operatorname{Pr}_W(\mathbf{v}),\mathbf{u}_j\rangle=0$ for each j. For uniqueness, what is $W\cap W^\perp$?
> *Hint 2.* For best approximation, write $\mathbf{v}-\mathbf{w}=(\mathbf{v}-\operatorname{Pr}_W(\mathbf{v}))+(\operatorname{Pr}_W(\mathbf{v})-\mathbf{w})$ and use Pythagoras. In (c), compute $\|\mathbf{e}_2-t\mathbf{e}_1\|_\infty$.
> *Sketch.* (b) $\mathbf{v}-\operatorname{Pr}_W(\mathbf{v})$ is orthogonal to each $\mathbf{u}_j$, hence to W; two decompositions differ by a vector of $W\cap W^\perp=\{\mathbf{0}\}$. For $\mathbf{w}\in W$, Pythagoras gives $\|\mathbf{v}-\mathbf{w}\|^2=\|\mathbf{v}-\operatorname{Pr}_W(\mathbf{v})\|^2+\|\operatorname{Pr}_W(\mathbf{v})-\mathbf{w}\|^2$. (c) $\|\mathbf{e}_2-t\mathbf{e}_1\|_\infty=\max(|t|,1)$, so every $t\mathbf{e}_1$ with $|t|\le1$ is closest. The Pythagorean step fails: $\|\cdot\|_\infty$ violates the parallelogram law, so it comes from no inner product.

(a) *Orthogonal decomposition* (Theorem 6.4.24): every $\mathbf{v}\in V$ can be written in exactly one way as $\mathbf{v}=\mathbf{w}+\mathbf{w}'$ with $\mathbf{w}\in W$ and $\mathbf{w}'\in W^\perp$, namely with $\mathbf{w}=\operatorname{Pr}_W(\mathbf{v})$. *Best approximation* (Theorem 6.4.33): $\|\mathbf{v}-\operatorname{Pr}_W(\mathbf{v})\|\le\|\mathbf{v}-\mathbf{w}\|$ for every $\mathbf{w}\in W$, with equality only for $\mathbf{w}=\operatorname{Pr}_W(\mathbf{v})$.

(b) *Existence.* Let $\mathbf{p}=\operatorname{Pr}_W(\mathbf{v})=\sum_i\langle\mathbf{v},\mathbf{u}_i\rangle\mathbf{u}_i\in W$. For each j, since $\langle\mathbf{u}_i,\mathbf{u}_j\rangle=\delta_{ij}$,

$$\langle\mathbf{v}-\mathbf{p},\mathbf{u}_j\rangle=\langle\mathbf{v},\mathbf{u}_j\rangle-\sum_i\langle\mathbf{v},\mathbf{u}_i\rangle\langle\mathbf{u}_i,\mathbf{u}_j\rangle=0.$$

A vector orthogonal to each $\mathbf{u}_j$ is orthogonal to their span W (Proposition 6.3.9(c)), so $\mathbf{v}-\mathbf{p}\in W^\perp$, and $\mathbf{v}=\mathbf{p}+(\mathbf{v}-\mathbf{p})$ is a decomposition.

Uniqueness. If $\mathbf{w}_1+\mathbf{w}_1'=\mathbf{w}_2+\mathbf{w}_2'$ with $\mathbf{w}_i\in W$ and $\mathbf{w}_i'\in W^\perp$, then $\mathbf{z}=\mathbf{w}_1-\mathbf{w}_2=\mathbf{w}_2'-\mathbf{w}_1'$ lies in $W\cap W^\perp$, so $\langle\mathbf{z},\mathbf{z}\rangle=0$ and $\mathbf{z}=\mathbf{0}$.

Best approximation. Let $\mathbf{w}\in W$. Then $\mathbf{v}-\mathbf{w}=(\mathbf{v}-\mathbf{p})+(\mathbf{p}-\mathbf{w})$, where the first piece lies in $W^\perp$ and the second in W. By Theorem 6.3.3,

$$\|\mathbf{v}-\mathbf{w}\|^2=\|\mathbf{v}-\mathbf{p}\|^2+\|\mathbf{p}-\mathbf{w}\|^2\ge\|\mathbf{v}-\mathbf{p}\|^2,$$

with equality if and only if $\|\mathbf{p}-\mathbf{w}\|=0$, that is, $\mathbf{w}=\mathbf{p}$.

(c) The points of W are $t\mathbf{e}_1=(t,0)$, and $\|\mathbf{e}_2-t\mathbf{e}_1\|_\infty=\|(-t,1)\|_\infty=\max(|t|,1)\ge1$, with equality if and only if $|t|\le1$. So the distance from $\mathbf{e}_2$ to W is 1, and it is attained at every point $(t,0)$ with $-1\le t\le1$.

The step with no analogue is the Pythagorean identity, which made the closest point unique: it adds the positive number $\|\mathbf{p}-\mathbf{w}\|^2$ for every $\mathbf{w}\ne\mathbf{p}$. The norm $\|\cdot\|_\infty$ does not come from any inner product, since it violates the parallelogram law (Proposition 6.1.30): $\|\mathbf{e}_1+\mathbf{e}_2\|_\infty^2+\|\mathbf{e}_1-\mathbf{e}_2\|_\infty^2=2$, while $2\|\mathbf{e}_1\|_\infty^2+2\|\mathbf{e}_2\|_\infty^2=4$. So there is no orthogonality to make the error perpendicular to W (compare Aside (Norms that do not come from inner products)). Closest points still exist here, but uniqueness, and the linear formula for Pr_W, are features of inner products.

Solution to Problem 6.3

Uses. Theorem 6.6.21 (Kernels and images of adjoints), Definition 6.6.8 (Adjoint), Proposition 6.1.37 (Inner products distinguish vectors), Proposition 6.6.18 (Properties of the adjoint), Theorem 6.4.26 (The double complement), Example 6.6.10 (The adjoint of a matrix), Corollary 6.6.22 (T and T^* have the same rank), Theorem 4.2.11 (Rank–nullity)
Techniques. Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. For (b), $\mathbf{w} \in \ker T^*$ if and only if $\langle \mathbf{v}, T^*\mathbf{w}\rangle = 0$ for every $\mathbf{v}$. Move T^* across.
Hint 2. For the second statement, apply the first one to T^*, whose adjoint is T, and take perps.
Hint 3. For (c), $A\mathbf{x} = \mathbf{b}$ is solvable if and only if $\mathbf{b} \in \operatorname{im} A = (\ker A^*)^\perp$. For the example, look for a combination of the three equations whose left side is 0.
Sketch. (b) $T^*\mathbf{w} = \mathbf{0}$ iff $\langle \mathbf{v}, T^*\mathbf{w}\rangle = 0$ for all $\mathbf{v}$ iff $\langle T\mathbf{v}, \mathbf{w}\rangle = 0$ for all $\mathbf{v}$ iff $\mathbf{w} \perp \operatorname{im} T$. Applied to T^*: $\ker T = (\operatorname{im} T^*)^\perp$, and the double perp gives $\operatorname{im} T^* = (\ker T)^\perp$. (c) $\mathbf{b} \in \operatorname{im} A$ iff $\mathbf{b} \perp \ker A^*$; the second alternative says exactly that this fails. For the example, $\mathbf{y} = (1, -1, -1)$ spans $\ker A^{\mathsf{T}}$; $\mathbf{y}^{\mathsf{T}}(1,1,1) = -1$, so no solution, while $\mathbf{y}^{\mathsf{T}}(1,1,0) = 0$, and $(1,0,1)$ is a solution.

(a) Let $T \in \mathrm{L}(V, W)$ have an adjoint. Then $\ker T^* = (\operatorname{im} T)^\perp$ and $\ker T = (\operatorname{im} T^*)^\perp$. If V is finite-dimensional, also $\operatorname{im} T^* = (\ker T)^\perp$ and $\operatorname{im} T = (\ker T^*)^\perp$, so that $V = \ker T \oplus \operatorname{im} T^*$ and $W = \operatorname{im} T \oplus \ker T^*$ (Theorem 6.6.21).

(b) Let $\mathbf{w} \in W$. A vector of V is $\mathbf{0}$ if and only if it is orthogonal to every vector of V (Proposition 6.1.37). So

$$\begin{aligned} T^*\mathbf{w} = \mathbf{0} &\iff \langle \mathbf{v}, T^*\mathbf{w}\rangle = 0 \text{ for all } \mathbf{v} \in V \\ &\iff \langle T\mathbf{v}, \mathbf{w}\rangle = 0 \text{ for all } \mathbf{v} \in V \iff \mathbf{w} \in (\operatorname{im} T)^\perp, \end{aligned}$$

using Definition 6.6.8 in the middle. This proves $\ker T^* = (\operatorname{im} T)^\perp$.

By Proposition 6.6.18, T^* has the adjoint T. Applying what we just proved to T^* gives $\ker T = (\operatorname{im} T^*)^\perp$. If V is finite-dimensional, so is its subspace $\operatorname{im} T^*$, and Theorem 6.4.26 gives $(\ker T)^\perp = \left((\operatorname{im} T^*)^\perp\right)^\perp = \operatorname{im} T^*$.

(c) With dot products, the map $\mathbf{x} \mapsto A\mathbf{x}$ from $\mathbb{F}^n$ to $\mathbb{F}^m$ has the adjoint $\mathbf{y} \mapsto A^*\mathbf{y}$ (Example 6.6.10). Applying (b) to the adjoint, with the finite-dimensional space $\mathbb{F}^m$ in the role of V, gives $\operatorname{im} A = (\ker A^*)^\perp$. So $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{y}^*\mathbf{b} = \langle \mathbf{b}, \mathbf{y}\rangle = 0$ for every $\mathbf{y} \in \ker A^*$. The second alternative says precisely that this fails for some $\mathbf{y}$, so exactly one of the two holds.

For the example, $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ is real, so $A^* = A^{\mathsf{T}}$. Subtracting the second equation from the first gives $x_1 - x_3$, so $\mathbf{y} = (1, -1, -1)$ satisfies $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$, as one checks column by column. With $\mathbf{b} = (1,1,1)$, $\mathbf{y}^{\mathsf{T}}\mathbf{b} = 1 - 1 - 1 = -1 \neq 0$, so the system has no solution. With $\mathbf{b}' = (1,1,0)$, $\mathbf{y}^{\mathsf{T}}\mathbf{b}' = 0$. The first two rows of A are independent and the third is their difference, so $\operatorname{rank} A = 2$, $\operatorname{rank} A^{\mathsf{T}} = 2$ (Corollary 6.6.22), and $\ker A^{\mathsf{T}} = \operatorname{span}(\mathbf{y})$ by Theorem 4.2.11. So $\mathbf{b}' \perp \ker A^{\mathsf{T}}$, and the system is solvable; indeed $\mathbf{x} = (1,0,1)$ works.

The vector $\mathbf{y}$ is a certificate of inconsistency that can be checked without solving anything. Row reduction finds it too (as the combination of rows that produces $0 = -1$), but (b) explains why it must exist: the right-hand side has to be orthogonal to every relation among the rows.

Solution to Problem 6.4

Uses. Proposition 6.1.29 (Expanding a norm), Definition 6.3.28 (Gram matrix), Proposition 6.3.30 (Gram matrices detect independence), Theorem 6.4.6 (The projection is characterized by orthogonality), Theorem 6.4.33 (Best approximation)
Techniques. Gram matrices (Technique 6.3.31), Recognize a projection by its two conditions (Technique 6.4.7)
Hint 1. Write $\mathbf{c} = \mathbf{p}_0 + \mathbf{z}$. Expand $\|\mathbf{z} - \mathbf{v}_i\|^2 = \|\mathbf{z}\|^2$. What linear condition on $\mathbf{z}$ do you get?
Hint 2. The conditions are $\langle \mathbf{z}, \mathbf{v}_i\rangle = \|\mathbf{v}_i\|^2/2$. With $\mathbf{z} = \sum_j t_j\mathbf{v}_j$ they become a linear system whose matrix is the Gram matrix of $\mathbf{v}_1, \ldots, \mathbf{v}_k$.
Sketch. $\|\mathbf{c} - \mathbf{p}_i\| = \|\mathbf{c} - \mathbf{p}_0\|$ iff $\langle \mathbf{c} - \mathbf{p}_0, \mathbf{v}_i\rangle = \|\mathbf{v}_i\|^2/2$. For $\mathbf{c} - \mathbf{p}_0 = \sum_j t_j\mathbf{v}_j$ this is $G\mathbf{t} = \mathbf{d}$ with G the invertible Gram matrix, so $\mathbf{c}$ is unique. For any $\mathbf{q}$ the same conditions hold with $\mathbf{q}$; subtracting gives $\mathbf{q} - \mathbf{c} \perp \mathbf{v}_i$, so $\mathbf{c} - \mathbf{p}_0$ is the projection of $\mathbf{q} - \mathbf{p}_0$. (c) $G = \begin{pmatrix} 4 & 2 \\ 2 & 9 \end{pmatrix}$, $\mathbf{d} = (2, 9/2)$, $\mathbf{t} = (9/32, 7/16)$, $\mathbf{c} = (1, 7/8, 7/8)$, distance $9\sqrt{2}/8$.

(a) Every point of A is $\mathbf{c} = \mathbf{p}_0 + \mathbf{z}$ with $\mathbf{z} \in U = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. Since $\mathbf{c} - \mathbf{p}_i = \mathbf{z} - \mathbf{v}_i$, Proposition 6.1.29 gives

$$\|\mathbf{c} - \mathbf{p}_i\|^2 = \|\mathbf{z}\|^2 - 2\langle \mathbf{z}, \mathbf{v}_i\rangle + \|\mathbf{v}_i\|^2 = \|\mathbf{c} - \mathbf{p}_0\|^2 - 2\langle \mathbf{z}, \mathbf{v}_i\rangle + \|\mathbf{v}_i\|^2.$$

So $\mathbf{c}$ is equidistant from all the points if and only if

$$\langle \mathbf{z}, \mathbf{v}_i \rangle = \frac{1}{2}\|\mathbf{v}_i\|^2 \quad (1 \le i \le k).$$

Write $\mathbf{z} = \sum_j t_j \mathbf{v}_j$. Then $\langle \mathbf{z}, \mathbf{v}_i \rangle = \sum_j t_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = (G\mathbf{t})_i$, where G is the Gram matrix of $\mathbf{v}_1, \dots, \mathbf{v}_k$ (Definition 6.3.28). So the conditions say $G\mathbf{t} = \mathbf{d}$ with $d_i = \|\mathbf{v}_i\|^2/2$. The $\mathbf{v}_i$ are independent, so G is invertible (Proposition 6.3.30), and there is exactly one $\mathbf{t}$, hence (by independence again) exactly one $\mathbf{z} \in U$ and one $\mathbf{c} \in A$.

(b) The computation in (a) did not use $\mathbf{z} \in U$: for every $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{q}$ is equidistant from the points if and only if $\mathbf{z} = \mathbf{q} - \mathbf{p}_0$ satisfies these k conditions. Since $\mathbf{c} - \mathbf{p}_0$ satisfies it, subtracting shows that $\mathbf{q}$ is equidistant if and only if $\langle \mathbf{q} - \mathbf{c}, \mathbf{v}_i \rangle = 0$ for all i.

For such $\mathbf{q}$, the vector $\mathbf{c} - \mathbf{p}_0$ lies in U, and $(\mathbf{q} - \mathbf{p}_0) - (\mathbf{c} - \mathbf{p}_0) = \mathbf{q} - \mathbf{c}$ is orthogonal to U. By Theorem 6.4.6, $\mathbf{c} - \mathbf{p}_0 = \mathrm{Pr}_U(\mathbf{q} - \mathbf{p}_0)$. For any $\mathbf{a} = \mathbf{p}_0 + \mathbf{u} \in A$, Theorem 6.4.33 gives

$$\|\mathbf{q} - \mathbf{a}\| = \|(\mathbf{q} - \mathbf{p}_0) - \mathbf{u}\| \ge \|(\mathbf{q} - \mathbf{p}_0) - \mathrm{Pr}_U(\mathbf{q} - \mathbf{p}_0)\| = \|\mathbf{q} - \mathbf{c}\|,$$

with equality only for $\mathbf{a} = \mathbf{c}$. If $k = n$, then $U = \mathbb{R}^n$, so $U^\perp = \{\mathbf{0}\}$ and $\mathbf{q} = \mathbf{c}$.

(c) Here $\mathbf{v}_1 = (2, 0, 0)$ and $\mathbf{v}_2 = (1, 2, 2)$, so

$$G = \begin{pmatrix} 4 & 2 \\ 2 & 9 \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} 2 \\ 9/2 \end{pmatrix}, \quad \mathbf{t} = G^{-1}\mathbf{d} = \frac{1}{32}\begin{pmatrix} 9 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 9/2 \end{pmatrix} = \begin{pmatrix} 9/32 \\ 7/16 \end{pmatrix}.$$

So $\mathbf{c} = (9/32)(2, 0, 0) + (7/16)(1, 2, 2) = (1, 7/8, 7/8)$. Check: $\|\mathbf{c}\|^2 = 1 + 49/32 = 81/32$, $\|\mathbf{c} - \mathbf{p}_1\|^2 = 1 + 49/32 = 81/32$, and $\|\mathbf{c} - \mathbf{p}_2\|^2 = 0 + 2 \cdot 81/64 = 81/32$. The common distance is $9/\sqrt{32} = 9\sqrt{2}/8$.

By (b), the points of $\mathbb{R}^3$ equidistant from the three vertices form the line through $\mathbf{c}$ perpendicular to the plane of the triangle; the circumcenter is where that line meets the plane.

Solution to Problem 6.5

Uses. Definition 6.1.2 (Inner product), Lemma 3.4.23 (Too many roots), Proposition 6.3.18 (Coordinates in an orthonormal basis), Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.30 (Gram matrices detect independence), Theorem 6.4.33 (Best approximation), Proposition 6.5.21 (Polynomial fits)
Techniques. Take an inner product to isolate a coefficient (Technique 6.3.12), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. For definiteness, a polynomial of degree at most n with $n+1$ roots is zero. For the Lagrange polynomials, compute $\langle L_i, L_j \rangle$ from $L_i(x_k) = \delta_{ik}$.
Hint 2. In (c), $\|p - q\|^2 = \sum_k (y_k - q(x_k))^2$ for every q. In (d), 1 and x are orthogonal for this inner product.
Sketch. (a) Positivity is a sum of squares, and $\sum_k p(x_k)^2 = 0$ forces $n+1$ roots, so $p = 0$. $\langle L_i, L_j \rangle = \sum_k \delta_{ik}\delta_{jk} = \delta_{ij}$, and $n+1$ orthonormal vectors in an $(n+1)$-dimensional space form a basis; $\langle p, L_j \rangle = p(x_j)$. (b) It is the Gram matrix of the basis $1, x, \dots, x^n$. (c) $\|p - q\|^2$ is the sum of squared residuals of q, so its minimizer is the projection. (d) $p = 1 + 3x/2 + x^2/2$; with $\|1\|^2 = 3$, $\|x\|^2 = 2$, $\langle p, 1 \rangle = 4$, $\langle p, x \rangle = 3$, the line is $y = 4/3 + 3x/2$.

(a) The formula is linear in p and symmetric, since evaluation at a point is linear. Also $\langle p, p \rangle = \sum_k p(x_k)^2 \ge 0$, and if it is 0, then every $p(x_k) = 0$ (Technique 6.1.9), so p, which has degree at most n and $n+1$ distinct roots, is 0 by Lemma 3.4.23. So this is an inner product.

Since $L_i(x_k) = \delta_{ik}$,

$$\langle L_i, L_j \rangle = \sum_{k=0}^{n} L_i(x_k) L_j(x_k) = \sum_{k=0}^{n} \delta_{ik}\delta_{jk} = \delta_{ij}.$$

So $L_0, \dots, L_n$ is orthonormal, and since $\dim V = n + 1$ it is an orthonormal basis (Corollary 6.3.14). By Proposition 6.3.18 the coordinates of p are $\langle p, L_j \rangle = \sum_k p(x_k)\delta_{jk} = p(x_j)$, that is, $p = \sum_j p(x_j) L_j$. The Lagrange interpolation formula is the coordinate formula for this orthonormal basis.

(b) The (i,j) entry is $\sum_k x_k^j x_k^i = \langle x^j, x^i \rangle$, so the matrix is the Gram matrix of $1, x, \ldots, x^n$ (Definition 6.3.28). These polynomials are independent, so the Gram matrix is invertible by Proposition 6.3.30. Problem 5.9 computes its determinant as $\prod_{i<j} (x_j - x_i)^2$; here invertibility came from the positivity of an inner product instead.

(c) For every $q \in \mathbb{R}[x]_{\le d}$,

$$\|p - q\|^2 = \sum_{k=0}^{n} (p(x_k) - q(x_k))^2 = \sum_{k=0}^{n} (y_k - q(x_k))^2,$$

the sum of squared residuals of q. By Theorem 6.4.33, this is smallest exactly for $q = \mathrm{Pr}_{\mathbb{R}[x]_{\le d}}(p)$, so that projection is the least-squares polynomial of Proposition 6.5.21, and in particular it is unique.

(d) $p = 1 + (3/2)x + (1/2)x^2$ satisfies $p(-1) = 0$, $p(0) = 1$, $p(1) = 3$. For this inner product $\langle 1, x \rangle = -1 + 0 + 1 = 0$, $\|1\|^2 = 3$, and $\|x\|^2 = 2$, while $\langle p, 1 \rangle = 0 + 1 + 3 = 4$ and $\langle p, x \rangle = 0 + 0 + 3 = 3$. By Corollary 6.4.10,

$$\mathrm{Pr}_{\mathbb{R}[x]_{\le 1}}(p) = \frac{4}{3} + \frac{3}{2}x.$$

This agrees with Proposition 6.5.17: the mean of the nodes is 0, the mean of the values is $4/3$, and the slope is $\sum_k x_k y_k / \sum_k x_k^2 = 3/2$.

Solution to Problem 6.6

> *Uses.* Proposition 6.1.41 (Over $\mathbb{C}$, the values $\langle T\mathbf{v}, \mathbf{v} \rangle$ determine T), Example 6.1.39 (A rotation), Proposition 6.6.18 (Properties of the adjoint), Proposition 6.6.30 (Self-adjoint means real values)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18), Plug in clever test vectors (Technique 6.1.40)
> *Hint 1.* Which hypotheses does Proposition 6.1.41 have? Over $\mathbb{R}$, compute $\langle S\mathbf{v}, \mathbf{v} \rangle$ for $S = T - T^*$ without using any hypothesis on T.
> *Hint 2.* Try $T = I_2 + J$ with J the rotation by $\pi/2$.
> *Sketch.* Proposition 6.1.41 needs complex scalars. Over $\mathbb{R}$, $\langle T^*\mathbf{v}, \mathbf{v} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle$ for every T, so the argument proves nothing, and $T = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ has $\langle T\mathbf{v}, \mathbf{v} \rangle = \|\mathbf{v}\|^2$ but $T^{\mathsf{T}} \neq T$. (a) Over $\mathbb{C}$ the conjugation step uses that $\langle T\mathbf{v}, \mathbf{v} \rangle$ is real, and Proposition 6.1.41 applies. (b) $\langle T^*\mathbf{v}, \mathbf{v} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle$ over $\mathbb{R}$, so averaging gives the claim.

The error. The last step cites Proposition 6.1.41, which is about complex inner product spaces. Over $\mathbb{R}$, $S = T - T^*$ satisfies $S^* = -S$ (Proposition 6.6.18), and every such operator has $\langle S\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, S^*\mathbf{v} \rangle = -\langle \mathbf{v}, S\mathbf{v} \rangle = -\langle S\mathbf{v}, \mathbf{v} \rangle$, hence $\langle S\mathbf{v}, \mathbf{v} \rangle = 0$, whether or not $S = 0$. In fact over $\mathbb{R}$ the first display holds for every T, so the argument never used the hypothesis, which is a sign that something is wrong.

The claim is false. Let $T = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ on $\mathbb{R}^2$ with the dot product. Then

$$\langle T\mathbf{v}, \mathbf{v} \rangle = (v_1 - v_2)v_1 + (v_1 + v_2)v_2 = v_1^2 + v_2^2 \ge 0,$$

but $T^* = T^{\mathsf{T}} = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right) \neq T$. Here $T = I_2 + J$ with J the rotation of Example 6.1.39, and J contributes nothing to $\langle T\mathbf{v}, \mathbf{v} \rangle$.

(a) Over $\mathbb{C}$ every step is correct. The conjugation step is the definition of the adjoint followed by conjugate symmetry, and the hypothesis is used exactly where it should be: a nonnegative number equals its conjugate. Then $\langle S\mathbf{v}, \mathbf{v} \rangle = 0$ for every $\mathbf{v}$, and Proposition 6.1.41 gives $S = 0$. (This is one half of Proposition 6.6.30.)

(b) Over $\mathbb{R}$, $\langle T^*\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, T\mathbf{v} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle$ by symmetry of the real inner product. So

$$\langle H\mathbf{v}, \mathbf{v} \rangle = \frac{1}{2}(\langle T\mathbf{v}, \mathbf{v} \rangle + \langle T^*\mathbf{v}, \mathbf{v} \rangle) = \langle T\mathbf{v}, \mathbf{v} \rangle.$$

The numbers $\langle T\mathbf{v}, \mathbf{v} \rangle$ see only the self-adjoint part H of a real operator, and adding any K with $K^* = -K$ to T does not change them. This is why Section 8.3 requires positive operators to be self-adjoint by definition.

Solution to Problem 6.7

Uses. Theorem 6.4.31 (The matrix of an orthogonal projection), Theorem 5.4.10 (Multiplicativity), Exercise 6.3.57, Lemma 6.5.4 (The orthogonal complement of a column space), Theorem 6.4.6 (The projection is characterized by orthogonality), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 4.2.11 (Rank–nullity), Exercise 6.4.59
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Move an operator across the inner product (Technique 6.6.20)
Hint 1. (a) Try $m > n$. (b) $\det(A^*) = \overline{\det A}$. (c) First show $\|P\mathbf{v}\| = \|P^{\mathsf{T}}\mathbf{v}\|$, and deduce $\ker P = \ker P^{\mathsf{T}}$.
Hint 2. (d) Apply rank–nullity to Pr_W restricted to U. (e) One direction is Hadamard's inequality from Exercise 6.4.59(c); for the other, try a diagonal matrix.
Sketch. (a) False: $A = \mathbf{e}_1 \in \mathbb{R}^{2\times 1}$ gives $AA^{\mathsf{T}} = \mathrm{diag}(1,0)$; in general AA^{T} is the projection onto $\mathrm{im}\, A$. (b) True: $\det(A^*A) = |\det A|^2$. (c) True: normality gives $\ker P = \ker P^{\mathsf{T}} = (\mathrm{im}\, P)^{\perp}$, so $\mathbf{v} - P\mathbf{v} \in \ker P \perp \mathrm{im}\, P$. (d) True: the restriction of Pr_W to U has kernel $U \cap W^{\perp}$ and image in W. (e) "Only if" is true by Hadamard, since the columns $A\mathbf{e}_j$ have length at most 1; "if" is false for $\mathrm{diag}(2, 1/2)$.

(a) False. $A = \begin{pmatrix}1\\0\end{pmatrix} \in \mathbb{R}^{2\times 1}$ has one orthonormal column, but $AA^{\mathsf{T}} = \begin{pmatrix}1&0\\0&0\end{pmatrix} \neq I_2$. In general AA^{T} is the matrix of $\mathrm{Pr}_{\mathrm{im}\, A}$ (Theorem 6.4.31), which is I_m only when $\mathrm{im}\, A = \mathbb{R}^m$, that is, when $m = n$.

(b) True. By Theorem 5.4.10 and Exercise 6.3.57(a), $\det(A^*A) = \det(A^*)\det A = \overline{\det A}\det A = |\det A|^2 \geq 0$.

(c) True. Let $\mathbf{v} \in \mathbb{R}^n$. Using the dot product, $\|P\mathbf{v}\|^2 = \langle P^{\mathsf{T}}P\mathbf{v}, \mathbf{v}\rangle = \langle PP^{\mathsf{T}}\mathbf{v}, \mathbf{v}\rangle = \|P^{\mathsf{T}}\mathbf{v}\|^2$, so $\ker P = \ker P^{\mathsf{T}}$, and $\ker P^{\mathsf{T}} = (\mathrm{im}\, P)^{\perp}$ by Lemma 6.5.4. Now $P\mathbf{v} \in \mathrm{im}\, P$, and $P(\mathbf{v} - P\mathbf{v}) = P\mathbf{v} - P^2\mathbf{v} = \mathbf{0}$, so $\mathbf{v} - P\mathbf{v} \in \ker P = (\mathrm{im}\, P)^{\perp}$. By Theorem 6.4.6, $P\mathbf{v} = \mathrm{Pr}_{\mathrm{im}\, P}(\mathbf{v})$ for every $\mathbf{v}$.

(d) True. Let $R : U \to W$ be the restriction of Pr_W to U; it maps into W and is linear. Its kernel is $U \cap \ker \mathrm{Pr}_W = U \cap W^{\perp}$ (Proposition 6.4.29(a)), and its image is a subspace of W. By Theorem 4.2.11,

$$\dim U = \dim\bigl(U \cap W^{\perp}\bigr) + \dim \mathrm{im}\, R \leq \dim\bigl(U \cap W^{\perp}\bigr) + \dim W.$$

In words: if U is bigger than W, then U contains $\dim U - \dim W$ independent directions perpendicular to all of W.

(e) "Only if" is true. The columns of A are $\mathbf{a}_j = A\mathbf{e}_j$, with $\|\mathbf{a}_j\| \leq \|\mathbf{e}_j\| = 1$. By Hadamard's inequality (Exercise 6.4.59(c)), $|\det A| \leq \|\mathbf{a}_1\|\cdots\|\mathbf{a}_n\| \leq 1$. "If" is false: $A = \begin{pmatrix}2&0\\0&1/2\end{pmatrix}$ has $\det A = 1$ but $\|A\mathbf{e}_1\| = 2 > 1$. The determinant measures the change of volume, and a map can stretch one direction while shrinking another.

Solution to Problem 6.8

Uses. Proposition 6.1.29 (Expanding a norm), Proposition 6.1.27 (Basic properties of the norm), Corollary 6.3.14 (Counting orthogonal vectors), Proposition 6.3.18 (Coordinates in an orthonormal basis), Proposition 2.3.17 (Every linear map $F^n \to F^m$ is some T_A), Theorem 6.3.29 (Orthonormal columns)
Techniques. Polarization (Technique 6.1.33), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. (a) Put $\mathbf{y} = \mathbf{0}$, then expand $\|f(\mathbf{x}) - f(\mathbf{y})\|^2$ and compare with $\|\mathbf{x} - \mathbf{y}\|^2$.
Hint 2. (b) Expand $\|f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x}) - f(\mathbf{y})\|^2$ into norms and inner products of values of f, and use (a) on every term.
Hint 3. For the second proof, $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ is orthonormal. Compute the coordinates of $f(\mathbf{x})$ in it. For (d), try complex conjugation on $\mathbb{C}^n$.
Sketch. (a) $\|f(\mathbf{x})\| = \|f(\mathbf{x}) - f(\mathbf{0})\| = \|\mathbf{x}\|$, and polarization turns equal distances and equal lengths into equal inner products. (b) $\|f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x}) - f(\mathbf{y})\|^2$ expands into terms that (a) lets us replace by the same terms without f, giving $\|\mathbf{0}\|^2 = 0$; likewise for $f(t\mathbf{x}) - tf(\mathbf{x})$. With an orthonormal basis: $f(\mathbf{e}_i)$ is an orthonormal basis and $f(\mathbf{x}) = \sum_i \langle \mathbf{x}, \mathbf{e}_i\rangle f(\mathbf{e}_i)$. (c) Apply (b) to $g - g(\mathbf{0})$. (d) The expansions involve only real parts, so the first proof gives real linearity; conjugation preserves distances and is not $\mathbb{C}$-linear.

(a) Taking $\mathbf{y} = \mathbf{0}$ gives $\|f(\mathbf{x})\| = \|f(\mathbf{x}) - f(\mathbf{0})\| = \|\mathbf{x}\|$. By Proposition 6.1.29 over $\mathbb{R}$,

$$\begin{aligned}\|f(\mathbf{x})\|^2 - 2\langle f(\mathbf{x}), f(\mathbf{y})\rangle + \|f(\mathbf{y})\|^2 &= \|f(\mathbf{x}) - f(\mathbf{y})\|^2 \\ &= \|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2,\end{aligned}$$

and the squared lengths agree, so $\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$.

(b) *First proof.* Let $\mathbf{x}, \mathbf{y} \in V$ and $\mathbf{z} = f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x}) - f(\mathbf{y})$. Expanding,

$$\|\mathbf{z}\|^2 = \|f(\mathbf{x}+\mathbf{y})\|^2 + \|f(\mathbf{x})\|^2 + \|f(\mathbf{y})\|^2 \\ -2\langle f(\mathbf{x}+\mathbf{y}), f(\mathbf{x})\rangle - 2\langle f(\mathbf{x}+\mathbf{y}), f(\mathbf{y})\rangle + 2\langle f(\mathbf{x}), f(\mathbf{y})\rangle.$$

By (a), each term equals the same term with f removed, and the result is the expansion of $\|(\mathbf{x}+\mathbf{y})-\mathbf{x}-\mathbf{y}\|^2 = 0$. So $\mathbf{z} = \mathbf{0}$ (Proposition 6.1.27). Similarly, for $t \in \mathbb{R}$,

$$\|f(t\mathbf{x}) - tf(\mathbf{x})\|^2 = \|f(t\mathbf{x})\|^2 - 2t\langle f(t\mathbf{x}), f(\mathbf{x})\rangle + t^2\|f(\mathbf{x})\|^2 = \|t\mathbf{x} - t\mathbf{x}\|^2 = 0.$$

So f is additive and homogeneous, that is, linear. This proof works verbatim for $f : V \to W$ between any two real inner product spaces.

Second proof, for $\dim V = n < \infty$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal basis. By (a), $\langle f(\mathbf{e}_i), f(\mathbf{e}_j)\rangle = \delta_{ij}$, so $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ is an orthonormal list of length n in V, hence an orthonormal basis (Corollary 6.3.14). By Proposition 6.3.18 and (a),

$$f(\mathbf{x}) = \sum_{i=1}^{n} \langle f(\mathbf{x}), f(\mathbf{e}_i)\rangle f(\mathbf{e}_i) = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{e}_i\rangle f(\mathbf{e}_i),$$

and each $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{e}_i\rangle$ is linear. This proof needs the values $f(\mathbf{e}_i)$ to span the target, which is why it asks for $f : V \to V$ with V finite-dimensional; the first proof needs nothing of the kind.

(c) Let $f(\mathbf{x}) = g(\mathbf{x}) - g(\mathbf{0})$. Then $f(\mathbf{0}) = \mathbf{0}$ and $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|g(\mathbf{x}) - g(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$, so f is linear by (b), say $f(\mathbf{x}) = Q\mathbf{x}$ (Proposition 2.3.17). By (a), $(Q\mathbf{e}_j)\cdot(Q\mathbf{e}_i) = \delta_{ij}$, so the columns of Q are orthonormal and $Q^{\mathsf{T}}Q = I_n$ (Theorem 6.3.29). With $\mathbf{b} = g(\mathbf{0})$, $g(\mathbf{x}) = Q\mathbf{x} + \mathbf{b}$.

(d) Over $\mathbb{C}$, Proposition 6.1.29 contains $2\,\mathrm{Re}\langle\cdot,\cdot\rangle$ instead of $2\langle\cdot,\cdot\rangle$, so (a) becomes $\mathrm{Re}\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \mathrm{Re}\langle \mathbf{x}, \mathbf{y}\rangle$, and the expansions in the first proof of (b) involve only lengths and real parts of inner products (for the second one, with t real). So they go through unchanged and show that f is additive and $f(t\mathbf{x}) = tf(\mathbf{x})$ for $t \in \mathbb{R}$. But f need not be $\mathbb{C}$-linear: on $\mathbb{C}^n$ with the dot product, let f conjugate every entry. Then $f(\mathbf{0}) = \mathbf{0}$ and $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\overline{\mathbf{x}-\mathbf{y}}\| = \|\mathbf{x}-\mathbf{y}\|$, while $f(i\mathbf{e}_1) = -i\mathbf{e}_1 \neq if(\mathbf{e}_1)$.

For surjective distance-preserving maps between real normed spaces, the conclusion of (b) still holds (the Mazur–Ulam theorem), but the proof is much harder: without an inner product there is no polarization.

Solution to Problem 6.9

> *Uses.* Proposition 6.3.30 (Gram matrices detect independence), Proposition 6.1.27 (Basic properties of the norm), Corollary 2.2.6 (Agreeing on a spanning list), Theorem 6.3.29 (Orthonormal columns), Theorem 6.4.20 (Orthonormal bases exist), Corollary 6.4.25 (Dimension of the orthogonal complement), Corollary 6.3.14 (Counting orthogonal vectors), Definition 6.3.33 (Unitary and orthogonal matrices), Exercise 6.3.37
> *Techniques.* Gram matrices (Technique 6.3.31), Show a vector is zero via its norm (Technique 6.1.36)
> *Hint 1.* (a) is Proposition 6.3.30(a) applied to both lists. Then use Technique 6.1.36.
> *Hint 2.* (b) Define $S\left(\sum_i c_i\mathbf{v}_i\right) = \sum_i c_i\mathbf{w}_i$. Why is this well defined?
> *Hint 3.* (c) The columns of A and of B have the same Gram matrix. Extend the map S of (b) from $\mathrm{im}\,A$ to all of $\mathbb{F}^m$ by matching orthonormal bases of $(\mathrm{im}\,A)^{\perp}$ and $(\mathrm{im}\,B)^{\perp}$.
> *Sketch.* (a) Both squared norms equal $\mathbf{c}^*G\mathbf{c}$. (b) Well defined because two expressions of the same vector differ by a relation, which (a) transports; inner products are preserved because they are computed from G; injective by (a). (c) A^*A is the Gram matrix of the columns. The isometry $S : \mathrm{im}\,A \to \mathrm{im}\,B$ sends an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_r$ of $\mathrm{im}\,A$ to one of $\mathrm{im}\,B$; completing both with orthonormal bases of the complements (both of dimension $m-r$) gives two unitary matrices U_1, U_2, and $U = U_2U_1^*$ agrees with S on $\mathrm{im}\,A$, so $UA = B$.

(a) Let G be the common Gram matrix and $\mathbf{c} = (c_1, \ldots, c_k)$. By Proposition 6.3.30(a), $\left\|\sum_i c_i\mathbf{v}_i\right\|^2 = \mathbf{c}^*G\mathbf{c} = \left\|\sum_i c_i\mathbf{w}_i\right\|^2$. Since a vector is $\mathbf{0}$ exactly when its norm is 0 (Proposition 6.1.27), the second statement follows.

(b) *Uniqueness.* A linear map on $\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ is determined by its values on the spanning list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ (Corollary 2.2.6).

Existence. Every vector of $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ is $\sum_i c_i \mathbf{v}_i$ for some scalars, and we set $S\big(\sum_i c_i \mathbf{v}_i\big) = \sum_i c_i \mathbf{w}_i$. If $\sum_i c_i \mathbf{v}_i = \sum_i d_i \mathbf{v}_i$, then $\sum_i (c_i - d_i)\mathbf{v}_i = \mathbf{0}$, so $\sum_i (c_i - d_i)\mathbf{w}_i = \mathbf{0}$ by (a), and the two definitions agree. So S is well defined, and it is linear because $\sum_i c_i \mathbf{v}_i + a \sum_i d_i \mathbf{v}_i = \sum_i (c_i + a d_i)\mathbf{v}_i$ is sent to $\sum_i (c_i + a d_i)\mathbf{w}_i$. Clearly $S\mathbf{v}_i = \mathbf{w}_i$.

S preserves inner products. For $\mathbf{x} = \sum_j c_j \mathbf{v}_j$ and $\mathbf{y} = \sum_i d_i \mathbf{v}_i$,

$$\langle S\mathbf{x}, S\mathbf{y} \rangle = \sum_{i,j} c_j \overline{d_i} \langle \mathbf{w}_j, \mathbf{w}_i \rangle = \sum_{i,j} c_j \overline{d_i} \langle \mathbf{v}_j, \mathbf{v}_i \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$$

S is bijective. If $S\mathbf{x} = \mathbf{0}$, then $\|\mathbf{x}\| = \|S\mathbf{x}\| = 0$, so S is injective. Its image contains every $\mathbf{w}_i$, hence their span, so S is surjective.

(c) If $B = UA$ with $U^*U = I_m$, then $B^*B = A^*U^*UA = A^*A$.

Conversely, suppose $A^*A = B^*B$. By Theorem 6.3.29, A^*A is the Gram matrix of the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ of A, and B^*B that of the columns $\mathbf{b}_1, \dots, \mathbf{b}_n$ of B. By (b) there is a bijective linear $S : \operatorname{im} A \to \operatorname{im} B$ with $S\mathbf{a}_j = \mathbf{b}_j$ that preserves inner products. In particular $r = \dim \operatorname{im} A = \dim \operatorname{im} B$.

Choose an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_r$ of $\operatorname{im} A$ (Theorem 6.4.20). Then $S\mathbf{u}_1, \dots, S\mathbf{u}_r$ is orthonormal in $\operatorname{im} B$, hence an orthonormal basis of it (Corollary 6.3.14). By Corollary 6.4.25, $(\operatorname{im} A)^\perp$ and $(\operatorname{im} B)^\perp$ both have dimension $m - r$; choose orthonormal bases $\mathbf{u}_{r+1}, \dots, \mathbf{u}_m$ and $\mathbf{u}'_{r+1}, \dots, \mathbf{u}'_m$ of them. Vectors of $\operatorname{im} A$ are orthogonal to vectors of $(\operatorname{im} A)^\perp$, so $\mathbf{u}_1, \dots, \mathbf{u}_m$ is an orthonormal list of length m, hence an orthonormal basis of $\mathbb{F}^m$, and likewise for $S\mathbf{u}_1, \dots, S\mathbf{u}_r, \mathbf{u}'_{r+1}, \dots, \mathbf{u}'_m$. Let U_1 and U_2 be the unitary matrices with these columns (Definition 6.3.33), and $U = U_2 U_1^*$, which is unitary (Exercise 6.3.37). Since $U_1^* U_1 = I_m$, $U_1^* \mathbf{u}_i = \mathbf{e}_i$, so $U\mathbf{u}_i = U_2 \mathbf{e}_i = S\mathbf{u}_i$ for $i \le r$. The linear maps $\mathbf{x} \mapsto U\mathbf{x}$ and S agree on a basis of $\operatorname{im} A$, so they agree on $\operatorname{im} A$, and $U\mathbf{a}_j = S\mathbf{a}_j = \mathbf{b}_j$ for every j. That is, $UA = B$.

Over $\mathbb{R}$, (c) says that two lists of vectors in $\mathbb{R}^m$ with the same lengths and the same pairwise angles differ by an orthogonal matrix. The unitary U is not unique when $r < m$: it can be changed freely on $(\operatorname{im} A)^\perp$.

Solution to Problem 6.10

> *Uses.* Proposition 6.4.29 (The orthogonal projection as a linear map), Proposition 6.6.9 (Adjoints are unique and linear), Definition 6.6.8 (Adjoint), Corollary 6.6.22 (T and T^* have the same rank), Corollary 4.2.18 (Injective if and only if surjective)
> *Techniques.* Move an operator across the inner product (Technique 6.6.20), Dimension counting (Technique 4.2.24)
> *Hint 1.* For (a), move Pr_W across: $\langle \operatorname{Pr}_W \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \operatorname{Pr}_W \mathbf{w} \rangle$, and $\operatorname{Pr}_W \mathbf{w} = \mathbf{w}$.
> *Hint 2.* For (b), compute $\ker A$. Then compare the ranks of A and A^*.
> *Sketch.* (a) $\langle \operatorname{Pr}_W \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \operatorname{Pr}_W \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle$, and symmetrically with U. So B satisfies the defining identity of A^*. (b) $\ker A = U \cap W^\perp$ and $\ker B = W \cap U^\perp$; with $\dim U = \dim W = k$, injective means bijective. $\operatorname{rank} A^* = \operatorname{rank} A$, so A is bijective iff B is. (c) $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \mathbb{R}^2$.

(a) By Proposition 6.4.29(c) and (b), $\langle \operatorname{Pr}_W \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \operatorname{Pr}_W \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle$, since $\mathbf{w} \in W$. In the same way $\langle \mathbf{u}, \operatorname{Pr}_U \mathbf{w} \rangle = \langle \operatorname{Pr}_U \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle$, since $\mathbf{u} \in U$. So $\langle A\mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{u}, B\mathbf{w} \rangle$ for all $\mathbf{u} \in U$ and $\mathbf{w} \in W$, which says that B is an adjoint of A (Definition 6.6.8); by Proposition 6.6.9, $B = A^*$.

(b) Let $k = \dim U = \dim W$. By Proposition 6.4.29(a), $\ker \operatorname{Pr}_W = W^\perp$, so $\ker A = U \cap W^\perp$; likewise $\ker B = W \cap U^\perp$. So (i) says that A is injective, and (ii) that B is injective. Since A maps between spaces of the same dimension k, it is injective if and only if it is bijective (Corollary 4.2.18), and the same holds for B. So (i) $\iff$ (iii) and (ii) $\iff$ (iv). Finally, by (a) and Corollary 6.6.22, $\operatorname{rank} B = \operatorname{rank} A^* = \operatorname{rank} A$, and a map between k-dimensional spaces is bijective exactly when its rank is k. So (iii) $\iff$ (iv).

(c) In $V = \mathbb{R}^2$, let $U = \operatorname{span}(\mathbf{e}_1)$ and $W = \mathbb{R}^2$. Then $W^\perp = \{\mathbf{0}\}$, so $U \cap W^\perp = \{\mathbf{0}\}$, but $W \cap U^\perp = \operatorname{span}(\mathbf{e}_2) \neq \{\mathbf{0}\}$.

Condition (i) says that no nonzero vector of U is perpendicular to all of W. Part (b) says that for subspaces of equal dimension this relation is symmetric, which is not obvious from the definition; the proof gets the symmetry from $\operatorname{rank} A^* = \operatorname{rank} A$. For two lines in $\mathbb{R}^3$, both conditions say that the lines are not

perpendicular. In orthonormal bases of U and W, the matrix of A has entries $\langle \mathbf{u}_j, \mathbf{w}_i \rangle$, and its singular values (Section 9.1) are the cosines of the *principal angles* between U and W.

Solution to Problem 6.11

Uses. Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.3.9 (Properties of $A^\perp$), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 6.1.37 (Inner products distinguish vectors), Exercise 2.4.43, Theorem 6.4.26 (The double complement), Lemma 2.4.24 (Matrix units), Example 2.6.28 (The trace)
Techniques. Cyclicity of trace (Technique 2.4.29), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. (a) $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, and $\langle A, I_n \rangle = \operatorname{tr} A$.
Hint 2. (b) Write $\langle XY - YX, M \rangle = \operatorname{tr}(M^* XY) - \operatorname{tr}(M^* YX)$ and move X cyclically in the second term. If the result vanishes for every Y, test with $Y = X^*M - MX^*$.
Hint 3. (c) (b) shows $C^\perp \subseteq \operatorname{span}(I_n)$. Take perps. (d) $U_{ij} = U_{ii}U_{ij} - U_{ij}U_{ii}$ for $i \neq j$.
Sketch. (a) Commutators have trace 0, and $\operatorname{tr} A = \langle A, I_n \rangle$. (b) By cyclicity, $\langle XY - YX, M \rangle = \operatorname{tr}((M^*X - XM^*)Y) = \langle Y, X^*M - MX^* \rangle$; for $M \in C^\perp$, choosing $Y = X^*M - MX^*$ gives $X^*M = MX^*$ for all X, so M is scalar. (c) $C^\perp = \operatorname{span}(I_n)$, so $C = \left(C^\perp\right)^\perp = \operatorname{span}(I_n)^\perp = Z$. (d) $U_{ij} = U_{ii}U_{ij} - U_{ij}U_{ii}$ and $U_{ii} - U_{nn} = U_{in}U_{ni} - U_{ni}U_{in}$.

(a) By Proposition 2.4.28, $\operatorname{tr}(XY - YX) = 0$, so every commutator lies in the subspace Z, and so does their span C. For every A, $\langle A, I_n \rangle = \operatorname{tr}(I_n^* A) = \operatorname{tr} A$, so $Z = \{I_n\}^\perp = \operatorname{span}(I_n)^\perp$ (Proposition 6.3.9(c)).

(b) By Proposition 2.4.28, $\operatorname{tr}(M^*YX) = \operatorname{tr}(XM^*Y)$, so

$$\langle XY - YX, M \rangle = \operatorname{tr}(M^*XY) - \operatorname{tr}(XM^*Y) = \operatorname{tr}((M^*X - XM^*)Y).$$

Let $N = X^*M - MX^*$. By Proposition 6.3.27, $N^* = M^*X - XM^*$, so the right side is $\operatorname{tr}(N^*Y) = \langle Y, N \rangle$, as claimed.

Now let $M \in C^\perp$ and fix X. Then $\langle Y, N \rangle = 0$ for every Y, and $Y = N$ gives $N = 0$ (Proposition 6.1.37), that is, $X^*M = MX^*$. As X runs over $\mathbb{F}^{n \times n}$ so does X^*, since $X = (X^*)^*$. So M commutes with every matrix.

(c) By (b) and Exercise 2.4.43 (a matrix that commutes with every upper triangular matrix is a scalar multiple of I_n; for $n = 1$ every matrix is), $C^\perp \subseteq \operatorname{span}(I_n)$. Conversely $I_n \in C^\perp$, since $\langle XY - YX, I_n \rangle = \operatorname{tr}(XY - YX) = 0$. So $C^\perp = \operatorname{span}(I_n)$, and by Theorem 6.4.26 and (a),

$$C = \left(C^\perp\right)^\perp = \operatorname{span}(I_n)^\perp = Z.$$

(d) Write $U_{ij} = \mathbf{e}_i \mathbf{e}_j^{\mathsf{T}}$, so that $U_{ij}U_{kl} = \delta_{jk}U_{il}$ (Lemma 2.4.24). For $i \neq j$,

$$U_{ii}U_{ij} - U_{ij}U_{ii} = U_{ij} - 0 = U_{ij},$$

and for $i \neq n$,

$$U_{in}U_{ni} - U_{ni}U_{in} = U_{ii} - U_{nn}.$$

By Example 2.6.28, these matrices span Z, so $Z \subseteq C$, and $C \subseteq Z$ holds over every field by Proposition 2.4.28.

The second proof is shorter and works over every field. The first illustrates a pattern: to show that a subspace C is all of a given subspace, compute $C^\perp$, and test a vector of $C^\perp$ against everything. It needs an inner product, so it works only over $\mathbb{R}$ and $\mathbb{C}$. (In fact every matrix of trace 0 over $\mathbb{R}$ or $\mathbb{C}$ is a *single* commutator, a theorem of Shoda that is much harder.)

Solution to Problem 6.12

Uses. Theorem 2.4.34 (Block multiplication), Theorem 6.3.29 (Orthonormal columns), Exercise 6.4.59, Theorem 6.4.15 (Gram–Schmidt is correct)
Techniques. Gram matrices (Technique 6.3.31), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. For (b), Exercise 6.4.59 gives $\det G = \|\mathbf{f}_1\|^2 \cdots \|\mathbf{f}_n\|^2$, where $\mathbf{f}_j$ is the Gram–Schmidt residual of $\mathbf{c}_j$.
Hint 2. $\mathbf{f}_j$ is $\mathbf{c}_j$ minus a combination of $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$. What is its jth coordinate?
Sketch. (a) $C^*C = I_n^* I_n + A^*A$ by block multiplication. (b) The columns are independent. The residual $\mathbf{f}_j = \mathbf{c}_j - \sum_{i<j} \beta_i \mathbf{c}_i$ has jth coordinate 1, since $\mathbf{c}_i$ has jth coordinate 0 for $i < j$; so $\|\mathbf{f}_j\| \geq 1$ and $\det G \geq 1$. Equality forces $\mathbf{f}_j = \mathbf{e}_j$ for all j, so $\operatorname{span}(\mathbf{c}_1, \ldots, \mathbf{c}_n) = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, and the lower block A is 0.

(a) $C^* = (I_n\ A^*)$, so by Theorem 2.4.34 $C^*C = I_nI_n + A^*A = I_n + A^*A$, and C^*C is the Gram matrix of the columns of C (Theorem 6.3.29).

(b) Write $\mathbf{c}_j = (\mathbf{e}_j, \mathbf{a}_j)$, where $\mathbf{a}_j$ is column j of A. Coordinate i of $\mathbf{c}_j$ is δ_{ij} for $i \le n$, so the columns are independent. Run Gram–Schmidt on them, with residuals $\mathbf{f}_1, \dots, \mathbf{f}_n$. By Exercise 6.4.59(b),

$$\det(I_n + A^*A) = \|\mathbf{f}_1\|^2 \cdots \|\mathbf{f}_n\|^2.$$

The residual $\mathbf{f}_j$ is $\mathbf{c}_j$ minus its projection onto $\operatorname{span}(\mathbf{c}_1, \dots, \mathbf{c}_{j-1})$, so $\mathbf{f}_j = \mathbf{c}_j - \sum_{i<j} \beta_i \mathbf{c}_i$ for some scalars β_i. Coordinate j of $\mathbf{c}_j$ is 1, and that of $\mathbf{c}_i$ is 0 for $i < j$, so coordinate j of $\mathbf{f}_j$ is 1. Since $\|\mathbf{f}_j\|^2$ is the sum of the squared moduli of its coordinates, $\|\mathbf{f}_j\|^2 \ge 1$, and $\det(I_n + A^*A) \ge 1$.

If equality holds, then every $\|\mathbf{f}_j\| = 1$, so $\mathbf{f}_j$ has no coordinate other than the jth, and $\mathbf{f}_j = \mathbf{e}_j \in \mathbb{F}^{n+m}$. By Theorem 6.4.15(a), $\operatorname{span}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \operatorname{span}(\mathbf{f}_1, \dots, \mathbf{f}_n)$, which consists of vectors whose last m coordinates vanish. So every $\mathbf{a}_j = \mathbf{0}$, and $A = 0$. Conversely $\det(I_n + 0) = 1$.

Solution to Problem 6.12 BEYOND 21-242

Uses. Exercise 5.5.52, Exercise 6.3.57, Corollary 5.4.15 (Rows), Proposition 5.3.15 (Triangular matrices)
Techniques. Gram matrices (Technique 6.3.31), A sum of nonnegative terms (Technique 6.1.9)
Hint 1. Apply Cauchy–Binet to C^*C: it is a sum over the n-element sets S of rows of C.
Hint 2. The set $S = \{1, \dots, n\}$ contributes 1. For the equality case, find a set S whose minor is $\pm a_{ij}$.
Sketch. $\det(C^*C) = \sum_S |\det C_S|^2$, where C_S consists of the rows of C indexed by S; the term $S = [n]$ is $\det I_n = 1$. The rows $[n] \setminus \{j\}$ together with row $n + i$ give a minor $\pm a_{ij}$, so equality forces $A = 0$.

For an n-element set $S \subseteq [n+m]$, let C_S be the $n \times n$ matrix formed by the rows of C with indices in S. The columns of C^* with indices in S form C_S^*, so by Exercise 5.5.52 and Exercise 6.3.57(a),

$$\det(C^*C) = \sum_S \det(C_S^*) \det(C_S) = \sum_S |\det C_S|^2.$$

The term $S = \{1, \dots, n\}$ is $|\det I_n|^2 = 1$, and the others are nonnegative, so $\det(I_n + A^*A) \ge 1$. For equality, every other minor must vanish. Take $S = ([n] \setminus \{j\}) \cup \{n + i\}$. Then C_S consists of the rows $\mathbf{e}_k^{\mathsf{T}}$ for $k \ne j$, in order, followed by row i of A. Moving that last row up to position j changes $\det C_S$ by a sign, and gives I_n with row j replaced by row i of A. Subtracting a_{ik} times row k from row j for every $k \ne j$ does not change the determinant (Corollary 5.4.15) and leaves a diagonal matrix whose only diagonal entry other than 1 is a_{ij}. So $\det C_S = \pm a_{ij}$. So $|a_{ij}|^2$ occurs in the sum, and equality forces every $a_{ij} = 0$.

The first approach is the geometric one: each column $\mathbf{c}_j$ sticks out of the span of the previous ones by at least 1, because of its identity part. The second expands the determinant into squares, one of which is the 1 coming from I_n. Either way, $I_n + A^*A$ is invertible for every A, which Exercise 5.6.12(c) proved for real A through the kernel; in terms of the singular values of Section 9.1, $\det(I_n + A^*A) = \prod_i (1 + \sigma_i^2)$.

Solution to Problem 6.13

Uses. Corollary 6.4.25 (Dimension of the orthogonal complement), Proposition 6.1.29 (Expanding a norm), Exercise 6.4.62
Techniques. Orthogonal decomposition (Technique 6.2.4), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Induct on n. For the step, project $\mathbf{v}_1, \dots, \mathbf{v}_{m-1}$ onto $\mathbf{v}_m^{\perp}$, which has dimension $n - 1$.
Hint 2. The projections $\mathbf{w}_i = \mathbf{v}_i - \left(\langle \mathbf{v}_i, \mathbf{v}_m \rangle / \|\mathbf{v}_m\|^2\right)\mathbf{v}_m$ still have $\langle \mathbf{w}_i, \mathbf{w}_j \rangle \le 0$. How many of them can be $\mathbf{0}$?
Sketch. For $n = 1$, three nonzero real multiples of one vector include two of the same sign. For the step, with $\mathbf{u} = \mathbf{v}_m$, the parts $\mathbf{w}_i$ of $\mathbf{v}_i$ orthogonal to $\mathbf{u}$ satisfy $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - \langle \mathbf{v}_i, \mathbf{u} \rangle \langle \mathbf{v}_j, \mathbf{u} \rangle / \|\mathbf{u}\|^2 \le 0$. If two were $\mathbf{0}$, the corresponding $\mathbf{v}_i, \mathbf{v}_j$ would be negative multiples of $\mathbf{u}$ with $\langle \mathbf{v}_i, \mathbf{v}_j \rangle > 0$. So at least $m - 2$ nonzero vectors in the $(n-1)$-dimensional $\mathbf{u}^{\perp}$, and $m - 2 \le 2(n-1)$. Example: $\pm\mathbf{e}_1, \dots, \pm\mathbf{e}_n$.

We induct on n.

$n = 1$. Then $V = \operatorname{span}(\mathbf{u})$ for any nonzero $\mathbf{u}$, and $\mathbf{v}_i = t_i\mathbf{u}$ with $t_i \ne 0$, so $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = t_i t_j \|\mathbf{u}\|^2$. If $m \ge 3$, two of t_1, t_2, t_3 have the same sign, and their vectors have a positive inner product. So $m \le 2$.

$n \ge 2$. We may assume $m \ge 2$. Let $\mathbf{u} = \mathbf{v}_m$, and for $i < m$ let

$$\mathbf{w}_i = \mathbf{v}_i - \lambda_i \mathbf{u}, \quad \lambda_i = \frac{\langle \mathbf{v}_i, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \leq 0,$$

the part of $\mathbf{v}_i$ orthogonal to $\mathbf{u}$ (Technique 6.2.4). For $i \neq j$, both less than m,

$$\langle \mathbf{w}_i, \mathbf{w}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - \lambda_i \langle \mathbf{u}, \mathbf{v}_j \rangle - \lambda_j \langle \mathbf{v}_i, \mathbf{u} \rangle + \lambda_i \lambda_j \|\mathbf{u}\|^2 = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - \lambda_i \lambda_j \|\mathbf{u}\|^2 \leq 0,$$

since $\lambda_i \lambda_j \geq 0$. If $\mathbf{w}_i = \mathbf{0}$, then $\mathbf{v}_i = \lambda_i \mathbf{u}$ with $\lambda_i < 0$ (it is ≤ 0, and nonzero because $\mathbf{v}_i \neq \mathbf{0}$). If this happened for two indices $i \neq j$, then $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \lambda_i \lambda_j \|\mathbf{u}\|^2 > 0$, a contradiction. So at least $m - 2$ of the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_{m-1}$ are nonzero. They lie in $\mathbf{u}^\perp$, a real inner product space of dimension $n - 1 \geq 1$ (Corollary 6.4.25), and have pairwise nonpositive inner products. By the induction hypothesis $m - 2 \leq 2(n-1)$, that is, $m \leq 2n$.

Example. In $\mathbb{R}^n$, the $2n$ vectors $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$ have pairwise inner products 0 or -1.

With strict inequalities $\langle \mathbf{v}_i, \mathbf{v}_j \rangle < 0$, the bound drops to $n + 1$ (Exercise 6.4.62). The difference is visible in the proof: with strict inequalities, the projected vectors still have strictly negative inner products, and none of them can vanish once $m \geq 3$.

Solution to Problem 6.14

Uses. Proposition 6.5.27 (The least-squares solution of minimal norm), Theorem 6.6.21 (Kernels and images of adjoints), Example 6.6.10 (The adjoint of a matrix), Theorem 6.5.5 (The normal equations), Proposition 6.5.9 ($\operatorname{rank}(A^*A) = \operatorname{rank} A$)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. (a) $\mathbf{w}^+$ lies in $(\ker \Phi)^\perp$. Which subspace is that?
Hint 2. (b) Write both conditions with the normal equations: $\Phi^{\mathsf{T}}(K\boldsymbol{\alpha} - \mathbf{y}) = \mathbf{0}$ and $K(K\boldsymbol{\alpha} - \mathbf{y}) = \mathbf{0}$. Then compare $\ker \Phi^{\mathsf{T}}$ with $\ker K$.
Hint 3. (c) $k(s,t) = 1 + 2st + s^2t^2 = (1+st)^2$, and K turns out to be invertible.
Sketch. (a) $\mathbf{w}^+ \in (\ker \Phi)^\perp = \operatorname{im} \Phi^{\mathsf{T}}$. (b) By the normal equations the two conditions say $K\boldsymbol{\alpha} - \mathbf{y} \in \ker \Phi^{\mathsf{T}}$ and $K\boldsymbol{\alpha} - \mathbf{y} \in \ker K$, and $\ker K = \ker \Phi^{\mathsf{T}}$ by Proposition 6.5.9(a) for Φ^{T}. So $\Phi^{\mathsf{T}}\boldsymbol{\alpha}$ is a least-squares solution in $\operatorname{im} \Phi^{\mathsf{T}} = (\ker \Phi)^\perp$, which is $\mathbf{w}^+$. (c) $K = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 4 \end{pmatrix}$, $\boldsymbol{\alpha} = (-1, 4, 5)/8$, and $\sum_i \alpha_i (1 + tt_i)^2 = 1 + 3t/2 + t^2/2$.

(a) By Proposition 6.5.27, $\mathbf{w}^+$ is the only least-squares solution in $(\ker \Phi)^\perp$. With dot products, Φ^{T} is the adjoint of Φ (Example 6.6.10), so $(\ker \Phi)^\perp = \operatorname{im} \Phi^{\mathsf{T}}$ by Theorem 6.6.21(c).

(b) By Theorem 6.5.5, $\Phi^{\mathsf{T}}\boldsymbol{\alpha}$ is a least-squares solution of $\Phi\mathbf{w} = \mathbf{y}$ if and only if $\Phi^{\mathsf{T}}\left(\Phi\Phi^{\mathsf{T}}\boldsymbol{\alpha} - \mathbf{y}\right) = \mathbf{0}$, that is, $K\boldsymbol{\alpha} - \mathbf{y} \in \ker \Phi^{\mathsf{T}}$. Since $K^{\mathsf{T}} = K$, $\boldsymbol{\alpha}$ is a least-squares solution of $K\boldsymbol{\alpha} = \mathbf{y}$ if and only if $K(K\boldsymbol{\alpha} - \mathbf{y}) = \mathbf{0}$, that is, $K\boldsymbol{\alpha} - \mathbf{y} \in \ker K$. By Proposition 6.5.9(a) applied to Φ^{T}, $\ker K = \ker\left(\left(\Phi^{\mathsf{T}}\right)^{\mathsf{T}}\Phi^{\mathsf{T}}\right) = \ker \Phi^{\mathsf{T}}$. So the two conditions agree.

If $\boldsymbol{\alpha}$ is a least-squares solution of $K\boldsymbol{\alpha} = \mathbf{y}$, then $\Phi^{\mathsf{T}}\boldsymbol{\alpha}$ is a least-squares solution of $\Phi\mathbf{w} = \mathbf{y}$ that lies in $\operatorname{im} \Phi^{\mathsf{T}} = (\ker \Phi)^\perp$, so it is $\mathbf{w}^+$ by the uniqueness in Proposition 6.5.27. Since $\Phi^{\mathsf{T}}\boldsymbol{\alpha} = \sum_i \alpha_i \boldsymbol{\varphi}_i$, we get $\boldsymbol{\varphi} \cdot \mathbf{w}^+ = \sum_i \alpha_i \, \boldsymbol{\varphi} \cdot \boldsymbol{\varphi}_i$.

(c) $k(s,t) = 1 + 2st + s^2t^2 = (1+st)^2$. For $t_1, t_2, t_3 = -1, 0, 1$,

$$K = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 4 \end{pmatrix}, \quad \det K = 8, \quad K^{-1} = \frac{1}{8}\begin{pmatrix} 3 & -4 & 1 \\ -4 & 16 & -4 \\ 1 & -4 & 3 \end{pmatrix}.$$

K is invertible, so the least-squares solution of $K\boldsymbol{\alpha} = \mathbf{y}$ is $\boldsymbol{\alpha} = K^{-1}\mathbf{y} = (1/8)(-1, 4, 5)$. By (b), the fitted function is

$$\sum_i \alpha_i k(t, t_i) = \frac{1}{8}\left(-(1-t)^2 + 4 + 5(1+t)^2\right) = 1 + \frac{3}{2}t + \frac{1}{2}t^2,$$

which takes the values $0, 1, 3$ at $-1, 0, 1$. Explicitly, $\mathbf{w}^+ = \Phi^{\mathsf{T}}\boldsymbol{\alpha} = \left(1, 3\sqrt{2}/4, 1/2\right)$, and $\boldsymbol{\varphi}(t) \cdot \mathbf{w}^+$ is the same polynomial.

The point is that N never appeared in the computation. The fit needs the $m \times m$ matrix K and the values $k(\varphi, \varphi_i)$ only, so the feature space can be huge, or even infinite-dimensional, as long as its inner products are cheap to compute. Section 16.1 develops this.

Solution to Problem 6.15

> *Uses.* Proposition 6.1.13 (Integral inner products), Lemma 6.1.14 (Nonnegative functions with zero integral), Proposition 6.1.29 (Expanding a norm)
> *Techniques.* Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21), A sum of nonnegative terms (Technique 6.1.9)
> *Hint 1.* Find an inner product on $C[0,1]$ for which $J(f) = \|f\|^2$. Then write $I(f)$ as an inner product of f with a fixed function.
> *Hint 2.* With $\langle f, g\rangle = \int_0^1 x f(x) g(x)\,dx$ and $h(x) = x$, $I(f) - J(f) = \langle f, h\rangle - \|f\|^2$. Complete the square.
> *Sketch.* $\langle f, g\rangle = \int_0^1 xfg$ is an inner product, $J(f) = \|f\|^2$, and $I(f) = \langle f, h\rangle$ for $h(x) = x$. Then $I(f) - J(f) = \|h\|^2/4 - \|f - h/2\|^2 \le \|h\|^2/4 = (1/4)\int_0^1 x^3\,dx = 1/16$, with equality exactly for $f(x) = x/2$.

Let $\langle f, g\rangle = \int_0^1 x f(x) g(x)\,dx$ on $C[0,1]$. It is bilinear and symmetric, and $\langle f, f\rangle = \int_0^1 x f(x)^2\,dx \ge 0$. If $\langle f, f\rangle = 0$, then the continuous nonnegative function $x f(x)^2$ is 0 (Lemma 6.1.14, as in the proof of Proposition 6.1.13), so $f(x) = 0$ for $0 < x \le 1$, and $f(0) = 0$ by continuity. So this is an inner product.

With $h(x) = x$, $I(f) = \int_0^1 x \cdot x f(x)\,dx = \langle f, h\rangle$ and $J(f) = \|f\|^2$. By Proposition 6.1.29,

$$\left\| f - \frac{h}{2} \right\|^2 = \|f\|^2 - \langle f, h\rangle + \frac{1}{4}\|h\|^2,$$

so

$$I(f) - J(f) = \frac{1}{4}\|h\|^2 - \left\| f - \frac{h}{2} \right\|^2 \le \frac{1}{4}\|h\|^2 = \frac{1}{4}\int_0^1 x^3\,dx = \frac{1}{16},$$

with equality if and only if $f = h/2$, that is, $f(x) = x/2$. The largest value is $1/16$.

Pointwise, the identity is $x^2 f - x f^2 = x^3/4 - x(f - x/2)^2$, and integrating it gives the same proof without naming an inner product. The inner product explains where the identity comes from: the maximum of $\langle f, h\rangle - \|f\|^2$ over a whole inner product space is found by completing the square, and the weight x in J tells us which inner product to use. One can also use Cauchy–Schwarz, $I(f) \le \|f\|\|h\|$, and then maximize $s\|h\| - s^2$ over $s = \|f\| \ge 0$.

This is Problem B5 of the 2006 Putnam Competition.

Solution to Problem 6.16

> *Uses.* Corollary 6.2.26 (Engel's form), Theorem 6.2.1 (Cauchy–Schwarz inequality)
> *Techniques.* Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21), A sum of nonnegative terms (Technique 6.1.9)
> *Hint 1.* Substitute $x = 1/a$, $y = 1/b$, $z = 1/c$, so that $xyz = 1$. What does the first term become?
> *Hint 2.* The sum becomes $x^2/(y+z) + y^2/(z+x) + z^2/(x+y)$. Apply Corollary 6.2.26, and then show $x + y + z \ge 3$ when $xyz = 1$.
> *Hint 3.* For the last step, with $x = p^3$, $y = q^3$, $z = r^3$: $p^3 + q^3 + r^3 - 3pqr = \frac{1}{2}(p+q+r)\big((p-q)^2 + (q-r)^2 + (r-p)^2\big)$.
> *Sketch.* With $x = 1/a$ etc., $1/\big(a^3(b+c)\big) = x^3 yz/(y+z) = x^2/(y+z)$. By Engel's form of Cauchy–Schwarz, the sum is at least $(x+y+z)^2/(2(x+y+z)) = (x+y+z)/2$, and $x + y + z \ge 3\sqrt[3]{xyz} = 3$ by the sum-of-squares identity for $p^3 + q^3 + r^3 - 3pqr$. Equality iff $a = b = c = 1$.

Let $x = 1/a$, $y = 1/b$, $z = 1/c$, positive with $xyz = 1$. Then

$$\frac{1}{a^3(b+c)} = \frac{x^3}{1/y + 1/z} = \frac{x^3 yz}{y+z} = \frac{x^2}{y+z},$$

and similarly for the other terms. By Corollary 6.2.26, with numerators x, y, z and denominators $y+z$, $z+x$, $x+y$,

$$\frac{x^2}{y+z}+\frac{y^2}{z+x}+\frac{z^2}{x+y}\geq\frac{(x+y+z)^2}{2(x+y+z)}=\frac{x+y+z}{2}.$$

That is Cauchy–Schwarz for the vectors $\left(x/\sqrt{y+z},\ldots\right)$ and $\left(\sqrt{y+z},\ldots\right)$, chosen so that their inner product is $x+y+z$ (Technique 6.2.21).

It remains to show $x+y+z\geq 3$. Let $p,q,r>0$ with $p^3=x, q^3=y, r^3=z$, so that $pqr=1$. Multiplying out,

$$p^3+q^3+r^3-3pqr=(p+q+r)\big(p^2+q^2+r^2-pq-qr-rp\big)=\frac{1}{2}(p+q+r)\Big((p-q)^2+(q-r)^2+(r-p)^2\Big),$$

which is ≥ 0, with equality only if $p=q=r$ (Technique 6.1.9). So $x+y+z\geq 3pqr=3$, and the sum is at least $3/2$.

If equality holds, then $x+y+z=3$, which forces $p=q=r$, hence $x=y=z=1$ and $a=b=c=1$. Conversely, at $a=b=c=1$ every term is $1/2$.

This is Problem 2 of the 1995 International Mathematical Olympiad. The substitution is the real idea; after it, the problem is the same Engel-form estimate as Nesbitt's inequality (Exercise 6.2.47).

Solution to Problem 6.17

> *Uses.* Lemma 6.2.6 (Nonnegative quadratics), Proposition 6.3.30 (Gram matrices detect independence), Theorem 6.3.22 (Parseval's identity), Theorem 6.3.29 (Orthonormal columns), Proposition 6.5.9 ($\operatorname{rank}(A^*A)=\operatorname{rank}A$), Theorem 6.4.20 (Orthonormal bases exist)
> *Techniques.* A nonnegative quadratic has nonpositive discriminant (Technique 6.2.7), Pass to orthonormal coordinates (Technique 6.7.2)
> *Hint 1.* (a) Run the second proof of Theorem 6.2.1: rotate the phase, then look at $\beta(\lambda\mathbf{c}-t\mathbf{d},\lambda\mathbf{c}-t\mathbf{d})$ for real t. Treat $\beta(\mathbf{d},\mathbf{d})=0$ separately: a linear function of t that is never negative is constant.
> *Hint 2.* (b) If $\beta(\mathbf{c},\mathbf{c})=0$, (a) gives $\beta(\mathbf{c},\mathbf{d})=0$ for every $\mathbf{d}$. Take $\mathbf{d}=G\mathbf{c}$.
> *Hint 3.* (c) For well-definedness in the second slot, use $G^*=G$. For (d), choose an orthonormal basis of V and record the coordinates of the classes $[\mathbf{e}_j]$.
> *Sketch.* (a) With $|\lambda|=1$ and $\lambda\beta(\mathbf{c},\mathbf{d})=|\beta(\mathbf{c},\mathbf{d})|$, $0\leq\beta(\lambda\mathbf{c}-t\mathbf{d},\lambda\mathbf{c}-t\mathbf{d})=\beta(\mathbf{c},\mathbf{c})-2t|\beta(\mathbf{c},\mathbf{d})|+t^2\beta(\mathbf{d},\mathbf{d})$ for real t; the discriminant gives (a) if $\beta(\mathbf{d},\mathbf{d})>0$, and a linear function bounded below gives $\beta(\mathbf{c},\mathbf{d})=0$ otherwise. (b) Then $(G\mathbf{c})^*G\mathbf{c}=0$. (c) Changing representatives changes $\mathbf{d}^*G\mathbf{c}$ by terms with $G\mathbf{z}=\mathbf{0}$; definiteness is (b). (d) An orthonormal basis of V turns $[\mathbf{e}_1],\ldots,[\mathbf{e}_k]$ into the columns of $B\in\mathbb{F}^{r\times k}$ with $B^*B=G$, and $r=\operatorname{rank}B=\operatorname{rank}B^*B=\operatorname{rank}G$.

The function β is linear in $\mathbf{c}$, and $\beta(\mathbf{d},\mathbf{c})=\mathbf{c}^*G\mathbf{d}=(\mathbf{d}^*G^*\mathbf{c})^*=\overline{\beta(\mathbf{c},\mathbf{d})}$, using $G^*=G$. So β satisfies every axiom of an inner product except possibly definiteness.

(a) The second proof of Theorem 6.2.1 uses definiteness only to know that $\|\mathbf{y}\|^2>0$, so we follow it and handle that case separately. Choose $\lambda\in\mathbb{F}$ with $|\lambda|=1$ and $\lambda\beta(\mathbf{c},\mathbf{d})=|\beta(\mathbf{c},\mathbf{d})|$. For real t, expanding as in Proposition 6.1.29,

$$0\leq\beta(\lambda\mathbf{c}-t\mathbf{d},\lambda\mathbf{c}-t\mathbf{d})=\beta(\mathbf{c},\mathbf{c})-2t|\beta(\mathbf{c},\mathbf{d})|+t^2\beta(\mathbf{d},\mathbf{d}).$$

If $\beta(\mathbf{d},\mathbf{d})>0$, Lemma 6.2.6 gives $|\beta(\mathbf{c},\mathbf{d})|^2\leq\beta(\mathbf{c},\mathbf{c})\beta(\mathbf{d},\mathbf{d})$. If $\beta(\mathbf{d},\mathbf{d})=0$, the right side is $\beta(\mathbf{c},\mathbf{c})-2t|\beta(\mathbf{c},\mathbf{d})|$, which is nonnegative for every real t only if $|\beta(\mathbf{c},\mathbf{d})|=0$ (otherwise take t large). In both cases (a) holds. (The first proof of Theorem 6.2.1 divides by $\|\mathbf{y}\|^2$, and the third normalizes, so neither survives as it stands.)

(b) If $G\mathbf{c}=\mathbf{0}$, then $\beta(\mathbf{c},\mathbf{c})=\mathbf{c}^*G\mathbf{c}=0$. Conversely, if $\beta(\mathbf{c},\mathbf{c})=0$, then (a) gives $\beta(\mathbf{c},\mathbf{d})=0$ for every $\mathbf{d}$. With $\mathbf{d}=G\mathbf{c}$ this reads $(G\mathbf{c})^*(G\mathbf{c})=0$, the squared length of $G\mathbf{c}$ for the dot product, so $G\mathbf{c}=\mathbf{0}$.

(c) Let $[\mathbf{c}]=[\mathbf{c}']$ and $[\mathbf{d}]=[\mathbf{d}']$, that is, $G(\mathbf{c}-\mathbf{c}')=\mathbf{0}$ and $G(\mathbf{d}-\mathbf{d}')=\mathbf{0}$. Then

$$\mathbf{d}^*G\mathbf{c}-\mathbf{d}'^*G\mathbf{c}'=\mathbf{d}^*G(\mathbf{c}-\mathbf{c}')+(\mathbf{d}-\mathbf{d}')^*G\mathbf{c}'=0+(G(\mathbf{d}-\mathbf{d}'))^*\mathbf{c}'=0,$$

using $(\mathbf{d}-\mathbf{d}')^*G=(G^*(\mathbf{d}-\mathbf{d}'))^*=(G(\mathbf{d}-\mathbf{d}'))^*$. So the formula is well defined. Linearity in the first slot and conjugate symmetry come from β, and $\langle[\mathbf{c}],[\mathbf{c}]\rangle=\beta(\mathbf{c},\mathbf{c})\geq 0$. By (b), $\langle[\mathbf{c}],[\mathbf{c}]\rangle=0$ if and only if $\mathbf{c}\in\ker G$, that is, $[\mathbf{c}]=\mathbf{0}$. So this is an inner product on V, and $\langle[\mathbf{e}_j],[\mathbf{e}_i]\rangle=\mathbf{e}_i^*G\mathbf{e}_j=g_{ij}$.

(d) If G is the Gram matrix of $\mathbf{v}_1, \dots, \mathbf{v}_k$, then $\mathbf{c}^* G \mathbf{c} = \left\| \sum_i c_i \mathbf{v}_i \right\|^2 \geq 0$ (Proposition 6.3.30(a)). Conversely, if $\mathbf{c}^* G \mathbf{c} \geq 0$ for all $\mathbf{c}$, then by (c) G is the Gram matrix of $[\mathbf{e}_1], \dots, [\mathbf{e}_k]$ in V. These classes span V, so V is finite-dimensional; let $\mathcal{O}$ be an orthonormal basis (Theorem 6.4.20), $r = \dim V$, and $B \in \mathbb{F}^{r \times k}$ the matrix with columns $\left[[\mathbf{e}_j]\right]_{\mathcal{O}}$. By Theorem 6.3.22, $(B^* B)_{ij} = \langle [\mathbf{e}_j], [\mathbf{e}_i] \rangle = g_{ij}$, so $G = B^* B$ (Technique 6.7.2). Finally, if $G = B^* B$ for any B, then G is the Gram matrix of the columns of B (Theorem 6.3.29), closing the cycle. For the B constructed above, the columns span $\mathbb{F}^r$ (they are the coordinates of a spanning list of V), so $\operatorname{rank} B = r$, and $\operatorname{rank} G = \operatorname{rank}(B^* B) = \operatorname{rank} B = r$ by Proposition 6.5.9(b).

Matrices with these properties are called *positive semidefinite*. Section 8.3 produces the factorization $G = B^* B$ differently, with B a square root of G, and Exercise 6.7.12 is the case $G = (1 - x) I_n + x J$. The quotient construction, dividing a semi-inner product by its null vectors, is also how inner product spaces of functions are built in analysis, where functions that differ only on a negligible set are identified.

Solution to Problem 6.18

> *Uses.* Proposition 6.1.10 (The dot product is an inner product), Definition 6.6.8 (Adjoint), Proposition 6.6.9 (Adjoints are unique and linear), Proposition 6.6.18 (Properties of the adjoint), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.4.20 (Orthonormal bases exist), Corollary 6.4.25 (Dimension of the orthogonal complement), Proposition 2.4.28 (Cyclicity of the trace), Theorem 6.3.22 (Parseval's identity), Theorem 6.3.29 (Orthonormal columns)
> *Techniques.* Move an operator across the inner product (Technique 6.6.20), Cyclicity of trace (Technique 2.4.29)
> *Hint 1.* (a) and (b): for fixed $N \in \mathcal{G}$, the maps $M \mapsto MN$, $M \mapsto NM$, and $M \mapsto M^{-1}$ are bijections of $\mathcal{G}$, so sums over $\mathcal{G}$ can be reindexed by them.
> *Hint 2.* For self-adjointness, first show that the adjoint of $N \in \mathcal{G}$ for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ is N^{-1}.
> *Hint 3.* (c) P is an orthogonal projection for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$. Take an orthonormal basis for that inner product whose first vectors span $\operatorname{im} P$, and compute the trace of P in it. (d) Use the same kind of basis as the columns of S.
> *Sketch.* (a) Positivity: $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{G}} \geq \|I_n \mathbf{x}\|^2$. Invariance: reindex by $M \mapsto MN$. (b) $NP = P$ by reindexing with $M \mapsto NM$, so $P^2 = |\mathcal{G}|^{-1} \sum_N NP = P$. For $N \in \mathcal{G}$, $\langle N\mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}} = \langle N\mathbf{x}, NN^{-1}\mathbf{y} \rangle_{\mathcal{G}} = \langle \mathbf{x}, N^{-1}\mathbf{y} \rangle_{\mathcal{G}}$, so $N^* = N^{-1}$ and $P^* = |\mathcal{G}|^{-1} \sum N^{-1} = P$. (c) $P = \operatorname{Pr}_{\operatorname{im} P}$; in an adapted orthonormal basis P is $\operatorname{diag}(1, \dots, 1, 0, \dots, 0)$, so $\operatorname{tr} P = \operatorname{rank} P$ by cyclicity of trace. If $\operatorname{tr} P = 0$, then $P = 0$. (d) The columns of S form an orthonormal basis for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$; $S^{-1} M S$ is the matrix of M in it, and M preserves $\langle \cdot, \cdot \rangle_{\mathcal{G}}$.

Fix $N \in \mathcal{G}$. The map $M \mapsto MN$ sends $\mathcal{G}$ into $\mathcal{G}$ (closure) and is injective ($MN = M'N$ implies $M = M'$), so it is a bijection of the finite set $\mathcal{G}$. The same holds for $M \mapsto NM$ and for $M \mapsto M^{-1}$, which is its own inverse. So a sum over $\mathcal{G}$ can be reindexed by any of these maps.

(a) Each term $(M\mathbf{x}) \cdot (M\mathbf{y})$ is linear in $\mathbf{x}$ and symmetric, and so is the sum. Since $I_n \in \mathcal{G}$, $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{G}} = \sum_M \|M\mathbf{x}\|^2 \geq \|\mathbf{x}\|^2$, which is positive for $\mathbf{x} \neq \mathbf{0}$ (Proposition 6.1.10). So $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ is an inner product. Reindexing by $M \mapsto MN$,

$$\langle N\mathbf{x}, N\mathbf{y} \rangle_{\mathcal{G}} = \sum_{M \in \mathcal{G}} (MN\mathbf{x}) \cdot (MN\mathbf{y}) = \sum_{M' \in \mathcal{G}} (M'\mathbf{x}) \cdot (M'\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}}.$$

(b) Reindexing by $M \mapsto NM$, $NP = |\mathcal{G}|^{-1} \sum_M NM = |\mathcal{G}|^{-1} \sum_{M'} M' = P$. Hence

$$P^2 = |\mathcal{G}|^{-1} \sum_{N \in \mathcal{G}} NP = |\mathcal{G}|^{-1} \sum_{N \in \mathcal{G}} P = P.$$

For $N \in \mathcal{G}$ and all $\mathbf{x}, \mathbf{y}$, by (a) applied with N,

$$\langle N\mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}} = \langle N\mathbf{x}, N(N^{-1}\mathbf{y}) \rangle_{\mathcal{G}} = \langle \mathbf{x}, N^{-1}\mathbf{y} \rangle_{\mathcal{G}},$$

so the adjoint of $\mathbf{x} \mapsto N\mathbf{x}$ for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$ is $\mathbf{x} \mapsto N^{-1}\mathbf{x}$ (Definition 6.6.8, Proposition 6.6.9). Adjoints of sums and real multiples are the sums and multiples of the adjoints (Proposition 6.6.18), so the adjoint of P is $|\mathcal{G}|^{-1} \sum_N N^{-1}$, which is P after reindexing by $N \mapsto N^{-1}$.

(c) By (b) and Corollary 6.6.29, P is the orthogonal projection onto $\operatorname{im} P$ for the inner product $\langle \cdot, \cdot \rangle_{\mathcal{G}}$. Let $r = \operatorname{rank} P$. By Theorem 6.4.20 and Corollary 6.4.25, there is a basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of $\mathbb{R}^n$, orthonormal for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$,

with $\mathbf{u}_1, \ldots, \mathbf{u}_r$ a basis of $\operatorname{im} P$ and $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ a basis of $(\operatorname{im} P)^{\perp} = \ker P$ (Proposition 6.4.29(a)). So $P\mathbf{u}_i = \mathbf{u}_i$ for $i \le r$ and $P\mathbf{u}_i = \mathbf{0}$ for $i > r$. Let S be the invertible matrix with columns $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then $PS = SD$ with $D = \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ (r ones), so $P = SDS^{-1}$ and, by Proposition 2.4.28, $\operatorname{tr} P = \operatorname{tr}\big(DS^{-1}S\big) = \operatorname{tr} D = r$.

If $\sum_M \operatorname{tr} M = 0$, then $\operatorname{tr} P = |\mathcal{G}|^{-1} \sum_M \operatorname{tr} M = 0$, so $\operatorname{rank} P = 0$, $P = 0$, and $\sum_M M = |\mathcal{G}| P = 0$.

(d) Take S with columns $\mathbf{u}_1, \ldots, \mathbf{u}_n$ forming any basis that is orthonormal for $\langle \cdot, \cdot \rangle_{\mathcal{G}}$, as in (c). For $\mathbf{x} \in \mathbb{R}^n$, $S^{-1}\mathbf{x}$ is the coordinate vector of $\mathbf{x}$ in this basis, since $S\big(S^{-1}\mathbf{x}\big) = \mathbf{x}$. By Theorem 6.3.22, $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}} = \big(S^{-1}\mathbf{x}\big) \cdot \big(S^{-1}\mathbf{y}\big)$. Let $M \in \mathcal{G}$ and $Q = S^{-1}MS$. Then for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, with $\mathbf{x} = S\mathbf{a}$ and $\mathbf{y} = S\mathbf{b}$,

$$(Q\mathbf{a}) \cdot (Q\mathbf{b}) = \big(S^{-1}M\mathbf{x}\big) \cdot \big(S^{-1}M\mathbf{y}\big) = \langle M\mathbf{x}, M\mathbf{y} \rangle_{\mathcal{G}} = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{G}} = \mathbf{a} \cdot \mathbf{b}.$$

With $\mathbf{a} = \mathbf{e}_j$ and $\mathbf{b} = \mathbf{e}_i$, the columns of Q are orthonormal, so Q is orthogonal (Theorem 6.3.29).

By (c), P is the orthogonal projection onto the vectors fixed by all of $\mathcal{G}$: if $M\mathbf{x} = \mathbf{x}$ for all M then $P\mathbf{x} = \mathbf{x}$, and conversely $MP\mathbf{y} = P\mathbf{y}$ by (b). The idempotence $P^2 = P$ alone already gives $\operatorname{tr} P = \operatorname{rank} P$, by the decomposition $\mathbb{R}^n = \operatorname{im} P \oplus \ker P$ of Section 4.4, so the inner product is not needed for the conclusion of (c). It explains what P is, and it gives (d): every finite group of real matrices is conjugate to a group of orthogonal matrices. Averaging over the group in this way is the basic tool of the representation theory of finite groups.

This is Problem B6 of the 1985 Putnam Competition, in a slightly restricted form: the original allows any finite set of real matrices that forms a group under multiplication, whose identity element need not be I_n.

Section 7.1

Solution to Exercise 7.1.8

Uses. Proposition 7.1.3 ($F[x]$ is a commutative ring), Definition 7.1.1 (Product of polynomials)
Hint 1. For the first product in (a), write the factors as $\big(x^2 + 1\big) - x$ and $\big(x^2 + 1\big) + x$. In the second, most terms cancel in pairs.
Hint 2. For (b), expand as over $\mathbb{Z}$, then reduce the coefficients modulo 2 or 3.
Sketch. (a) $\big(x^2 + 1\big)^2 - x^2 = x^4 + x^2 + 1$, and the second product telescopes to $1 - x^n$. (b) $\big(1 + x + x^2\big)^2 = 1 + 2x + 3x^2 + 2x^3 + x^4$, which is $1 + x^2 + x^4$ over $\mathbb{F}_2$; $(1 + x)^3 = 1 + 3x + 3x^2 + x^3$, which is $1 + x^3$ over $\mathbb{F}_3$.

(a) Since $F[x]$ is a commutative ring, $(u - v)(u + v) = u^2 - v^2$. With $u = x^2 + 1$ and $v = x$,

$$\begin{aligned}\big(x^2 - x + 1\big)\big(x^2 + x + 1\big) &= \big(x^2 + 1\big)^2 - x^2 \\ &= x^4 + x^2 + 1.\end{aligned}$$

For the second product,

$$(1 - x) \sum_{k=0}^{n-1} x^k = \sum_{k=0}^{n-1} x^k - \sum_{k=0}^{n-1} x^{k+1} = 1 - x^n,$$

since every x^k with $1 \le k \le n - 1$ occurs once in each sum and cancels.

(b) Expanding, $\big(1 + x + x^2\big)^2 = 1 + 2x + 3x^2 + 2x^3 + x^4$. In $\mathbb{F}_2$, $2 = 0$ and $3 = 1$, so the square is $1 + x^2 + x^4$. Similarly $(1 + x)^3 = 1 + 3x + 3x^2 + x^3$, and in $\mathbb{F}_3$ this is $1 + x^3$.

Solution to Exercise 7.1.18

Uses. Theorem 7.1.14 (Division with remainder), Algorithm 7.1.15 (Long division), Definition 7.1.16 (Divisibility)
Hint 1. Run Algorithm 7.1.15 over $\mathbb{Q}$. Since s is monic, no division by a leading coefficient is needed.
Hint 2. The computation over $\mathbb{Q}$ uses only integer arithmetic, so it can be reduced modulo 2 and modulo 3. Why does reducing give the right answer?
Sketch. Over $\mathbb{Q}$: $p - x^2 s = x^3 - x^2 + x + 2$, minus xs gives $-x^2 + 2$, plus s gives 3. So $q = x^2 + x - 1$ and $r = 3$. Reducing $p = sq + 3$ modulo 2 and 3 gives a valid division there, so by uniqueness $r = 1$ over $\mathbb{F}_2$ and $r = \mathbf{0}$ over $\mathbb{F}_3$. Only over $\mathbb{F}_3$ does s divide p.

Over $\mathbb{Q}$, long division gives

$$\begin{aligned} p - x^2 s &= x^3 - x^2 + x + 2, \\ p - (x^2 + x)s &= -x^2 + 2, \\ p - (x^2 + x - 1)s &= 3. \end{aligned}$$

So $p = (x^2+1)(x^2+x-1) + 3$, with quotient $x^2 + x - 1$ and remainder 3. Since $3 \neq 0$, s does not divide p in $\mathbb{Q}[x]$.

Every coefficient in this identity is an integer, and the identity was obtained by adding and multiplying integers, so it remains true when every coefficient is read in $\mathbb{F}_2$ or $\mathbb{F}_3$. Over $\mathbb{F}_2$, it reads $p = (x^2+1)(x^2+x+1) + 1$, and $\deg 1 = 0 < 2$, so by the uniqueness in Theorem 7.1.14 the quotient is $x^2 + x + 1$ and the remainder is 1. Over $\mathbb{F}_3$, $3 = 0$, so $p = (x^2+1)(x^2+x+2)$ (as $-1 = 2$), the remainder is $\mathbf{0}$, and $s \mid p$. So s divides p only over $\mathbb{F}_3$.

Solution to Exercise 7.1.28

> *Uses.* Theorem 7.1.20 (Factor theorem), Lemma 7.1.25 (Reading off a multiplicity), Exercise 2.1.49, Proposition 7.1.10 (Evaluation respects sums and products)
> *Hint 1.* Both conditions include that c is a root. So write $p = (x-c)q$ and compute p' with the product rule (Exercise 2.1.49).
> *Hint 2.* You should find $p'(c) = q(c)$. Compare with Lemma 7.1.25 for $m = 1$.
> *Sketch.* If c is not a root, both sides are false. If it is, $p = (x-c)q$, and the product rule gives $p' = q + (x-c)q'$, so $p'(c) = q(c)$. By Lemma 7.1.25, c is simple exactly when $q(c) \neq 0$, that is, when $p'(c) \neq 0$.

If $p(c) \neq 0$, then c is not a root, so it is not a repeated root, and both sides of the equivalence are false. So suppose $p(c) = 0$, and write $p = (x-c)q$ (Theorem 7.1.20). The formal derivative satisfies the product rule (Exercise 2.1.49(a)), and $(x-c)' = 1$, so

$$p' = q + (x-c)q', \quad \text{hence} \quad p'(c) = q(c) + 0 \cdot q'(c) = q(c)$$

by Proposition 7.1.10. By Lemma 7.1.25 with $m = 1$, the root c is simple if and only if $q(c) \neq 0$. So c is repeated if and only if $q(c) = 0$, that is, if and only if $p'(c) = 0$.

Solution to Exercise 7.1.49

> *Uses.* Theorem 7.1.20 (Factor theorem), Example 7.1.26 (Multiplicity depends on the field), Proposition 7.1.34 (Roots of a split polynomial)
> *Hint 1.* Check that 1 is a root, and split off $x - 1$ with Horner's scheme.
> *Hint 2.* The quotient is the polynomial of Example 7.1.26. The resulting identity has integer coefficients, so it can be read in $\mathbb{F}_3$ and $\mathbb{F}_2$.
> *Sketch.* $p = (x-1)(x^3 - x^2 - x + 1) = (x-1)^3(x+1)$. Over $\mathbb{Q}$: 1 with multiplicity 3 and -1 with multiplicity 1. Over $\mathbb{F}_3$, $x + 1 = x - 2$: 1 with multiplicity 3 and 2 with multiplicity 1. Over $\mathbb{F}_2$, $x + 1 = x - 1$: 1 with multiplicity 4.

Since $p(1) = 1 - 2 + 2 - 1 = 0$, Horner's scheme with $c = 1$ applies to the coefficients $1, -2, 0, 2, -1$: it gives $1, -2 + 1 = -1, 0 - 1 = -1, 2 - 1 = 1$, and remainder $-1 + 1 = 0$. So $p = (x-1)(x^3 - x^2 - x + 1)$, and by Example 7.1.26, $x^3 - x^2 - x + 1 = (x-1)^2(x+1)$. Hence

$$p = (x-1)^3(x+1).$$

This identity was derived with integer arithmetic, so it holds over every field when the integers are read in that field. By Proposition 7.1.34, the roots and multiplicities can be read off from a factorization into linear factors.

Over $\mathbb{Q}$, the roots are 1, with multiplicity 3, and -1, with multiplicity 1. Over $\mathbb{F}_3$, $x + 1 = x - 2$, so $p = (x-1)^3(x-2)$: the roots are 1, with multiplicity 3, and 2, with multiplicity 1. Over $\mathbb{F}_2$, $x + 1 = x - 1$, so $p = (x-1)^4$, and 1 is the only root, with multiplicity 4. In each case the multiplicities add up to $4 = \deg p$.

Solution to Exercise 7.1.50

Uses. Theorem 7.1.14 (Division with remainder), Proposition 7.1.10 (Evaluation respects sums and products), Corollary 7.1.17 (Division does not depend on the field), Exercise 2.1.49
Hint 1. The remainder has the form $ax + b$. Evaluate $x^{100} = sq + ax + b$ at the roots of s to get equations for a and b.
Hint 2. For (b), the roots $\pm i$ are not rational. Why may you divide in $\mathbb{C}[x]$ instead? For (c), there is only one root; differentiate to get a second equation.
Sketch. (a) At ± 1: $1 = a + b$ and $1 = -a + b$, so the remainder is 1. (b) Over $\mathbb{C}$, at $\pm i$: $i^{100} = 1 = \pm ai + b$, so the remainder is 1; by Corollary 7.1.17 it is also the remainder over $\mathbb{Q}$. (c) At 1: $a + b = 1$; differentiating, $100x^{99} = 2(x-1)q + (x-1)^2 q' + a$, so $a = 100$ and the remainder is $100x - 99$.

In each case s has degree 2, so $x^{100} = sq + ax + b$ for some q and some $a, b \in \mathbb{Q}$ (Theorem 7.1.14). We evaluate with Proposition 7.1.10.

(a) With $s = x^2 - 1$, evaluating at 1 and at -1 gives $1 = a + b$ and $1 = -a + b$. So $a = 0$, $b = 1$, and the remainder is 1.

(b) By Corollary 7.1.17, the quotient and remainder over $\mathbb{Q}$ are also those over $\mathbb{C}$, so we may evaluate at the complex roots $\pm i$ of $x^2 + 1$. Since $i^4 = 1$, $(\pm i)^{100} = 1$, so $1 = ai + b$ and $1 = -ai + b$. Adding and subtracting gives $b = 1$ and $a = 0$. The remainder is 1.

(c) With $s = (x-1)^2$, evaluating at 1 gives $1 = a + b$. To get a second equation, apply the formal derivative, using the product rule (Exercise 2.1.49):

$$100x^{99} = 2(x-1)q + (x-1)^2 q' + a.$$

Evaluating at 1 gives $100 = a$, so $b = -99$, and the remainder is $100x - 99$.

Solution to Exercise 7.1.51

Uses. Theorem 7.1.14 (Division with remainder), Proposition 1.2.18 (Degree and evaluation versus the operations), Proposition 7.1.4 (Degree of a product), Corollary 7.1.7 (No zero divisors, cancellation, and units), Theorem 4.2.11 (Rank–nullity), Example 4.1.7 (Dimensions of the standard spaces)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), add two divisions $p_1 = sq_1 + r_1$ and $p_2 = sq_2 + r_2$ (one multiplied by c), and use uniqueness in Theorem 7.1.14.
Hint 2. For (b), what is $R(r)$ when $\deg r < m$? For the kernel, $R(p) = \mathbf{0}$ means $s \mid p$; bound the degree of the quotient.
Hint 3. For (c), $q \mapsto sq$ is an injective linear map from $F[x]_{\leq n-m}$ onto $\ker R$.
Sketch. (a) $cp_1 + p_2 = s(cq_1 + q_2) + (cr_1 + r_2)$ with $\deg(cr_1 + r_2) < m$, so by uniqueness $R(cp_1 + p_2) = cR(p_1) + R(p_2)$. (b) $R(r) = r$ for $\deg r < m$. $R(p) = \mathbf{0}$ iff $p = sq$, and then $\deg q = \deg p - m \leq n - m$. (c) $\dim \ker R = n - m + 1$ and $\dim \operatorname{im} R = m$ add up to $n + 1$.

(a) Let $p_1, p_2 \in F[x]_{\leq n}$ and $c \in F$, with $p_1 = sq_1 + r_1$ and $p_2 = sq_2 + r_2$, where $r_i = R(p_i)$. Then

$$cp_1 + p_2 = s(cq_1 + q_2) + (cr_1 + r_2),$$

and $\deg(cr_1 + r_2) < m$ by Proposition 1.2.18. By the uniqueness in Theorem 7.1.14, $R(cp_1 + p_2) = cr_1 + r_2 = cR(p_1) + R(p_2)$.

(b) If $\deg r \leq m - 1$, then $r = s \cdot \mathbf{0} + r$, so $R(r) = r$; thus R is surjective onto $F[x]_{\leq m-1}$ (note $m - 1 \leq n$). Next, $R(p) = \mathbf{0}$ if and only if $p = sq$ for some q. If $p = sq \neq \mathbf{0}$ has degree at most n, then $\deg q = \deg p - m \leq n - m$ by Proposition 7.1.4. Conversely, if $\deg q \leq n - m$, then $\deg(sq) \leq n$ and $R(sq) = \mathbf{0}$. So $\ker R = \{sq \mid q \in F[x]_{\leq n-m}\}$.

(c) The map $q \mapsto sq$ from $F[x]_{\leq n-m}$ to $\ker R$ is linear, onto by (b), and injective by Corollary 7.1.7(b). So $\dim \ker R = \dim F[x]_{\leq n-m} = n - m + 1$ (Example 4.1.7), while $\dim \operatorname{im} R = \dim F[x]_{\leq m-1} = m$. Their sum is $n + 1 = \dim F[x]_{\leq n}$, as Theorem 4.2.11 requires.

Solution to Exercise 7.1.52

Uses. Proposition 7.1.22 (Counting roots), Proposition 1.1.17 (No zero divisors), Definition 1.1.19 (Integers modulo n)
Hint 1. Square each odd residue modulo 8.
Hint 2. Follow the proof with $c_1 = 1$ and $c_2 = 3$. After splitting off $x - 1$, what is q, and what is $q(3)$?

Sketch. (a) $1^2, 3^2, 5^2, 7^2$ are all 1 modulo 8, and $(x-3)(x+3) = x^2 - 9 = x^2 - 1$. (b) With $c_1 = 1$, $q = x + 1$, and $p(3) = (3-1)q(3) = 2 \cdot 4 = 0$ although $q(3) = 4 \neq 0$. The step "the product of the differences is nonzero, so $q(c_k) = 0$" uses Proposition 1.1.17, which fails in $\mathbb{Z}/8\mathbb{Z}$.

(a) Modulo 8, $1^2 = 1$, $3^2 = 9 = 1$, $5^2 = 25 = 1$, and $7^2 = 49 = 1$, so $1, 3, 5, 7$ are roots of $x^2 - 1$. Also $(x-3)(x+3) = x^2 - 9 = x^2 - 1$, since $9 = 1$ in $\mathbb{Z}/8\mathbb{Z}$. So $x^2 - 1$ has two different factorizations into monic linear factors.

(b) Take $c_1 = 1$ and $c_2 = 3$. The factor theorem still holds, because dividing by the monic polynomial $x - 1$ never requires inverting a scalar, and it gives $x^2 - 1 = (x-1)q$ with $q = x + 1$. The proof then evaluates at $c_2 = 3$:

$$0 = p(3) = (3-1)q(3) = 2 \cdot 4.$$

The next step concludes $q(3) = 0$ because $c_2 - c_1 = 2$ is nonzero, using Proposition 1.1.17. In $\mathbb{Z}/8\mathbb{Z}$ this is false: $2 \cdot 4 = 0$ while $2 \neq 0$ and $q(3) = 4 \neq 0$. That is the step that fails. (The degree formula of Proposition 7.1.4 also fails over $\mathbb{Z}/8\mathbb{Z}$, for instance $(4x+1)(2x+1) = 6x + 1$, but the proof only uses it for monic factors, where it still holds.)

Solution to Exercise 7.1.53

Uses. Corollary 7.1.23 (Polynomials and their values), Definition 7.1.16 (Divisibility), Corollary 7.1.43 (Odd degree gives a real root), Theorem 7.1.39 (Factorization over $\mathbb{R}$)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Try small cases and small fields (Technique 1.2.38)
Hint 1. Try small fields for (a), constants for (b), and repeated factors for (c).
Hint 2. For (d) and (e), use Theorem 7.1.39. What does it say about the number of real roots counted with multiplicity?
Sketch. (a) False over $\mathbb{F}_2$: $x^2 + x$ and $\mathbf{0}$; true over infinite fields. (b) False: x and $2x$ over $\mathbb{Q}$; true for monic polynomials. (c) False: $x^2 \mid x \cdot x$. (d) True by Corollary 7.1.43 (a polynomial with no real root is nonzero). (e) True: in Theorem 7.1.39, k counts the real roots with multiplicity and $k + 2l = \deg p$.

(a) False. Over $\mathbb{F}_2$, $x^2 + x$ and $\mathbf{0}$ both vanish at 0 and 1 (Aside (Polynomials versus polynomial functions)). It is true when F is infinite, by Corollary 7.1.23(b).

(b) False. Over $\mathbb{Q}$, $x \mid 2x$ and $2x \mid x$, since $x = \frac{1}{2} \cdot 2x$. Polynomials that divide each other differ by a nonzero constant factor, so the statement is true if p and q are monic (see the remarks after Definition 7.1.16).

(c) False. x^2 divides $x \cdot x$ but not x, by degrees. It is true when p is irreducible (Exercise 7.1.68).

(d) True. A polynomial with no real root is nonzero, since every real number is a root of $\mathbf{0}$. If its degree were odd, it would have a real root by Corollary 7.1.43. So its degree is even (possibly 0).

(e) True. In Theorem 7.1.39, $\lambda_1, \ldots, \lambda_k$ are the real roots of p listed with multiplicity, so the number of real roots counted with multiplicity is $k = \deg p - 2l$, which has the parity of $\deg p$.

Solution to Exercise 7.1.54

Uses. Exercise 7.1.28, Corollary 7.1.35 (Factorization over $\mathbb{C}$), Proposition 7.1.34 (Roots of a split polynomial), the binomial theorem in a commutative ring
Hint 1. For (a), use Exercise 7.1.28: can $x^n - 1$ and its derivative have a common root?
Hint 2. For (b), expand $(x-1)^p$ with the binomial theorem. Which binomial coefficients $\binom{p}{k}$ are divisible by p?
Sketch. (a) A repeated root c would satisfy $c^n = 1$ and $nc^{n-1} = 0$, which forces $c = 0$, but $0^n \neq 1$. So all n roots given by Corollary 7.1.35 are simple, hence distinct. (b) For $0 < k < p$, p divides $\binom{p}{k} = p!/(k!(p-k)!)$, so $(x-1)^p = x^p + (-1)^p = x^p - 1$ in $\mathbb{F}_p[x]$ (also for $p = 2$, where $1 = -1$).

(a) The derivative of $f = x^n - 1$ is $f' = nx^{n-1}$, where $n = 1 + \cdots + 1 \neq 0$ in $\mathbb{C}$. If c were a repeated root of f, then $f'(c) = nc^{n-1} = 0$ by Exercise 7.1.28. For $n = 1$ this is impossible, as $f' = 1$; for $n \geq 2$ it forces $c = 0$, but $f(0) = -1 \neq 0$. So every root of f is simple. By Corollary 7.1.35, f has n roots counted with multiplicity, and since each multiplicity is 1, these are n distinct roots.

(b) $\mathbb{F}_p[x]$ is a commutative ring, so the binomial theorem holds there, with the usual inductive proof:

$$(x-1)^p = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} x^k.$$

For $0 < k < p$, the integer $\binom{p}{k} = p!/(k!(p-k)!)$ is divisible by p, because the prime p divides the numerator but none of the factors $1, \dots, k$ and $1, \dots, p-k$ of the denominator. So these coefficients are 0 in $\mathbb{F}_p$, and $(x-1)^p = x^p + (-1)^p$. For odd p, $(-1)^p = -1$, and for $p = 2$, $(-1)^2 = 1 = -1$ in $\mathbb{F}_2$. Hence $(x-1)^p = x^p - 1$, and by Proposition 7.1.34, 1 is the only root, with multiplicity p. The test of Exercise 7.1.28 does not separate roots here, since the derivative px^{p-1} is $\mathbf{0}$ in $\mathbb{F}_p[x]$.

Solution to Exercise 7.1.55

Uses. Exercise 1.1.51, Proposition 7.1.22 (Counting roots), Proposition 7.1.4 (Degree of a product), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 7.1.10 (Evaluation respects sums and products), Lemma 1.1.21 (Arithmetic in $\mathbb{Z}/n\mathbb{Z}$)
Techniques. Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. By Fermat's little theorem (Exercise 1.1.51), every $a \in \mathbb{F}_p$ is a root of $x^p - x$. What does Proposition 7.1.22 then say?
Hint 2. Both sides of (a) are monic of degree p, and one divides the other.
Hint 3. For (b), cancel the factor x from both sides of (a) and compare constant terms.
Sketch. Each $a \in \mathbb{F}_p$ is a root of $x^p - x$ (for $a \neq 0$, $a^{p-1} = 1$), so $\omega = \prod_a (x-a)$ divides $x^p - x$ by Proposition 7.1.22. Both are monic of degree p, so the quotient is 1. Cancelling x gives $x^{p-1} - 1 = \prod_{a \neq 0}(x-a)$; at 0 this reads $-1 = (-1)^{p-1}(p-1)!$ in $\mathbb{F}_p$, and $(-1)^{p-1} = 1$ for odd p, while $p = 2$ is checked directly.

(a) Every $a \in \mathbb{F}_p$ is a root of $x^p - x$: for $a = 0$ this is clear, and for $a \neq 0$, $a^{p-1} = 1$ by Exercise 1.1.51, so $a^p = a$. The p elements of $\mathbb{F}_p$ are distinct roots, so by Proposition 7.1.22, $\omega = \prod_{a \in \mathbb{F}_p}(x-a)$ divides $x^p - x$, say $x^p - x = \omega h$. Both ω and $x^p - x$ are monic of degree p, so by Proposition 7.1.4 h has degree 0 and leading coefficient 1, that is, $h = 1$.

(b) Write $\omega = x\prod_{a \neq 0}(x-a)$ and $x^p - x = x\left(x^{p-1} - 1\right)$. Cancelling x (Corollary 7.1.7(b)) gives

$$x^{p-1} - 1 = \prod_{a \in \mathbb{F}_p, a \neq 0} (x-a).$$

Evaluating at 0 (Proposition 7.1.10), $-1 = \prod_{a \neq 0}(-a) = (-1)^{p-1} \cdot (1 \cdot 2 \cdots (p-1))$, where the product is computed in $\mathbb{F}_p$, so it is $(p-1)!$ mod p (Lemma 1.1.21). If p is odd, $(-1)^{p-1} = 1$, so $(p-1)! \equiv -1 \pmod p$. If $p = 2$, then $(p-1)! = 1 \equiv -1 \pmod 2$ directly.

Solution to Exercise 7.1.56

Uses. Corollary 7.1.12 (Substituting twice), Definition 1.1.28 (Characteristic), Corollary 7.1.23 (Polynomials and their values), Exercise 7.1.55, Proposition 7.1.10 (Evaluation respects sums and products)
Hint 1. Evaluate $f(x+1) = f$ at $t \in F$ with Corollary 7.1.12. What does this say about $f(0), f(1), f(1+1), \dots$?
Hint 2. In characteristic 0, the elements $k \cdot 1$ for $k = 0, 1, 2, \dots$ are distinct. For (b), substitute $x+1$ into the factorization of Exercise 7.1.55.
Sketch. (a) $f(t+1) = f(t)$ for all t, so $f(k \cdot 1) = f(0)$ for all $k \geq 0$. In characteristic 0 these are infinitely many distinct points, so $f - f(0)$ is $\mathbf{0}$. (b) $f = x^p - x = \prod_a (x-a)$; substituting $x+1$ gives $\prod_a (x-(a-1))$, the same product reordered.

(a) By Corollary 7.1.12 with $r = x+1$, for every $t \in F$ we have $f(t+1) = (f(x+1))(t) = f(t)$. By induction, $f(k \cdot 1) = f(0)$ for every integer $k \geq 0$, where $k \cdot 1 = 1 + \cdots + 1$. In characteristic 0 these elements are distinct: if $j \cdot 1 = k \cdot 1$ with $j < k$, then $(k-j) \cdot 1 = 0$ with $k - j \geq 1$, which is impossible by Definition 1.1.28. So $f - f(0)$ has infinitely many roots, hence $f - f(0) = \mathbf{0}$ by Corollary 7.1.23(b), and f is the constant $f(0)$.

(b) Let $f = x^p - x$, which has degree $p \geq 2$. By Exercise 7.1.55, $f = \prod_{a \in \mathbb{F}_p}(x-a)$. Substituting $x+1$ and using Proposition 7.1.10 for the product,

$$f(x+1) = \prod_{a \in \mathbb{F}_p} (x+1-a) = \prod_{a \in \mathbb{F}_p} (x-(a-1))$$

$$= \prod_{b \in \mathbb{F}_p} (x - b) = f,$$

because $a \mapsto a - 1$ is a bijection of $\mathbb{F}_p$, so the last two products have the same factors in a different order. (Alternatively, $(x+1)^p = x^p + 1$ in $\mathbb{F}_p[x]$ by the binomial argument of Exercise 7.1.54.)

Solution to Exercise 7.1.57

Uses. Exercise 7.1.28, Proposition 7.1.27 (Multiplicities add up to at most the degree), Corollary 4.2.18 (Injective if and only if surjective), Example 2.1.9 (Maps of polynomials), Example 4.1.7 (Dimensions of the standard spaces)
Techniques. Uniqueness gives existence (Technique 3.4.21), Dimension counting (Technique 4.2.24)
Hint 1. The map $f \mapsto (f(0), f'(0), f(1), f'(1))$ from $F[x]_{\le 3}$ to F^4 is linear, and both spaces have dimension 4. So it is enough to show that it is injective.
Hint 2. If all four values are 0, then 0 and 1 are repeated roots of f (Exercise 7.1.28). Count multiplicities.
Hint 3. For (b), 1 is a repeated root, so $f = (x-1)^2(ux+v)$.
Sketch. The map $T(f) = (f(0), f'(0), f(1), f'(1))$ is linear from a 4-dimensional space to F^4. If $T(f) = \mathbf{0}$ and $f \neq \mathbf{0}$, then 0 and 1 are roots of multiplicity at least 2, so $4 \le \deg f \le 3$ by Proposition 7.1.27, which is impossible. So T is injective, hence bijective. For (b), $f = (x-1)^2(ux+v)$ with $v = 1$ and $u - 2v = 0$, so $f = (x-1)^2(2x+1) = 2x^3 - 3x^2 + 1$.

(a) Let $T : F[x]_{\le 3} \to F^4$, $T(f) = (f(0), f'(0), f(1), f'(1))$. Evaluation and the formal derivative are linear (Example 2.1.9), so T is linear. Suppose $T(f) = \mathbf{0}$ and $f \neq \mathbf{0}$. By Exercise 7.1.28, 0 and 1 are repeated roots of f, so each has multiplicity at least 2, and Proposition 7.1.27 gives $2 + 2 \le \deg f \le 3$, which is impossible. So $\ker T = \{\mathbf{0}\}$, and T is injective. Since $\dim F[x]_{\le 3} = 4 = \dim F^4$ (Example 4.1.7), T is bijective by Corollary 4.2.18. So each (a, b, c, d) is $T(f)$ for exactly one f.

(b) Here $f(1) = f'(1) = 0$, so 1 is a repeated root (or $f = \mathbf{0}$, which is excluded by $f(0) = 1$), and $(x-1)^2$ divides f. Since $\deg f \le 3$, $f = (x-1)^2(ux+v)$ for some $u, v \in \mathbb{Q}$. Then $f(0) = v = 1$, and by the product rule $f' = 2(x-1)(ux+v) + u(x-1)^2$, so $f'(0) = -2v + u = 0$ and $u = 2$. Hence

$$f = (x-1)^2(2x+1) = 2x^3 - 3x^2 + 1.$$

Check: $f(0) = 1$, $f(1) = 0$, and $f' = 6x^2 - 6x$ vanishes at 0 and 1.

Solution to Exercise 7.1.58

Uses. Proposition 3.4.27 (Lagrange's formula), Theorem 3.4.24 (Polynomial interpolation), Proposition 7.1.29 (Lagrange polynomials in closed form), Exercise 2.1.49, Proposition 7.1.10 (Evaluation respects sums and products)
Techniques. Uniqueness via two candidates (Technique 1.1.8)
Hint 1. For (a), both sides of each identity have degree at most n. Compare their values at the nodes.
Hint 2. For (b), write $\omega = (x - c_i)\omega_i$ and differentiate with the product rule.
Sketch. (a) By Proposition 3.4.27, $\sum_i d_i L_i$ is the interpolant of the values d_i. The constant 1 and the polynomial x (of degree $1 \le n$) interpolate the values 1 and c_i, so they equal $\sum_i L_i$ and $\sum_i c_i L_i$. (b) $\omega' = \omega_i + (x - c_i)\omega_{i'}$, so $\omega'(c_i) = \omega_i(c_i) = \prod_{j \neq i}(c_i - c_j)$, and Proposition 7.1.29 gives $L_i = \omega_i / \omega'(c_i)$.

(a) By Proposition 3.4.27, for any values $d_0, \ldots, d_n$, the polynomial $\sum_i d_i L_i$ is the unique element of $F[x]_{\le n}$ taking the value d_k at c_k. The constant polynomial 1 lies in $F[x]_{\le n}$ and takes the value 1 at every node, so $\sum_i L_i = 1$. Since $n \ge 1$, the polynomial x lies in $F[x]_{\le n}$ and takes the value c_k at c_k, so $\sum_i c_i L_i = x$.

(b) Write $\omega = (x - c_i)\omega_i$. By the product rule (Exercise 2.1.49), $\omega' = \omega_i + (x - c_i)\omega_{i'}$, so by Proposition 7.1.10

$$\omega'(c_i) = \omega_i(c_i) + 0 = \prod_{j \neq i}(c_i - c_j).$$

This is nonzero, and by Proposition 7.1.29, $L_i = \omega_i / \prod_{j \neq i}(c_i - c_j) = \omega_i / \omega'(c_i)$.

Solution to Exercise 7.1.59

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Proposition 7.1.36 (Sum and product of the roots)
Hint 1. Multiply out one factor at a time.
Hint 2. By Corollary 7.1.35, $x^3 - 2x + 5 = (x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$. Compare coefficients. For the last sum, square $\lambda_1 + \lambda_2 + \lambda_3$.

Sketch. (a) $x^3 - e_1x^2 + e_2x - e_3$ with $e_1 = \sum \lambda_i$, $e_2 = \sum_{i<j} \lambda_i\lambda_j$, $e_3 = \lambda_1\lambda_2\lambda_3$. (b) Comparing with $x^3 - 2x + 5$: $e_1 = 0$, $e_2 = -2$, $e_3 = -5$, and $\sum \lambda_i^2 = e_1^2 - 2e_2 = 4$.

(a) First $(x - \lambda_1)(x - \lambda_2) = x^2 - (\lambda_1 + \lambda_2)x + \lambda_1\lambda_2$. Multiplying by $x - \lambda_3$,

$$(x - \lambda_1)(x - \lambda_2)(x - \lambda_3) = x^3 - e_1x^2 + e_2x - e_3,$$

where $e_1 = \lambda_1 + \lambda_2 + \lambda_3$, $e_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$, and $e_3 = \lambda_1\lambda_2\lambda_3$. The first and last agree with Proposition 7.1.36.

(b) By Corollary 7.1.35, $x^3 - 2x + 5 = (x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$, so comparing coefficients with (a) gives $e_1 = 0$, $e_2 = -2$, and $-e_3 = 5$, that is, $e_3 = -5$. Finally, expanding the square gives $e_1^2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + 2e_2$, so

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = e_1^2 - 2e_2 = 0 - 2(-2) = 4.$$

Solution to Exercise 7.1.60

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Theorem 7.1.39 (Factorization over $\mathbb{R}$), Corollary 7.1.42 (Irreducible polynomials over $\mathbb{C}$ and $\mathbb{R}$), Proposition 7.1.34 (Roots of a split polynomial)
Hint 1. For $x^4 + 1$, find a complex number whose square is i; its fourth power is -1.
Hint 2. Over $\mathbb{R}$, pair conjugate roots as in the proof of Theorem 7.1.39. For $x^6 - 1$, start from $\big(x^3 - 1\big)\big(x^3 + 1\big)$.
Sketch. The roots of $x^4 + 1$ are $(\pm 1 \pm i)/\sqrt{2}$, and pairing conjugates gives $x^4 + 1 = \big(x^2 - \sqrt{2}x + 1\big)\big(x^2 + \sqrt{2}x + 1\big)$ over $\mathbb{R}$. Also $x^6 - 1 = (x - 1)(x + 1)\big(x^2 + x + 1\big)\big(x^2 - x + 1\big)$ over $\mathbb{R}$, and the quadratics split over $\mathbb{C}$ with roots $\big(-1 \pm i\sqrt{3}\big)/2$ and $\big(1 \pm i\sqrt{3}\big)/2$.

By Corollary 7.1.42, the irreducible factors over $\mathbb{C}$ are linear, and those over $\mathbb{R}$ are linear or quadratics with negative discriminant.

$x^4 + 1$. Let $\zeta = (1 + i)/\sqrt{2}$. Then $\zeta^2 = (1 + 2i - 1)/2 = i$, so $\zeta^4 = -1$. The same computation shows that all four numbers $(\pm 1 \pm i)/\sqrt{2}$, that is, $\pm\zeta$ and $\pm\overline{\zeta}$, have squares $\pm i$ and are roots. They are distinct, so by Corollary 7.1.35

$$x^4 + 1 = (x - \zeta)\big(x - \overline{\zeta}\big)(x + \zeta)\big(x + \overline{\zeta}\big).$$

Pairing conjugates, $(x - \zeta)\big(x - \overline{\zeta}\big) = x^2 - 2(\operatorname{Re}\zeta)x + |\zeta|^2 = x^2 - \sqrt{2}x + 1$, and similarly for $-\zeta$. So over $\mathbb{R}$,

$$x^4 + 1 = \big(x^2 - \sqrt{2}x + 1\big)\big(x^2 + \sqrt{2}x + 1\big),$$

and both quadratics have discriminant $2 - 4 < 0$. Check: the product is $\big(x^2 + 1\big)^2 - 2x^2 = x^4 + 1$.

$x^6 - 1$. Since $x^3 - 1 = (x - 1)\big(x^2 + x + 1\big)$ and $x^3 + 1 = (x + 1)\big(x^2 - x + 1\big)$,

$$x^6 - 1 = (x - 1)(x + 1)\big(x^2 + x + 1\big)\big(x^2 - x + 1\big)$$

over $\mathbb{R}$, and the two quadratics have discriminant $1 - 4 < 0$, so they are irreducible. Over $\mathbb{C}$, by the quadratic formula, $x^2 + x + 1 = (x - \omega)(x - \overline{\omega})$ with $\omega = \big(-1 + i\sqrt{3}\big)/2$, and $x^2 - x + 1 = (x + \omega)(x + \overline{\omega})$, so

$$x^6 - 1 = (x - 1)(x + 1)(x - \omega)(x - \overline{\omega})(x + \omega)(x + \overline{\omega}).$$

The six roots $\pm 1, \pm\omega, \pm\overline{\omega}$ are distinct, consistent with Exercise 7.1.54(a).

Solution to Exercise 7.1.61

Uses. Definition 7.1.41 (Irreducible polynomial), Theorem 7.1.20 (Factor theorem), Proposition 7.1.4 (Degree of a product), Example 7.1.40 ($x^4 + 4$), Exercise 7.1.8
Hint 1. If $p = gh$ with both factors of degree at least 1 and $\deg g + \deg h \le 3$, then one factor has degree 1.
Hint 2. For (b), look for a product of two quadratics without roots. Example 7.1.40 and Exercise 7.1.8 have candidates.
Sketch. (a) A root c gives $p = (x - c)q$ with $\deg q \ge 1$. Conversely, in a factorization $p = gh$ the degrees add to 2 or 3, so one factor is $ux + v$ with $u \ne 0$, and $-v/u$ is a root of p. (b) $x^4 + 4 = \big(x^2 - 2x + 2\big)\big(x^2 + 2x + 2\big)$ over $\mathbb{Q}$ and $x^4 + x^2 + 1 = \big(x^2 + x + 1\big)^2$ over $\mathbb{F}_2$. (c) The first two have no root in $\mathbb{F}_2$, so they are irreducible; $x^3 + 1$ has the root 1.

(a) If p has a root c, then $p = (x - c)q$ with $\deg q = \deg p - 1 \ge 1$ (Theorem 7.1.20), so p is reducible. Conversely, suppose $p = gh$ with $\deg g, \deg h \ge 1$. Since $\deg g + \deg h = \deg p \le 3$ (Proposition 7.1.4), one

of the factors, say g, has degree 1: $g = ux + v$ with $u \neq 0$. Then $c = -v/u$ satisfies $g(c) = 0$, so $p(c) = g(c)h(c) = 0$, and p has a root.

(b) Over $\mathbb{Q}$, $x^4 + 4 = (x^2 - 2x + 2)(x^2 + 2x + 2)$ (Example 7.1.40) is reducible, and it has no root even in $\mathbb{R}$, since $t^4 + 4 \geq 4$. Over $\mathbb{F}_2$, $(x^2 + x + 1)^2 = x^4 + x^2 + 1$ (Exercise 7.1.8), and $x^4 + x^2 + 1$ takes the value 1 at both 0 and 1, so it has no root.

(c) The polynomial $x^2 + x + 1$ takes the values 1 and $1 + 1 + 1 = 1$ at 0 and 1, and $x^3 + x + 1$ takes the values 1 and 1, so by (a) both are irreducible over $\mathbb{F}_2$. But 1 is a root of $x^3 + 1$, so it is reducible: $x^3 + 1 = (x + 1)(x^2 + x + 1)$.

Solution to Exercise 7.1.62

Uses. Exercise 7.1.61, Theorem 7.1.20 (Factor theorem), Proposition 7.1.34 (Roots of a split polynomial)
Hint 1. Count all monic quadratics, then count the reducible ones. By Exercise 7.1.61, a reducible monic quadratic has a root.
Hint 2. A reducible monic quadratic is $(x - a)(x - b)$. By Proposition 7.1.34, how many choices of a, b give the same polynomial?
Sketch. There are p^2 monic quadratics $x^2 + bx + c$. The reducible ones are the products $(x - a)(x - b)$, and by Proposition 7.1.34 two such products are equal exactly when the pairs agree up to order. There are $p + p(p-1)/2 = p(p+1)/2$ unordered pairs, so $p^2 - p(p+1)/2 = p(p-1)/2$ are irreducible. For $p = 3$: $x^2 + 1$, $x^2 + x + 2$, $x^2 + 2x + 2$.

A monic quadratic is $x^2 + bx + c$ with $b, c \in \mathbb{F}_p$, so there are p^2 of them. If one is reducible, it has a root a (Exercise 7.1.61), and then it equals $(x - a)q$ with q monic of degree 1 (Theorem 7.1.20, comparing leading coefficients), so it is $(x - a)(x - b)$ for some $a, b \in \mathbb{F}_p$. Conversely every such product is a reducible monic quadratic. By Proposition 7.1.34, $(x - a)(x - b) = (x - a')(x - b')$ if and only if the lists a, b and a', b' agree up to order. So the reducible monic quadratics correspond to unordered pairs $\{a, b\}$ with repetition allowed: p with $a = b$, and $p(p-1)/2$ with $a \neq b$. That makes $p(p+1)/2$ reducible ones, and

$$p^2 - p\frac{p+1}{2} = p\frac{p-1}{2}$$

irreducible ones.

For $p = 3$ there are 3. Checking values at $0, 1, 2$: $x^2 + 1$ takes the values $1, 2, 2$; $x^2 + x + 2$ takes $2, 1, 2$; and $x^2 + 2x + 2$ takes $2, 2, 1$. None has a root, so these are the three monic irreducible quadratics over $\mathbb{F}_3$.

Solution to Exercise 7.1.63

Uses. Lemma 7.1.37 (Conjugating polynomials), Corollary 7.1.23 (Polynomials and their values)
Hint 1. Compare f with the conjugate polynomial $\overline{f}$ at the points t_k.
Hint 2. For the example, take a polynomial that vanishes at all the t_k.
Sketch. For real t_k, $\overline{f}(t_k) = \overline{f}(\overline{t_k}) = \overline{f(t_k)} = f(t_k)$. So f and $\overline{f}$, both of degree at most n, agree at $n + 1$ points, and $f = \overline{f} \in \mathbb{R}[x]$. Without the degree bound, $i(x - t_0)\cdots(x - t_n)$ is real (zero) at every t_k.

The conjugate polynomial $\overline{f}$ has the same degree as f, so it lies in $\mathbb{C}[x]_{\leq n}$. For each k, since t_k is real and $f(t_k)$ is real, Lemma 7.1.37(b) gives

$$\overline{f}(t_k) = \overline{f}(\overline{t_k}) = \overline{f(t_k)} = f(t_k).$$

So f and $\overline{f}$ agree at the $n + 1$ distinct points $t_0, \ldots, t_n$, and $f = \overline{f}$ by Corollary 7.1.23(a). By Lemma 7.1.37(c), $f \in \mathbb{R}[x]$.

The degree bound is needed: $g = i(x - t_0)(x - t_1)\cdots(x - t_n)$ has $g(t_k) = 0 \in \mathbb{R}$ for every k, but its leading coefficient i is not real.

Solution to Exercise 7.1.64

Uses. Theorem 7.1.39 (Factorization over $\mathbb{R}$), Lemma 7.1.37 (Conjugating polynomials), Lemma 7.1.25 (Reading off a multiplicity), continuity of polynomial functions
Hint 1. First show that every real root of p has even multiplicity: a root of odd multiplicity is a point where p changes sign.

Hint 2. Then factor p as in Theorem 7.1.39, show $a > 0$, and split each quadratic factor over $\mathbb{C}$ as $(x-\mu)(x-\overline{\mu})$.
Hint 3. Find $h \in \mathbb{C}[x]$ with $p = h\overline{h}$, and write $h = f + ig$ with $f, g \in \mathbb{R}[x]$.
Sketch. If a real root λ had odd multiplicity m, then $p = (x-\lambda)^m q$ with $q(\lambda) \neq 0$, and p would change sign at λ. So $p = a\prod (x-\lambda_i)^{2k_i} \prod (x-\mu_j)(x-\overline{\mu_j})$ with $a > 0$. Let $h = \sqrt{a}\prod (x-\lambda_i)^{k_i} \prod (x-\mu_j)$. Then $\overline{h} = \sqrt{a}\prod (x-\lambda_i)^{k_i} \prod (x-\overline{\mu_j})$, so $p = h\overline{h}$, and with $h = f + ig$ this is $(f+ig)(f-ig) = f^2 + g^2$.

If $p = \mathbf{0}$, take $f = g = \mathbf{0}$. Let $p \neq \mathbf{0}$.

Real roots have even multiplicity. Let $\lambda \in \mathbb{R}$ be a root of multiplicity m, and write $p = (x-\lambda)^m q$ with $q(\lambda) \neq 0$ (Lemma 7.1.25). The function $t \mapsto q(t)$ is continuous, so it has constant sign on some interval $(\lambda - \delta, \lambda + \delta)$. If m were odd, $(t-\lambda)^m$ would be negative for $t < \lambda$ and positive for $t > \lambda$, so $p(t) = (t-\lambda)^m q(t)$ would take both signs near λ, contradicting $p \geq 0$. So m is even.

Factor. By Theorem 7.1.39, whose real linear factors are the real roots listed with multiplicity, and grouping equal linear factors,

$$p = a\prod_{i=1}^{r} (x-\lambda_i)^{2k_i} \prod_{j=1}^{l} (x^2 + b_j x + c_j),$$

with distinct real λ_i, each of even multiplicity $2k_i$ by the first step, and $b_j^2 < 4c_j$. Each quadratic factor is positive on $\mathbb{R}$, since $t^2 + b_j t + c_j = \left(t + b_j/2\right)^2 + \left(c_j - b_j^2/4\right)$. Choose $t \in \mathbb{R}$ that is not a root of p (there are finitely many roots). Then $p(t) > 0$ and every factor other than a is positive at t, so $a > 0$.

Write $p = h\overline{h}$. As in the proof of Theorem 7.1.39, $x^2 + b_j x + c_j = \left(x - \mu_j\right)\left(x - \overline{\mu_j}\right)$ for some $\mu_j \in \mathbb{C}$. Let

$$h = \sqrt{a}\prod_{i=1}^{r} (x-\lambda_i)^{k_i} \prod_{j=1}^{l} (x-\mu_j) \in \mathbb{C}[x].$$

By Lemma 7.1.37(a), and since $\sqrt{a}$ and the λ_i are real, $\overline{h} = \sqrt{a}\prod_i (x-\lambda_i)^{k_i} \prod_j \left(x - \overline{\mu_j}\right)$. Hence $h\overline{h} = p$.

Real and imaginary parts. Write $h = f + ig$, where $f, g \in \mathbb{R}[x]$ have as coefficients the real and imaginary parts of the coefficients of h. Then $\overline{h} = f - ig$, and

$$p = h\overline{h} = (f+ig)(f-ig) = f^2 + g^2.$$

Both sides lie in $\mathbb{R}[x]$, and the identity holds there because the operations of $\mathbb{R}[x]$ are those of $\mathbb{C}[x]$.

Solution to Exercise 7.1.65

Uses. Definition 7.1.9 (Substitution), Definition 7.1.1 (Product of polynomials), Proposition 7.1.4 (Degree of a product)
Techniques. Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. Try small cases: which monomials ax^n work? Compare leading coefficients.
Hint 2. Suppose $p = x^n + a_k x^k + \cdots$, where $k < n$, $a_k \neq 0$, and the omitted terms have degree less than k. Compute the coefficient of x^{n+k} on both sides.
Hint 3. On the right it is $2a_k$. On the left, $p\left(x^2\right)$ has only even powers x^{2j}, and x^{n+k} would need $j = (n+k)/2$, strictly between k and n.
Sketch. (a) The solutions are $\mathbf{0}$ and x^n, $n \geq 0$. For $p \neq \mathbf{0}$ of degree n with leading coefficient a, comparing leading coefficients gives $a = a^2$, so $a = 1$. If p had another nonzero coefficient, let a_k be the one with the largest $k < n$. The coefficient of x^{n+k} is $2a_k \neq 0$ in p^2, but 0 in $p\left(x^2\right)$. (b) Over $\mathbb{F}_2$, $p^2 = \sum a_i^2 x^{2i} + 2\sum_{i<j} a_i a_j x^{i+j} = \sum a_i x^{2i}$.

(a) The solutions are $p = \mathbf{0}$ and $p = x^n$ for $n \geq 0$; these work, since $\left(x^2\right)^n = x^{2n} = (x^n)^2$. Conversely, let $p = \sum_{j=0}^{n} a_j x^j$ with $a_n \neq 0$ satisfy $p\left(x^2\right) = p^2$. The polynomial $p\left(x^2\right) = \sum_j a_j x^{2j}$ has leading coefficient a_n, and p^2 has leading coefficient a_n^2 (Proposition 7.1.4). So $a_n^2 = a_n$, and $a_n = 1$.

Suppose $p \neq x^n$, and let $k < n$ be the largest index below n with $a_k \neq 0$. In p^2, the coefficient of x^{n+k} is $\sum_{i+j=n+k} a_i a_j$. A nonzero term needs $a_i \neq 0$ and $a_j \neq 0$, so each of i, j is n or at most k. If $i \leq k$, then $j = n + k - i \geq n$, so $j = n$ and $i = k$; similarly if $j \leq k$. So only the pairs (n,k) and (k,n) contribute, and the coefficient is $2a_n a_k = 2a_k$, which is nonzero because $1 + 1 \neq 0$ and F has no zero divisors. In $p\left(x^2\right) = \sum_j a_j x^{2j}$, the

coefficient of x^{n+k} is 0 if $n+k$ is odd, and $a_{(n+k)/2}$ if it is even. Since $k < (n+k)/2 < n$, this is 0 by the choice of k. The two sides differ, a contradiction. So $p = x^n$.

(b) Let $p = \sum_i a_i x^i \in \mathbb{F}_2[x]$. Expanding the square and grouping the pairs (i,j) and (j,i) with $i \neq j$,

$$\begin{aligned} p^2 &= \sum_{i,j} a_i a_j x^{i+j} = \sum_i a_i^2 x^{2i} + \sum_{i<j} 2a_i a_j x^{i+j} \\ &= \sum_i a_i x^{2i} = p\big(x^2\big), \end{aligned}$$

because $2 = 0$ and $a^2 = a$ for both elements a of $\mathbb{F}_2$.

Solution to Exercise 7.1.66

Uses. Corollary 7.1.35 (Factorization over $\mathbb{C}$), Exercise 2.1.49, Proposition 7.1.10 (Evaluation respects sums and products), Notation 6.1.1 (Real and complex scalars)
Hint 1. If w is a root of p, the claim is easy. Otherwise, compute p' from $p = a\prod(x-\lambda_i)$ with the product rule.
Hint 2. You should find $p'(w) = p(w)\sum_i 1/(w-\lambda_i)$, so $\sum_i 1/(w-\lambda_i) = 0$.
Hint 3. Write $1/(w-\lambda_i) = \overline{w-\lambda_i}/\left|w-\lambda_i\right|^2$, and conjugate the equation.
Sketch. If $p(w) = 0$, then $w = \lambda_i$ for some i. Otherwise $p' = a\sum_i \prod_{j\neq i}(x-\lambda_j)$ gives $0 = p'(w) = p(w)\sum_i 1/(w-\lambda_i)$, so $\sum_i \overline{w-\lambda_i}/\left|w-\lambda_i\right|^2 = 0$. Conjugating, $\sum_i s_i(w-\lambda_i) = 0$ with $s_i = \left|w-\lambda_i\right|^{-2} > 0$, so $w = \sum_i\big(s_i/\sum_j s_j\big)\lambda_i$.

If $p(w) = 0$, then $w = \lambda_i$ for some i (Corollary 7.1.35), and $t_i = 1$, $t_j = 0$ for $j \neq i$ works. So suppose $p(w) \neq 0$.

Write $p = a(x-\lambda_1)\cdots(x-\lambda_n)$. Applying the product rule (Exercise 2.1.49) $n-1$ times, by induction on the number of factors,

$$p' = a\sum_{i=1}^{n}\prod_{j\neq i}(x-\lambda_j).$$

Evaluating at w (Proposition 7.1.10) and using $a\prod_{j\neq i}(w-\lambda_j) = p(w)/(w-\lambda_i)$, which is valid since $w \neq \lambda_i$,

$$0 = p'(w) = p(w)\sum_{i=1}^{n}\frac{1}{w-\lambda_i}.$$

Since $p(w) \neq 0$, $\sum_i 1/(w-\lambda_i) = 0$. For $z \neq 0$, $1/z = \overline{z}/\left|z\right|^2$ (Notation 6.1.1), so with $s_i = 1/\left|w-\lambda_i\right|^2 > 0$ we get $\sum_i s_i\overline{w-\lambda_i} = 0$. Conjugating, $\sum_i s_i(w-\lambda_i) = 0$, that is,

$$w = \sum_{i=1}^{n}\left(\frac{s_i}{s_1+\cdots+s_n}\right)\lambda_i.$$

The coefficients $t_i = s_i/(s_1+\cdots+s_n)$ are positive and add up to 1.

Geometrically, the roots of p' lie in the smallest convex polygon containing the roots of p. For real-rooted p this is the interval between the smallest and the largest root.

Solution to Exercise 7.1.67

Uses. Algorithm 7.1.46 (Euclidean algorithm), Theorem 7.1.45 (Bézout's identity), Definition 7.1.44 (Greatest common divisor)
Hint 1. Over $\mathbb{Q}$ the algorithm takes two divisions.
Hint 2. Over $\mathbb{F}_2$, $p = x^4+x^3+x+1$. Divide it by $s = x^3+1$.
Sketch. (a) $p = (x+1)s + 2\big(x^2+x+1\big)$ and $s = \big(\frac{1}{2}x - \frac{1}{2}\big)\cdot 2\big(x^2+x+1\big)$, so $\gcd(p,s) = x^2+x+1 = \frac{1}{2}p - \frac{1}{2}(x+1)s$. (b) Over $\mathbb{F}_2$, $p = (x+1)s$ exactly, so $\gcd(p,s) = s = x^3+1$. Reducing modulo 2 is not an inclusion of fields, so the remainders, and with them the gcd, can change.

(a) Dividing, $p = (x+1)\big(x^3-1\big) + \big(2x^2+2x+2\big)$: indeed $(x+1)\big(x^3-1\big) = x^4+x^3-x-1$, and p minus this is $2x^2+2x+2$. Next, $x^3-1 = \big(\frac{1}{2}x-\frac{1}{2}\big)\big(2x^2+2x+2\big)$ exactly, since $(x-1)\big(x^2+x+1\big) = x^3-1$. The last nonzero remainder is $2x^2+2x+2$, so $\gcd(p,s) = x^2+x+1$, and from the first division

$$x^2 + x + 1 = \frac{1}{2}(p - (x+1)s) = \frac{1}{2}p - \frac{1}{2}(x+1)s.$$

Consistently, $p = \left(x^2+1\right)\left(x^2+x+1\right)$ and $s = (x-1)\left(x^2+x+1\right)$.

(b) Over $\mathbb{F}_2$, $p = x^4 + x^3 + x + 1$ and $s = x^3 + 1$, and $(x+1)\left(x^3+1\right) = x^4 + x^3 + x + 1 = p$. So $s \mid p$, the algorithm stops after one division, and $\gcd(p, s) = x^3 + 1$. The answer has larger degree than over $\mathbb{Q}$, because in $\mathbb{F}_2[x]$ the factor $x^2 + 1$ of p becomes $(x+1)^2$, which shares the factor $x + 1 = x - 1$ with s. Reducing the coefficients modulo 2 is not a passage to a larger field, so nothing like Corollary 7.1.17 applies; already the remainders change, as in Exercise 7.1.18. For a subfield $F \subseteq K$, on the other hand, the gcd computed in $K[x]$ is the one computed in $F[x]$, since the Euclidean algorithm only takes remainders.

Solution to Exercise 7.1.68

Uses. Theorem 7.1.45 (Bézout's identity), Definition 7.1.41 (Irreducible polynomial), Definition 7.1.16 (Divisibility), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 7.1.4 (Degree of a product)
Techniques. Strong induction and well-ordering (Technique 5.2.12)
Hint 1. For (a), let $d = \gcd(p, g)$. Since p is irreducible, d is 1 or p made monic. In the first case use Bézout and multiply by h.
Hint 2. For existence in (b), use strong induction on the degree. For uniqueness, apply (a) repeatedly to see that p_1 divides some q_j, and cancel.
Sketch. (a) $d = \gcd(p, g)$ divides the irreducible p, so $d = 1$ or d is p made monic. In the second case $p \mid g$. In the first, $up + vg = 1$, so $h = uph + vgh$ is divisible by p. (b) A reducible polynomial is a product of two of smaller degree, so strong induction gives existence. For uniqueness, compare leading coefficients, then $p_1 \mid q_1 \cdots q_l$, so $p_1 \mid q_j$ for some j by (a); both are monic irreducible, so $p_1 = q_j$; cancel and induct on k.

(a) Let $d = \gcd(p, g)$ (Theorem 7.1.45). Since $d \mid p$, write $p = de$. As p is irreducible, $\deg d = 0$ or $\deg e = 0$. If $\deg d = 0$, then $d = 1$, being monic, and Bézout gives $up + vg = 1$ for some u, v. Multiplying by h, $h = uph + vgh$. Both terms are divisible by p, the second because $p \mid gh$, so $p \mid h$. If $\deg e = 0$, then e is a nonzero constant and $p = ed$ divides g, because d does.

(b) *Existence*, by strong induction on $\deg f$ for f of degree at least 1. If f is irreducible, it is a product of one irreducible. Otherwise $f = gh$ with $1 \le \deg g, \deg h < \deg f$, and by the induction hypothesis both g and h are products of irreducibles, hence so is f. Writing each irreducible factor as a constant times a monic one, $f = ap_1 \cdots p_k$ with p_i monic irreducible.

Uniqueness. The product of monic polynomials is monic, so comparing leading coefficients gives $a = b$, and cancelling a leaves $p_1 \cdots p_k = q_1 \cdots q_l$. We induct on k. If $k = 0$, the left side is 1, so $l = 0$ by degrees. If $k \ge 1$, then $l \ge 1$, since otherwise $p_1 \cdots p_k = 1$ would have degree $0 < k$. So p_1 divides $q_1 \cdots q_l$, and by (a) applied $l - 1$ times, $p_1 \mid q_j$ for some j. Since q_j is irreducible and $\deg p_1 \ge 1$, $q_j = p_1 c$ with c constant, and as both are monic, $c = 1$. Cancelling $p_1 = q_j$ from both sides (Corollary 7.1.7(b)) leaves products of $k - 1$ and $l - 1$ monic irreducibles, and the induction hypothesis finishes the proof.

Solution to Exercise 7.1.69

Uses. Theorem 7.1.45 (Bézout's identity), Corollary 7.1.48 (Coprime polynomials), Theorem 7.1.14 (Division with remainder), Theorem 7.1.20 (Factor theorem), Proposition 7.1.4 (Degree of a product)
Techniques. Compare degrees and leading coefficients (Technique 7.1.6)
Hint 1. Use $um_1 + vm_2 = 1$. The polynomial vm_2 is 1 "modulo m_1" and 0 "modulo m_2".
Hint 2. Take $P = r_1 vm_2 + r_2 um_1$, then reduce it modulo $m_1 m_2$. For uniqueness, show that $m_1 m_2$ divides the difference of two solutions.
Sketch. With $um_1 + vm_2 = 1$, $P = r_1 vm_2 + r_2 um_1$ satisfies $P - r_1 = (r_2 - r_1)um_1$ and $P - r_2 = (r_1 - r_2)vm_2$. Its remainder p on division by $m_1 m_2$ still works and has degree less than $d_1 + d_2$. If p, p' both work, then $p - p' = m_1 g$ and $m_2 \mid m_1 g$, so $m_2 \mid g = um_1 g + vm_2 g$; hence $m_1 m_2 \mid p - p'$, and degrees force $p = p'$. For (b), $x - c_i$ divides $p - r_i$ iff $p(c_i) = r_i(c_i)$.

(a) By Theorem 7.1.45 there are u, v with $um_1 + vm_2 = 1$.

Existence. Let $P = r_1 vm_2 + r_2 um_1$. Using $vm_2 = 1 - um_1$ and $um_1 = 1 - vm_2$,

$$P - r_1 = r_1(vm_2 - 1) + r_2 um_1 = (r_2 - r_1)um_1,$$

$$P - r_2 = r_1 v m_2 + r_2(u m_1 - 1) = (r_1 - r_2) v m_2.$$

So $m_1 \mid P - r_1$ and $m_2 \mid P - r_2$. Divide P by $m_1 m_2$, which has degree $d_1 + d_2$ (Proposition 7.1.4): $P = m_1 m_2 q + p$ with $\deg p < d_1 + d_2$ (Theorem 7.1.14). Then $p - r_1 = (P - r_1) - m_1 m_2 q$ is divisible by m_1, and likewise $m_2 \mid p - r_2$.

Uniqueness. Suppose p and p' both have the property. Then m_1 and m_2 both divide $e = p - p'$, since $e = (p - r_i) - (p' - r_i)$. Write $e = m_1 g$. Then

$$g = (u m_1 + v m_2) g = u e + v m_2 g,$$

and m_2 divides both terms, so $m_2 \mid g$ and $m_1 m_2 \mid e$. But $\deg e < d_1 + d_2 = \deg(m_1 m_2)$, so $e = \mathbf{0}$ by Proposition 7.1.4. Hence $p = p'$.

(b) Let $c_1 \neq c_2$ and $y_1, y_2 \in F$, and let $m_i = x - c_i$ and $r_i = y_i$ (constants). These are coprime, since $(x - c_1) - (x - c_2) = c_2 - c_1$ is a nonzero constant and Corollary 7.1.48(a) applies after dividing by it. By Theorem 7.1.20, $x - c_i$ divides $p - y_i$ if and only if $p(c_i) = y_i$. So (a) says that there is exactly one p of degree less than 2 with $p(c_1) = y_1$ and $p(c_2) = y_2$, which is Theorem 3.4.24 for two nodes. With more coprime moduli, the same argument by induction gives interpolation at any number of nodes. Moduli $(x - c)^2$ give interpolation with prescribed values and first derivatives, as in Exercise 7.1.57; higher powers $(x - c)^k$ prescribe higher derivatives, at least in characteristic 0.

Section 7.2

Solution to Exercise 7.2.11

Uses. Definition 7.2.6 (Eigenvalue, eigenvector)
Hint 1. Multiply each vector by A and compare with the vector itself.
Hint 2. $A\mathbf{v}$ must be a scalar multiple of $\mathbf{v}$. Look at a nonzero coordinate of $\mathbf{v}$ to find the only possible scalar, then check the other coordinates.
Sketch. $A(1,1,0) = 3(1,1,0)$; $A(1,-1,0) = -(1,-1,0)$; $A(0,0,1) = -(0,0,1)$; $A(1,0,0) = (1,2,0)$, not an eigenvector; $A(1,-1,1) = -(1,-1,1)$; $A(1,1,1) = (3,3,-1)$, not an eigenvector.

We compute $A\mathbf{v}$ in each case:

$$\begin{aligned} A(1,1,0) &= (3,3,0) = 3(1,1,0), & A(1,-1,0) &= (-1,1,0) = -(1,-1,0), \\ A(0,0,1) &= (0,0,-1) = -(0,0,1), & A(1,0,0) &= (1,2,0), \\ A(1,-1,1) &= (-1,1,-1) = -(1,-1,1), & A(1,1,1) &= (3,3,-1). \end{aligned}$$

So $(1,1,0)$ is an eigenvector with eigenvalue 3, and $(1,-1,0)$, $(0,0,1)$, $(1,-1,1)$ are eigenvectors with eigenvalue -1. The vector $(1,0,0)$ is not an eigenvector: a multiple of it has second coordinate 0. Neither is $(1,1,1)$: a multiple $c(1,1,1)$ has three equal coordinates.

The sum of the two eigenvectors $(1,-1,0)$ and $(0,0,1)$ for the eigenvalue -1 is again an eigenvector for -1, but the sum $(1,1,1)$ of the eigenvectors $(1,1,0)$ and $(0,0,1)$, which have different eigenvalues, is not. The definition below and Theorem 7.2.21 explain both observations.

Solution to Exercise 7.2.12

Uses. Theorem 7.2.9 (Eigenvalues and invertibility), Definition 2.5.12 (Invertible), Definition 7.2.6 (Eigenvalue, eigenvector)
Hint 1. An invertible map is injective. What does Theorem 7.2.9 say about $\lambda = 0$?
Hint 2. Apply T^{-1} to both sides of $T\mathbf{v} = \lambda\mathbf{v}$, and divide by λ.
Sketch. $T - 0\,\mathrm{id} = T$ is injective, so 0 is not an eigenvalue. If $T\mathbf{v} = \lambda\mathbf{v}$, applying T^{-1} gives $\mathbf{v} = \lambda T^{-1}\mathbf{v}$, so $T^{-1}\mathbf{v} = \lambda^{-1}\mathbf{v}$. The converse is the same argument for T^{-1}, whose inverse is T.

Since T has a left inverse, it is injective: $T\mathbf{v} = \mathbf{0}$ gives $\mathbf{v} = T^{-1}T\mathbf{v} = \mathbf{0}$. So $T - 0\,\mathrm{id} = T$ is injective, and 0 is not an eigenvalue by Theorem 7.2.9.

Let $\lambda \neq 0$ and suppose $\mathbf{v} \neq \mathbf{0}$ and $T\mathbf{v} = \lambda\mathbf{v}$. Applying the linear map T^{-1} gives $\mathbf{v} = T^{-1}(\lambda\mathbf{v}) = \lambda T^{-1}\mathbf{v}$, and multiplying by λ^{-1} gives $T^{-1}\mathbf{v} = \lambda^{-1}\mathbf{v}$. Conversely, T^{-1} is invertible with inverse T, so the same argument with the roles of T and T^{-1} exchanged shows that $T^{-1}\mathbf{v} = \lambda^{-1}\mathbf{v}$ implies $T\mathbf{v} = \left(\lambda^{-1}\right)^{-1}\mathbf{v} = \lambda\mathbf{v}$.

Solution to Exercise 7.2.25

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 1.5.5 (Independence is decided at zero)
Hint 1. Suppose $T(\sum \mathbf{v}_l) = \mu \sum \mathbf{v}_l$. Expand both sides in terms of the $\mathbf{v}_l$.
Hint 2. You get a linear relation with coefficients $\lambda_l - \mu$. What does independence say about them?
Sketch. If $T(\sum \mathbf{v}_l) = \mu \sum \mathbf{v}_l$, then $\sum(\lambda_l - \mu)\mathbf{v}_l = \mathbf{0}$, so $\lambda_l = \mu$ for all l by independence. This contradicts $\lambda_1 \neq \lambda_2$.

Let $\mathbf{w} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$, and suppose $T\mathbf{w} = \mu\mathbf{w}$ for some $\mu \in F$. Then

$$\mathbf{0} = T\mathbf{w} - \mu\mathbf{w} = \sum_{l=1}^{m} \lambda_l \mathbf{v}_l - \sum_{l=1}^{m} \mu\mathbf{v}_l = \sum_{l=1}^{m} (\lambda_l - \mu)\mathbf{v}_l.$$

By Theorem 7.2.21 the list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is independent, so $\lambda_l - \mu = 0$ for every l (Theorem 1.5.5). Then $\lambda_1 = \mu = \lambda_2$, contradicting distinctness. So $\mathbf{w}$ is not an eigenvector. (It is nonzero, again by independence, so the only way to fail is the one we ruled out.)

Solution to Exercise 7.2.29

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 2.5.12 (Invertible)
Hint 1. T is a root of $x^2 - 5x + 6$. Factor it and apply Proposition 7.2.28.
Hint 2. For (b), divide $x^2 - 5x + 6$ by $x - 4$ with remainder, and evaluate the resulting identity at T.
Sketch. (a) $p = x^2 - 5x + 6 = (x-2)(x-3)$ satisfies $p(T) = 0$. (b) $x^2 - 5x + 6 = (x-4)(x-1) + 2$, so $0 = (T - 4\,\mathrm{id})(T - \mathrm{id}) + 2\,\mathrm{id}$, and $-(1/2)(T - \mathrm{id})$ is a two-sided inverse of $T - 4\,\mathrm{id}$ (polynomials in T commute). Over a field with $2 = 0$, (b) fails.

(a) Let $p = x^2 - 5x + 6 = (x-2)(x-3)$. By hypothesis $p(T) = T^2 - 5T + 6\,\mathrm{id} = 0$, so by Proposition 7.2.28 every eigenvalue λ satisfies $(\lambda - 2)(\lambda - 3) = 0$, that is, $\lambda \in \{2, 3\}$.

(b) In $F[x]$, $(x-4)(x-1) = x^2 - 5x + 4$, so $p = (x-4)(x-1) + 2$. Evaluating at T with Proposition 7.2.27(a),

$$0 = p(T) = (T - 4\,\mathrm{id})(T - \mathrm{id}) + 2\,\mathrm{id}, \quad \text{so} \quad (T - 4\,\mathrm{id})(T - \mathrm{id}) = -2\,\mathrm{id}.$$

Since $2 \neq 0$ in F, we may let $S = -2^{-1}(T - \mathrm{id})$. Then $(T - 4\,\mathrm{id})S = \mathrm{id}$, and $S(T - 4\,\mathrm{id}) = \mathrm{id}$ as well, because polynomials in T commute (Proposition 7.2.27(b)). So $T - 4\,\mathrm{id}$ is invertible with inverse $-2^{-1}(T - \mathrm{id})$.

The hypothesis $1 + 1 \neq 0$ is needed. Over $\mathbb{F}_2$, $5 = 1$, $6 = 0$, and $4 = 0$, so the relation reads $T^2 = T$ and $T - 4\,\mathrm{id} = T$, which is not invertible when T is, say, the zero operator on $V \neq \{\mathbf{0}\}$.

Solution to Exercise 7.2.40

Uses. Proposition 7.2.37 (Chains are independent), Corollary 7.2.38 (The span of a chain), Theorem 4.1.10 (Independent lists are short), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
Hint 1. $N\mathbf{e}_3$ is the third column of N.
Hint 2. For (b), N^3 kills a basis, so $N^3 = 0$. If $A\mathbf{w} = \mu\mathbf{w}$, what is $N^3\mathbf{w}$?
Sketch. (a) $N\mathbf{e}_3 = (1,1,0)$, $N^2\mathbf{e}_3 = (1,0,0)$, $N^3\mathbf{e}_3 = \mathbf{0}$; by Proposition 7.2.37 with $k = 2$ the chain is independent, hence a basis of F^3. (b) N^3 kills the basis, so $N^3 = 0$, and an eigenvalue μ satisfies $(\mu - 1)^3 = 0$. $E_1 = \ker N = \{y + z = 0, z = 0\} = \mathrm{span}(\mathbf{e}_1)$.

(a) $N = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, so $N(x, y, z) = (y + z, z, 0)$. Hence

$$N\mathbf{e}_3 = (1,1,0), \quad N^2\mathbf{e}_3 = N(1,1,0) = (1,0,0), \quad N^3\mathbf{e}_3 = N(1,0,0) = \mathbf{0}.$$

By Proposition 7.2.37 with $\lambda = 1$ and $k = 2$, the list $\mathbf{e}_3, (1,1,0), (1,0,0)$ is independent. It has length $3 = \dim F^3$, so it is a basis (Theorem 4.1.10(b)).

(b) N^3 sends each of the three basis vectors $N^j\mathbf{e}_3$ to $N^{3+j}\mathbf{e}_3 = \mathbf{0}$, so $N^3 = 0$. That is, $p(A) = 0$ for $p = (x-1)^3$, and by Proposition 7.2.28 every eigenvalue μ of A satisfies $(\mu-1)^3 = 0$, so $\mu = 1$. And 1 is an eigenvalue, with eigenvector $N^2\mathbf{e}_3 = \mathbf{e}_1$ (Corollary 7.2.38(a)). Finally $E_1 = \ker N = \{(x,y,z) \mid y+z=0, z=0\} = \operatorname{span}(\mathbf{e}_1)$.

Solution to Exercise 7.2.41

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Proposition 1.2.35 (Zero products)
Hint 1. Write out $B(x,y) = \lambda(x,y)$ as two equations and eliminate x.
Hint 2. You should find $(\lambda^2-2)y = 0$ and $x = \lambda y$. So the eigenvalues are the square roots of 2 in the field. List the squares in each $\mathbb{F}_p$.
Sketch. $B(x,y) = (2y,x)$, so $B\mathbf{v} = \lambda\mathbf{v}$ means $2y = \lambda x$ and $x = \lambda y$, hence $(\lambda^2-2)y = 0$, and $y = 0$ forces $\mathbf{v} = \mathbf{0}$. So λ is an eigenvalue iff $\lambda^2 = 2$, with eigenvector $(\lambda, 1)$. $\mathbb{R}$: $\pm\sqrt{2}$. $\mathbb{F}_2$: $\lambda = 0$. $\mathbb{F}_3$, $\mathbb{F}_5$: none. $\mathbb{F}_7$: 3 and 4.

$B(x,y) = (2y,x)$, so $B(x,y) = \lambda(x,y)$ means $2y = \lambda x$ and $x = \lambda y$. Substituting the second into the first gives $(\lambda^2-2)y = 0$. If $y = 0$, then $x = \lambda y = 0$. So a nonzero solution exists only if $\lambda^2 = 2$, and conversely, if $\lambda^2 = 2$, then $B(\lambda,1) = (2,\lambda) = (\lambda^2,\lambda) = \lambda(\lambda,1)$. So the eigenvalues are the $\lambda \in F$ with $\lambda^2 = 2$, with eigenvectors $(\lambda,1)$.

- $\mathbb{R}$: $\lambda = \pm\sqrt{2}$, eigenvectors $(\pm\sqrt{2},1)$.
- $\mathbb{F}_2$: $2 = 0$, so $\lambda^2 = 0$ and $\lambda = 0$, with eigenvector $(0,1)$. (Here $B = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right)$.)
- $\mathbb{F}_3$: the squares are $0, 1, 4 = 1$, and 2 is not one of them. No eigenvalues.
- $\mathbb{F}_5$: the squares are $0, 1, 4, 9 = 4, 16 = 1$. No eigenvalues.
- $\mathbb{F}_7$: $3^2 = 9 = 2$ and $4^2 = 16 = 2$. Eigenvalues 3 and 4, with eigenvectors $(3,1)$ and $(4,1)$.

Solution to Exercise 7.2.42

Uses. Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 5.1.13 (The 2×2 determinant), Definition 7.2.6 (Eigenvalue, eigenvector)
Hint 1. By Theorem 7.2.9, λ is an eigenvalue iff $\det(R_\theta - \lambda I_2) = 0$. Compute this 2×2 determinant.
Hint 2. $(\cos\theta-\lambda)^2 + \sin^2\theta$ is a sum of two real squares. When is it 0?
Sketch. (a) $\det(R_\theta - \lambda I_2) = (\cos\theta-\lambda)^2 + \sin^2\theta$, which for real λ vanishes iff $\sin\theta = 0$ and $\lambda = \cos\theta$. So $\theta \in \pi\mathbb{Z}$: $R_\theta = I_2$ (eigenvalue 1, $E_1 = \mathbb{R}^2$) or $R_\theta = -I_2$ (eigenvalue -1, $E_{-1} = \mathbb{R}^2$). (b) Direct multiplication: $R_\theta(1,-i) = (\cos\theta + i\sin\theta, \sin\theta - i\cos\theta) = (\cos\theta + i\sin\theta)(1,-i)$, and similarly for $(1,i)$.

(a) By Theorem 7.2.9 and Theorem 5.1.13, $\lambda \in \mathbb{R}$ is an eigenvalue if and only if

$$\det(R_\theta - \lambda I_2) = (\cos\theta-\lambda)^2 + \sin^2\theta = 0.$$

A sum of two squares of real numbers is 0 only if both are 0, so this happens if and only if $\sin\theta = 0$ and $\lambda = \cos\theta$. Thus R_θ has a real eigenvalue exactly when θ is an integer multiple of π. If $\theta \in 2\pi\mathbb{Z}$, then $R_\theta = I_2$, the only eigenvalue is 1, and $E_1 = \mathbb{R}^2$. If $\theta \in \pi + 2\pi\mathbb{Z}$, then $R_\theta = -I_2$, the only eigenvalue is -1, and $E_{-1} = \mathbb{R}^2$. Geometrically: a rotation that is not by a multiple of π moves every line through $\mathbf{0}$.

(b) Write $c = \cos\theta$, $s = \sin\theta$. Then

$$R_\theta\begin{pmatrix}1\\-i\end{pmatrix} = \begin{pmatrix}c+is\\s-ic\end{pmatrix} = (c+is)\begin{pmatrix}1\\-i\end{pmatrix},$$

because $(c+is)(-i) = -ic - i^2 s = s - ic$. Similarly

$$R_\theta\begin{pmatrix}1\\i\end{pmatrix} = \begin{pmatrix}c-is\\s+ic\end{pmatrix} = (c-is)\begin{pmatrix}1\\i\end{pmatrix},$$

because $(c-is)i = ic + s$. Both vectors are nonzero, so they are eigenvectors with eigenvalues $c+is = e^{i\theta}$ and $c-is = e^{-i\theta}$. These are distinct unless $s = 0$, in agreement with (a).

Solution to Exercise 7.2.43

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Example 7.2.15 (Diagonal matrices)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)

Hint 1. For (a) and (b), try diagonal matrices whose eigenvectors are different standard basis vectors.
Hint 2. For (d), take S a projection and T with $TS \neq ST$, and look for $\mathbf{v} \in \ker(ST)$. For (e), what commutes with everything?
Sketch. (a), (b) false: $S = \operatorname{diag}(1,0)$, $T = \operatorname{diag}(0,1)$ have eigenvalue 1, but $S+T = I_2$ has only $1 \neq 2$, and $ST = 0$ has only $0 \neq 1$. (c) true: $(S+T)\mathbf{v} = (\lambda+\mu)\mathbf{v}$, $ST\mathbf{v} = \mu\lambda\mathbf{v}$. (d) false: $S = \operatorname{diag}(1,0)$, $T = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{v} = (1,-1)$: $ST\mathbf{v} = \mathbf{0}$ but $TS\mathbf{v} = (1,0)$. (e) false: $S = \mathrm{id}$.

(a) False. On F^2 let $S = \operatorname{diag}(1,0)$ and $T = \operatorname{diag}(0,1)$. Both have the eigenvalue 1 (with eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$). But $S+T = I_2$, whose only eigenvalue is $1 \neq 1+1$ when $1+1 \neq 1$, which holds in every field.

(b) False, with the same S and T: $ST = 0$, whose only eigenvalue is 0, while $1 \cdot 1 = 1$.

(c) True. If $S\mathbf{v} = \lambda\mathbf{v}$ and $T\mathbf{v} = \mu\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $(S+T)\mathbf{v} = (\lambda+\mu)\mathbf{v}$ and $ST\mathbf{v} = S(\mu\mathbf{v}) = \mu\lambda\mathbf{v}$.

(d) False. Let $S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, so $ST = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $TS = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. For $\mathbf{v} = (1,-1)$, $ST\mathbf{v} = \mathbf{0} = 0\mathbf{v}$, so $\mathbf{v}$ is an eigenvector of ST. But $TS\mathbf{v} = (1,0)$, which is not a multiple of $(1,-1)$. (What is true: if $ST\mathbf{v} = \lambda\mathbf{v}$ and $T\mathbf{v} \neq \mathbf{0}$, then $T\mathbf{v}$ is an eigenvector of TS for λ.)

(e) False. $S = \mathrm{id}_V$ commutes with every T, and every nonzero vector is an eigenvector of S, but the rotation $T = R_{\pi/2}$ of $\mathbb{R}^2$ has no eigenvectors at all (Example 7.2.18). By Proposition 7.2.30, what is true is that T maps each eigenspace of S into itself.

Solution to Exercise 7.2.44

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 4.3.8 (Row rank equals column rank)
Hint 1. For (a), compute $A(1,1,\ldots,1)$.
Hint 2. For (b), use the rank criterion in Theorem 7.2.9, and $(A-\lambda I_n)^{\mathsf{T}} = A^{\mathsf{T}} - \lambda I_n$.
Sketch. (a) $A(1,\ldots,1)$ is the vector of row sums, $s(1,\ldots,1)$. (b) $\operatorname{rank}\left(A^{\mathsf{T}} - \lambda I_n\right) = \operatorname{rank}\left((A-\lambda I_n)^{\mathsf{T}}\right) = \operatorname{rank}(A-\lambda I_n)$, so one is less than n iff the other is. (c) Apply (a) to A^{T} and use (b). $(1,\ldots,1)$ need not be an eigenvector: $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(a) The i-th coordinate of $A(1,\ldots,1)$ is $\sum_j a_{ij}$, the i-th row sum, which is s. So $A(1,\ldots,1) = s(1,\ldots,1)$, and $(1,\ldots,1) \neq \mathbf{0}$ is an eigenvector with eigenvalue s.

(b) For $\lambda \in F$, $(A-\lambda I_n)^{\mathsf{T}} = A^{\mathsf{T}} - \lambda I_n$, so by Theorem 4.3.8, $\operatorname{rank}\left(A^{\mathsf{T}} - \lambda I_n\right) = \operatorname{rank}(A - \lambda I_n)$. By Theorem 7.2.9, λ is an eigenvalue of A if and only if this rank is less than n, if and only if λ is an eigenvalue of A^{T}.

(c) The rows of A^{T} are the columns of A, so every row of A^{T} sums to s, and s is an eigenvalue of A^{T} by (a), hence of A by (b). The vector $(1,\ldots,1)$ need not be an eigenvector of A: for $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, every column sums to 1, but $A(1,1,1) = (2,1,0)$, which is not a multiple of $(1,1,1)$ because its last two coordinates differ. An eigenvector of A for the eigenvalue 1 is $\mathbf{e}_1$. (A 2×2 example such as $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ works only when $1+1 \neq 0$: over $\mathbb{F}_2$ it sends $(1,1)$ to $\mathbf{0}$.) (Matrices with nonnegative entries and column sums 1 are the transition matrices of Markov chains, and this exercise shows they always have the eigenvalue 1; see Section 11.1.)

Solution to Exercise 7.2.45

Uses. Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Corollary 7.2.22 (At most $\dim V$ eigenvalues), Proposition 4.1.21 (Dimensions of subspaces), Example 7.2.18 (Rotations), Theorem 2.2.1 (Linear maps are determined by a basis)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. By Proposition 7.2.7, an invariant line is a line contained in some eigenspace. How many lines can an eigenspace contain?
Hint 2. If some E_λ is all of V, then $T = \lambda\,\mathrm{id}$. Otherwise each E_λ is $\{\mathbf{0}\}$ or a line, and there are at most 2 eigenvalues.
Sketch. An invariant line is $\operatorname{span}(\mathbf{v})$ for an eigenvector $\mathbf{v}$, so it lies in E_λ for its eigenvalue. If $T = c\,\mathrm{id}$, every line is invariant: infinitely many over $\mathbb{R}$, all 3 lines of $\mathbb{F}_2^2$. Otherwise every eigenspace is a proper subspace of a 2-dimensional

space, hence $\{\mathbf{0}\}$ or a line, and invariant lines correspond to eigenvalues, of which there are at most 2. Examples: rotation, shear, $\operatorname{diag}(1,2)$, id over $\mathbb{R}$; over $\mathbb{F}_2$: $\left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, $\operatorname{diag}(0,1)$, id.

Let $V = F^2$ with $F = \mathbb{R}$ or $\mathbb{F}_2$, and $T \in \mathrm{L}(V)$. By Proposition 7.2.7, a line ℓ is invariant if and only if it is spanned by an eigenvector, and then $\ell \subseteq E_\lambda$ for the eigenvalue λ of that eigenvector. Conversely every line inside an eigenspace is invariant.

If $E_\lambda = V$ for some λ, then $T = \lambda\,\mathrm{id}$, and every line is invariant. Over $\mathbb{R}$ there are infinitely many lines. Over $\mathbb{F}_2$ there are exactly 3: each line $\{\mathbf{0}, \mathbf{v}\}$ contains one nonzero vector, and $\mathbb{F}_2^2$ has 3 nonzero vectors.

Otherwise each E_λ is a proper subspace of the 2-dimensional V, so it is $\{\mathbf{0}\}$ or a line (Proposition 4.1.21), and a line contains exactly one line. So the invariant lines are exactly the eigenspaces E_λ of the eigenvalues λ, and their number is the number of eigenvalues, which is at most 2 by Corollary 7.2.22. This proves the restrictions in (a) and (b). Examples:

(a) $R_{\pi/2}$ has no eigenvalues (Example 7.2.18): 0 lines. The shear $A = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ satisfies $(A - I_2)^2 = 0$, so 1 is its only possible eigenvalue (Proposition 7.2.28), and $E_1 = \ker(A - I_2) = \operatorname{span}(\mathbf{e}_1)$: 1 line. $\operatorname{diag}(1,2)$ has $E_1 = \operatorname{span}(\mathbf{e}_1)$ and $E_2 = \operatorname{span}(\mathbf{e}_2)$ (Example 7.2.15): 2 lines. The identity: infinitely many.

(b) Over $\mathbb{F}_2$ the only possible eigenvalues are 0 and 1. For $A = \left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right)$, $A(x,y) = (y, x+y)$; $A\mathbf{v} = \mathbf{0}$ forces $y = 0$ and then $x = 0$, and $A\mathbf{v} = \mathbf{v}$ means $y = x$ and $x + y = y$, so $x = 0$ and $y = 0$. So A has no eigenvalues: 0 lines. The shear $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ is invertible, so 0 is not an eigenvalue, and $E_1 = \ker\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right) = \operatorname{span}(\mathbf{e}_1)$: 1 line. $\operatorname{diag}(0,1)$: 2 lines. The identity: 3 lines.

Solution to Exercise 7.2.46

Uses. Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 1.5.24 (Distinct degrees), Theorem 4.1.10 (Independent lists are short), Example 4.1.7 (Dimensions of the standard spaces), Example 2.1.9 (Maps of polynomials)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Let $U \neq \{\mathbf{0}\}$ be invariant and $p \in U$ of degree n. What are the degrees of $p, Dp, D^2p, \ldots, D^np$ over $\mathbb{R}$?
Hint 2. These $n+1$ polynomials have distinct degrees $n, n-1, \ldots, 0$, so they are independent, and they lie in $U \cap \mathbb{R}[x]_{\leq n}$. Count dimensions. Then split into the cases where the degrees of elements of U are bounded or not.
Sketch. If $p \in U$ has degree n, then D^jp has degree $n-j$ for $0 \leq j \leq n$ (leading coefficient $n(n-1)\cdots(n-j+1)a_n \neq 0$ in $\mathbb{R}$). These $n+1$ polynomials are independent and lie in $U \cap \mathbb{R}[x]_{\leq n}$, so $\mathbb{R}[x]_{\leq n} \subseteq U$. If degrees in U are bounded with maximum n, then $U = \mathbb{R}[x]_{\leq n}$; if unbounded, U contains every $\mathbb{R}[x]_{\leq n}$, so $U = \mathbb{R}[x]$. Over $\mathbb{F}_p$, $\operatorname{span}(x^p)$ is invariant since $D(x^p) = \mathbf{0}$.

(a) Each listed subspace is invariant: D lowers degrees (Example 7.2.2). Conversely, let U be D-invariant with $U \neq \{\mathbf{0}\}$.

Claim. If $p \in U$ has degree n, then $\mathbb{R}[x]_{\leq n} \subseteq U$. Write $p = a_nx^n + (\text{lower terms})$ with $a_n \neq 0$. By the formula for D (Example 2.1.9), $D^jp = n(n-1)\cdots(n-j+1)a_nx^{n-j} + (\text{lower terms})$ for $0 \leq j \leq n$, and the leading coefficient is nonzero in $\mathbb{R}$. So $p, Dp, \ldots, D^np$ have the distinct degrees $n, n-1, \ldots, 0$, hence are independent (Proposition 1.5.24). They lie in U by Proposition 7.2.27(d), and in $\mathbb{R}[x]_{\leq n}$, which has dimension $n+1$ (Example 4.1.7). So they form a basis of $\mathbb{R}[x]_{\leq n}$ (Theorem 4.1.10(b)), and $\mathbb{R}[x]_{\leq n} = \operatorname{span}(p, \ldots, D^np) \subseteq U$.

If the degrees of the nonzero elements of U are bounded, let n be the largest. Then $U \subseteq \mathbb{R}[x]_{\leq n}$, and the claim gives equality. If they are unbounded, then for every m there is $p \in U$ with $\deg p \geq m$, and the claim gives $\mathbb{R}[x]_{\leq m} \subseteq \mathbb{R}[x]_{\leq \deg p} \subseteq U$. Every polynomial lies in some $\mathbb{R}[x]_{\leq m}$, so $U = \mathbb{R}[x]$.

(b) Over $\mathbb{F}_p$, $D(x^p) = px^{p-1} = \mathbf{0}$, so the line $\operatorname{span}(x^p)$ is D-invariant. It is not $\{\mathbf{0}\}$, it is not $\mathbb{F}_p[x]$, and it is not $\mathbb{F}_p[x]_{\leq n}$ for any n, since it does not contain the constant 1. The proof of (a) breaks down because the leading coefficient p of $D(x^p)$ vanishes in $\mathbb{F}_p$.

Solution to Exercise 7.2.47

Uses. Corollary 7.2.23 (Eigenspaces are independent), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Theorem 7.2.21 (Eigenvectors for distinct

eigenvalues are independent), Theorem 4.1.10 (Independent lists are short), Theorem 3.4.24 (Polynomial interpolation), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Corollary 2.2.6 (Agreeing on a spanning list)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Each eigenspace $E_{\lambda_i}(T)$ is S-invariant. What is its dimension?
Hint 2. The dimensions of the n eigenspaces are each at least 1 and add up to at most n.
Hint 3. For (b), if $S\mathbf{v}_i = \mu_i\mathbf{v}_i$, choose p with $p(\lambda_i) = \mu_i$ for all i, and compare S and $p(T)$ on the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
Sketch. (a) By Corollary 7.2.23, $\sum \dim E_{\lambda_i} \leq n$ with each term ≥ 1, so each $E_{\lambda_i} = \operatorname{span}(\mathbf{v}_i)$ is a line. It is S-invariant by Proposition 7.2.30, so $\mathbf{v}_i$ is an eigenvector of S, say $S\mathbf{v}_i = \mu_i\mathbf{v}_i$. (b) Interpolate: $p(\lambda_i) = \mu_i$ with $\deg p < n$. Then $p(T)\mathbf{v}_i = \mu_i\mathbf{v}_i = S\mathbf{v}_i$, and the $\mathbf{v}_i$ form a basis.

(a) Each $E_{\lambda_i}(T)$ contains $\mathbf{v}_i \neq \mathbf{0}$, so $\dim E_{\lambda_i} \geq 1$, and by Corollary 7.2.23 the n dimensions add up to at most n. So each $\dim E_{\lambda_i} = 1$, and $E_{\lambda_i} = \operatorname{span}(\mathbf{v}_i)$. By Proposition 7.2.30(b) (with the roles of S and T exchanged), $E_{\lambda_i}(T)$ is S-invariant, so $\mathbf{v}_i$ is an eigenvector of S by Proposition 7.2.7. Write $S\mathbf{v}_i = \mu_i\mathbf{v}_i$.

(b) By Theorem 3.4.24 there is $p \in F[x]_{\leq n-1}$ with $p(\lambda_i) = \mu_i$ for $i = 1, \ldots, n$, since the λ_i are distinct. By Proposition 7.2.28, $p(T)\mathbf{v}_i = p(\lambda_i)\mathbf{v}_i = \mu_i\mathbf{v}_i = S\mathbf{v}_i$. The list $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is independent (Theorem 7.2.21) of length n, so it is a basis (Theorem 4.1.10(b)), and $S = p(T)$ by Corollary 2.2.6. Conversely every $p(T)$ commutes with T, so the operators commuting with T are exactly the polynomials in T.

Solution to Exercise 7.2.48

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 4.1.10 (Independent lists are short), Definition 3.4.26 (Lagrange polynomials), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 7.2.4 (Invariance is checked on a spanning list)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors for the distinct eigenvalues $\lambda_1, \ldots, \lambda_n$; they form a basis. Guess the invariant subspaces.
Hint 2. For an invariant U and $\mathbf{u} = \sum c_i\mathbf{v}_i \in U$, find a polynomial q with $q(T)\mathbf{u} = c_i\mathbf{v}_i$. Lagrange polynomials do this.
Hint 3. Conclude that $U = \operatorname{span}(\mathbf{v}_i \mid \mathbf{v}_i \in U)$, so invariant subspaces correspond to subsets of $\{1, \ldots, n\}$.
Sketch. With an eigenbasis $\mathbf{v}_i$, each $U_I = \operatorname{span}(\mathbf{v}_i \mid i \in I)$ is invariant, and distinct I give distinct U_I. Conversely, if U is invariant and $\mathbf{u} = \sum c_i\mathbf{v}_i \in U$, then $L_i(T)\mathbf{u} = c_i\mathbf{v}_i \in U$, where L_i is the Lagrange polynomial with $L_i(\lambda_j) = [i = j]$. So $U = U_I$ for $I = \{i \mid \mathbf{v}_i \in U\}$. Hence 2^n.

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues and $\mathbf{v}_i$ an eigenvector for λ_i. By Theorem 7.2.21 and Theorem 4.1.10(b), $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of V. For $I \subseteq \{1, \ldots, n\}$ let $U_I = \operatorname{span}(\mathbf{v}_i \mid i \in I)$ (with $U_\emptyset = \{\mathbf{0}\}$). Each U_I is invariant by Proposition 7.2.4, since $T\mathbf{v}_i = \lambda_i\mathbf{v}_i \in U_I$ for $i \in I$. Different subsets give different subspaces: if $j \in I$ and $j \notin I'$, then $\mathbf{v}_j \in U_I$ but $\mathbf{v}_j \notin U_{I'}$, because $\mathbf{v}_j$ is not a combination of the other basis vectors.

Conversely, let U be T-invariant and $I = \{i \mid \mathbf{v}_i \in U\}$. Clearly $U_I \subseteq U$. Let $\mathbf{u} \in U$ and write $\mathbf{u} = \sum_i c_i\mathbf{v}_i$. For each i, let L_i be the Lagrange polynomial for the nodes $\lambda_1, \ldots, \lambda_n$ with $L_i(\lambda_i) = 1$ and $L_i(\lambda_j) = 0$ for $j \neq i$ (Definition 3.4.26). By Proposition 7.2.28,

$$L_i(T)\mathbf{u} = \sum_j c_j L_i(\lambda_j)\mathbf{v}_j = c_i\mathbf{v}_i,$$

and $L_i(T)\mathbf{u} \in U$ by Proposition 7.2.27(d). So $c_i\mathbf{v}_i \in U$ for every i. If $c_i \neq 0$, then $\mathbf{v}_i = c_i^{-1}(c_i\mathbf{v}_i) \in U$, so $i \in I$. Hence $\mathbf{u} = \sum_{i \in I} c_i\mathbf{v}_i \in U_I$, and $U = U_I$.

So the invariant subspaces are exactly the U_I, one for each of the 2^n subsets I. Without distinct eigenvalues the count can be infinite: every line in $\mathbb{R}^2$ is invariant under id.

Solution to Exercise 7.2.49

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Proposition 7.1.22 (Counting roots), Proposition 7.2.34 (The smallest invariant subspace containing a vector), Theorem 4.1.10 (Independent lists are short)
Hint 1. A relation $\sum_{j<n} c_j T^j\mathbf{w} = \mathbf{0}$ says $p(T)\mathbf{w} = \mathbf{0}$ for $p = \sum c_j x^j$. Compute $p(T)\mathbf{w}$ with Proposition 7.2.28.
Hint 2. $p(T)\mathbf{w} = \sum_i p(\lambda_i)\mathbf{v}_i$. Independence of the $\mathbf{v}_i$ gives n roots of p, which has degree less than n.

Sketch. If $p(T)\mathbf{w} = \mathbf{0}$ with $\deg p < n$, then $\sum_i p(\lambda_i)\mathbf{v}_i = \mathbf{0}$, so $p(\lambda_i) = 0$ for all n distinct λ_i, and $p = \mathbf{0}$ (Proposition 7.1.22). So the n vectors are independent. They lie in the invariant subspace $W = \operatorname{span}(\mathbf{v}_i)$, which has dimension n, so they span it; and $W \subseteq K_T(\mathbf{w}) \subseteq W$.

Suppose $\sum_{j=0}^{n-1} c_j T^j \mathbf{w} = \mathbf{0}$, and let $p = \sum_{j=0}^{n-1} c_j x^j$. By Proposition 7.2.28,

$$\mathbf{0} = p(T)\mathbf{w} = \sum_{i=1}^{n} p(T)\mathbf{v}_i = \sum_{i=1}^{n} p(\lambda_i)\mathbf{v}_i.$$

The $\mathbf{v}_i$ are independent by Theorem 7.2.21, so $p(\lambda_i) = 0$ for every i. Thus p has at least n distinct roots but degree less than n (if nonzero), so $p = \mathbf{0}$ by Proposition 7.1.22, and all $c_j = 0$.

Let $W = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. It is T-invariant (Proposition 7.2.4) and contains $\mathbf{w}$, so $K_T(\mathbf{w}) \subseteq W$ (Proposition 7.2.34). On the other hand, $\mathbf{w}, \ldots, T^{n-1}\mathbf{w}$ are n independent vectors of $K_T(\mathbf{w}) \subseteq W$, and $\dim W = n$, so they span W (Theorem 4.1.10(b)) and $W \subseteq K_T(\mathbf{w})$. Hence $K_T(\mathbf{w}) = W$.

The independence can also be read off a determinant: in the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of W, the coordinates of $T^j\mathbf{w}$ are $\left(\lambda_1^j, \ldots, \lambda_n^j\right)$, so the matrix with these columns is a Vandermonde matrix, which is invertible for distinct λ_i (Theorem 3.4.24).

Solution to Exercise 7.2.50

Uses. Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.1.10 (Independent lists are short), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 4.1.21 (Dimensions of subspaces), Theorem 1.5.5 (Independence is decided at zero)
Techniques. Dimension counting (Technique 4.2.24)
Hint 1. Choose $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$. The chain $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is a basis. Compute N^j on a general combination of it.
Hint 2. For (b), let U be invariant and j the least integer with $N^j(U) = \{\mathbf{0}\}$. Find a chain of length j inside U.
Hint 3. Compare $\dim U$ with $\dim \ker N^j = j$.
Sketch. (a) With $\mathbf{v}$ as in the hint, the chain is independent (Proposition 4.1.26) of length n, hence a basis, and $N^j\left(\sum_i c_i N^i \mathbf{v}\right) = \sum_{i \le n-1-j} c_i N^{i+j}\mathbf{v}$ vanishes iff $c_0 = \cdots = c_{n-1-j} = 0$; so $\ker N^j = \operatorname{span}\left(N^{n-j}\mathbf{v}, \ldots, N^{n-1}\mathbf{v}\right)$. (b) Kernels of powers are invariant. Conversely, if j is least with $U \subseteq \ker N^j$ and $j \ge 1$, some $\mathbf{u} \in U$ has $N^{j-1}\mathbf{u} \neq \mathbf{0} = N^j\mathbf{u}$, giving j independent vectors in U; so $j \le \dim U \le \dim \ker N^j = j$, and $U = \ker N^j$.

(a) Choose $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$; then $N^n\mathbf{v} = \mathbf{0}$. By Proposition 4.1.26, $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is independent, and it has length $n = \dim V$, so it is a basis (Theorem 4.1.10(b)). Let $0 \le j \le n$. For $\mathbf{x} = \sum_{i=0}^{n-1} c_i N^i \mathbf{v}$,

$$N^j\mathbf{x} = \sum_{i=0}^{n-1} c_i N^{i+j}\mathbf{v} = \sum_{i=0}^{n-1-j} c_i N^{i+j}\mathbf{v},$$

since $N^{i+j}\mathbf{v} = \mathbf{0}$ for $i + j \ge n$. The vectors $N^{i+j}\mathbf{v}$, $0 \le i \le n-1-j$, are distinct members of the basis, so $N^j\mathbf{x} = \mathbf{0}$ if and only if $c_0 = \cdots = c_{n-1-j} = 0$. Hence $\ker N^j = \operatorname{span}\left(N^{n-j}\mathbf{v}, \ldots, N^{n-1}\mathbf{v}\right)$, which has dimension j.

(b) Each $\ker N^j$ is N-invariant by Proposition 7.2.30(a). Conversely, let U be N-invariant. Since $N^n = 0$, there is a least $j \in \{0, \ldots, n\}$ with $N^j\mathbf{u} = \mathbf{0}$ for all $\mathbf{u} \in U$, that is, $U \subseteq \ker N^j$. If $j = 0$, then $U = \{\mathbf{0}\} = \ker N^0$. If $j \ge 1$, minimality gives some $\mathbf{u} \in U$ with $N^{j-1}\mathbf{u} \neq \mathbf{0}$, while $N^j\mathbf{u} = \mathbf{0}$. By Proposition 4.1.26, $\mathbf{u}, N\mathbf{u}, \ldots, N^{j-1}\mathbf{u}$ is independent, and it lies in U because U is invariant. So

$$j \le \dim U \le \dim \ker N^j = j,$$

and $U = \ker N^j$ by Proposition 4.1.21. By (a) the $n+1$ subspaces $\ker N^0 \subsetneq \ker N^1 \subsetneq \cdots \subsetneq \ker N^n = V$ have different dimensions, so they are distinct.

Solution to Exercise 7.2.51

Uses. Theorem 1.5.5 (Independence is decided at zero), Notation 2.1.18, Proposition 7.2.37 (Chains are independent)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. For one chain this is Proposition 7.2.37. Imitate its proof by the lowest nonzero coefficient. A term $a_{ij}N^j\mathbf{u}_i$ sits $k_i - j$ steps below the top of its chain.

Hint 2. Given a nontrivial relation, let h be the largest value of $k_i - j$ over the nonzero coefficients a_{ij}, and apply N^h.
Hint 3. Terms with $k_i - j < h$ are pushed past the top and die; terms with $k_i - j = h$ land on the tops. Use independence of the tops.
Sketch. Suppose $\sum_{i,j} a_{ij} N^j \mathbf{u}_i = \mathbf{0}$ is nontrivial, and let $h = \max\{k_i - j \mid a_{ij} \neq 0\}$. Applying N^h kills every term with $k_i - j < h$ (as $j + h \geq k_i + 1$) and sends each term with $k_i - j = h$ to $a_{i,k_i-h} N^{k_i} \mathbf{u}_i$; for each i there is at most one such term. The result is a relation among the tops with at least one nonzero coefficient, a contradiction.

The hypothesis implies that each top $N^{k_i}\mathbf{u}_i$ is nonzero. Suppose, for contradiction, that

$$\sum_{i=1}^{r} \sum_{j=0}^{k_i} a_{ij} N^j \mathbf{u}_i = \mathbf{0}$$

with not all a_{ij} zero. Call $k_i - j$ the depth of the term (i, j), and let h be the largest depth among the terms with $a_{ij} \neq 0$. Then $0 \leq h$, and apply N^h to the relation. For a term with $a_{ij} \neq 0$:

- if its depth is less than h, then $j + h \geq k_i + 1$, so $N^h N^j \mathbf{u}_i = N^{j+h-k_i-1} N^{k_i+1} \mathbf{u}_i = \mathbf{0}$;
- if its depth is h, then $N^h N^j \mathbf{u}_i = N^{k_i} \mathbf{u}_i$.

Terms with $a_{ij} = 0$ contribute nothing. For each i at most one j has depth h, namely $j = k_i - h$. So, with $I = \{i \mid k_i \geq h \text{ and } a_{i,k_i-h} \neq 0\}$,

$$\sum_{i \in I} a_{i,k_i-h} N^{k_i} \mathbf{u}_i = \mathbf{0}.$$

By the choice of h, I is nonempty, so this is a nontrivial relation among the tops, contradicting their independence. Hence all $a_{ij} = 0$, and the list is independent by Theorem 1.5.5.

For $r = 1$ the depth of the deepest nonzero term is $k - m$, where m is the lowest index with $a_m \neq 0$, and the argument is the second proof of Proposition 7.2.37. The statement is the independence half of the Jordan form (Section 10.3): to build a basis out of chains, it suffices to choose their tops independent.

Solution to Exercise 7.2.52

Uses. Example 2.1.9 (Maps of polynomials), Definition 7.2.13 (Eigenspace), Definition 1.2.13 (Polynomials)
Hint 1. Compute $T(x^k)$.
Hint 2. For $p = \sum a_k x^k$, the equation $T(p) = \lambda p$ says $(k - \lambda)a_k = 0$ for every k. In $\mathbb{F}_3$, which $k \in \{0, \ldots, 5\}$ are equal to a given λ?
Sketch. $T(x^k) = kx^k$, so $T(\sum a_k x^k) = \lambda \sum a_k x^k$ iff $(k - \lambda)a_k = 0$ for all k. Over $\mathbb{R}$: eigenvalues $0, 1, \ldots, n$ with $E_k = \operatorname{span}(x^k)$. Over $\mathbb{F}_3$, $n = 5$: eigenvalues $0, 1, 2$ with $E_0 = \operatorname{span}(1, x^3)$, $E_1 = \operatorname{span}(x, x^4)$, $E_2 = \operatorname{span}(x^2, x^5)$.

For $k \geq 1$, $T(x^k) = x \cdot kx^{k-1} = kx^k$, and $T(1) = \mathbf{0} = 0 \cdot 1$. By linearity, for $p = \sum_{k=0}^{n} a_k x^k$,

$$T(p) - \lambda p = \sum_{k=0}^{n} (k - \lambda) a_k x^k,$$

which is $\mathbf{0}$ if and only if $(k - \lambda)a_k = 0$ for each k, where k means $1 + \cdots + 1$ in F. So $E_\lambda = \operatorname{span}(x^k \mid 0 \leq k \leq n, k = \lambda \text{ in } F)$, and λ is an eigenvalue exactly when λ equals some $k \in \{0, \ldots, n\}$ in F.

(a) Over $\mathbb{R}$ the integers $0, 1, \ldots, n$ are distinct, so the eigenvalues are $0, 1, \ldots, n$, with $E_k = \operatorname{span}(x^k)$. These are $n + 1 = \dim \mathbb{R}[x]_{\leq n}$ distinct eigenvalues.

(b) Over $\mathbb{F}_3$, $k = \lambda$ in $\mathbb{F}_3$ when $k \equiv \lambda \pmod 3$. For $k \in \{0, \ldots, 5\}$ the residues are $0, 1, 2, 0, 1, 2$. So the eigenvalues are $0, 1, 2$, with

$$E_0 = \operatorname{span}(1, x^3), \quad E_1 = \operatorname{span}(x, x^4), \quad E_2 = \operatorname{span}(x^2, x^5).$$

The dimensions add up to $6 = \dim \mathbb{F}_3[x]_{\leq 5}$, as Corollary 7.2.23 allows, and the monomials still form a basis of eigenvectors.

Solution to Exercise 7.2.53

Uses. Proposition 7.2.27 (Evaluating polynomials at an operator), Definition 7.2.6 (Eigenvalue, eigenvector)

Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Let $\mathbf{w} = (T + \lambda\,\mathrm{id})\mathbf{v}$. The display says $(T - \lambda\,\mathrm{id})\mathbf{w} = \mathbf{0}$. Does that say anything about $\mathbf{v}$ when $\mathbf{w} \neq \mathbf{0}$?
Hint 2. Test the claim on $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, for which $T^2 = I_2$.
Sketch. The display shows $(T - \lambda\,\mathrm{id})\mathbf{w} = \mathbf{0}$ for $\mathbf{w} = (T + \lambda\,\mathrm{id})\mathbf{v}$. If $\mathbf{w} \neq \mathbf{0}$, this makes $\mathbf{w}$, not $\mathbf{v}$, an eigenvector for λ; the "either/or" step confuses $\mathbf{w}$ with $\mathbf{v}$. The claim is false: $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has $T^2 = I_2$, so $\mathbf{e}_1$ is an eigenvector of T^2, but $T\mathbf{e}_1 = \mathbf{e}_2$.

The first display is correct, by Proposition 7.2.27(a). Write $\mathbf{w} = (T + \lambda\,\mathrm{id})\mathbf{v}$; the display says $(T - \lambda\,\mathrm{id})\mathbf{w} = \mathbf{0}$. The next sentence treats this as if it were a statement about $\mathbf{v}$: from "$AB\mathbf{v} = \mathbf{0}$" it concludes "$B\mathbf{v} = \mathbf{0}$ or $A\mathbf{v} = \mathbf{0}$". What actually follows is: either $\mathbf{w} = \mathbf{0}$, in which case $\mathbf{v}$ is an eigenvector for $-\lambda$, or $\mathbf{w} \neq \mathbf{0}$ and $\mathbf{w}$ is an eigenvector of T for λ. In the second case nothing has been shown about $\mathbf{v}$.

The claim is false. Let $T = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathbb{C}^{2\times 2}$, so $T^2 = I_2$ and every nonzero vector is an eigenvector of T^2. But $T\mathbf{e}_1 = \mathbf{e}_2$ is not a multiple of $\mathbf{e}_1$. Following the corrected argument with $\mu = 1$, $\lambda = 1$, $\mathbf{v} = \mathbf{e}_1$: $\mathbf{w} = (T + \mathrm{id})\mathbf{e}_1 = (1, 1)$, which is indeed an eigenvector of T for 1. Over $\mathbb{R}$ the argument fails even earlier, since μ may have no real square root: $R_{\pi/2}^2 = -\mathrm{id}$ (Example 7.2.18).

Solution to Exercise 7.2.54

Uses. Definition 7.2.1 (Invariant subspace), Corollary 4.2.18 (Injective if and only if surjective), Example 7.2.18 (Rotations)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (a) is a generic-element check. For (b), find a T with $T^2 = -\mathrm{id}$.
Hint 2. For (c), $T|_U$ is an injective operator on the finite-dimensional U. For (d), look for an invertible shift on sequences indexed by all of $\mathbb{Z}$.
Sketch. (a) True: $T(\mathbf{u} + \mathbf{w}) = T\mathbf{u} + T\mathbf{w} \in U + W$, and T maps $U \cap W$ into both. (b) False: $R_{\pi/2}^2 = -\mathrm{id}$ leaves every line invariant, but $R_{\pi/2}$ leaves none. (c) True: $T|_U \in \mathrm{L}(U)$ is injective, hence surjective, so each $\mathbf{u} \in U$ is $T\mathbf{u}'$ with $\mathbf{u}' \in U$, and $T^{-1}\mathbf{u} = \mathbf{u}' \in U$. (d) False: on $F^{\mathbb{Z}}$ the right shift is invertible, and the sequences vanishing at negative indices form a subspace invariant under it but not under its inverse.

(a) True. If $\mathbf{u} \in U$ and $\mathbf{w} \in W$, then $T(\mathbf{u} + \mathbf{w}) = T\mathbf{u} + T\mathbf{w} \in U + W$. If $\mathbf{x} \in U \cap W$, then $T\mathbf{x} \in U$ and $T\mathbf{x} \in W$.

(b) False. Let $T = R_{\pi/2}$ on $\mathbb{R}^2$. Then $T^2 = -\mathrm{id}$, under which every subspace is invariant, but no line is T-invariant (Example 7.2.18). Take $U = \mathrm{span}(\mathbf{e}_1)$.

(c) True. Since T is injective, so is $T|_U \in \mathrm{L}(U)$, and U is finite-dimensional, so $T|_U$ is surjective onto U (Corollary 4.2.18). Given $\mathbf{u} \in U$, choose $\mathbf{u}' \in U$ with $T\mathbf{u}' = \mathbf{u}$. Then $T^{-1}\mathbf{u} = \mathbf{u}' \in U$.

(d) False. Let $V = F^{\mathbb{Z}}$ be the space of two-sided sequences $(x_n)_{n\in\mathbb{Z}}$, and let T be the right shift, $(T\mathbf{x})_n = x_{n-1}$. It is invertible, with inverse the left shift $\left(T^{-1}\mathbf{x}\right)_n = x_{n+1}$. The subspace $U = \{\mathbf{x} \mid x_n = 0 \text{ for all } n < 0\}$ is T-invariant: if $x_n = 0$ for $n < 0$, then $(T\mathbf{x})_n = x_{n-1} = 0$ for $n < 0$. But the sequence $\mathbf{x}$ with $x_0 = 1$ and all other terms 0 lies in U, while $\left(T^{-1}\mathbf{x}\right)_{-1} = x_0 = 1$, so $T^{-1}\mathbf{x} \notin U$. The proof of (c) fails because $T|_U$ is injective but not surjective onto U.

Solution to Exercise 7.2.55

Uses. Corollary 7.1.48 (Coprime polynomials), Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Theorem 4.4.7 (Criteria for direct sums), Definition 4.4.4 (Direct sum)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. By Bézout (Corollary 7.1.48), $up + vq = 1$ for some $u, v \in F[x]$. Evaluate at T: $\mathrm{id} = u(T)p(T) + v(T)q(T)$.
Hint 2. Apply this to $\mathbf{v}$. Which kernel does $u(T)p(T)\mathbf{v}$ lie in, and which does $v(T)q(T)\mathbf{v}$ lie in?
Hint 3. For the intersection, apply the same identity to a vector killed by both $p(T)$ and $q(T)$.
Sketch. From $up + vq = 1$, $\mathrm{id} = u(T)p(T) + v(T)q(T)$. For $\mathbf{x} \in V$, $q(T)u(T)p(T)\mathbf{x} = u(T)(pq)(T)\mathbf{x} = \mathbf{0}$, so the first term lies in $\ker q(T)$, and likewise the second in $\ker p(T)$: $V = \ker p(T) + \ker q(T)$. If $\mathbf{x}$ is in both kernels, the identity gives $\mathbf{x} = \mathbf{0}$. Invariance is Proposition 7.2.30. For $T^2 = \mathrm{id}$, take $p = x - 1$, $q = x + 1$, coprime since $1 = ((x + 1) - (x - 1))/2$.

By Corollary 7.1.48 there are $u, v \in F[x]$ with $up + vq = 1$. By Proposition 7.2.27(a),

$$\mathrm{id} = u(T)p(T) + v(T)q(T). \tag{C.1.9}$$

Let $\mathbf{x} \in V$, and set $\mathbf{y} = v(T)q(T)\mathbf{x}$ and $\mathbf{z} = u(T)p(T)\mathbf{x}$, so that $\mathbf{x} = \mathbf{y} + \mathbf{z}$. Since polynomials in T commute (Proposition 7.2.27(b)),

$$p(T)\mathbf{y} = v(T)p(T)q(T)\mathbf{x} = \mathbf{0}, \quad q(T)\mathbf{z} = u(T)p(T)q(T)\mathbf{x} = \mathbf{0}.$$

So $\mathbf{y} \in \ker p(T)$ and $\mathbf{z} \in \ker q(T)$, and $V = \ker p(T) + \ker q(T)$. If $\mathbf{x} \in \ker p(T) \cap \ker q(T)$, then (C.1.9) gives $\mathbf{x} = u(T)p(T)\mathbf{x} + v(T)q(T)\mathbf{x} = \mathbf{0}$. So the sum is direct (Theorem 4.4.7). Both kernels are T-invariant by Proposition 7.2.30(a).

If $T^2 = \mathrm{id}$ and $2 \neq 0$, take $p = x - 1$ and $q = x + 1$. Then $pq = x^2 - 1$, so $p(T)q(T) = T^2 - \mathrm{id} = 0$, and $2^{-1}q - 2^{-1}p = 1$, so p and q are coprime (Corollary 7.1.48). Hence $V = \ker(T - \mathrm{id}) \oplus \ker(T + \mathrm{id}) = E_1 \oplus E_{-1}$. Over $\mathbb{F}_2$ this fails: the swap of Example 7.2.17 satisfies $\sigma^2 = \mathrm{id}$, but E_1 is a line in $\mathbb{F}_2^2$ and $E_{-1} = E_1$.

Solution to Exercise 7.2.56

Uses. Definition 7.2.6 (Eigenvalue, eigenvector), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Definition 7.2.13 (Eigenspace)
Hint 1. $Z^3 = \mathrm{id}$, so every eigenvalue is a cube root of 1 (Proposition 7.2.28). Conversely, for $\lambda^3 = 1$, solve $Z\mathbf{v} = \lambda\mathbf{v}$ starting from $x = 1$.
Hint 2. The equations are $z = \lambda x$, $x = \lambda y$, $y = \lambda z$. Try $\mathbf{v} = (1, \lambda^2, \lambda)$.
Sketch. $Z\mathbf{v} = \lambda\mathbf{v}$ means $z = \lambda x$, $y = \lambda z = \lambda^2 x$, $x = \lambda y = \lambda^3 x$, so $\mathbf{v} = x(1, \lambda^2, \lambda)$ and $(\lambda^3 - 1)x = 0$. Eigenvalues are the cube roots of 1 in F, each with eigenspace $\mathrm{span}((1, \lambda^2, \lambda))$. $\mathbb{C}$: $1, \omega, \omega^2$. $\mathbb{R}$: 1. $\mathbb{F}_7$: $1, 2, 4$. $\mathbb{F}_3$: 1 only.

$Z\mathbf{v} = \lambda\mathbf{v}$ means $(z, x, y) = (\lambda x, \lambda y, \lambda z)$, that is, $z = \lambda x$, $x = \lambda y$, $y = \lambda z$. Then $y = \lambda z = \lambda^2 x$ and $x = \lambda y = \lambda^3 x$. So every solution is $\mathbf{v} = x(1, \lambda^2, \lambda)$, and it is nonzero only if $x \neq 0$ and $\lambda^3 = 1$. Conversely, if $\lambda^3 = 1$, then

$$Z(1, \lambda^2, \lambda) = (\lambda, 1, \lambda^2) = (\lambda, \lambda^3, \lambda^2) = \lambda(1, \lambda^2, \lambda).$$

So the eigenvalues are the cube roots of 1 in F, and $E_\lambda = \mathrm{span}((1, \lambda^2, \lambda))$ for each.

- Over $\mathbb{C}$: $1, \omega, \omega^2$ with $\omega = e^{2\pi i/3} = (-1 + i\sqrt{3})/2$, and eigenvectors $(1, 1, 1)$, $(1, \omega^2, \omega)$, $(1, \omega, \omega^2)$.
- Over $\mathbb{R}$: only 1, since $t \mapsto t^3$ is injective on $\mathbb{R}$. Eigenvector $(1, 1, 1)$.
- Over $\mathbb{F}_7$: $2^3 = 8 = 1$ and $4^3 = 64 = 1$, and $3^3 = 6$, $5^3 = 6$, $6^3 = 6$ are not 1. So the eigenvalues are $1, 2, 4$, with eigenvectors $(1, 1, 1)$, $(1, 4, 2)$, and $(1, 2, 4)$.
- Over $\mathbb{F}_3$: $0^3 = 0$, $1^3 = 1$, $2^3 = 8 = 2$. Only 1, with $E_1 = \mathrm{span}((1, 1, 1))$.

Over $\mathbb{C}$ and $\mathbb{F}_7$ there are 3 distinct eigenvalues, so the three eigenvectors form a basis (Theorem 7.2.21). Over $\mathbb{F}_3$ there is only one eigenvector line.

Solution to Exercise 7.2.57

Uses. Theorem 4.1.16 (Extending to a basis), Exercise 7.2.54, Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Exercise 2.1.48
Hint 1. Show that every line $\mathrm{span}(\mathbf{v})$ is invariant. Intersections of invariant subspaces are invariant.
Hint 2. For $\mathbf{v} \neq \mathbf{0}$ and $\mathbf{w} \notin \mathrm{span}(\mathbf{v})$, build a k-dimensional subspace containing $\mathbf{v}$ but not $\mathbf{w}$ by extending $\mathbf{v}, \mathbf{w}$ to a basis.
Hint 3. So $\mathrm{span}(\mathbf{v})$ is the intersection of all k-dimensional subspaces that contain $\mathbf{v}$. Finish with Exercise 2.1.48.
Sketch. Fix $\mathbf{v} \neq \mathbf{0}$. For $\mathbf{w} \notin \mathrm{span}(\mathbf{v})$, extend $\mathbf{v}, \mathbf{w}$ to a basis $\mathbf{v}, \mathbf{w}, \mathbf{b}_3, \ldots, \mathbf{b}_n$ and let $U = \mathrm{span}(\mathbf{v}, \mathbf{b}_3, \ldots, \mathbf{b}_{k+1})$, of dimension k, containing $\mathbf{v}$ but not $\mathbf{w}$. So $\mathrm{span}(\mathbf{v})$ is the intersection of the k-dimensional subspaces containing $\mathbf{v}$, all invariant, hence invariant. Every vector is then an eigenvector or $\mathbf{0}$, so T is scalar.

Fix $\mathbf{v} \neq \mathbf{0}$, and let X be the intersection of all k-dimensional subspaces of V that contain $\mathbf{v}$. There is at least one such subspace, since $k \geq 1$ (extend $\mathbf{v}$ to a basis and take the span of $\mathbf{v}$ and $k - 1$ further basis vectors). So X is a subspace containing $\mathrm{span}(\mathbf{v})$. We show $X = \mathrm{span}(\mathbf{v})$. Let $\mathbf{w} \notin \mathrm{span}(\mathbf{v})$. Then $\mathbf{v}, \mathbf{w}$ is independent, and it extends to a basis $\mathbf{v}, \mathbf{w}, \mathbf{b}_3, \ldots, \mathbf{b}_n$ of V (Theorem 4.1.16). Since $k - 1 \leq n - 2$, the list $\mathbf{v}, \mathbf{b}_3, \ldots, \mathbf{b}_{k+1}$ makes sense; its span U has dimension k and contains $\mathbf{v}$, but not $\mathbf{w}$, because $\mathbf{w}$ is not a combination of the other basis vectors. So $\mathbf{w} \notin X$, and $X = \mathrm{span}(\mathbf{v})$.

Each of these k-dimensional subspaces is T-invariant by hypothesis, so X is T-invariant: if $\mathbf{x} \in X$, then $T\mathbf{x}$ lies in each of them, as in Exercise 7.2.54(a). So $T\mathbf{v} \in \mathrm{span}(\mathbf{v})$ for every $\mathbf{v} \neq \mathbf{0}$, and trivially for $\mathbf{v} = \mathbf{0}$.

By Exercise 2.1.48, $T = \lambda\,\mathrm{id}$ for some λ. (The converse is clear, and the hypothesis $k < n$ is needed: every operator leaves the one n-dimensional subspace V invariant.)

Solution to Exercise 7.2.58

> *Uses.* Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 2.5.28 (Inverse of a 2×2 matrix)
> *Techniques.* Construct a map by choosing it on a basis (Technique 2.2.4)
> *Hint 1.* $(1,2), (1,1)$ is a basis, so Theorem 2.2.1 gives exactly one such map.
> *Hint 2.* With $S = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, the conditions say $AS = S\,\mathrm{diag}(3,-1)$. Solve for A.
> *Sketch.* $AS = SD$ with $S = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ and $D = \mathrm{diag}(3,-1)$, so $A = SDS^{-1} = \begin{pmatrix} 3 & -1 \\ 6 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ -8 & 7 \end{pmatrix}$. It is unique because a linear map is determined by its values on a basis.

The vectors $(1,2)$ and $(1,1)$ form a basis of $\mathbb{Q}^2$, so by Theorem 2.2.1 there is exactly one linear map with the prescribed values, and A is unique. To find it, let $S = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, whose columns are the two vectors, and $D = \mathrm{diag}(3,-1)$. The conditions say $AS = SD$, column by column. By Proposition 2.5.28, $\det S = -1$ and $S^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$, so

$$A = SDS^{-1} = \begin{pmatrix} 3 & -1 \\ 6 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} -5 & 4 \\ -8 & 7 \end{pmatrix}.$$

Check: $A(1,2) = (-5+8, -8+14) = (3,6)$ and $A(1,1) = (-1,-1)$.

Section 7.3

Solution to Exercise 7.3.9

> *Uses.* Example 7.3.5 (Small cases), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial)
> *Hint 1.* $p_A = x^2 - (\mathrm{tr}\,A)x + \det A$. Then find the roots field by field.
> *Hint 2.* For a root λ, the first row of $A - \lambda I_2$ is $(1-\lambda, 1)$, so $(1, \lambda - 1)$ is in the kernel. Since $\lambda^2 = \lambda + 1$, this is $(1, 1/\lambda)$, or more simply the multiple $(\lambda, 1)$.
> *Sketch.* $p_A = x^2 - x - 1$. Over $\mathbb{R}$: $\left(1 \pm \sqrt{5}\right)/2$. Over $\mathbb{F}_2$: $x^2 + x + 1$ has no root. Over $\mathbb{F}_5$: $(x-3)^2$, one eigenvalue 3. Over $\mathbb{F}_{11}$: roots 4 and 8. Eigenvectors $(\lambda, 1)$ in each case.

By Example 7.3.5(a), $p_A = x^2 - x - 1$. If λ is a root, then $(A - \lambda I_2)(\lambda, 1) = \left(\lambda - \lambda^2 + 1, \lambda - \lambda\right) = (0,0)$, because $\lambda^2 = \lambda + 1$. So $(\lambda, 1)$ is an eigenvector for every root λ, and by Theorem 7.3.6 the eigenvalues are exactly the roots.

Over $\mathbb{R}$, the roots are $\left(1 \pm \sqrt{5}\right)/2$, the golden ratio and its conjugate.

Over $\mathbb{F}_2$, $p_A = x^2 + x + 1$ takes the value 1 at both 0 and 1, so A has no eigenvalues.

Over $\mathbb{F}_5$, $p_A = x^2 - x - 1 = x^2 + 4x + 4 = (x+2)^2 = (x-3)^2$, so 3 is the only eigenvalue, with eigenvector $(3,1)$. Indeed $A(3,1) = (4,3) = 3(3,1)$, as $9 = 4$ in $\mathbb{F}_5$.

Over $\mathbb{F}_{11}$, $p_A(4) = 16 - 4 - 1 = 11 = 0$ and $p_A(8) = 64 - 8 - 1 = 55 = 0$, so the eigenvalues are 4 and 8 (there are at most two), with eigenvectors $(4,1)$ and $(8,1)$. These are the two "square roots of 5 formulas" $(1 \pm 4)/2$ read in $\mathbb{F}_{11}$, since $4^2 = 5$ there.

Solution to Exercise 7.3.17

> *Uses.* Proposition 7.3.2 (Determinants over a commutative ring), Lemma 7.1.37 (Conjugating polynomials), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Definition 6.3.26 (Conjugate transpose)
> *Hint 1.* For (a), what is $(xI_n - A)^{\mathsf{T}}$? Use Proposition 7.3.2(d).
> *Hint 2.* For (b), coefficientwise conjugation $\mathbb{C}[x] \to \mathbb{C}[x]$ respects sums, products, and 1 (Lemma 7.1.37). Apply Proposition 7.3.2(g) to it.
> *Sketch.* (a) $(xI_n - A)^{\mathsf{T}} = xI_n - A^{\mathsf{T}}$, and transposing does not change det. (b) Conjugating every entry of $xI_n - A$ gives $xI_n - \overline{A}$, so $p_{\overline{A}} = \overline{p_A}$ by Proposition 7.3.2(g). With (a), $p_{A^{\mathsf{H}}} = \overline{p_A}$, whose roots are the conjugates of the roots of p_A by Lemma 7.1.37(b).

(a) The transpose of $xI_n - A$ is $xI_n - A^{\mathsf{T}}$, since xI_n is symmetric. By Proposition 7.3.2(d), $p_{A^{\mathsf{T}}} = \det\left((xI_n - A)^{\mathsf{T}}\right) = \det(xI_n - A) = p_A$.

(b) Let $\varphi : \mathbb{C}[x] \to \mathbb{C}[x]$, $\varphi(f) = \overline{f}$. By Lemma 7.1.37(a), φ respects sums and products, and $\varphi(1) = 1$. Applying φ to every entry of $xI_n - A$ gives $xI_n - \overline{A}$, because $\overline{x} = x$ and the constant $-a_{ij}$ becomes $-\overline{a_{ij}}$. By Proposition 7.3.2(g), $p_{\overline{A}} = \det \varphi(xI_n - A) = \varphi(p_A) = \overline{p_A}$.

Now $A^{\mathsf{H}} = \overline{A}^{\mathsf{T}}$, so by (a), $p_{A^{\mathsf{H}}} = p_{\overline{A}} = \overline{p_A}$. By Lemma 7.1.37(b), $\overline{p_A}\big(\overline{\lambda}\big) = \overline{p_A(\lambda)}$, so $\overline{\lambda}$ is a root of $\overline{p_A}$ if and only if λ is a root of p_A. By Theorem 7.3.6, the eigenvalues of A^{H} are exactly the conjugates $\overline{\lambda}$ of the eigenvalues λ of A.

Solution to Exercise 7.3.20

Uses. Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Remark 7.3.19 (Trace and determinant from the roots), Theorem 7.2.9 (Eigenvalues and invertibility)
Hint 1. Since $A - I_3$ is not invertible, 1 is an eigenvalue. Write $p_A = (x-1)(x-r)(x-s)$.
Hint 2. Remark 7.3.19 gives $1 + r + s = 2$ and $rs = -2$.
Sketch. 1 is a root, $r + s = 1$, $rs = -2$, so r, s are the roots of $t^2 - t - 2 = (t-2)(t+1)$. Hence $p_A = (x-1)(x-2)(x+1)$.

By Theorem 7.2.9, 1 is an eigenvalue of A, so $p_A(1) = 0$ by Theorem 7.3.6, and since p_A splits we can write $p_A = (x-1)(x-r)(x-s)$ with $r, s \in \mathbb{R}$. By Remark 7.3.19, $1 + r + s = \operatorname{tr} A = 2$ and $1 \cdot rs = \det A = -2$. So $r + s = 1$ and $rs = -2$, which makes r and s the roots of $(t-r)(t-s) = t^2 - t - 2 = (t-2)(t+1)$. Therefore $p_A = (x-1)(x-2)(x+1)$, and the eigenvalues are 1, 2, -1. (The matrix $\operatorname{diag}(1, 2, -1)$ shows that such an A exists.)

Solution to Exercise 7.3.37

Uses. Algorithm 7.3.7 (Eigenvalues by the characteristic polynomial), Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.3.15 (Block-triangular matrices), Exercise 7.3.17, Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial)
Hint 1. In (a), add the second and third rows of $xI_3 - A$ to the first; a common factor appears. In (b), A is the transpose of a companion matrix. In (c), A is block diagonal.
Hint 2. The answers are $(x-5)(x+1)^2$, $(x-1)^2(x-2)$, and $(x-2)\big(x^2 - 2x + 2\big)$. Now row reduce $A - \lambda I_3$ for each root.
Sketch. (a) $p_A = (x-5)(x+1)^2$, $E_5 = \operatorname{span}((1,1,1))$, $E_{-1} = \operatorname{span}((1,-1,0),(1,0,-1))$. (b) $p_A = x^3 - 4x^2 + 5x - 2 = (x-1)^2(x-2)$, $E_1 = \operatorname{span}((1,1,1))$, $E_2 = \operatorname{span}((1,2,4))$. (c) $p_A = (x-2)\big(x^2 - 2x + 2\big)$; over $\mathbb{R}$ only $E_2 = \operatorname{span}(\mathbf{e}_3)$; over $\mathbb{C}$ also $E_{1+i} = \operatorname{span}((i,1,0))$ and $E_{1-i} = \operatorname{span}((-i,1,0))$.

(a) In $xI_3 - A$, add rows 2 and 3 to row 1; this does not change the determinant (Proposition 7.3.2(b), (d)). Row 1 becomes $(x-5, x-5, x-5)$, and pulling out the factor $x - 5$ and then adding 2 times row 1 to rows 2 and 3,

$$p_A = (x-5)\det\begin{pmatrix} 1 & 1 & 1 \\ -2 & x-1 & -2 \\ -2 & -2 & x-1 \end{pmatrix} = (x-5)\det\begin{pmatrix} 1 & 1 & 1 \\ 0 & x+1 & 0 \\ 0 & 0 & x+1 \end{pmatrix} = (x-5)(x+1)^2.$$

Next, $A - 5I_3 = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$ has rank 2 and kills $(1,1,1)$, so $E_5 = \operatorname{span}((1,1,1))$. Every row of $A + I_3$ is $(2,2,2)$, so $E_{-1} = \{(x,y,z) \mid x + y + z = 0\} = \operatorname{span}((1,-1,0),(1,0,-1))$. As a check, $A = 2J - I_3$ with J the all-ones matrix, and J acts as 3 on $(1,1,1)$ and as 0 on the plane $x + y + z = 0$.

(b) $A^{\mathsf{T}} = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -5 \\ 0 & 1 & 4 \end{pmatrix}$ is the companion matrix of $p = x^3 - 4x^2 + 5x - 2$, so $p_A = p_{A^{\mathsf{T}}} = p$ by Exercise 7.3.17 and Theorem 7.3.24. (Expanding $\det(xI_3 - A)$ along the first row gives the same.) Since $p(1) = 0$, dividing gives $p = (x-1)\big(x^2 - 3x + 2\big) = (x-1)^2(x-2)$. Then

$$A - I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & -5 & 3 \end{pmatrix}, \quad A - 2I_3 = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 2 & -5 & 2 \end{pmatrix}.$$

The first two rows of each are independent, so both have rank 2. The first two rows of $A - I_3$ force $y = x$ and $z = y$, and $(1,1,1)$ satisfies the third row; so $E_1 = \operatorname{span}((1,1,1))$. Likewise $E_2 = \operatorname{span}((1,2,4))$. Although 1 is a double root of p_A, its eigenspace is only a line.

(c) A is block diagonal with blocks $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and (2), so by Proposition 7.3.15 and Example 7.3.5(a), $p_A = (x^2 - 2x + 2)(x - 2)$. Over $\mathbb{R}$, $x^2 - 2x + 2 = (x-1)^2 + 1$ has no root, so the only eigenvalue is 2, and $A - 2I_3 = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ gives $E_2 = \operatorname{span}(\mathbf{e}_3)$. Over $\mathbb{C}$, the quadratic has roots $1 \pm i$. For $\lambda = 1 + i$, the first row of $A - \lambda I_3$ is $(-i, -1, 0)$, so $(i, 1, 0)$ spans E_{1+i}; indeed $A(i,1,0) = (i-1, i+1, 0) = (1+i)(i,1,0)$. Since A is real, conjugating $A\mathbf{v} = \lambda\mathbf{v}$ entrywise gives $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$, so $E_{1-i} = \operatorname{span}((-i, 1, 0))$, and $E_2 = \operatorname{span}(\mathbf{e}_3)$ as before.

Solution to Exercise 7.3.38

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.1.10 (Evaluation respects sums and products), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Corollary 3.5.25 (The invertible matrix theorem), Proposition 4.5.24 (Similarity is an equivalence relation)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. For (a) and (b), look at 2×2 matrices with a repeated eigenvalue.
Hint 2. For (c), $xI_n - (A + cI_n) = (x - c)I_n - A$. Which map $F[x] \to F[x]$ turns $xI_n - A$ into this matrix?
Hint 3. For (d), pick an eigenvector of A for each root. They form a basis, in which A becomes $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Do the same for B separately, in its own basis of eigenvectors.
Sketch. (a) False: I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. (b) False: 0 and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. (c) True: substitute $x - c$ for x entrywise and use Proposition 7.3.2(g). (d) True: n distinct eigenvalues give a basis of eigenvectors, so A and B are both similar to $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

(a) False. I_2 and $N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ both have characteristic polynomial $(x-1)^2$ (Example 7.3.5(d)). But $S^{-1}I_2S = I_2 \neq N$ for every invertible S, so they are not similar.

(b) False. $0_{2\times 2}$ and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ both have characteristic polynomial x^2, but their ranks are 0 and 1.

(c) True. Substituting $x - c$ for x respects sums, products, and 1 (Proposition 7.1.10), and applied to every entry of $xI_n - A$ it gives $(x-c)I_n - A = xI_n - (A + cI_n)$, since the constant entries are unchanged. By Proposition 7.3.2(g), $p_{A+cI_n} = p_A(x - c)$. In particular the eigenvalues of $A + cI_n$ are the $\lambda + c$ for the eigenvalues λ of A.

(d) True. Let $\lambda_1, \ldots, \lambda_n$ be the distinct roots. By Theorem 7.3.6 each λ_i is an eigenvalue of A; choose eigenvectors $\mathbf{v}_i$. By Theorem 7.2.21 they are independent, so they form a basis of F^n, and the matrix S with columns $\mathbf{v}_i$ is invertible (Corollary 3.5.25). Column i of AS is $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, which is column i of SD for $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. So $S^{-1}AS = D$. The same argument for B gives $R^{-1}BR = D$. Hence $A \sim D \sim B$, and $A \sim B$ by Proposition 4.5.24.

Solution to Exercise 7.3.39

Uses. Lemma 2.4.24 (Matrix units), Definition 2.3.1 (Matrix of a linear map), Proposition 7.3.15 (Block-triangular matrices), Corollary 7.3.13 (Eigenvalues of an operator)
Hint 1. For (a), order the matrix units so that U_{ij} and U_{ji} are adjacent. What does the matrix of τ look like?
Hint 2. It is block diagonal, with n blocks (1) and $n(n-1)/2$ blocks $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. For (b), compute D on each basis function; the matrix is block upper triangular.
Sketch. (a) $\tau(U_{ii}) = U_{ii}$ and τ swaps U_{ij} and U_{ji}; so $p_\tau = (x-1)^n(x^2-1)^{n(n-1)/2}$. (b) The matrix is $\begin{pmatrix} R & I_2 \\ 0 & R \end{pmatrix}$ with $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so $p_D = (x^2+1)^2$, which has no real root.

(a) By Lemma 2.4.24, the matrix units U_{ij} form a basis of $F^{n\times n}$. Order it as $U_{11}, \ldots, U_{nn}$, followed by the pairs U_{ij}, U_{ji} for $i < j$. Since $\tau(U_{ii}) = U_{ii}$, $\tau(U_{ij}) = U_{ji}$, and $\tau(U_{ji}) = U_{ij}$, the matrix of τ is block diagonal with n blocks (1) and $n(n-1)/2$ blocks $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, whose characteristic polynomial is $x^2 - 1 = (x-1)(x+1)$. Applying Proposition 7.3.15 once for each block,

$$p_\tau = (x-1)^n\left(x^2-1\right)^{n(n-1)/2} = (x-1)^{n+n(n-1)/2}(x+1)^{n(n-1)/2},$$

and $n + n(n-1)/2 = n(n+1)/2$. This fits Example 7.2.17: $\tau^2 = \mathrm{id}$, the symmetric matrices form E_1, of dimension $n(n+1)/2$, and the antisymmetric ones form E_{-1}. When $1+1=0$ in F, the formula reads $p_\tau = (x-1)^{n^2}$.

(b) Let $\mathbf{b}_1 = \sin t$, $\mathbf{b}_2 = \cos t$, $\mathbf{b}_3 = t\sin t$, $\mathbf{b}_4 = t\cos t$. Then

$$D\mathbf{b}_1 = \mathbf{b}_2, \quad D\mathbf{b}_2 = -\mathbf{b}_1, \quad D\mathbf{b}_3 = \mathbf{b}_1 + \mathbf{b}_4, \quad D\mathbf{b}_4 = \mathbf{b}_2 - \mathbf{b}_3,$$

so D maps W to W, and its matrix (Definition 2.3.1) is

$$[D] = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This is block upper triangular with both diagonal blocks equal to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, so $p_D = \left(x^2+1\right)^2$ by Proposition 7.3.15. It has no real root, so D has no real eigenvalue (Corollary 7.3.13): no nonzero $f \in W$ satisfies $f' = \lambda f$ with $\lambda \in \mathbb{R}$. The upper left block reflects the invariant plane $\mathrm{span}(\sin t, \cos t)$.

Solution to Exercise 7.3.40

Uses. Proposition 7.3.22 (AB and BA), Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Proposition 7.3.15 (Block-triangular matrices), Theorem 2.4.34 (Block multiplication), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial)
Techniques. Compare AB and BA (Technique 7.3.21)
Hint 1. Imitate the first proof of Proposition 7.3.22 with $S = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$. The zero blocks now have different sizes.
Hint 2. For (b), evaluate the identity of (a) at $\lambda \neq 0$.
Hint 3. For (c), why is $B\mathbf{v} \neq \mathbf{0}$? Apply A to it.
Sketch. (a) $X = \left(\begin{smallmatrix} AB & 0 \\ B & 0_{n\times n} \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} 0_{m\times m} & 0 \\ B & BA \end{smallmatrix}\right)$ satisfy $XS = SY$ with $S = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$, and $p_X = p_{AB}x^n$, $p_Y = x^m p_{BA}$. (b) For $\lambda \neq 0$, $\lambda^n p_{AB}(\lambda) = \lambda^m p_{BA}(\lambda)$. (c) $A(B\mathbf{v}) = \lambda\mathbf{v} \neq \mathbf{0}$, and $BA(B\mathbf{v}) = B(\lambda\mathbf{v})$. (d) $A = (1,0)$, $B = (1,0)^{\mathsf{T}}$.

(a) Let $S = \left(\begin{smallmatrix} I_m & A \\ 0 & I_n \end{smallmatrix}\right)$, which is invertible with inverse $\left(\begin{smallmatrix} I_m & -A \\ 0 & I_n \end{smallmatrix}\right)$, and let

$$X = \begin{pmatrix} AB & 0 \\ B & 0_{n\times n} \end{pmatrix}, \quad Y = \begin{pmatrix} 0_{m\times m} & 0 \\ B & BA \end{pmatrix}.$$

By block multiplication (Theorem 2.4.34), $XS = \left(\begin{smallmatrix} AB & ABA \\ B & BA \end{smallmatrix}\right) = SY$. So X and Y are similar and $p_X = p_Y$ (Proposition 7.3.11). Both are block lower triangular, so by Proposition 7.3.15, $p_X = p_{AB}p_{0_{n\times n}} = p_{AB}x^n$ and $p_Y = x^m p_{BA}$.

(b) Let $\lambda \neq 0$. Evaluating (a) at λ gives $\lambda^n p_{AB}(\lambda) = \lambda^m p_{BA}(\lambda)$, and $\lambda^n, \lambda^m \neq 0$. So $p_{AB}(\lambda) = 0$ if and only if $p_{BA}(\lambda) = 0$, and Theorem 7.3.6 finishes.

(c) Suppose $AB\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$ and $\lambda \neq 0$. Then $A(B\mathbf{v}) = \lambda\mathbf{v} \neq \mathbf{0}$, so $B\mathbf{v} \neq \mathbf{0}$, and $(BA)(B\mathbf{v}) = B(AB\mathbf{v}) = \lambda B\mathbf{v}$. So every nonzero eigenvalue of AB is one of BA, and exchanging the roles of A and B gives the converse.

(d) Let $A = \begin{pmatrix} 1 & 0 \end{pmatrix} \in F^{1\times 2}$ and $B = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \in F^{2\times 1}$. Then $AB = (1)$ has the single eigenvalue 1, while $BA = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ also has the eigenvalue 0. Consistently with (a), $x^2(x-1) = x \cdot x(x-1)$.

Solution to Exercise 7.3.41

Uses. Exercise 7.3.40, Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 4.5.35 (Preserved subspaces give block triangular matrices), Proposition 4.5.27 (Similarity invariants), Theorem 4.2.11 (Rank–nullity), Example 7.3.5 (Small cases)
Techniques. Compare AB and BA (Technique 7.3.21), Choose a basis adapted to a preserved subspace (Technique 4.5.36)

Hint 1. Write $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ as a product of an $n \times 1$ and a $1 \times n$ matrix, and use Exercise 7.3.40.
Hint 2. Alternatively, if $\mathbf{v} \neq \mathbf{0}$, the hyperplane $\{\mathbf{w} \mid \mathbf{v}^{\mathsf{T}}\mathbf{w} = 0\}$ lies in the kernel. Choose a basis starting with a basis of it; the matrix becomes triangular. Its last diagonal entry is the trace.
Sketch. (a) With $A = \mathbf{u} \in F^{n\times 1}$ and $B = \mathbf{v}^{\mathsf{T}} \in F^{1\times n}$, Exercise 7.3.40 gives $x p_{\mathbf{u}\mathbf{v}^{\mathsf{T}}} = x^n p_{\mathbf{v}^{\mathsf{T}}\mathbf{u}} = x^n\big(x - \mathbf{v}^{\mathsf{T}}\mathbf{u}\big)$; cancel x. (b) $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ and $\mathbf{1}^{\mathsf{T}}\mathbf{1} = n$, so $p_J = x^{n-1}(x-n)$: eigenvalues $0, n$ over $\mathbb{Q}$, and only 0 over $\mathbb{F}_p$ when $p \mid n$.

(a) *First proof.* Let $A = \mathbf{u} \in F^{n\times 1}$ and $B = \mathbf{v}^{\mathsf{T}} \in F^{1\times n}$, so $AB = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ and $BA = \big(\mathbf{v}^{\mathsf{T}}\mathbf{u}\big)$ is 1×1, with $p_{BA} = x - \mathbf{v}^{\mathsf{T}}\mathbf{u}$. By Exercise 7.3.40, $x p_{\mathbf{u}\mathbf{v}^{\mathsf{T}}} = x^n\big(x - \mathbf{v}^{\mathsf{T}}\mathbf{u}\big)$. Cancelling x in $F[x]$ (Corollary 7.1.7) gives the claim.

Second proof. Let $M = \mathbf{u}\mathbf{v}^{\mathsf{T}}$. If $\mathbf{v} = \mathbf{0}$, then $M = 0$ and $p_M = x^n$, as claimed. Otherwise $H = \{\mathbf{w} \in F^n \mid \mathbf{v}^{\mathsf{T}}\mathbf{w} = 0\}$ is the kernel of a nonzero linear map $F^n \to F$, so $\dim H = n - 1$ (Theorem 4.2.11), and $M\mathbf{w} = \mathbf{u}\big(\mathbf{v}^{\mathsf{T}}\mathbf{w}\big) = \mathbf{0}$ for $\mathbf{w} \in H$. Extend a basis $\mathbf{w}_1, \ldots, \mathbf{w}_{n-1}$ of H to a basis $\mathcal{B}$ of F^n. The first $n-1$ columns of $[T_M]_{\mathcal{B}\mathcal{B}}$ are $\mathbf{0}$, so this matrix is upper triangular with diagonal entries $0, \ldots, 0, c$, and $p_M = x^{n-1}(x-c)$ by Example 7.3.5(d). Since the trace is a similarity invariant (Proposition 4.5.27), $c = \operatorname{tr} M = \sum_i u_i v_i = \mathbf{v}^{\mathsf{T}}\mathbf{u}$.

(b) $J = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ with $\mathbf{1} = (1, \ldots, 1)$, and $\mathbf{1}^{\mathsf{T}}\mathbf{1} = n$. By (a), $p_J = x^{n-1}(x-n)$. Over $\mathbb{Q}$ the eigenvalues are 0 and n. Over $\mathbb{F}_p$ with $p \mid n$, we have $n = 0$, so $p_J = x^n$ and 0 is the only eigenvalue; indeed $J\mathbf{1} = n\mathbf{1} = \mathbf{0}$ there.

Solution to Exercise 7.3.42

Uses. Definition 7.3.1 (Determinant over a commutative ring), Definition 7.1.1 (Product of polynomials), Proposition 7.3.18 (The first and last coefficients), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial)
Hint 1. For (a), every entry of $xI_n - A$ is a polynomial with integer coefficients. What happens to such polynomials under sums and products?
Hint 2. For (b), write the eigenvalue as a/b in lowest terms with $b \geq 1$, substitute into p_A, and multiply by b^n. Which terms are divisible by b?
Sketch. (a) The Leibniz formula is a signed sum of products of polynomials with integer coefficients. (b) If $p_A(a/b) = 0$ with $\gcd(a,b) = 1$, then $a^n = -b\big(c_{n-1}a^{n-1} + \cdots + c_0 b^{n-1}\big)$, so every prime factor of b divides a; hence $b = 1$. Then a divides the constant term $(-1)^n \det A$. (c) $p_A = x^2 - 5x - 2$ has no root among $\pm 1, \pm 2$.

(a) Every entry of $xI_n - A$ is $x - a_{ii}$ or $-a_{ij}$, a polynomial with integer coefficients. By the formula for the product (Definition 7.1.1), sums and products of polynomials with integer coefficients have integer coefficients. So every Leibniz term, and hence p_A, has integer coefficients.

(b) Write $p_A = x^n + c_{n-1}x^{n-1} + \cdots + c_0$ with $c_i \in \mathbb{Z}$, and suppose $p_A(a/b) = 0$ with $a, b \in \mathbb{Z}$, $b \geq 1$, and $\gcd(a,b) = 1$. Multiplying by b^n,

$$a^n = -b\big(c_{n-1}a^{n-1} + c_{n-2}a^{n-2}b + \cdots + c_0 b^{n-1}\big).$$

If $b > 1$, a prime q dividing b divides a^n, hence divides a, contradicting $\gcd(a,b) = 1$. So $b = 1$ and the eigenvalue a is an integer. Then $0 = a^n + c_{n-1}a^{n-1} + \cdots + c_1 a + c_0$ gives $c_0 = -a\big(a^{n-1} + c_{n-1}a^{n-2} + \cdots + c_1\big)$, so a divides $c_0 = (-1)^n \det A$ (Proposition 7.3.18), and therefore divides $\det A$.

(c) Here $p_A = x^2 - 5x - 2$ and $\det A = -2$. By (b), a rational eigenvalue would be one of $\pm 1, \pm 2$, but $p_A(1) = -6$, $p_A(-1) = 4$, $p_A(2) = -8$, and $p_A(-2) = 12$. So there is none, and the eigenvalues $\big(5 \pm \sqrt{33}\big)/2$ are irrational.

Solution to Exercise 7.3.43

Uses. Theorem 5.4.10 (Multiplicativity), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Corollary 7.1.35 (Factorization over $\mathbb{C}$), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Proposition 7.1.38 (Conjugate roots)
Hint 1. For (a), take determinants of both sides. What sign does a real square have?
Hint 2. For (b), an eigenvalue λ of A over $\mathbb{C}$ satisfies $\lambda^2 = -1$. Then $p_A = (x-i)^a(x+i)^b$. Use that p_A has real coefficients.
Sketch. (a) $(\det A)^2 = \det(-I_n) = (-1)^n$, and a nonzero real square is positive, so $(-1)^n = 1$. (b) All complex eigenvalues are $\pm i$, so $p_A = (x-i)^a(x+i)^b$ with $a + b = n$; real coefficients force $a = b$ (Proposition 7.1.38). (c) Block-diagonal rotations by $\pi/2$ over $\mathbb{R}$; iI_n over $\mathbb{C}$.

(a) By Theorem 5.4.10, $(\det A)^2 = \det\big(A^2\big) = \det(-I_n) = (-1)^n$. Since $\det A \in \mathbb{R}$ and $(\det A)^2 = (-1)^n \neq 0$, we get $(-1)^n = (\det A)^2 > 0$, so n is even.

(b) Regard A as a complex matrix. If $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $-\mathbf{v} = A^2\mathbf{v} = \lambda^2\mathbf{v}$ (Proposition 7.2.28), so $\lambda^2 = -1$ and $\lambda = \pm i$. By Corollary 7.1.35, p_A is a product of n factors $x - \lambda_k$, and each λ_k is a root, hence an eigenvalue (Theorem 7.3.6), hence $\pm i$. So $p_A = (x-i)^a(x+i)^b$ with $a + b = n$. The entries of A are real, so $p_A \in \mathbb{R}[x]$, and by Proposition 7.1.38 the roots i and $-i = \bar{i}$ have the same multiplicity: $a = b$. Hence $n = 2a$ is even.

(c) Over $\mathbb{C}$, $(iI_n)^2 = -I_n$. Over $\mathbb{R}$ with $n = 2m$, the block-diagonal matrix with m blocks $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ works, since $R^2 = -I_2$.

Solution to Exercise 7.3.44

Uses. Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Definition 7.3.23 (Companion matrix)
Hint 1. Evaluate p at each c_j.
Hint 2. Every monic polynomial is a characteristic polynomial (Theorem 7.3.24).
Sketch. (a) $p(c_j) = 0 + 1 = 1$. (b) $C_p \in F^{q\times q}$ has $p_{C_p} = p$, which has no root. (c) Over $\mathbb{F}_2$, $p = x(x-1) + 1 = x^2 + x + 1$ and $C_p = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$; neither C_p nor $C_p - I_2$ is singular.

(a) For each j, the product $(c_j - c_1)\cdots(c_j - c_q)$ contains the factor $c_j - c_j = 0$, so $p(c_j) = 0 + 1 = 1 \neq 0$.

(b) p is monic of degree q, so by Theorem 7.3.24 the companion matrix $C_p \in F^{q\times q}$ has $p_{C_p} = p$. By (a) and Theorem 7.3.6, C_p has no eigenvalue in F.

(c) Over $\mathbb{F}_2$, $p = x(x-1) + 1 = x^2 - x + 1 = x^2 + x + 1$, and $C_p = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Its determinant is $-1 = 1 \neq 0$, so 0 is not an eigenvalue, and $C_p - I_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ has determinant $-1 = 1 \neq 0$, so 1 is not one either.

Solution to Exercise 7.3.45

Uses. Definition 7.3.23 (Companion matrix), Theorem 4.2.11 (Rank–nullity), Remark 7.1.11 (Polynomials in a matrix), Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Proposition 7.1.22 (Counting roots)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. For (a), look at the first $n-1$ columns of $C - \lambda I_n$.
Hint 2. For (b), $C^j\mathbf{e}_1 = \mathbf{e}_{j+1}$ for $j \leq n-1$. Compute $q(C)\mathbf{e}_1$ and $p(C)\mathbf{e}_1$.
Hint 3. For (c), a basis of eigenvectors has at most one vector in each eigenspace, by (a).
Sketch. (a) Columns $k < n$ of $C - \lambda I_n$ are $\mathbf{e}_{k+1} - \lambda\mathbf{e}_k$, which are independent (look at the last coordinate first). (b) $q(C)\mathbf{e}_1$ is the given vector, and $(C - \lambda I_n)q(C)\mathbf{e}_1 = p(C)\mathbf{e}_1 = C\mathbf{e}_n + \sum a_j\mathbf{e}_{j+1} = \mathbf{0}$. (c) Distinct roots give independent eigenvectors; conversely a basis of eigenvectors needs n different eigenvalues, all roots of p.

(a) For $k \leq n-1$, column k of $C - \lambda I_n$ is $\mathbf{e}_{k+1} - \lambda\mathbf{e}_k$. Suppose $\sum_{k=1}^{n-1} c_k(\mathbf{e}_{k+1} - \lambda\mathbf{e}_k) = \mathbf{0}$. The coordinate at $\mathbf{e}_n$ is c_{n-1}, so $c_{n-1} = 0$. If $c_{k+1} = \cdots = c_{n-1} = 0$, the coordinate at $\mathbf{e}_{k+1}$ is $c_k - \lambda c_{k+1} = c_k$, so $c_k = 0$. Hence these $n-1$ columns are independent, $\operatorname{rank}(C - \lambda I_n) \geq n-1$, and $\dim E_\lambda = n - \operatorname{rank}(C - \lambda I_n) \leq 1$ by Theorem 4.2.11.

(b) Since $C\mathbf{e}_k = \mathbf{e}_{k+1}$ for $k < n$, we have $C^j\mathbf{e}_1 = \mathbf{e}_{j+1}$ for $0 \leq j \leq n-1$. So

$$\mathbf{w} = q(C)\mathbf{e}_1 = b_0\mathbf{e}_1 + b_1\mathbf{e}_2 + \cdots + b_{n-2}\mathbf{e}_{n-1} + \mathbf{e}_n = (b_0, \ldots, b_{n-2}, 1),$$

which is nonzero. Also $C^n\mathbf{e}_1 = C\mathbf{e}_n = -\sum_{j=0}^{n-1} a_j\mathbf{e}_{j+1}$, so $p(C)\mathbf{e}_1 = C^n\mathbf{e}_1 + \sum_j a_j\mathbf{e}_{j+1} = \mathbf{0}$. By Remark 7.1.11, $(C - \lambda I_n)\mathbf{w} = ((x-\lambda)q)(C)\mathbf{e}_1 = p(C)\mathbf{e}_1 = \mathbf{0}$.

(c) By Theorem 7.3.24 and Theorem 7.3.6, the eigenvalues of C are the roots of p in F. If there are n distinct ones, eigenvectors for them are independent (Theorem 7.2.21), hence a basis of F^n. Conversely, suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of eigenvectors. Two of them with the same eigenvalue would be two independent vectors in one eigenspace, which is a line by (a). So their eigenvalues are n distinct roots of p, and p has at least, hence exactly, n distinct roots (Proposition 7.1.22). In the language of Section 7.4, C_p is diagonalizable over F exactly when p splits with distinct roots.

Solution to Exercise 7.3.46

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Proposition 7.3.26 (Krylov subspaces and the characteristic polynomial), Remark 7.1.11 (Polynomials in a matrix)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. $\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}$ turn out to be independent. Solve $A^3\mathbf{v} = \alpha\mathbf{v} + \beta A\mathbf{v} + \gamma A^2\mathbf{v}$.
Hint 2. $p = x^3 - 4x^2 + 6x - 4$ has the root 2. For a root λ, apply $q = p/(x - \lambda)$ at A to $\mathbf{v}$.
Sketch. (a) $\mathbf{v} = (1, 0, 1)$, $A\mathbf{v} = (1, 1, 2)$, $A^2\mathbf{v} = (0, 2, 4)$, $A^3\mathbf{v} = (-2, 2, 8) = 4\mathbf{v} - 6A\mathbf{v} + 4A^2\mathbf{v}$, so $p = x^3 - 4x^2 + 6x - 4$. (b) $p = (x - 2)(x - 1 - i)(x - 1 + i)$, with eigenvectors $(0, 0, 2)$, $(-1 - i, -1 + i, 0)$, and $(-1 + i, -1 - i, 0)$. (c) $A^2\mathbf{e}_1 = 2A\mathbf{e}_1 - 2\mathbf{e}_1$, so $p = x^2 - 2x + 2$, which misses 2; it is a proper divisor of p_A.

(a) We get $\mathbf{v} = (1, 0, 1)$, $A\mathbf{v} = (1, 1, 2)$, $A^2\mathbf{v} = (0, 2, 4)$, and $A^3\mathbf{v} = (-2, 2, 8)$. The first three are independent: the matrix with these columns has determinant $1 \cdot (4 - 4) - 1 \cdot (0 - 2) + 0 = 2 \neq 0$. Solving $A^3\mathbf{v} = \alpha\mathbf{v} + \beta A\mathbf{v} + \gamma A^2\mathbf{v}$ coordinatewise gives $\alpha + \beta = -2$, $\beta + 2\gamma = 2$, $\alpha + 2\beta + 4\gamma = 8$, so $\gamma = 4$, $\beta = -6$, $\alpha = 4$. Hence $d = 3$ and $p = x^3 - 4x^2 + 6x - 4$.

(b) $p(2) = 0$, and dividing, $p = (x - 2)\big(x^2 - 2x + 2\big) = (x - 2)(x - (1 + i))(x - (1 - i))$. For each root λ let $q = p/(x - \lambda)$; then $q(A)\mathbf{v}$ is an eigenvector for λ.

- $\lambda = 2$: $q = x^2 - 2x + 2$, and $q(A)\mathbf{v} = (0, 2, 4) - 2(1, 1, 2) + 2(1, 0, 1) = (0, 0, 2)$.
- $\lambda = 1 + i$: $q = (x - 2)(x - 1 + i) = x^2 - (3 - i)x + (2 - 2i)$, and

$$q(A)\mathbf{v} = (0, 2, 4) - (3 - i)(1, 1, 2) + (2 - 2i)(1, 0, 1) = (-1 - i, -1 + i, 0) = (-1 + i)(i, 1, 0).$$

- $\lambda = 1 - i$: conjugating the previous computation, since A and $\mathbf{v}$ are real, gives $(-1 + i, -1 - i, 0) = (-1 - i)(-i, 1, 0)$.

Check: $A(0, 0, 1) = 2(0, 0, 1)$ and $A(i, 1, 0) = (i - 1, i + 1, 0) = (1 + i)(i, 1, 0)$.

(c) For $\mathbf{v} = \mathbf{e}_1$: $A\mathbf{e}_1 = (1, 1, 0)$, which is not a multiple of $\mathbf{e}_1$, and $A^2\mathbf{e}_1 = (0, 2, 0) = 2A\mathbf{e}_1 - 2\mathbf{e}_1$. So $d = 2$ and $p = x^2 - 2x + 2$, which produces the eigenvectors $(A - (1 - i)I_3)\mathbf{e}_1 = (i, 1, 0)$ and $(A - (1 + i)I_3)\mathbf{e}_1 = (-i, 1, 0)$ but misses the eigenvalue 2. This is consistent with Proposition 7.3.26: p divides $p_A = (x - 2)\big(x^2 - 2x + 2\big)$, but need not equal it, because $K_A(\mathbf{e}_1) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$ is a proper invariant subspace.

Solution to Exercise 7.3.47

Uses. Definition 7.3.4 (Characteristic polynomial), Proposition 7.3.2 (Determinants over a commutative ring), Remark 7.1.11 (Polynomials in a matrix)
Hint 1. What kind of object is each side? $p_A(A)$ is a matrix. What is det(...)?
Hint 2. Proposition 7.3.2(g) lets us substitute A for x in every entry of $xI_n - A$. What does that array look like? Is it the matrix $A - A$?
Hint 3. Apply the same reasoning to the single entry $x - a_{11}$ of $xI_n - A$, instead of to its determinant.
Sketch. The two sides have different types: $p_A(A)$ is a matrix and a determinant is a scalar. Substituting A for x entrywise turns $xI_n - A$ into an $n \times n$ array of matrices, with $A - a_{ii}I_n$ on the diagonal and $-a_{ij}I_n$ elsewhere, not into $A - A = 0$. The same move would show $A = a_{11}I_n$ for every A. The claim is true (Cayley–Hamilton) but needs a real proof.

The claim is true; it is the Cayley–Hamilton theorem (Theorem 7.6.20), and Exercise 7.3.51 proves it. The argument is not a proof.

First, the types do not match: $p_A(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_0 I_n$ is an $n \times n$ matrix, while a determinant is a scalar.

Second, the substitution is done wrongly. The identity $p_A = \det(xI_n - A)$ holds in $F[x]$, and Proposition 7.3.2(g) does allow us to apply a map like "substitute A for x" to it, but entry by entry. The map $q \mapsto q(A)$ takes values in the commutative ring of polynomials in A (Remark 7.1.11), and applied to the entries of $xI_n - A$ it gives the $n \times n$ array whose (i, i) entry is the matrix $A - a_{ii}I_n$ and whose (i, j) entry is the matrix $-a_{ij}I_n$. The determinant of that array, computed in this ring, is indeed $p_A(A)$. But the array is not the zero matrix "$AI_n - A$": the scalar entries a_{ij} became the matrices $a_{ij}I_n$, not the scalars a_{ij}. For $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, the array is $\left(\begin{smallmatrix} A & -I_2 \\ 0 & A \end{smallmatrix}\right)$, whose determinant in this ring is $A^2 = 0$, although no entry is 0.

Third, the argument proves too much. Apply the same move to the polynomial $f = x - a_{11}$, the $(1, 1)$ entry of $xI_n - A$: "substituting $x = A$" gives $f(A) = (AI_n - A)_{11} = 0$, that is, $A = a_{11}I_n$ for every A, which is false.

Solution to Exercise 7.3.48

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Theorem 7.1.20 (Factor theorem), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 7.3.18 (The first and last coefficients), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Corollary 7.1.35 (Factorization over $\mathbb{C}$)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. For (a), multiply out; the sum telescopes.
Hint 2. For (b), take determinants: $p_A g = x^{kn}$ for some $g \in F[x]$. So p_A divides a power of x.
Hint 3. Write $p_A = x^m h$ with $h(0) \neq 0$. Show that h divides 1, by evaluating at 0 and cancelling x repeatedly.
Sketch. (a) Telescoping: $\sum_j x^{k-j}A^j - \sum_j x^{k-1-j}A^{j+1} = x^k I_n - A^k$. (b) Taking determinants, $p_A \cdot g = x^{kn}$. Write $p_A = x^m h$ with $h(0) \neq 0$; cancelling x^m gives $hg = x^{kn-m}$, and evaluating at 0 repeatedly shows that g absorbs every factor x, so $hg' = 1$ and h is a constant, hence 1. So $p_A = x^n$. (c) Over $\mathbb{C}$, every eigenvalue satisfies $\lambda^k = 0$, and p_A splits.

(a) Let $N = \sum_{j=0}^{k-1} x^{k-1-j}A^j$. By Proposition 7.3.2(a), and since the scalar matrices $x^m I_n$ commute with every matrix,

$$(xI_n - A)N = \sum_{j=0}^{k-1} x^{k-j}A^j - \sum_{j=0}^{k-1} x^{k-1-j}A^{j+1} = \sum_{j=0}^{k-1} x^{k-j}A^j - \sum_{j=1}^{k} x^{k-j}A^j = x^k I_n - A^k,$$

which is $x^k I_n$ because $A^k = 0$.

(b) Let $g = \det N \in F[x]$. Taking determinants in (a) with Proposition 7.3.2(c) and (e), $p_A g = \det\left(x^k I_n\right) = x^{kn}$. Since $p_A \neq \mathbf{0}$, write $p_A = x^m h$ with $m \geq 0$ as large as possible, so that x does not divide h, that is, $h(0) \neq 0$ (Theorem 7.1.20). Comparing degrees in $x^m hg = x^{kn}$ gives $m \leq kn$, and cancelling x^m (Corollary 7.1.7(b)) gives $hg = x^{kn-m}$.

We show by induction on $r = kn - m$ that an equation $hg = x^r$ forces h to be a nonzero constant. If $r = 0$, then $hg = 1$, so h is invertible in $F[x]$, hence a nonzero constant (Corollary 7.1.7(c)). If $r \geq 1$, evaluating at 0 gives $h(0)g(0) = 0$, so $g(0) = 0$ and $g = xg_1$ (Theorem 7.1.20). Then $x(hg_1) = x \cdot x^{r-1}$, and cancelling x gives $hg_1 = x^{r-1}$, to which the induction hypothesis applies.

So h is a constant, and $h = 1$ because p_A is monic. Thus $p_A = x^m$, and $m = \deg p_A = n$ (Proposition 7.3.18). The coefficient of x^{n-1} is $0 = -\operatorname{tr} A$, and 0 is the only root.

(c) Over $\mathbb{C}$: if $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{0} = A^k\mathbf{v} = \lambda^k\mathbf{v}$ (Proposition 7.2.28), so $\lambda = 0$. By Corollary 7.1.35, p_A is a product of factors $x - \lambda_j$ whose roots λ_j are eigenvalues, hence 0; so $p_A = x^n$. Over a general field this argument only shows that 0 is the only root of p_A in F; it cannot rule out an irreducible factor of degree at least 2, because p_A need not split. The proof of (b) never needs p_A to split.

Solution to Exercise 7.3.49

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.1.10 (Evaluation respects sums and products), Remark 7.3.19 (Trace and determinant from the roots)
Techniques. Polynomial identity arguments (Technique 5.5.24)
Hint 1. Factor $x^2 I_n - A^2$ in $F[x]^{n\times n}$.
Hint 2. $\det\left(x^2 I_n - A^2\right)$ is p_{A^2} with x^2 substituted, by Proposition 7.3.2(g). Similarly, what is $\det(xI_n + A)$ in terms of p_A?
Hint 3. For (b), $(x-\lambda)(x+\lambda) = x^2 - \lambda^2$. A polynomial r is determined by $r\left(x^2\right)$.
Sketch. (a) $x^2 I_n - A^2 = (xI_n - A)(xI_n + A)$. Substituting x^2 into $xI_n - A^2$ gives $x^2 I_n - A^2$, so its determinant is $p_{A^2}\left(x^2\right)$; and $\det(xI_n + A) = (-1)^n \det(-xI_n - A) = (-1)^n p_A(-x)$. (b) $(-1)^n p_A(-x) = \prod(x + \lambda_i)$, so the right side is $\prod\left(x^2 - \lambda_i^2\right) = q\left(x^2\right)$ with $q = \prod\left(x - \lambda_i^2\right)$; compare coefficients. (c) $\operatorname{tr}\left(A^2\right) = \sum \lambda_i^2 = 0$ with real λ_i.

(a) In $F[x]^{n\times n}$, $(xI_n - A)(xI_n + A) = x^2 I_n + xA - xA - A^2 = x^2 I_n - A^2$. Substituting x^2 for x respects sums, products, and 1 (Proposition 7.1.10), and applied to every entry of $xI_n - A^2$ it gives $x^2 I_n - A^2$; so by Proposition 7.3.2(g), $\det\left(x^2 I_n - A^2\right) = p_{A^2}\left(x^2\right)$. In the same way, substituting $-x$ into $xI_n - A$ gives $-xI_n - A$, so $\det(-xI_n - A) = p_A(-x)$, and $\det(xI_n + A) = (-1)^n \det(-xI_n - A) = (-1)^n p_A(-x)$ by multilinearity. Taking determinants of the factorization (Proposition 7.3.2(c)) gives $p_{A^2}\left(x^2\right) = p_A(x) \cdot (-1)^n p_A(-x)$.

(b) Substituting $-x$ into the factorization gives $p_A(-x) = \prod_i(-x - \lambda_i) = (-1)^n \prod_i(x + \lambda_i)$. So by (a),

$$p_{A^2}(x^2) = \prod_i (x-\lambda_i)(x+\lambda_i) = \prod_i \left(x^2-\lambda_i^2\right) = q(x^2), \quad q = \prod_i \left(x-\lambda_i^2\right).$$

For any $r = \sum_k r_k x^k$, the coefficient of x^{2k} in $r(x^2) = \sum_k r_k x^{2k}$ is r_k. So $r(x^2) = s(x^2)$ implies $r = s$, and $p_{A^2} = q$.

(c) Let $p_A = \prod(x-\lambda_i)$ with $\lambda_i \in \mathbb{R}$. By (b) and Remark 7.3.19 applied to A^2, $\operatorname{tr}(A^2) = \lambda_1^2 + \cdots + \lambda_n^2$. If this is 0, then every $\lambda_i = 0$, since a sum of squares of reals vanishes only when each term does. So $p_A = x^n$.

Solution to Exercise 7.3.50

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Proposition 7.3.26 (Krylov subspaces and the characteristic polynomial), Theorem 4.5.25 (Similar matrices are one operator in two bases), Proposition 4.5.24 (Similarity is an equivalence relation), Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Exercise 7.3.17
Hint 1. For (a), if $K_T(\mathbf{v}) = V$, what is d in Proposition 7.2.35? Then use Proposition 7.3.26.
Hint 2. For (b), both matrices are similar to the same companion matrix.
Hint 3. For (c), C_p has the cyclic vector $\mathbf{e}_1$. For C_p^{T}, try $\mathbf{e}_n$ and show that $\left(C_p^{\mathsf{T}}\right)^j \mathbf{e}_n$ is $\mathbf{e}_{n-j}$ plus a combination of $\mathbf{e}_{n-j+1}, \ldots, \mathbf{e}_n$.
Sketch. (a) $d = n$, and in the basis $\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ the matrix is C_p with $p = p_T$ (Proposition 7.3.26). (b) Both are similar to C_f for $f = p_A = p_B$; the converse is Proposition 7.3.11. (c) $\dim K(\mathbf{v}) \le 1$ for I_2. C_p is cyclic with $\mathbf{e}_1$, C_p^{T} with $\mathbf{e}_n$ (triangular argument), and $p_{C_p} = p_{C_p^{\mathsf{T}}}$.

(a) Let $\mathbf{v}$ be cyclic, so $\mathbf{v} \neq \mathbf{0}$ (as $V \neq \{\mathbf{0}\}$). In Proposition 7.2.35, $d = \dim K_T(\mathbf{v}) = n$, so $\mathbf{v}, T\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is a basis of V, and by Proposition 7.3.26 the matrix of T in it is C_p, with $p_T = p$. So the matrix is C_{p_T}.

(b) If $A \sim B$, then $p_A = p_B$ by Proposition 7.3.11. Conversely, let $f = p_A = p_B$. By (a), T_A has matrix C_f in some basis, so $A \sim C_f$ by Theorem 4.5.25, and likewise $B \sim C_f$. By Proposition 4.5.24, $A \sim B$.

(c) For every $\mathbf{v} \in F^2$, $K_{I_2}(\mathbf{v}) = \operatorname{span}(\mathbf{v})$ has dimension at most 1, so I_2 has no cyclic vector. (Indeed I_2 and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ have the same characteristic polynomial and are not similar, so (b) needs the hypothesis.)

Now let $C = C_p$. Since $C^j \mathbf{e}_1 = \mathbf{e}_{j+1}$ for $j < n$, $\mathbf{e}_1$ is cyclic for C. For C^{T}: column j of C^{T} is row j of C, which is $-a_0 \mathbf{e}_n^{\mathsf{T}}$ for $j = 1$ and $\mathbf{e}_{j-1}^{\mathsf{T}} - a_{j-1}\mathbf{e}_n^{\mathsf{T}}$ for $j \ge 2$. So $C^{\mathsf{T}}\mathbf{e}_j = \mathbf{e}_{j-1} - a_{j-1}\mathbf{e}_n$ for $j \ge 2$. By induction on j, $\left(C^{\mathsf{T}}\right)^j \mathbf{e}_n = \mathbf{e}_{n-j} + \mathbf{r}_j$ with $\mathbf{r}_j \in \operatorname{span}(\mathbf{e}_{n-j+1}, \ldots, \mathbf{e}_n)$ for $0 \le j \le n-1$ (with $\mathbf{r}_0 = \mathbf{0}$): applying C^{T} lowers each index by one and adds a multiple of $\mathbf{e}_n$. So $\left(C^{\mathsf{T}}\right)^j \mathbf{e}_n$ has coordinate 1 at $\mathbf{e}_{n-j}$ and coordinate 0 at $\mathbf{e}_1, \ldots, \mathbf{e}_{n-j-1}$. Hence the matrix with columns $\left(C^{\mathsf{T}}\right)^{n-1}\mathbf{e}_n, \ldots, C^{\mathsf{T}}\mathbf{e}_n, \mathbf{e}_n$ is lower triangular with 1's on the diagonal, so these vectors are independent (Proposition 5.3.15, Theorem 5.4.7), and $\mathbf{e}_n$ is cyclic for C^{T}. Since $p_{C^{\mathsf{T}}} = p_C$ (Exercise 7.3.17), (b) shows $C \sim C^{\mathsf{T}}$.

Solution to Exercise 7.3.51

Uses. Proposition 7.3.26 (Krylov subspaces and the characteristic polynomial), Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 4.1.31 (Every square matrix satisfies a polynomial)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. It is enough to show $p_T(T)\mathbf{v} = \mathbf{0}$ for each $\mathbf{v}$. Fix $\mathbf{v} \neq \mathbf{0}$ and its Krylov polynomial p.
Hint 2. p divides p_T, so $p_T = gp$. Evaluate at T and apply to $\mathbf{v}$.
Sketch. For $\mathbf{v} \neq \mathbf{0}$ with Krylov polynomial p, $p_T = gp$ by Proposition 7.3.26(c), so $p_T(T)\mathbf{v} = g(T)p(T)\mathbf{v} = \mathbf{0}$. For matrices, apply this to T_A and use $[q(T)] = q([T])$.

(a) Let $\mathbf{v} \in V$. If $\mathbf{v} = \mathbf{0}$, then $p_T(T)\mathbf{v} = \mathbf{0}$. Otherwise let p be the polynomial of Proposition 7.2.35 for $\mathbf{v}$, so $p(T)\mathbf{v} = \mathbf{0}$. By Proposition 7.3.26(c), $p_T = gp$ for some $g \in F[x]$, and by Proposition 7.2.27(a),

$$p_T(T)\mathbf{v} = g(T)(p(T)\mathbf{v}) = g(T)\mathbf{0} = \mathbf{0}.$$

So $p_T(T)$ sends every vector to $\mathbf{0}$, that is, $p_T(T) = 0$.

(b) Apply (a) to T_A on F^n. Its matrix in the standard basis is A, so $p_{T_A} = p_A$, and by Proposition 7.2.27(c), $p_A(A)$ is the matrix of $p_A(T_A) = 0$. So $p_A(A) = 0$, and p_A is monic of degree n.

Solution to Exercise 7.3.52

Uses. Example 7.3.5 (Small cases), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Proposition 7.1.22 (Counting roots)
Techniques. Double counting (Technique 4.1.36)
Hint 1. Write $A = \left(\begin{smallmatrix} a & b \\ c & e \end{smallmatrix}\right)$. The conditions are $a+e=t$ and $ae-bc=d$. Choose a first.
Hint 2. Then $e = t-a$ and $bc = a(t-a)-d = -f(a)$. How many pairs (b,c) have a given product m? Distinguish $m=0$ and $m \neq 0$.
Hint 3. If f has r roots in $\mathbb{F}_q$, the total is $r(2q-1)+(q-r)(q-1) = q(q-1)+rq$.
Sketch. For each a, $e = t-a$ is forced and $bc = -f(a)$; there are $2q-1$ pairs (b,c) if $f(a)=0$ and $q-1$ otherwise. So the count is $q(q-1)+rq$, where $r \in \{0,1,2\}$ is the number of roots. There are $\left(q^2-q\right)/2$ monic quadratics without roots, each realized by q^2-q matrices.

(a) Write $A = \left(\begin{smallmatrix} a & b \\ c & e \end{smallmatrix}\right)$. By Example 7.3.5(a), $p_A = f$ if and only if $a+e=t$ and $ae-bc=d$. Choose $a \in \mathbb{F}_q$ arbitrarily (q ways). Then $e = t-a$ is determined, and the second condition becomes

$$bc = a(t-a)-d = -\left(a^2-ta+d\right) = -f(a).$$

The number of pairs (b,c) with $bc=m$ is $q-1$ if $m \neq 0$ (any $b \neq 0$, and then $c = m/b$), and $2q-1$ if $m=0$ ($b=0$ with c arbitrary, or $b \neq 0$ with $c=0$). So if f has r roots in $\mathbb{F}_q$, the number of matrices is

$$r(2q-1)+(q-r)(q-1) = q(q-1)+rq,$$

which is q^2+q, q^2, q^2-q for $r=2,1,0$.

(b) By Theorem 7.3.6, A has no eigenvalue exactly when p_A has no root. There are q^2 monic quadratics x^2-tx+d. Those with two distinct roots are the products $(x-\alpha)(x-\beta)$ with $\alpha \neq \beta$, and there are $\binom{q}{2} = \left(q^2-q\right)/2$ of them; those with a repeated root are the $(x-\alpha)^2$, and there are q of them (a monic quadratic is determined by its roots, by Proposition 7.1.22 and the factor theorem). So $q^2 - \left(q^2-q\right)/2 - q = \left(q^2-q\right)/2$ have no root. Each is realized by q^2-q matrices, which gives $\left(q^2-q\right)^2/2$ matrices without eigenvalues. The total is

$$\frac{q^2-q}{2}\left(q^2+q\right) + q\cdot q^2 + \frac{q^2-q}{2}\left(q^2-q\right) = \frac{q^2-q}{2}\cdot 2q^2 + q^3 = q^4 - q^3 + q^3 = q^4,$$

as it should be. For $q=2$ the two matrices without eigenvalues are $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Solution to Exercise 7.3.53

Uses. Proposition 7.3.2 (Determinants over a commutative ring), Proposition 7.3.18 (The first and last coefficients)
Techniques. Expand by multilinearity (Technique 5.1.7)
Hint 1. Column j of $xI_n - A$ is $x\mathbf{e}_j - \mathbf{a}_j$, a sum of two vectors. Expand det by multilinearity in all columns at once.
Hint 2. You get 2^n terms indexed by the set S of columns where $-\mathbf{a}_j$ was chosen. In the term for S, the columns $x\mathbf{e}_j$ with $j \notin S$ have a single nonzero entry. Expand along them.
Sketch. Multilinearity gives $p_A = \sum_S \det M_S$, where column j of M_S is $-\mathbf{a}_j$ for $j \in S$ and $x\mathbf{e}_j$ otherwise. Expanding along the columns $x\mathbf{e}_j$, $j \notin S$, each contributes a factor x and deletes row and column j, leaving $-A_S$. So $\det M_S = x^{n-|S|}(-1)^{|S|}\det A_S$.

Column j of $xI_n - A$ is $x\mathbf{e}_j - \mathbf{a}_j$, where $\mathbf{a}_j$ is column j of A. Expanding by linearity in each of the n columns (Proposition 7.3.2(b), Technique 5.1.7),

$$p_A = \sum_{S \subseteq \{1,\ldots,n\}} \det M_S,$$

where column j of M_S is $-\mathbf{a}_j$ for $j \in S$ and $x\mathbf{e}_j$ for $j \notin S$.

We claim that $\det M_S = x^{n-|S|}\det(-A_S)$ (where $\det\left(-A_\varnothing\right) = 1$), by induction on $n-|S|$. If $S = \{1,\ldots,n\}$, then $M_S = -A$. If $n=1$ and $S=\varnothing$, then $M_S = (x)$ and the claim holds. Otherwise $n \geq 2$; pick $j \notin S$. Column j of M_S is $x\mathbf{e}_j$, with the single nonzero entry x in position (j,j), so cofactor expansion along it (Proposition 7.3.2(f)) gives $\det M_S = x(-1)^{j+j}\det M'$, where M' is M_S with row j and column j deleted. In M', each remaining column $x\mathbf{e}_{j'}$ still has its single entry x on the diagonal, and each column $-\mathbf{a}_{j'}$ has lost its j-th

entry. So M' is the matrix of the same form for the $(n-1)\times(n-1)$ matrix obtained from A by deleting row and column j, with the same set S, and the induction hypothesis applies. Since $\det(-A_S) = (-1)^{|S|}\det A_S$ by multilinearity,

$$p_A = \sum_S (-1)^{|S|}x^{n-|S|}\det A_S = \sum_{k=0}^{n}(-1)^k e_k(A)x^{n-k}.$$

For $k=1$, the principal 1×1 minors are the diagonal entries, so $e_1(A) = \operatorname{tr} A$; for $k=n$, $e_n(A) = \det A$. This matches Proposition 7.3.18. For $n=3$, the coefficient of x is

$$e_2(A) = (a_{11}a_{22}-a_{12}a_{21}) + (a_{11}a_{33}-a_{13}a_{31}) + (a_{22}a_{33}-a_{23}a_{32}).$$

Solution to Exercise 7.3.54

Uses. Proposition 7.3.30 (Odd dimension), Definition 7.2.1 (Invariant subspace), Proposition 7.3.15 (Block-triangular matrices), Proposition 7.2.4 (Invariance is checked on a spanning list)
Hint 1. If U is invariant of odd dimension, then $T{\restriction_U}$ is an operator on an odd-dimensional real space.
Hint 2. For (b), try two copies of the rotation $R_{\pi/2}$. Look for planes spanned by $\mathbf{w}$ and $T\mathbf{w}$.
Sketch. (a) An invariant U of odd dimension gives an eigenvector of $T{\restriction_U}$ (Proposition 7.3.30), which is an eigenvector of T. (b) $T = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix}$ with $R = R_{\pi/2}$ has $p_T = \left(x^2+1\right)^2$; for any $\mathbf{w}\neq\mathbf{0}$, $\operatorname{span}(\mathbf{w}, T\mathbf{w})$ is invariant since $T^2 = -\mathrm{id}$.

(a) Let U be T-invariant. If $\dim U = 0$ it is even. If $\dim U$ is odd, then $T{\restriction_U}\in \mathrm{L}(U)$ has an eigenvalue by Proposition 7.3.30: there are $\mathbf{u}\in U$, $\mathbf{u}\neq\mathbf{0}$, and $\lambda\in\mathbb{R}$ with $T\mathbf{u} = \lambda\mathbf{u}$. Then λ is an eigenvalue of T, a contradiction.

(b) Let $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix}$ on $\mathbb{R}^4$. By Proposition 7.3.15, $p_T = \left(x^2+1\right)^2$, which has no real root, so T has no eigenvalue. Since $R^2 = -I_2$, $T^2 = -I_4$. So for every $\mathbf{w}\neq\mathbf{0}$, $T(T\mathbf{w}) = -\mathbf{w}$, and $\operatorname{span}(\mathbf{w}, T\mathbf{w})$ is T-invariant (Proposition 7.2.4); it is a plane by (a). Three different ones: $\mathbf{w} = \mathbf{e}_1$ gives $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$; $\mathbf{w} = \mathbf{e}_3$ gives $\operatorname{span}(\mathbf{e}_3, \mathbf{e}_4)$; and $\mathbf{w} = \mathbf{e}_1 + \mathbf{e}_3$ gives $\operatorname{span}(\mathbf{e}_1+\mathbf{e}_3, \mathbf{e}_2+\mathbf{e}_4)$, which contains neither $\mathbf{e}_1$ nor $\mathbf{e}_3$.

Solution to Exercise 7.3.55

Uses. Corollary 7.2.22 (At most $\dim V$ eigenvalues), Exercise 7.3.48, Definition 7.2.6 (Eigenvalue, eigenvector)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. For (a), induct on k: write $A^{k+1}B - BA^{k+1} = A\left(A^kB - BA^k\right) + (AB-BA)A^k$.
Hint 2. For (b), consider the linear map $L(X) = XB - BX$ on $\mathbb{C}^{n\times n}$. What does (a) say about $L\left(A^k\right)$?
Hint 3. An operator on an n^2-dimensional space has at most n^2 eigenvalues, but $1, 2, 3, \ldots$ are infinitely many.
Sketch. (a) $A^{k+1}B - BA^{k+1} = A\left(A^kB - BA^k\right) + (AB-BA)A^k = kA^{k+1} + A^{k+1}$. (b) With $L(X) = XB - BX$, $L\left(A^k\right) = kA^k$, so if $A^k\neq 0$ then k is an eigenvalue of L. L has at most n^2 eigenvalues, so $A^k = 0$ for some $k\leq n^2+1$. (c) Exercise 7.3.48.

(a) For $k=1$ this is the hypothesis. If it holds for k, then

$$A^{k+1}B - BA^{k+1} = A\left(A^kB - BA^k\right) + (AB-BA)A^k = A\left(kA^k\right) + A\cdot A^k = (k+1)A^{k+1},$$

where the first equality holds because the terms $-ABA^k$ and $+ABA^k$ cancel.

(b) Let $L:\mathbb{C}^{n\times n}\to\mathbb{C}^{n\times n}$, $L(X) = XB - BX$. It is linear, and (a) says $L\left(A^k\right) = kA^k$. If $A^k\neq 0$, then k is an eigenvalue of L (Definition 7.2.6). By Corollary 7.2.22, L has at most n^2 eigenvalues, and $1, 2, \ldots, n^2+1$ are n^2+1 distinct complex numbers. So $A^k = 0$ for some $k\leq n^2+1$.

(c) By (b) and Exercise 7.3.48, $p_A = x^n$. The proof used that the integers $1, 2, 3, \ldots$ are distinct in $\mathbb{C}$; over $\mathbb{F}_p$ the scalars k repeat, and (b) breaks down.

Section 7.4

Solution to Exercise 7.4.10

Uses. Corollary 7.4.5 (Distinct eigenvalues give a diagonalization), Proposition 7.4.7 (One eigenvalue), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 5.3.15 (Triangular matrices)
Hint 1. Find the eigenvalues from $\det(A-\lambda I) = 0$. Triangular matrices show them on the diagonal.

Hint 2. Distinct eigenvalues settle the question at once. A matrix with a single eigenvalue is diagonalizable only if it is a scalar matrix.
Sketch. A_1 has distinct eigenvalues $1, 2$: diagonalizable over both fields. A_2 has only the eigenvalue 2 and is not $2I_2$: diagonalizable over neither. A_3 has $\det(A_3 - \lambda I) = \lambda^2 + 1$: no real eigenvalues, so not over $\mathbb{R}$; distinct eigenvalues $\pm i$ over $\mathbb{C}$. $A_4 = 7I_3$ is already diagonal.

The matrices A_1 and A_2 are upper triangular, so $A - \lambda I$ is triangular with diagonal entries $a_{ii} - \lambda$, and by Proposition 5.3.15 and Theorem 7.2.9 the eigenvalues are the diagonal entries. For A_1 they are 1 and 2, which are distinct, so A_1 is diagonalizable over $\mathbb{R}$, and over $\mathbb{C}$, by Corollary 7.4.5. For A_2 the only eigenvalue is 2, over either field, and $A_2 \neq 2I_2$, so A_2 is diagonalizable over neither field (Proposition 7.4.7).

For A_3, $\det(A_3 - \lambda I_2) = (1 - \lambda)(-1 - \lambda) + 2 = \lambda^2 + 1$. Over $\mathbb{R}$ it never vanishes, so A_3 has no eigenvalues and no eigenvectors, and is not diagonalizable. Over $\mathbb{C}$ it has the two distinct eigenvalues i and $-i$, so A_3 is diagonalizable over $\mathbb{C}$.

Finally $A_4 = 7I_3$ is diagonal, so it is diagonalizable over both fields (take $S = I_3$), even though it has only one eigenvalue.

Solution to Exercise 7.4.21

Uses. Definition 7.4.11 (Algebraic multiplicity), Lemma 7.4.13 (Geometric multiplicity through a rank), Theorem 7.4.16 (Criteria for diagonalizability), Proposition 5.3.15 (Triangular matrices)
Hint 1. A is upper triangular, so p_A can be read off the diagonal.
Hint 2. For $\mathrm{gm}(2)$, compute the rank of $A - 2I_4$.
Sketch. $p_A = (x-2)^3(x-3)$, so $\mathrm{am}(2) = 3$ and $\mathrm{am}(3) = 1$. $A - 2I_4$ has rank 2 (columns $\mathbf{e}_1$ and $\mathbf{e}_4$), so $\mathrm{gm}(2) = 2 < 3$: not diagonalizable. Changing $a_{12} = 1$ to 0 makes A diagonal. Any other change above the diagonal keeps A triangular, so $\mathrm{am}(2) = 3$, and keeps columns 2 and 4 of $A - 2I_4$ independent.

The matrix $xI_4 - A$ is upper triangular with diagonal $x-2, x-2, x-2, x-3$, so $p_A = (x-2)^3(x-3)$ by Proposition 5.3.15. Hence the eigenvalues are 2 and 3, with $\mathrm{am}(2) = 3$ and $\mathrm{am}(3) = 1$. Next,

$$A - 2I_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

has rank 2, so $\mathrm{gm}(2) = 4 - 2 = 2$ by Lemma 7.4.13, and $1 \leq \mathrm{gm}(3) \leq \mathrm{am}(3) = 1$ gives $\mathrm{gm}(3) = 1$. Since $\mathrm{gm}(2) = 2 \neq 3 = \mathrm{am}(2)$, A is not diagonalizable (Theorem 7.4.16); equivalently, $\mathrm{gm}(2) + \mathrm{gm}(3) = 3 < 4$.

The only nonzero off-diagonal entry is $a_{12} = 1$, and replacing it by 0 gives a diagonal matrix. Changing any other entry above the diagonal keeps A upper triangular with the same diagonal, so still $p_A = (x-2)^3(x-3)$ and $\mathrm{am}(2) = 3$. It also keeps $a_{12} = 1$, and column 2 of $A - 2I_4$ is still $\mathbf{e}_1$ (the only entry above the diagonal in column 2 is a_{12}), while column 4 still has last entry 1. These two columns are independent, so the rank stays at least 2, $\mathrm{gm}(2) \leq 2 < 3$, and the matrix stays non-diagonalizable.

Below the diagonal the answer is different, because the characteristic polynomial can change. Setting $a_{21} = 1$ makes the top left block $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, with eigenvalues 1 and 3 (Example 7.2.8); the new matrix has the eigenvalues $1, 3, 2, 3$ and is diagonalizable, as one checks with Theorem 7.4.16.

Solution to Exercise 7.4.27

Uses. Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix), Proposition 2.4.23 (Transpose of a product), Definition 7.4.1 (Diagonalizable), Theorem 7.4.16 (Criteria for diagonalizability)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. Write $A = SDS^{-1}$ and apply Proposition 7.4.23.
Hint 2. For the transpose, transpose the factorization; the inverse of S^{T} is $\left(S^{-1}\right)^{\mathsf{T}}$.
Sketch. $p(A) = Sp(D)S^{-1}$ with $p(D)$ diagonal; $A^{-1} = SD^{-1}S^{-1}$; $A^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}} DS^{\mathsf{T}} = RDR^{-1}$ with $R = \left(S^{-1}\right)^{\mathsf{T}}$. The eigenvalues of $p(A)$ are the diagonal entries $p(\lambda_i)$ of $p(D)$.

Write $A = SDS^{-1}$ with $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. By Proposition 7.4.23, $p(A) = S\operatorname{diag}(p(\lambda_1), \ldots, p(\lambda_n))S^{-1}$, which is a diagonalization of $p(A)$. If A is invertible, every λ_i is nonzero (otherwise D, and with it $A = SDS^{-1}$, would not be invertible), and Proposition 7.4.23 gives $A^{-1} = S\operatorname{diag}\left(\lambda_1^{-1}, \ldots, \lambda_n^{-1}\right)S^{-1}$. For the transpose, Proposition 2.4.23 gives

$$A^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}} D^{\mathsf{T}} S^{\mathsf{T}} = RDR^{-1}, \quad R = \left(S^{-1}\right)^{\mathsf{T}},$$

since $D^{\mathsf{T}} = D$ and $RS^{\mathsf{T}} = \left(SS^{-1}\right)^{\mathsf{T}} = I_n$, and similarly $S^{\mathsf{T}}R = I_n$.

By the last part of Theorem 7.4.16, the eigenvalues of a diagonalizable matrix are its diagonal entries in any diagonalization. So the eigenvalues of $p(A)$ are $p(\lambda_1), \ldots, p(\lambda_n)$, that is, the values $p(\lambda)$ for the eigenvalues λ of A. Distinct eigenvalues of A can have the same image: for $p = x^2$ and $A = \operatorname{diag}(1, -1)$, $p(A) = I_2$.

Solution to Exercise 7.4.35

Uses. Algorithm 7.4.18 (Diagonalizing a matrix), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 7.4.7 (One eigenvalue)
Hint 1. Follow Algorithm 7.4.18: find the eigenvalues from $\det(A - \lambda I) = 0$, then a basis of each eigenspace.
Hint 2. In (b) there is only one eigenvalue. In (c), A swaps $\mathbf{e}_1$ and $\mathbf{e}_3$ and fixes $\mathbf{e}_2$; guess the eigenvectors.
Sketch. (a) $(\lambda - 4)(\lambda + 1)$; $E_4 = \operatorname{span}((2, 3))$, $E_{-1} = \operatorname{span}((1, -1))$; $S = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$, $D = \operatorname{diag}(4, -1)$. (b) $(\lambda - 2)^2$, and $A \neq 2I$: not diagonalizable. (c) $E_1 = \operatorname{span}(\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_3)$, $E_{-1} = \operatorname{span}(\mathbf{e}_1 - \mathbf{e}_3)$.

(a) $\det(A - \lambda I_2) = (1 - \lambda)(2 - \lambda) - 6 = \lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1)$. Then $A - 4I_2 = \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix}$ has kernel $\operatorname{span}((2, 3))$, and $A + I_2 = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$ has kernel $\operatorname{span}((1, -1))$. The eigenvalues are distinct, so

$$S = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}.$$

Check: $AS = \begin{pmatrix} 8 & -1 \\ 12 & 1 \end{pmatrix} = SD$.

(b) $\det(A - \lambda I_2) = (3 - \lambda)(1 - \lambda) + 1 = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, so 2 is the only eigenvalue. Since $A \neq 2I_2$, A is not diagonalizable (Proposition 7.4.7). Indeed $A - 2I_2 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ has rank 1, so $E_2 = \operatorname{span}((1, -1))$ is a line.

(c) $A(x, y, z) = (z, y, x)$. The vectors $\mathbf{e}_2$ and $\mathbf{e}_1 + \mathbf{e}_3$ are fixed, and $A(\mathbf{e}_1 - \mathbf{e}_3) = \mathbf{e}_3 - \mathbf{e}_1$. The matrix S below, whose columns are these three vectors, has determinant 2 (expand along its second row), so they form an eigenbasis and

$$S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

To see that nothing was missed, $\det(A - \lambda I_3) = (1 - \lambda)\left(\lambda^2 - 1\right) = -(\lambda - 1)^2(\lambda + 1)$ by cofactor expansion along the middle row, and $\operatorname{gm}(1) = 2 = \operatorname{am}(1)$.

Solution to Exercise 7.4.36

Uses. Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22), Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Diagonalize: the eigenvalues are 1 and 2.
Hint 2. For (b), any B with $B^2 = A$ commutes with A. What does that say about the eigenlines of A?
Hint 3. B preserves each eigenline of A, so it is diagonal in the eigenbasis: $S^{-1}BS = \operatorname{diag}(\beta_1, \beta_2)$ with $\beta_1^2 = 1$, $\beta_2^2 = 2$.
Sketch. $A = S\operatorname{diag}(1, 2)S^{-1}$ with $S = \begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}$, $S^{-1} = \begin{pmatrix} -1 & 2 \\ 3 & -5 \end{pmatrix}$, so $A^k = \begin{pmatrix} 6 \cdot 2^k - 5 & 10 - 10 \cdot 2^k \\ 3 \cdot 2^k - 3 & 6 - 5 \cdot 2^k \end{pmatrix}$. A square root B commutes with A, so it preserves the eigenlines E_1, E_2 and is diagonal in the eigenbasis with entries $\pm 1, \pm\sqrt{2}$: exactly four square roots.

(a) $\det(A-\lambda I_2) = (7-\lambda)(-4-\lambda)+30 = \lambda^2-3\lambda+2 = (\lambda-1)(\lambda-2)$. Then $A - I_2 = \begin{pmatrix} 6 & -10 \\ 3 & -5 \end{pmatrix}$ has kernel $\operatorname{span}((5,3))$, and $A - 2I_2 = \begin{pmatrix} 5 & -10 \\ 3 & -6 \end{pmatrix}$ has kernel $\operatorname{span}((2,1))$. So $A = SDS^{-1}$ with

$$S = \begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} -1 & 2 \\ 3 & -5 \end{pmatrix},$$

and both eigenvalues are nonzero, so by Proposition 7.4.23, for every $k \in \mathbb{Z}$,

$$A^k = \begin{pmatrix} 5 & 2\cdot 2^k \\ 3 & 2^k \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 3 & -5 \end{pmatrix} = \begin{pmatrix} 6\cdot 2^k - 5 & 10 - 10\cdot 2^k \\ 3\cdot 2^k - 3 & 6 - 5\cdot 2^k \end{pmatrix}.$$

At $k = 1$ this is A, and at $k = -1$ it is $\begin{pmatrix} -2 & 5 \\ -3/2 & 7/2 \end{pmatrix}$, which one checks is A^{-1}.

(b) Let $\mathbf{s}_1 = (5,3)$ and $\mathbf{s}_2 = (2,1)$. If $B^2 = A$, then $BA = B^3 = AB$. By Proposition 7.2.30(b), each eigenspace $E_1(A) = \operatorname{span}(\mathbf{s}_1)$ and $E_2(A) = \operatorname{span}(\mathbf{s}_2)$ is B-invariant, so by Proposition 7.2.7, $B\mathbf{s}_1 = \beta_1\mathbf{s}_1$ and $B\mathbf{s}_2 = \beta_2\mathbf{s}_2$ for some $\beta_1, \beta_2 \in \mathbb{R}$. Then $A\mathbf{s}_j = B^2\mathbf{s}_j = \beta_j^2\mathbf{s}_j$, so $\beta_1^2 = 1$ and $\beta_2^2 = 2$. By Theorem 7.4.2, $B = S\operatorname{diag}(\beta_1,\beta_2)S^{-1}$ with $\beta_1 = \pm 1$ and $\beta_2 = \pm\sqrt{2}$. Conversely each of these four matrices squares to $S\operatorname{diag}(1,2)S^{-1} = A$. So A has exactly four real square roots; for $\beta_1 = 1$, $\beta_2 = \sqrt{2}$,

$$B = \begin{pmatrix} 5 & 2\sqrt{2} \\ 3 & \sqrt{2} \end{pmatrix}\begin{pmatrix} -1 & 2 \\ 3 & -5 \end{pmatrix} = \begin{pmatrix} 6\sqrt{2} - 5 & 10 - 10\sqrt{2} \\ 3\sqrt{2} - 3 & 6 - 5\sqrt{2} \end{pmatrix},$$

and the four square roots are $S\operatorname{diag}(\beta_1,\beta_2)S^{-1}$ with $(\beta_1,\beta_2) \in \{\pm 1\} \times \{\pm\sqrt{2}\}$, namely B, $-B$, and the matrices obtained from B and $-B$ by replacing $\sqrt{2}$ with $-\sqrt{2}$.

Solution to Exercise 7.4.37

Uses. Theorem 7.4.16 (Criteria for diagonalizability), Lemma 7.4.13 (Geometric multiplicity through a rank), Theorem 7.4.15 (Geometric multiplicity is at most algebraic multiplicity), Proposition 5.3.15 (Triangular matrices)
Hint 1. The characteristic polynomial does not depend on a, b, c. Which eigenvalue could cause trouble?
Hint 2. Compute $\operatorname{gm}(1) = 3 - \operatorname{rank}(A - I_3)$ as a function of a, b, c.
Sketch. $p_A = (x-1)^2(x-2)$, so A is diagonalizable iff $\operatorname{gm}(1) = 2$, iff $\operatorname{rank}(A - I_3) = 1$. The rows of $A - I_3$ are $(0,a,b)$, $(0,0,c)$, $(0,0,1)$; the last is nonzero and the first is independent of it iff $a \neq 0$. So A is diagonalizable iff $a = 0$, for all b, c.

A is upper triangular, so $p_A = (x-1)^2(x-2)$ (Proposition 5.3.15), which splits, with $\operatorname{am}(1) = 2$ and $\operatorname{am}(2) = 1$. Since $1 \leq \operatorname{gm}(2) \leq \operatorname{am}(2) = 1$ (Theorem 7.4.15), $\operatorname{gm}(2) = 1$ always. By Theorem 7.4.16, A is diagonalizable if and only if $\operatorname{gm}(1) = 2$, that is, $\operatorname{rank}(A - I_3) = 1$ (Lemma 7.4.13). The rows of

$$A - I_3 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 1 \end{pmatrix}$$

are $(0,a,b)$, $(0,0,c)$, $(0,0,1)$. The second is a multiple of the third, so the rank is the rank of $(0,a,b)$, $(0,0,1)$, which is 2 if $a \neq 0$ and 1 if $a = 0$. So A is diagonalizable if and only if $a = 0$, whatever b and c are. The entries b and c couple the eigenvalue 1 to the eigenvalue 2, which does no harm because the two eigenvalues are different; the entry a couples the eigenvalue 1 to itself, as in a shear.

Solution to Exercise 7.4.38

Uses. Example 7.4.8 (Shears and nilpotent matrices), Proposition 7.4.7 (One eigenvalue), Proposition 4.5.24 (Similarity is an equivalence relation), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. The shear and the nilpotent matrix $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ are the usual suspects.
Hint 2. For (d), look for two diagonalizable matrices whose sum is N. For (e), similarity is transitive.
Sketch. (a) False: 0. (b) False: the shear. (c) False: $N^2 = 0$ is diagonal, N is not diagonalizable. (d) False: $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = N$. (e) True: $B \sim A \sim D$.

The counterexamples below are 2×2. For $n > 2$, enlarge each to a block diagonal matrix by adding the block I_{n-2} in (b) and a zero block in (c) and (d). A block diagonal matrix with diagonalizable blocks is diagonalizable (combine eigenbases of the blocks), the enlarged shear still has 1 as its only eigenvalue without being I_n, and the enlarged nilpotent matrices are still nilpotent and nonzero.

(a) False: the zero matrix is diagonal, so diagonalizable, but not invertible.

(b) False: the shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is invertible but not diagonalizable (Example 7.4.8).

(c) False: $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has $N^2 = 0$, which is diagonal, while N is nonzero and nilpotent, so not diagonalizable (Example 7.4.8).

(d) False: $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ has the distinct eigenvalues 1 and 0 (it is triangular), so it is diagonalizable (Corollary 7.4.5), and $\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$ is diagonal; their sum is N.

(e) True: if A is similar to a diagonal D and B is similar to A, then B is similar to D by transitivity (Proposition 4.5.24).

Solution to Exercise 7.4.39

> *Uses.* Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Example 7.4.8 (Shears and nilpotent matrices)
> *Hint 1.* The λ_i in the factorization need not be distinct. What does Theorem 7.2.21 assume?
> *Hint 2.* Run the argument on the shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
> *Sketch.* The roots λ_i are listed with multiplicity, so they can repeat, and Theorem 7.2.21 only applies to distinct eigenvalues. For the shear, $\lambda_1 = \lambda_2 = 1$, and any two eigenvectors are multiples of $\mathbf{e}_1$. The argument proves only that matrices with n distinct eigenvalues are diagonalizable.

The factorization lists the roots with multiplicity, so the λ_i need not be distinct, while Theorem 7.2.21 applies only to eigenvectors for pairwise distinct eigenvalues. When $\lambda_i = \lambda_j$ for some $i \neq j$, nothing prevents $\mathbf{v}_j$ from being a multiple of $\mathbf{v}_i$, and then the list is dependent. For the shear $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $p_J = (x-1)^2$, so $\lambda_1 = \lambda_2 = 1$, and every eigenvector is a multiple of $\mathbf{e}_1$; no choice of $\mathbf{v}_1, \mathbf{v}_2$ is independent, and J is not diagonalizable (Example 7.4.8). The argument is correct when the λ_i are distinct, and then it is the proof of Corollary 7.4.5.

Solution to Exercise 7.4.40

> *Uses.* Proposition 4.3.16 (Rank-one matrices), Example 4.3.18 (Powers of a rank-one matrix), Theorem 4.2.11 (Rank–nullity), Corollary 7.2.23 (Eigenspaces are independent), Theorem 7.4.16 (Criteria for diagonalizability), Example 7.4.8 (Shears and nilpotent matrices)
> *Techniques.* Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
> *Hint 1.* Write $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$. Compute $A\mathbf{u}$, and recall $\mathbf{v}^{\mathsf{T}}\mathbf{u} = \operatorname{tr} A$.
> *Hint 2.* $\operatorname{gm}(0) = \dim \ker A = n - 1$. If $\operatorname{tr} A \neq 0$, you have one more eigenvector. If $\operatorname{tr} A = 0$, compute A^2.
> *Sketch.* $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{v}^{\mathsf{T}}\mathbf{u} = \operatorname{tr} A$, so $A\mathbf{u} = (\operatorname{tr} A)\mathbf{u}$, and $\operatorname{gm}(0) = n - 1$. If $\operatorname{tr} A \neq 0$, then $\operatorname{gm}(0) + \operatorname{gm}(\operatorname{tr} A) \geq n$: diagonalizable, with $D = \operatorname{diag}(\operatorname{tr} A, 0, \ldots, 0)$. If $\operatorname{tr} A = 0$, $A^2 = (\operatorname{tr} A)A = 0$ and $A \neq 0$: nilpotent, not diagonalizable.

By Proposition 4.3.16, $A = \mathbf{u}\mathbf{v}^{\mathsf{T}}$ with $\mathbf{u}, \mathbf{v}$ nonzero, and $\mathbf{v}^{\mathsf{T}}\mathbf{u} = \operatorname{tr} A$ (Example 4.3.18). Then

$$A\mathbf{u} = \mathbf{u}\left(\mathbf{v}^{\mathsf{T}}\mathbf{u}\right) = (\operatorname{tr} A)\mathbf{u}.$$

By rank–nullity, $\operatorname{gm}_A(0) = \dim \ker A = n - 1$.

Suppose $\operatorname{tr} A \neq 0$. Then $\mathbf{u} \neq \mathbf{0}$ is an eigenvector with the nonzero eigenvalue $\operatorname{tr} A$, so $\operatorname{gm}(0) + \operatorname{gm}(\operatorname{tr} A) \geq (n-1) + 1 = n$. By Corollary 7.2.23 the sum of all geometric multiplicities is at most n, so it equals n, and A is diagonalizable by Theorem 7.4.16, (III) $\Rightarrow$ (I). An eigenbasis is $\mathbf{u}$ followed by a basis of $\ker A$, so A is similar to $\operatorname{diag}(\operatorname{tr} A, 0, \ldots, 0)$.

Suppose $\operatorname{tr} A = 0$. Then $A^2 = (\operatorname{tr} A)A = 0$ (Example 4.3.18), and $A \neq 0$, so A is a nonzero nilpotent matrix, which is not diagonalizable (Example 7.4.8).

Solution to Exercise 7.4.41

Uses. Example 7.2.17 (Reflections, and the field $\mathbb{F}_2$), Theorem 7.4.16 (Criteria for diagonalizability), Proposition 7.4.7 (One eigenvalue), Example 4.5.31 (The trace of transposition)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. $\tau^2 = \mathrm{id}$, so every eigenvalue satisfies $\lambda^2 = 1$ (Example 7.2.17). What are E_1 and E_{-1}?
Hint 2. E_1 is the space of symmetric matrices and E_{-1} the space of skew-symmetric ones. Count their dimensions when $1 + 1 \neq 0$; when $1 + 1 = 0$, there is only one eigenvalue.
Sketch. Eigenvalues satisfy $\lambda^2 = 1$. If $1 + 1 \neq 0$: E_1 (symmetric, dimension $n(n+1)/2$) and E_{-1} (skew-symmetric, with zero diagonal, dimension $n(n-1)/2$) have dimensions adding up to n^2. If $1 + 1 = 0$: $1 = -1$ is the only eigenvalue and $\tau \neq \mathrm{id}$ since $n \geq 2$.

Since $\tau^2 = \mathrm{id}$, every eigenvalue of τ satisfies $\lambda^2 = 1$, so $\lambda = \pm 1$ (Example 7.2.17). The eigenspace E_1 is the space of symmetric matrices, $X^\mathsf{T} = X$, and E_{-1} is the space of skew-symmetric matrices, $X^\mathsf{T} = -X$.

Suppose $1 + 1 \neq 0$. A symmetric matrix is determined by its entries x_{ij} with $i \leq j$, which are arbitrary, so $\dim E_1 = n(n+1)/2$; a basis is U_{ii} and $U_{ij} + U_{ji}$ for $i < j$. A skew-symmetric matrix has $x_{ii} = -x_{ii}$, so $(1+1)x_{ii} = 0$ and $x_{ii} = 0$; it is determined by its entries with $i < j$, which are arbitrary, so $\dim E_{-1} = n(n-1)/2$, with basis $U_{ij} - U_{ji}$ for $i < j$. Since $n(n+1)/2 + n(n-1)/2 = n^2 = \dim F^{n\times n}$, τ is diagonalizable by Theorem 7.4.16, with $\mathrm{gm}(1) = n(n+1)/2$ and $\mathrm{gm}(-1) = n(n-1)/2$. (This recovers $\operatorname{tr}\tau = n$ from Example 4.5.31.)

Suppose $1 + 1 = 0$. Then $-1 = 1$, so 1 is the only eigenvalue of τ. Since $n \geq 2$, $\tau(U_{12}) = U_{21} \neq U_{12}$, so $\tau \neq \mathrm{id}$, and τ is not diagonalizable by Proposition 7.4.7.

Solution to Exercise 7.4.42

Uses. Theorem 7.4.16 (Criteria for diagonalizability), Theorem 4.4.7 (Criteria for direct sums), Theorem 4.2.11 (Rank–nullity), Theorem 4.4.19 (Dimension of a sum)
Hint 1. Split an eigenbasis into the vectors with eigenvalue 0 and the others. Where do the others lie?
Hint 2. If $T\mathbf{b} = \lambda\mathbf{b}$ with $\lambda \neq 0$, then $\mathbf{b} = T\left(\lambda^{-1}\mathbf{b}\right) \in \operatorname{im} T$.
Hint 3. For (b), invertible operators have $\ker T = \{\mathbf{0}\}$ and $\operatorname{im} T = V$.
Sketch. In an eigenbasis, the vectors with eigenvalue 0 span $\ker T = E_0$ and the others lie in $\operatorname{im} T$, since $\mathbf{b} = T\left(\lambda^{-1}\mathbf{b}\right)$. So $\ker T + \operatorname{im} T = V$, and rank–nullity forces the sum to be direct. (b) The invertible shear; and $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, with $\ker N = \operatorname{im} N = \operatorname{span}(\mathbf{e}_1)$.

(a) Let $\mathcal{B}$ be an eigenbasis, and split it into the vectors with eigenvalue 0 and the vectors $\mathbf{b}$ with eigenvalue $\lambda \neq 0$. The former lie in $\ker T$. For the latter, $\mathbf{b} = \lambda^{-1}T\mathbf{b} = T\left(\lambda^{-1}\mathbf{b}\right) \in \operatorname{im} T$. So the subspace $\ker T + \operatorname{im} T$ contains the basis $\mathcal{B}$, and $\ker T + \operatorname{im} T = V$. By Theorem 4.4.19 and rank–nullity,

$$\dim(\ker T \cap \operatorname{im} T) = \dim \ker T + \dim \operatorname{im} T - \dim V = 0,$$

so $\ker T \cap \operatorname{im} T = \{\mathbf{0}\}$, and the sum is direct (Theorem 4.4.7).

(b) The shear $J = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is invertible, so $\ker T_J = \{\mathbf{0}\}$ and $\operatorname{im} T_J = F^2$, and $F^2 = \{\mathbf{0}\} \oplus F^2$; but J is not diagonalizable. For the second, $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ has $\ker N = \operatorname{span}(\mathbf{e}_1) = \operatorname{im} N$.

Solution to Exercise 7.4.43

Uses. Theorem 7.4.16 (Criteria for diagonalizability), Example 7.4.14 (The multiplicities of a diagonal matrix), Proposition 7.1.34 (Roots of a split polynomial), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Proposition 4.5.24 (Similarity is an equivalence relation), Lemma 7.4.13 (Geometric multiplicity through a rank), Example 7.4.8 (Shears and nilpotent matrices)
Techniques. Rule out similarity with an invariant (Technique 4.5.29)
Hint 1. Similar matrices have the same characteristic polynomial (Section 7.3). For the converse, what does p_A say about the diagonal matrix D in $A = SDS^{-1}$?
Hint 2. $p_A = p_D = \prod(x - d_i)$, and the d_i are determined by p_A up to order (Proposition 7.1.34).
Sketch. If $A = SDS^{-1}$ and $B = RER^{-1}$, then $p_A = p_D$ and $p_B = p_E$, and p_D determines the diagonal of D up to order. So $p_A = p_B$ gives $E = P^{-1}DP$ for a permutation matrix P, and $A \sim D \sim E \sim B$. (b) I_2 and the shear both have characteristic polynomial $(x-1)^2$.

(a) Similar matrices have the same characteristic polynomial (Section 7.3), which proves "only if". Conversely, let $A = SDS^{-1}$ and $B = RER^{-1}$ with D, E diagonal, and suppose $p_A = p_B$. Since similar matrices have equal characteristic polynomials, $p_D = p_A = p_B = p_E$. By Example 7.4.14, $p_D = \prod_i (x - d_{ii})$, and by Proposition 7.1.34 each λ occurs on the diagonal of D as many times as its multiplicity as a root of p_D; likewise for E. So the diagonal of E is a rearrangement of the diagonal of D. Rearrange the columns of S in the same way, obtaining an invertible S' whose j-th column is an eigenvector of A with eigenvalue e_{jj}. By Theorem 7.4.2, $A = S'ES'^{-1}$. Hence $A \sim E \sim B$, and $A \sim B$ by Proposition 4.5.24.

(b) I_2 is diagonal and the shear $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable, and $p_{I_2} = p_J = (x-1)^2$. They are not similar, since the only matrix similar to I_2 is I_2. (Alternatively, $\mathrm{gm}_{I_2}(1) = 2 \neq 1 = \mathrm{gm}_J(1)$, by Lemma 7.4.13.)

Solution to Exercise 7.4.44

> *Uses.* Example 7.4.20 (The Fibonacci matrix over several fields), Theorem 7.4.30 (Recurrences with distinct characteristic roots), Exercise 1.1.51
> *Techniques.* Diagonalize, solve, conjugate back (Technique 7.4.22)
> *Hint 1.* Over $\mathbb{F}_{11}$, $x^2 - x - 1$ has the distinct roots 8 and 4. By Theorem 7.4.30, $f_k = a8^k + b4^k$; find a, b from f_0, f_1.
> *Hint 2.* For (b), Fermat's little theorem (Exercise 1.1.51) gives $8^{10} = 4^{10} = 1$.
> *Sketch.* The roots of $x^2 - x - 1$ in $\mathbb{F}_{11}$ are 8 and 4. With $f_k = a8^k + b4^k$, $f_0 = 0$ gives $b = -a$, and $f_1 = 1$ gives $4a = 1$, so $a = 3$. Since $8^{10} = 4^{10} = 1$, $f_{k+10} = f_k$.

(a) In $\mathbb{F}_{11}$, $x^2 - x - 1 = (x-4)(x+3) = (x-4)(x-8)$ (Example 7.4.20), with distinct roots 8 and 4. By Theorem 7.4.30, applied with $k = 2$ and $c_0 = c_1 = 1$ over $F = \mathbb{F}_{11}$, there are $a, b \in \mathbb{F}_{11}$ with $f_k = a8^k + b4^k$ for all k. From $f_0 = 0$, $a + b = 0$, and from $f_1 = 1$, $8a + 4b = 4a = 1$, so $a = 4^{-1} = 3$ (as $4 \cdot 3 = 12 = 1$) and $b = -3$. So $f_k = 3\left(8^k - 4^k\right)$. Check: $f_2 = 3(64 - 16) = 3 \cdot 48 = 144 = 1$ and $f_3 = 3(512 - 64) = 3 \cdot 448 = 1344 = 2$ in $\mathbb{F}_{11}$.

(b) By Fermat's little theorem (Exercise 1.1.51), $8^{10} = 4^{10} = 1$ in $\mathbb{F}_{11}$. So $f_{k+10} = 3\left(8^k 8^{10} - 4^k 4^{10}\right) = 3\left(8^k - 4^k\right) = f_k$. Indeed, modulo 11 the sequence is $0, 1, 1, 2, 3, 5, 8, 2, 10, 1, 0, 1, \ldots$, with period exactly 10.

Solution to Exercise 7.4.45

> *Uses.* Theorem 7.4.30 (Recurrences with distinct characteristic roots)
> *Hint 1.* The characteristic polynomial is $x^3 - 2x^2 - x + 2$. Look for integer roots.
> *Hint 2.* With roots $2, 1, -1$, write $s_j = a2^j + b + c(-1)^j$ and solve for a, b, c from s_0, s_1, s_2.
> *Sketch.* $x^3 - 2x^2 - x + 2 = (x-2)(x-1)(x+1)$. Solving $a + b + c = 3$, $2a + b - c = 2$, $4a + b + c = 6$ gives $a = b = c = 1$, so $s_j = 2^j + 1 + (-1)^j$.

The characteristic polynomial is $q = x^3 - 2x^2 - x + 2 = x^2(x-2) - (x-2) = (x-2)(x-1)(x+1)$, with distinct roots $2, 1, -1$. By Theorem 7.4.30, $s_j = a2^j + b + c(-1)^j$ for unique $a, b, c \in \mathbb{R}$. The initial values give

$$a + b + c = 3, \quad 2a + b - c = 2, \quad 4a + b + c = 6.$$

Subtracting the first equation from the third gives $3a = 3$, so $a = 1$; then $b + c = 2$ and $b - c = 0$, so $b = c = 1$. Hence $s_j = 2^j + 1 + (-1)^j$. Check: $s_3 = 8 + 1 - 1 = 8 = 2 \cdot 6 + 2 - 2 \cdot 3$.

Solution to Exercise 7.4.46

> *Uses.* Proposition 7.4.33 (Long-run behavior), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
> *Techniques.* Diagonalize, solve, conjugate back (Technique 7.4.22)
> *Hint 1.* The matrix is $A = \begin{pmatrix} 0.7 & 0.4 \\ 0.2 & 0.9 \end{pmatrix}$. Its eigenvalues are 1.1 and 0.5.
> *Hint 2.* Expand (x_0, y_0) in the eigenbasis $(1, 1), (-2, 1)$ and use (7.4.3).
> *Hint 3.* The $(1, 1)$-component is $(x_0 + 2y_0)/3$. When is it zero?
> *Sketch.* $A(1, 1) = 1.1(1, 1)$ and $A(-2, 1) = 0.5(-2, 1)$. With $c_1 = (x_0 + 2y_0)/3$, $c_2 = (y_0 - x_0)/3$, $(x_k, y_k) = c_1 1.1^k (1, 1) + c_2 0.5^k (-2, 1)$. If $x_0, y_0 \geq 0$, not both 0, then $c_1 > 0$ and the first term dominates. The state tends to $\mathbf{0}$ iff $c_1 = 0$, iff (x_0, y_0) is a multiple of $(-2, 1)$.

(a) The system is $\mathbf{x}_{k+1} = A\mathbf{x}_k$ with $A = \begin{pmatrix} 0.7 & 0.4 \\ 0.2 & 0.9 \end{pmatrix}$. Here $\det(A - \lambda I_2) = \lambda^2 - 1.6\lambda + 0.55 = (\lambda - 1.1)(\lambda - 0.5)$, and $A(1, 1) = (1.1, 1.1)$, $A(-2, 1) = (-1, 0.5)$, so $(1, 1)$ and $(-2, 1)$ form an eigenbasis with eigenvalues 1.1 and 0.5. Solving $(x_0, y_0) = c_1(1, 1) + c_2(-2, 1)$ gives $c_2 = (y_0 - x_0)/3$ and $c_1 = (x_0 + 2y_0)/3$. By (7.4.3),

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \frac{x_0 + 2y_0}{3}(1.1)^k \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{y_0 - x_0}{3}(0.5)^k \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

(b) If $x_0, y_0 \geq 0$ are not both 0, then $c_1 = (x_0 + 2y_0)/3 > 0$. By Proposition 7.4.33(b) with $\mu = 1.1$, $1.1^{-k}(x_k, y_k) \to c_1(1,1)$, so $1.1^{-k}x_k \to c_1$ and $1.1^{-k}y_k \to c_1$. Hence $x_k/y_k \to c_1/c_1 = 1$ (for large k, $y_k > 0$), and $x_{k+1}/x_k = 1.1 \cdot \left(1.1^{-k-1}x_{k+1}\right)/\left(1.1^{-k}x_k\right) \to 1.1$, and likewise for y_k.

(c) If $c_1 \neq 0$, the x-coordinate is $c_1(1.1)^k - 2c_2(0.5)^k$, whose absolute value tends to ∞. If $c_1 = 0$, then $(x_k, y_k) = c_2(0.5)^k(-2, 1) \to (0, 0)$. So the state tends to $(0, 0)$ exactly when $x_0 + 2y_0 = 0$, that is, when (x_0, y_0) lies on the eigenline $\operatorname{span}((-2, 1))$ of the eigenvalue 0.5.

Solution to Exercise 7.4.47

Uses. Lemma 7.4.13 (Geometric multiplicity through a rank), Theorem 3.3.17 (Uniqueness of the reduced row echelon form), Theorem 7.4.16 (Criteria for diagonalizability), Example 3.5.3 (A rank that depends on the field)
Hint 1. By Theorem 7.4.16(III), it suffices to compare $\operatorname{gm}(\lambda)$ over $\mathbb{R}$ and over $\mathbb{C}$ for each real eigenvalue λ.
Hint 2. $\operatorname{gm}(\lambda) = n - \operatorname{rank}(A - \lambda I_n)$. Is the rank of a real matrix the same over $\mathbb{R}$ and over $\mathbb{C}$?
Sketch. For real λ, $A - \lambda I_n$ is real, and Gauss–Jordan elimination over $\mathbb{R}$ produces an RREF that is also the RREF over $\mathbb{C}$ (uniqueness), so the rank, and hence $\operatorname{gm}(\lambda)$, is the same over both fields. The complex eigenvalues are exactly the real ones, so $\sum \operatorname{gm}_{\mathbb{R}}(\lambda) = \sum \operatorname{gm}_{\mathbb{C}}(\lambda) = n$.

Let $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ be the distinct eigenvalues of A over $\mathbb{C}$; by hypothesis they are real, so they are also its eigenvalues over $\mathbb{R}$, and A has no other real eigenvalues. For each i, $B_i = A - \lambda_i I_n$ is a real matrix. Row reducing B_i over $\mathbb{R}$ produces a matrix in reduced row echelon form by a sequence of row operations with real scalars. The same operations are complex row operations, so this real matrix is also the RREF of B_i over $\mathbb{C}$ (Theorem 3.3.17). The rank is the number of pivots, so B_i has the same rank over $\mathbb{R}$ and over $\mathbb{C}$, and by Lemma 7.4.13, $\dim_{\mathbb{R}} \ker_{\mathbb{R}} B_i = \dim_{\mathbb{C}} \ker_{\mathbb{C}} B_i$.

Since A is diagonalizable over $\mathbb{C}$, $\sum_i \dim_{\mathbb{C}} \ker_{\mathbb{C}} B_i = n$ by Theorem 7.4.16. Hence $\sum_i \dim_{\mathbb{R}} \ker_{\mathbb{R}} B_i = n$, and A is diagonalizable over $\mathbb{R}$ by the same theorem. (Without the hypothesis on the eigenvalues, the rotation of Example 7.4.9 is a counterexample. Contrast Example 3.5.3, where the rank of one integer matrix depends on the field because the fields $\mathbb{Q}$ and $\mathbb{F}_p$ do different arithmetic; $\mathbb{R} \subseteq \mathbb{C}$ do the same arithmetic on real numbers.)

Solution to Exercise 7.4.48

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Exercise 7.2.55, Theorem 7.4.16 (Criteria for diagonalizability), Proposition 7.4.7 (One eigenvalue), Proposition 4.5.27 (Similarity invariants), Proposition 5.4.17 (Similar matrices have equal determinants)
Hint 1. Every eigenvalue satisfies $\lambda^2 = 1$. Show that A is diagonalizable: $\mathbf{v} = (\mathbf{v} + A\mathbf{v})/2 + (\mathbf{v} - A\mathbf{v})/2$.
Hint 2. If $A \neq \pm I_2$, then A has both eigenvalues 1 and -1. What are its trace and determinant then?
Sketch. Every $\mathbf{v}$ splits as $(\mathbf{v} + A\mathbf{v})/2 \in E_1$ plus $(\mathbf{v} - A\mathbf{v})/2 \in E_{-1}$, so A is diagonalizable with eigenvalues in $\{1, -1\}$. So $A = \pm I_2$, or $A = S\operatorname{diag}(1, -1)S^{-1}$, which happens iff $\operatorname{tr} A = 0$ and $\det A = -1$: $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ with $a^2 + bc = 1$. Geometrically, the reflections across a line along another line.

If $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{v} = A^2\mathbf{v} = \lambda^2\mathbf{v}$, so $\lambda = \pm 1$. For every $\mathbf{v} \in \mathbb{R}^2$,

$$\mathbf{v} = \frac{1}{2}(\mathbf{v} + A\mathbf{v}) + \frac{1}{2}(\mathbf{v} - A\mathbf{v}),$$

and $A(\mathbf{v} + A\mathbf{v}) = A\mathbf{v} + \mathbf{v}$, $A(\mathbf{v} - A\mathbf{v}) = -(\mathbf{v} - A\mathbf{v})$. So $\mathbb{R}^2 = E_1 + E_{-1}$ (this is Exercise 7.2.55). Dropping a summand that is $\{\mathbf{0}\}$ if 1 or -1 is not an eigenvalue, $\mathbb{R}^2$ is the sum, hence the direct sum, of the eigenspaces of A, and A is diagonalizable by Theorem 7.4.16, (II) $\Rightarrow$ (I), with eigenvalues in $\{1, -1\}$.

If only one of ± 1 is an eigenvalue, then $A = \pm I_2$ by Proposition 7.4.7. Otherwise $A = S\operatorname{diag}(1, -1)S^{-1}$ for an invertible S, so $\operatorname{tr} A = \operatorname{tr}\operatorname{diag}(1, -1) = 0$ and $\det A = -1$ (Proposition 4.5.27(b), Proposition 5.4.17). Conversely, if $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ with $-a^2 - bc = \det A = -1$, then

$$A^2 = \begin{pmatrix} a^2+bc & ab-ab \\ ac-ac & bc+a^2 \end{pmatrix} = I_2.$$

So the solutions are $\pm I_2$ and the matrices $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ with $a^2+bc=1$. Geometrically, $A = S\,\mathrm{diag}(1,-1)S^{-1}$ fixes the line E_1 pointwise and negates the line E_{-1}, with $\mathbb{R}^2 = E_1 \oplus E_{-1}$: it is the reflection across E_1 along E_{-1}, and every pair of distinct lines through $\mathbf{0}$ occurs. The orthogonal reflections of Example 2.1.11 are the cases with $E_{-1} \perp E_1$.

Solution to Exercise 7.4.49

Uses. Definition 3.4.26 (Lagrange polynomials), Proposition 3.4.27 (Lagrange's formula), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 7.1.20 (Factor theorem), Proposition 7.1.27 (Multiplicities add up to at most the degree), Corollary 7.2.23 (Eigenspaces are independent), Theorem 7.4.16 (Criteria for diagonalizability), Exercise 1.1.51, Proposition 7.1.34 (Roots of a split polynomial)
Techniques. Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. $L_1 + \cdots + L_k$ and 1 both have degree less than k and take the value 1 at every node. Evaluate the identity at T: every $\mathbf{v}$ is $\sum_j L_j(T)\mathbf{v}$.
Hint 2. L_j vanishes at every r_i with $i \neq j$, so $(x-r_j)L_j$ is a multiple of q. Apply this at T to show $(T - r_j\,\mathrm{id})L_j(T) = 0$.
Hint 3. For (b), $x^m - 1$ has m distinct roots in $\mathbb{C}$, the m-th roots of unity. For (c), every $a \in \mathbb{F}_p$ is a root of $x^p - x$.
Sketch. By uniqueness of interpolation, $\sum_j L_j = 1$, so $\mathbf{v} = \sum_j L_j(T)\mathbf{v}$ for every $\mathbf{v}$. Since $L_j(r_i) = 0$ for $i \neq j$, $L_j = c_j \prod_{i\neq j}(x - r_i)$, so $(x-r_j)L_j = c_j q$ and $(T - r_j\,\mathrm{id})L_j(T) = c_j q(T) = 0$: $L_j(T)\mathbf{v} \in E_{r_j}$. So the eigenspaces span V. (b) $x^m - 1 = \prod\big(x - e^{2\pi i l/m}\big)$. (c) $x^p - x = \prod_{a\in\mathbb{F}_p}(x-a)$ by Fermat.

(a) By Proposition 3.4.27 with all values $d_j = 1$, the unique polynomial of degree at most $k-1$ with value 1 at every node is $L_1 + \cdots + L_k$; the constant 1 is such a polynomial, so $L_1 + \cdots + L_k = 1$. Evaluating at T (Proposition 7.2.27(a)),

$$\mathbf{v} = L_1(T)\mathbf{v} + \cdots + L_k(T)\mathbf{v} \quad \text{for every } \mathbf{v} \in V.$$

Next, L_j vanishes at the $k-1$ distinct nodes r_i, $i \neq j$, so by the factor theorem (applied $k-1$ times, as in Proposition 7.1.27) $\prod_{i\neq j}(x - r_i)$ divides L_j; since $\deg L_j \leq k-1$, $L_j = c_j \prod_{i\neq j}(x-r_i)$ for a scalar c_j. Hence $(x-r_j)L_j = c_j q$, and evaluating at T,

$$(T - r_j\,\mathrm{id})L_j(T) = c_j q(T) = 0.$$

So $L_j(T)\mathbf{v} \in \ker(T - r_j\,\mathrm{id}) = E_{r_j}$ for every $\mathbf{v}$. Thus $V = E_{r_1} + \cdots + E_{r_k}$. Dropping the r_j that are not eigenvalues (their eigenspaces are $\{\mathbf{0}\}$), V is the sum of the eigenspaces of T, so it is their direct sum (Corollary 7.2.23), and T is diagonalizable by Theorem 7.4.16, (II) $\Rightarrow$ (I).

(b) The m numbers $\omega_l = e^{2\pi i l/m}$, $l = 0, \ldots, m-1$, are distinct roots of $x^m - 1$, so by Proposition 7.1.34 (or the factor theorem, repeatedly) $x^m - 1 = \prod_l (x - \omega_l)$, comparing degrees and leading coefficients. Since $A^m - I_n = 0$, (a) applies with $q = x^m - 1$.

(c) By Fermat's little theorem (Exercise 1.1.51), $a^{p-1} = 1$ for $a \neq 0$ in $\mathbb{F}_p$, so $a^p = a$ for every $a \in \mathbb{F}_p$, including $a = 0$. So the p distinct elements of $\mathbb{F}_p$ are roots of $x^p - x$, and as in (b), $x^p - x = \prod_{a\in\mathbb{F}_p}(x-a)$. Since $A^p - A = 0$, (a) applies.

Solution to Exercise 7.4.50

Uses. Definition 3.4.26 (Lagrange polynomials), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 7.4.16 (Criteria for diagonalizability), Corollary 7.2.23 (Eigenspaces are independent), Theorem 4.4.13 (Bases and dimension of a sum), Proposition 7.4.26 (Polynomials that kill a diagonalizable operator), Exercise 7.4.49
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Write $\mathbf{u} \in U$ as $\mathbf{u} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$ with $\mathbf{v}_i \in E_{\lambda_i}$. You need each $\mathbf{v}_i$ to lie in U.
Hint 2. Find a polynomial L_i with $L_i(T)\mathbf{u} = \mathbf{v}_i$. Polynomials in T map U into U.
Hint 3. Alternatively, apply Exercise 7.4.49 to the restriction: it is killed by $\prod(x - \lambda_i)$.

Sketch. Let L_i be the Lagrange polynomials for the nodes $\lambda_1, \dots, \lambda_m$. For $\mathbf{u} = \sum \mathbf{v}_j$ with $\mathbf{v}_j \in E_{\lambda_j}$, $L_i(T)\mathbf{u} = \sum_j L_i(\lambda_j)\mathbf{v}_j = \mathbf{v}_i$, and $L_i(T)\mathbf{u} \in U$ since U is invariant under polynomials in T. So $\mathbf{v}_i \in U \cap E_{\lambda_i}$, and U is the (direct) sum of these. Concatenating bases gives an eigenbasis of U.

First proof. Let $L_1, \dots, L_m$ be the Lagrange polynomials for the distinct nodes $\lambda_1, \dots, \lambda_m$ (Definition 3.4.26). Let $\mathbf{u} \in U$. Since $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$ (Theorem 7.4.16), $\mathbf{u} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$ with $\mathbf{v}_j \in E_{\lambda_j}$, and by Proposition 7.2.28

$$L_i(T)\mathbf{u} = \sum_{j=1}^{m} L_i(\lambda_j)\mathbf{v}_j = \mathbf{v}_i.$$

Since U is T-invariant, it is $L_i(T)$-invariant (Proposition 7.2.27(d)), so $\mathbf{v}_i = L_i(T)\mathbf{u} \in U \cap E_{\lambda_i}$. Hence $U \subseteq \sum_i \big(U \cap E_{\lambda_i}\big)$, and the reverse inclusion holds because each $U \cap E_{\lambda_i}$ lies in U. The sum is direct because $E_{\lambda_1} + \cdots + E_{\lambda_m}$ is (Corollary 7.2.23). Concatenating bases of the spaces $U \cap E_{\lambda_i}$ gives a basis of U (Theorem 4.4.13) consisting of eigenvectors of T, which are eigenvectors of the restriction; so the restriction is diagonalizable.

Second proof. By Proposition 7.4.26, $q(T) = 0$ for $q = (x - \lambda_1)\cdots(x - \lambda_m)$. For $\mathbf{u} \in U$, $q(T\restriction_U)\mathbf{u} = q(T)\mathbf{u} = \mathbf{0}$, where $T\restriction_U$ is regarded as an operator on U. So the restriction is killed by a product of distinct linear factors, and it is diagonalizable by Exercise 7.4.49(a). Its eigenspaces are $\ker(T\restriction_U - \lambda_i \mathrm{id}_U) = U \cap E_{\lambda_i}$, and Theorem 7.4.16(II) for the restriction gives the decomposition of U.

Solution to Exercise 7.4.51

Uses. Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Exercise 7.4.50, Theorem 7.4.16 (Criteria for diagonalizability), Theorem 4.4.13 (Bases and dimension of a sum)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Each eigenspace $E_\lambda(T)$ is S-invariant (Proposition 7.2.30).
Hint 2. The restriction of S to $E_\lambda(T)$ is diagonalizable by Exercise 7.4.50. Its eigenvectors are eigenvectors of T too.
Hint 3. For the counterexample, a diagonal and a non-diagonal matrix with distinct eigenvalues.
Sketch. $V = \bigoplus E_{\lambda_i}(T)$, and each $E_{\lambda_i}(T)$ is S-invariant, so S restricted to it is diagonalizable (Exercise 7.4.50). An eigenbasis of each restriction consists of common eigenvectors; concatenate. Without commuting: $\mathrm{diag}(1,2)$ and $\left(\begin{smallmatrix}1 & 1\\ 0 & 2\end{smallmatrix}\right)$ are both diagonalizable but have no common eigenbasis, since common eigenvectors would make them commute.

Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T, so that $V = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_m}(T)$ by Theorem 7.4.16. Since $ST = TS$, each $E_{\lambda_i}(T)$ is S-invariant (Proposition 7.2.30(b)). By Exercise 7.4.50, the restriction of S to $E_{\lambda_i}(T)$ is diagonalizable, so $E_{\lambda_i}(T)$ has a basis $\mathcal{B}_i$ of eigenvectors of S. Each vector of $\mathcal{B}_i$ is nonzero and lies in $E_{\lambda_i}(T)$, so it is also an eigenvector of T. The concatenation $\mathcal{B}_1, \dots, \mathcal{B}_m$ is a basis of V (Theorem 4.4.13), consisting of common eigenvectors.

The hypothesis is needed. If $\mathcal{B}$ is a common eigenbasis, then $[S]_{\mathcal{B}\mathcal{B}}$ and $[T]_{\mathcal{B}\mathcal{B}}$ are diagonal, hence commute, so $ST = TS$. Thus $D = \mathrm{diag}(1,2)$ and $A = \left(\begin{smallmatrix}1 & 1\\ 0 & 2\end{smallmatrix}\right)$, which are both diagonalizable (distinct eigenvalues) but do not commute ($DA = \left(\begin{smallmatrix}1 & 1\\ 0 & 4\end{smallmatrix}\right)$, $AD = \left(\begin{smallmatrix}1 & 2\\ 0 & 4\end{smallmatrix}\right)$), have no common eigenbasis.

Solution to Exercise 7.4.52

Uses. Exercise 7.4.49, Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix), Proposition 4.5.27 (Similarity invariants), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22), Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. $x^3 - x = x(x-1)(x+1)$ has distinct roots, so A is diagonalizable (Exercise 7.4.49).
Hint 2. Write $A = SDS^{-1}$. Both sides are similarity invariants, so compute them for D.
Hint 3. The diagonal entries of D lie in $\{0, 1, -1\}$, and $d^2 = 1$ exactly when $d \neq 0$.
Sketch. A is killed by $x(x-1)(x+1)$, so $A = SDS^{-1}$ with D diagonal, entries in $\{0, \pm 1\}$. Then $\operatorname{rank} A = \operatorname{rank} D$ is the number of nonzero d_i, and $\operatorname{tr} A^2 = \operatorname{tr} D^2 = \sum d_i^2$ is the same number.

Since $A^3 - A = 0$ and $x^3 - x = x(x-1)(x+1)$ has the distinct roots $0, 1, -1$, A is diagonalizable by Exercise 7.4.49(a): $A = SDS^{-1}$ with $D = \mathrm{diag}(d_1, \dots, d_n)$. Each d_i is an eigenvalue, hence a root of $x^3 - x$ (Proposition 7.2.28), so $d_i \in \{0, 1, -1\}$. Rank and trace are similarity invariants (Proposition 4.5.27), and $A^2 = SD^2S^{-1}$ (Proposition 7.4.23), so

$$\operatorname{rank} A = \operatorname{rank} D = \#\{i \mid d_i \neq 0\}, \quad \operatorname{tr}\big(A^2\big) = \operatorname{tr}\big(D^2\big) = \sum_{i=1}^{n} d_i^2.$$

For $d_i \in \{0, 1, -1\}$, $d_i^2 = 1$ if $d_i \neq 0$ and $d_i^2 = 0$ otherwise. So both sides count the nonzero d_i, and $\operatorname{rank} A = \operatorname{tr}\big(A^2\big)$.

Solution to Exercise 7.4.53

Uses. Exercise 7.4.42, Proposition 4.4.16 (Complements exist), Definition 4.4.24 (Projection along a complement), Proposition 4.4.26 (Properties of projections), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Corollary 7.2.23 (Eigenspaces are independent), Theorem 7.4.16 (Criteria for diagonalizability)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For "only if", $T - \lambda\,\mathrm{id}$ is diagonalizable; use Exercise 7.4.42.
Hint 2. For "if", let U be the sum of all eigenspaces and suppose $U \neq V$. Take a complement W and the projection P onto W along U. The operator $PT|_W$ on W has an eigenvector $\mathbf{w}$ over $\mathbb{C}$: $T\mathbf{w} - \lambda\mathbf{w} \in U$.
Hint 3. Write $(T - \lambda\,\mathrm{id})\mathbf{w} = \sum_\mu \mathbf{u}_\mu$ with $\mathbf{u}_\mu \in E_\mu$. For $\mu \neq \lambda$, $\mathbf{u}_\mu = (T - \lambda\,\mathrm{id})\big(\mathbf{u}_\mu/(\mu - \lambda)\big)$. Subtract these from $\mathbf{w}$ and use the hypothesis.
Sketch. "Only if": $T - \lambda\,\mathrm{id}$ is diagonalizable, so its kernel and image meet in $\mathbf{0}$. "If": let $U = \bigoplus E_\mu$ and suppose $U \neq V$; choose $V = U \oplus W$. The operator $\mathbf{w} \mapsto PT\mathbf{w}$ on W has an eigenvector $\mathbf{w} \neq \mathbf{0}$, so $(T - \lambda\,\mathrm{id})\mathbf{w} = \sum_\mu \mathbf{u}_\mu \in U$. Put $\mathbf{w}' = \mathbf{w} - \sum_{\mu\neq\lambda} \mathbf{u}_\mu/(\mu - \lambda)$; then $(T - \lambda\,\mathrm{id})\mathbf{w}' = \mathbf{u}_\lambda \in \ker \cap\, \mathrm{im}$, so $\mathbf{u}_\lambda = \mathbf{0}$ and $\mathbf{w}' \in E_\lambda \subseteq U$. Then $\mathbf{w} \in U \cap W = \{\mathbf{0}\}$, a contradiction.

Only if. If $\mathcal{B}$ is an eigenbasis for T, it is also one for $T - \lambda\,\mathrm{id}$, since $T\mathbf{b} = \mu\mathbf{b}$ gives $(T - \lambda\,\mathrm{id})\mathbf{b} = (\mu - \lambda)\mathbf{b}$. So $T - \lambda\,\mathrm{id}$ is diagonalizable, and Exercise 7.4.42(a) gives $\ker(T - \lambda\,\mathrm{id}) \cap \mathrm{im}(T - \lambda\,\mathrm{id}) = \{\mathbf{0}\}$.

If. Let U be the sum of the eigenspaces of T; it is the direct sum of the eigenspaces of the finitely many eigenvalues (Corollary 7.2.23). Suppose, for contradiction, that $U \neq V$. Choose a complement W, so $V = U \oplus W$ and $W \neq \{\mathbf{0}\}$ (Proposition 4.4.16), and let $P = \mathrm{Pr}_{W,U}$ be the projection onto W along U (Definition 4.4.24). The map $R : W \to W$, $R\mathbf{w} = PT\mathbf{w}$, is an operator on the nonzero complex space W, so it has an eigenvector $\mathbf{w} \neq \mathbf{0}$ (Theorem 7.3.27): $PT\mathbf{w} = \lambda\mathbf{w}$ for some $\lambda \in \mathbb{C}$. Since $P\mathbf{w} = \mathbf{w}$, this says $P(T\mathbf{w} - \lambda\mathbf{w}) = \mathbf{0}$, and since $\ker P = U$ and P fixes W (Proposition 4.4.26(b)), $(T - \lambda\,\mathrm{id})\mathbf{w} \in U$. Write

$$(T - \lambda\,\mathrm{id})\mathbf{w} = \sum_{\mu} \mathbf{u}_\mu, \quad \mathbf{u}_\mu \in E_\mu,$$

the sum over the eigenvalues μ of T. For $\mu \neq \lambda$, $(T - \lambda\,\mathrm{id})\mathbf{u}_\mu = (\mu - \lambda)\mathbf{u}_\mu$, so $\mathbf{u}_\mu = (T - \lambda\,\mathrm{id})\big(\mathbf{u}_\mu/(\mu - \lambda)\big)$. Let

$$\mathbf{w}' = \mathbf{w} - \sum_{\mu\neq\lambda} \frac{\mathbf{u}_\mu}{\mu - \lambda}.$$

Then $(T - \lambda\,\mathrm{id})\mathbf{w}' = \sum_\mu \mathbf{u}_\mu - \sum_{\mu\neq\lambda} \mathbf{u}_\mu = \mathbf{u}_\lambda$, where $\mathbf{u}_\lambda = \mathbf{0}$ if λ is not an eigenvalue. In any case $(T - \lambda\,\mathrm{id})\mathbf{w}'$ lies in $\ker(T - \lambda\,\mathrm{id})$ (it is $\mathbf{0}$ or in E_λ) and in $\mathrm{im}(T - \lambda\,\mathrm{id})$, so it is $\mathbf{0}$ by hypothesis. Hence $\mathbf{w}' \in \ker(T - \lambda\,\mathrm{id}) \subseteq U$, and therefore $\mathbf{w} = \mathbf{w}' + \sum_{\mu\neq\lambda} \mathbf{u}_\mu/(\mu - \lambda) \in U$. But $\mathbf{w} \in W$ and $U \cap W = \{\mathbf{0}\}$, so $\mathbf{w} = \mathbf{0}$, a contradiction. Hence $U = V$, and T is diagonalizable by Theorem 7.4.16, (II) $\Rightarrow$ (I).

The hypothesis is needed only for the one λ that the argument produces, but that λ is not known in advance. Over $\mathbb{R}$ the statement fails: the rotation by a right angle has $\ker(T - \lambda\,\mathrm{id}) = \{\mathbf{0}\}$ for every real λ and is not diagonalizable over $\mathbb{R}$.

Section 7.5

Solution to Exercise 7.5.7

Uses. Definition 2.3.1 (Matrix of a linear map), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces)
Hint 1. The (i, j) entry of $[T]_{\mathcal{B}'}$ is the coefficient of $\mathbf{b}_{n+1-i}$ in $T\mathbf{b}_{n+1-j}$. Relate it to an entry of $[T]_{\mathcal{B}}$.
Hint 2. Reversing both the rows and the columns sends position (i, j) to $(n + 1 - i, n + 1 - j)$, and $i > j$ if and only if $n + 1 - i < n + 1 - j$.

Sketch. The (i,j) entry of $[T]_{\mathcal{B}'}$ is the $(n+1-i, n+1-j)$ entry of $[T]_{\mathcal{B}}$. Below-diagonal positions of one matrix correspond to above-diagonal positions of the other. For D, $[D]_{\mathcal{B}'}$ has $3, 2, 1$ just below the diagonal.

Write $\mathbf{b}'_j = \mathbf{b}_{n+1-j}$. The (i,j) entry of $[T]_{\mathcal{B}'}$ is the coefficient of $\mathbf{b}'_i = \mathbf{b}_{n+1-i}$ in $T\mathbf{b}'_j = T\mathbf{b}_{n+1-j}$ (Definition 2.3.1), which is the $(n+1-i, n+1-j)$ entry of $[T]_{\mathcal{B}}$. Now $i < j$ if and only if $n+1-i > n+1-j$. So the entries of $[T]_{\mathcal{B}'}$ above the diagonal are exactly the entries of $[T]_{\mathcal{B}}$ below the diagonal, and one set vanishes if and only if the other does.

For D, $D(x^3) = 3x^2$, $D(x^2) = 2x$, $D(x) = 1$, and $D(1) = \mathbf{0}$, so

$$[D]_{x^3,x^2,x,1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This is $[D]_{1,x,x^2,x^3}$ with the order of the rows and of the columns reversed.

Solution to Exercise 7.5.10

Uses. Proposition 7.5.8 (The diagonal of a triangular matrix), Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Hint 1. Each $T\mathbf{b}_j$ lies in $\operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_j)$, so $[T]_{\mathcal{B}}$ is upper triangular. Read off the diagonal.
Hint 2. For the eigenvectors, solve $(T - \lambda \operatorname{id})(x\mathbf{b}_1 + y\mathbf{b}_2 + z\mathbf{b}_3) = \mathbf{0}$ coefficient by coefficient.
Sketch. The diagonal is $3, -1, 3$, so the eigenvalues are 3 (multiplicity 2) and -1 (multiplicity 1), and T is invertible. $E_3 = \operatorname{span}(\mathbf{b}_1)$ and $E_{-1} = \operatorname{span}(\mathbf{b}_1 - 4\mathbf{b}_2)$. Only two independent eigenvectors exist, so T is not diagonalizable.

The matrix

$$[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

is upper triangular. By Proposition 7.5.8 and Proposition 7.5.9, the eigenvalues are 3, with algebraic multiplicity 2, and -1, with algebraic multiplicity 1. No diagonal entry is 0, so T is invertible.

For $\lambda = 3$: $(T - 3\operatorname{id})(x\mathbf{b}_1 + y\mathbf{b}_2 + z\mathbf{b}_3) = y(\mathbf{b}_1 - 4\mathbf{b}_2) + z(\mathbf{b}_1 + \mathbf{b}_2) = (y+z)\mathbf{b}_1 + (z - 4y)\mathbf{b}_2$. This vanishes if and only if $y + z = 0$ and $z = 4y$, that is, $y = z = 0$. So $E_3 = \operatorname{span}(\mathbf{b}_1)$.

For $\lambda = -1$: $(T + \operatorname{id})(x\mathbf{b}_1 + y\mathbf{b}_2 + z\mathbf{b}_3) = (4x + y + z)\mathbf{b}_1 + z\mathbf{b}_2 + 4z\mathbf{b}_3$, which vanishes if and only if $z = 0$ and $y = -4x$. So $E_{-1} = \operatorname{span}(\mathbf{b}_1 - 4\mathbf{b}_2)$. Check: $T(\mathbf{b}_1 - 4\mathbf{b}_2) = 3\mathbf{b}_1 - 4\mathbf{b}_1 + 4\mathbf{b}_2 = -(\mathbf{b}_1 - 4\mathbf{b}_2)$.

Every eigenvector lies in $E_3 \cup E_{-1}$, a union of two lines, so there is no basis of eigenvectors, and T is not diagonalizable (Theorem 7.4.2).

Solution to Exercise 7.5.21

Uses. Theorem 7.5.18 (Schur's theorem), Theorem 7.2.9 (Eigenvalues and invertibility), Proposition 6.3.18 (Coordinates in an orthonormal basis)
Hint 1. Find an eigenvalue and a unit eigenvector $\mathbf{u}_1$; complete it to an orthonormal basis of $\mathbb{C}^2$.
Hint 2. The entries of R are $r_{ij} = \langle A\mathbf{u}_j, \mathbf{u}_i \rangle$.
Sketch. $\det(A - \lambda I_2) = (\lambda - 2)^2$; $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1)$, $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-1,1)$; $A\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-4,0)$, so $R = \begin{pmatrix} 2 & -2 \\ 0 & 2 \end{pmatrix}$.

$\det(A - \lambda I_2) = (3-\lambda)(1-\lambda) + 1 = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, so 2 is the only eigenvalue. From $A - 2I_2 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$, a unit eigenvector is $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1,1)$; complete it with $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-1,1)$. Then $A\mathbf{u}_1 = 2\mathbf{u}_1$, and $A\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-4,0)$, so by Proposition 6.3.18

$$r_{12} = \langle A\mathbf{u}_2, \mathbf{u}_1 \rangle = \frac{1}{2}(-4) = -2, \quad r_{22} = \langle A\mathbf{u}_2, \mathbf{u}_2 \rangle = \frac{1}{2}(4) = 2.$$

Hence $A = URU^*$ with $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and $R = \begin{pmatrix} 2 & -2 \\ 0 & 2 \end{pmatrix}$. Check: $AU = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & -4 \\ 2 & 0 \end{pmatrix} = UR$.

Solution to Exercise 7.5.29

Uses. Theorem 7.5.13 (Upper-triangular form), Lemma 7.5.11 (Compressing an operator to a complement), Corollary 7.5.14 (Triangularizing a matrix)
Hint 1. $(2,1,1)$ is an eigenvector. Use the complement $\operatorname{span}(\mathbf{e}_2, \mathbf{e}_3)$ and compute the compression.
Hint 2. In the basis $(2,1,1), \mathbf{e}_2, \mathbf{e}_3$, the matrix is $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}$. Triangularize the bottom right 2×2 block, starting with an eigenvector for 1.
Sketch. $A(2,1,1) = (2,1,1)$. With M having columns $(2,1,1), \mathbf{e}_2, \mathbf{e}_3$, $M^{-1}AM = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}$. The block $\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$ has eigenvector $(-2,1)$ for 1, giving $\mathbf{b}_2 = (0,-2,1)$. With S having columns $(2,1,1), (0,-2,1), \mathbf{e}_3$, $S^{-1}AS = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

First, $A(2,1,1) = (4-2, 4-1-2, -2+3) = (2,1,1)$, so $\mathbf{b}_1 = (2,1,1)$ is an eigenvector for 1. Since its first coordinate is nonzero, $\mathbf{b}_1, \mathbf{e}_2, \mathbf{e}_3$ is a basis, and $W = \operatorname{span}(\mathbf{e}_2, \mathbf{e}_3)$ is a complement of $\operatorname{span}(\mathbf{b}_1)$. Writing $(x_1, x_2, x_3) = (x_1/2)\mathbf{b}_1 + (0, x_2 - x_1/2, x_3 - x_1/2)$, the projection onto W along $\operatorname{span}(\mathbf{b}_1)$ is $(x_1, x_2, x_3) \mapsto (0, x_2 - x_1/2, x_3 - x_1/2)$. Since $A\mathbf{e}_2 = (-2,-1,0)$ and $A\mathbf{e}_3 = (0,-2,3)$, the compression sends

$$\mathbf{e}_2 \mapsto (0, -1+1, 0+1) = \mathbf{e}_3, \quad \mathbf{e}_3 \mapsto (0,-2,3) = -2\mathbf{e}_2 + 3\mathbf{e}_3,$$

so its matrix is $\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$, with characteristic polynomial $x^2 - 3x + 2 = (x-1)(x-2)$. An eigenvector for 1 is $(-2,1)$, that is, $\mathbf{b}_2 = -2\mathbf{e}_2 + \mathbf{e}_3 = (0,-2,1)$. Complete with $\mathbf{b}_3 = \mathbf{e}_3$. Then

$$A\mathbf{b}_2 = (4,0,3) = 2\mathbf{b}_1 + \mathbf{b}_2,$$
$$A\mathbf{b}_3 = (0,-2,3) = \mathbf{b}_2 + 2\mathbf{b}_3.$$

So with

$$S = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

we have $AS = SR$ column by column, that is, $A = SRS^{-1}$. The diagonal $1, 1, 2$ agrees with $p_A = (x-1)^2(x-2)$: the first entry comes from $\mathbf{b}_1$ and the other two from the compression. (Starting instead with the eigenvector $(1,0,1)$ for 2 gives a triangular form with diagonal $2, 1, 1$.)

Solution to Exercise 7.5.30

Uses. Theorem 7.5.13 (Upper-triangular form), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Corollary 7.1.43 (Odd degree gives a real root)
Techniques. Build a flag of invariant subspaces (Technique 7.5.4)
Hint 1. For (a), a triangularizing basis comes with an invariant flag.
Hint 2. For (b), compute T^2. Why does it follow that T has no real eigenvalue?
Hint 3. If U is invariant of dimension 3, then $T|_U$ is an operator on a 3-dimensional real space. Its characteristic polynomial has odd degree.
Sketch. (a) The flag $\operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ of a triangularizing basis. (b) $T^2 = -\mathrm{id}$, so $T\mathbf{v} = \lambda\mathbf{v}$ forces $(\lambda^2 + 1)\mathbf{v} = \mathbf{0}$ and T has no eigenvector, hence no invariant line. An invariant U of dimension 3 would give an operator $T|_U$ whose characteristic polynomial has odd degree, hence a real root, hence an eigenvector of T.

(a) By Theorem 7.5.13 there is a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ with $[T]$ upper triangular, and by Proposition 7.5.2 the subspaces $\operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_k)$, of dimension k, are T-invariant ($k = 0$ gives $\{\mathbf{0}\}$).

(b) $T^2(x_1, x_2, x_3, x_4) = T(-x_2, x_1, -x_4, x_3) = (-x_1, -x_2, -x_3, -x_4)$, so $T^2 = -\mathrm{id}$. If $T\mathbf{v} = \lambda\mathbf{v}$ with $\lambda \in \mathbb{R}$, then $-\mathbf{v} = T^2\mathbf{v} = \lambda^2\mathbf{v}$, so $(\lambda^2 + 1)\mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$, as $\lambda^2 + 1 > 0$. So T has no eigenvector, and no invariant subspace of dimension 1 (Proposition 7.2.7).

Suppose U were invariant with $\dim U = 3$. Then $T|_U$ is an operator on U, and its characteristic polynomial has degree 3, so it has a real root (Corollary 7.1.43), which is an eigenvalue of $T|_U$ (Theorem 7.3.6). An

eigenvector of $T|_U$ is an eigenvector of T, which we just showed does not exist. So no such U exists. (Invariant subspaces of dimension 2 do exist, such as $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_2)$.)

Solution to Exercise 7.5.31

Uses. Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 4.1.16 (Extending to a basis), Theorem 7.2.9 (Eigenvalues and invertibility), Theorem 5.1.13 (The 2×2 determinant)
Techniques. Try small cases and small fields (Technique 1.2.38)
Hint 1. In dimension 2, triangularizable means: there is an eigenvector. Why is that enough?
Hint 2. λ is an eigenvalue if and only if $\det(A - \lambda I_2) = 0$. There are only two candidates, $\lambda = 0$ and $\lambda = 1$. Express both determinants in terms of $\operatorname{tr} A$ and $\det A$.
Sketch. A 2×2 matrix is triangularizable if and only if it has an eigenvector $\mathbf{b}_1$: extend it to a basis, and $T\mathbf{b}_2 \in \operatorname{span}(\mathbf{b}_1, \mathbf{b}_2) = \mathbb{F}_2^2$ automatically. Since $\det(A - \lambda I_2) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A$, the candidates fail exactly when $\det A = 1$ and $\operatorname{tr} A = 1$, which forces $b = c = 1$ and $\{a, d\} = \{0, 1\}$. So 14 are triangularizable, and the exceptions are $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$.

In dimension 2, A is triangularizable if and only if it has an eigenvector. If T_A is upper triangular with respect to $\mathbf{b}_1, \mathbf{b}_2$, then $\mathbf{b}_1$ is an eigenvector. Conversely, if $\mathbf{b}_1$ is an eigenvector, extend it to a basis $\mathbf{b}_1, \mathbf{b}_2$ (Theorem 4.1.16); then $T\mathbf{b}_1 \in \operatorname{span}(\mathbf{b}_1)$ and $T\mathbf{b}_2 \in \operatorname{span}(\mathbf{b}_1, \mathbf{b}_2)$, so T_A is upper triangular by Proposition 7.5.2.

For $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, Theorem 5.1.13 gives $\det(A - \lambda I_2) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$. By Theorem 7.2.9, A has no eigenvalue in $\mathbb{F}_2$ if and only if this is nonzero for $\lambda = 0$ and $\lambda = 1$, that is,

$$\det A = 1 \quad \text{and} \quad 1 + \operatorname{tr} A + \det A = 1,$$

using $-1 = 1$ in $\mathbb{F}_2$. Substituting $\det A = 1$ into the second condition gives $\operatorname{tr} A = 1$, since $1 + 1 = 0$. From $a + d = 1$, one of a, d is 0, so $ad = 0$ and $\det A = -bc = bc = 1$, giving $b = c = 1$. So the matrices with no eigenvalue are exactly

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

whose characteristic polynomial is $x^2 + x + 1$, the one irreducible quadratic over $\mathbb{F}_2$. The other 14 matrices are triangularizable.

Solution to Exercise 7.5.32

Uses. Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 7.2.9 (Eigenvalues and invertibility)
Hint 1. A triangularizable operator has an eigenvector. Look for a rational matrix whose eigenvalues are irrational.
Hint 2. For (b), if A is upper triangular in a real basis, its characteristic polynomial is a product of real linear factors (Proposition 7.5.9). Put a rotation block next to a 1×1 block.
Sketch. (a) $A = \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$ has eigenvalues $\pm\sqrt{2}$: no eigenvector in $\mathbb{Q}^2$, but an eigenvector $\left(\sqrt{2}, 1\right)$ in $\mathbb{R}^2$, which a second vector completes to a triangularizing basis. (b) $A = \left(\begin{smallmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ has eigenvalue 1, but $p_A = (x - 1)\left(x^2 + 1\right)$ is not a product of real linear factors.

(a) Let $A = \left(\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right)$. Then $\det(A - \lambda I_2) = \lambda^2 - 2$, which has no rational root, so A has no eigenvalue in $\mathbb{Q}$ (Theorem 7.2.9). The first vector of a triangularizing basis would be an eigenvector, so A is not triangularizable over $\mathbb{Q}$. Over $\mathbb{R}$, $A\left(\sqrt{2}, 1\right) = \left(2, \sqrt{2}\right) = \sqrt{2}\left(\sqrt{2}, 1\right)$, and in the basis $\left(\sqrt{2}, 1\right), \mathbf{e}_2$ the first vector spans an invariant line, so the flag $\{\mathbf{0}\} \subseteq \operatorname{span}\left(\left(\sqrt{2}, 1\right)\right) \subseteq \mathbb{R}^2$ is invariant and A is triangularizable over $\mathbb{R}$.

(b) Let $A = \left(\begin{smallmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Then $A\mathbf{e}_3 = \mathbf{e}_3$, so 1 is an eigenvalue. By cofactor expansion along the last row, $p_A = (x - 1)\left(x^2 + 1\right)$. If $A = SRS^{-1}$ with S, R real and R upper triangular, then $p_A = (x - r_{11})(x - r_{22})(x - r_{33})$ with real r_{jj} (Proposition 7.5.9). Evaluating at i gives $0 = p_A(i) = (i - r_{11})(i - r_{22})(i - r_{33})$, a product of nonzero complex numbers, which is impossible. So A is not triangularizable over $\mathbb{R}$.

Solution to Exercise 7.5.33

Uses. Definition 7.5.1 (Triangularizable), Exercise 7.5.7, Non-example 7.5.6 (A rotation), Proposition 4.5.27 (Similarity invariants)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), transpose $A = SRS^{-1}$. The result has a lower-triangular middle factor; reverse the order of a basis.
Hint 2. For (b), add an upper-triangular and a lower-triangular nilpotent matrix.
Hint 3. For (c) and (d), look at diagonal matrices, and at the identity.
Sketch. (a) True: $A^{\mathsf{T}} = \left(S^{\mathsf{T}}\right)^{-1} R^{\mathsf{T}} S^{\mathsf{T}}$, and R^{T} is lower triangular, so with the reversal permutation matrix P, $PR^{\mathsf{T}}P$ is upper triangular and similar to A^{T}. (b) False: $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right)$ is a rotation. (c) False: $\operatorname{diag}(1,2)$ with the bases $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_2, \mathbf{e}_1$. (d) False: I_2 and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

(a) True. Write $A = SRS^{-1}$ with R upper triangular. Then $A^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}} R^{\mathsf{T}} S^{\mathsf{T}}$, and R^{T} is lower triangular. Let P be the permutation matrix with $P\mathbf{e}_j = \mathbf{e}_{n+1-j}$, so $P^2 = I_n$. The (i,j) entry of $PR^{\mathsf{T}}P$ is the $(n+1-i, n+1-j)$ entry of R^{T}, which is $r_{n+1-j,n+1-i}$; it vanishes when $n+1-j > n+1-i$, that is, when $i > j$. So $R' = PR^{\mathsf{T}}P$ is upper triangular (this is Exercise 7.5.7 in matrix form), and with $M = \left(S^{-1}\right)^{\mathsf{T}} P$,

$$A^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}} PR'PS^{\mathsf{T}} = MR'M^{-1},$$

since $M^{-1} = P^{-1}S^{\mathsf{T}} = PS^{\mathsf{T}}$.

(b) False. $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ is upper triangular, and $B = \left(\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right)$ is lower triangular, hence triangularizable by (a). But $A + B = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ is the rotation by $-\pi/2$, which has no real eigenvalue and is not triangularizable over $\mathbb{R}$ (as in Non-example 7.5.6).

(c) False. $T = T_D$ with $D = \operatorname{diag}(1,2)$ is upper triangular (even diagonal) with respect to $\mathbf{e}_1, \mathbf{e}_2$ and with respect to $\mathbf{e}_2, \mathbf{e}_1$, and $\mathbf{e}_2$ is not a multiple of $\mathbf{e}_1$.

(d) False. I_2 and $J = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ have the same diagonal, but $S^{-1}I_2S = I_2 \neq J$ for every invertible S, so I_2 is similar only to itself.

Solution to Exercise 7.5.34

Uses. Corollary 7.5.14 (Triangularizing a matrix), Proposition 2.4.36 (Products of upper triangular matrices), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 7.5.22 (Trace and determinant from eigenvalues), Proposition 7.5.8 (The diagonal of a triangular matrix)
Techniques. Build a flag of invariant subspaces (Technique 7.5.4), Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. Write $A = SRS^{-1}$ with R upper triangular. Then $A^k = SR^kS^{-1}$ and $A^{-1} = SR^{-1}S^{-1}$.
Hint 2. Powers of R are upper triangular with the powers of the diagonal entries on the diagonal (Proposition 2.4.36). For R^{-1}, use flags: T_R maps each $\operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_j)$ into itself and is injective.
Sketch. (a) R^k is upper triangular with diagonal r_{jj}^k by Proposition 2.4.36 and induction, and $A^k = SR^kS^{-1}$, so Proposition 7.5.9 applies. (b) T_R maps $V_j = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_j)$ injectively into itself, hence onto it, so T_R^{-1} also preserves every V_j and R^{-1} is upper triangular; $RR^{-1} = I_n$ and Proposition 2.4.36 give diagonal entries r_{jj}^{-1}. Then take traces.

By Corollary 7.5.14, $A = SRS^{-1}$ with R upper triangular and diagonal entries $r_{jj} = \lambda_j$, after reordering the λ_j.

(a) $A^k = SR^kS^{-1}$, by induction on k, since $S^{-1}S = I_n$ cancels between consecutive factors. By Proposition 2.4.36 and induction on k, R^k is upper triangular with diagonal entries $\lambda_1^k, \ldots, \lambda_n^k$. So T_{A^k} is upper triangular with respect to the columns of S, with this diagonal, and by Proposition 7.5.9 these are the eigenvalues of A^k counted with multiplicity. By Theorem 7.5.22, $\operatorname{tr}\left(A^k\right) = \sum_j \lambda_j^k$.

(b) A is invertible, so $R = S^{-1}AS$ is too, and $A^{-1} = SR^{-1}S^{-1}$. Let $V_j = \operatorname{span}(\mathbf{e}_1, \ldots, \mathbf{e}_j)$. By Proposition 7.5.2, T_R maps V_j into V_j, and it is injective, so $T_R|_{V_j}$ is an injective map from V_j into itself, hence onto V_j (Corollary 4.2.18). So each $\mathbf{v} \in V_j$ is $T_R\mathbf{w}$ for some $\mathbf{w} \in V_j$, and $T_R^{-1}\mathbf{v} = \mathbf{w} \in V_j$. Thus $T_R^{-1} = T_{R^{-1}}$ preserves every V_j, and R^{-1} is upper triangular by Proposition 7.5.2. By Proposition 2.4.36, the diagonal entries of $RR^{-1} = I_n$ are $r_{jj}\left(R^{-1}\right)_{jj} = 1$, so $\left(R^{-1}\right)_{jj} = \lambda_j^{-1}$ (each $\lambda_j \neq 0$ by Proposition 7.5.8). Now argue as in (a).

Solution to Exercise 7.5.35

Uses. Theorem 7.5.22 (Trace and determinant from eigenvalues), Exercise 7.5.34, Proposition 7.5.9 (Multiplicities on the diagonal), Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Theorem 7.4.2 (Diagonalizable means an eigenbasis)
Hint 1. Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues. You know $\sum \lambda_i$, $\sum \lambda_i^2$, and $\prod \lambda_i$.
Hint 2. $(\lambda_1+\lambda_2+\lambda_3)^2 = \sum \lambda_i^2 + 2(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3)$. This gives every coefficient of $p_A = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$.
Sketch. $e_1 = \sum \lambda_i = 2$, $\sum \lambda_i^2 = 6$ gives $e_2 = \sum_{i<j} \lambda_i\lambda_j = (4-6)/2 = -1$, and $e_3 = \prod \lambda_i = -2$. So $p_A = x^3 - 2x^2 - x + 2 = (x-1)(x+1)(x-2)$, with three distinct eigenvalues, so A is diagonalizable.

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of A counted with multiplicity. By Theorem 7.5.22 and Exercise 7.5.34(a),

$$\lambda_1 + \lambda_2 + \lambda_3 = 2, \quad \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 6, \quad \lambda_1\lambda_2\lambda_3 = -2.$$

Let $e_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3$. Expanding the square, $(\lambda_1+\lambda_2+\lambda_3)^2 = \sum \lambda_i^2 + 2e_2$, so $4 = 6 + 2e_2$ and $e_2 = -1$. Expanding the product,

$$p_A = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3) = x^3 - 2x^2 + e_2 x - \lambda_1\lambda_2\lambda_3 = x^3 - 2x^2 - x + 2.$$

This factors as $x^2(x-2) - (x-2) = (x-2)(x-1)(x+1)$, so the eigenvalues are 1, -1, and 2. (Check: $1 + 1 + 4 = 6$.) Eigenvectors for the three distinct eigenvalues are linearly independent (Theorem 7.2.21), so they form a basis of $\mathbb{C}^3$, and A is diagonalizable (Theorem 7.4.2).

Solution to Exercise 7.5.36

Uses. Corollary 7.5.14 (Triangularizing a matrix), Proposition 7.5.8 (The diagonal of a triangular matrix), Example 2.8.3 (Strictly upper triangular matrices are nilpotent, three ways), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$)
Techniques. Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. If 0 is the only eigenvalue, a triangular form of A has zeros on the diagonal.
Hint 2. Strictly upper-triangular $n \times n$ matrices satisfy $N^n = 0$ (Example 2.8.3). For the converse, apply A^n to an eigenvector.
Hint 3. For (b), the real eigenvalues do not see a rotation block.
Sketch. (a) $A = SRS^{-1}$ with R upper triangular and diagonal the eigenvalues, all 0; so R is strictly upper triangular, $R^n = 0$, and $A^n = SR^nS^{-1} = 0$. Conversely $A\mathbf{v} = \lambda\mathbf{v}$ gives $\mathbf{0} = A^n\mathbf{v} = \lambda^n\mathbf{v}$, so $\lambda = 0$, and an eigenvalue exists over $\mathbb{C}$. (b) $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$.

(a) Suppose 0 is the only eigenvalue. By Corollary 7.5.14, $A = SRS^{-1}$ with R upper triangular, and the diagonal entries of R are eigenvalues of A (Proposition 7.5.8), hence all 0. So R is strictly upper triangular, and $R^n = 0$ by Example 2.8.3. Then $A^n = SR^nS^{-1} = 0$.

Conversely, suppose $A^n = 0$, and let $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Then $A^n\mathbf{v} = \lambda^n\mathbf{v}$ (induction), so $\lambda^n\mathbf{v} = \mathbf{0}$, $\lambda^n = 0$, and $\lambda = 0$. Since A has at least one eigenvalue (Theorem 7.3.27), 0 is the only one.

(b) Let $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$, which is block diagonal with blocks (0) and the rotation $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. A real eigenvector (x, y, z) with eigenvalue λ satisfies $0 = \lambda x$, $-z = \lambda y$, $y = \lambda z$, so $y = -\lambda^2 y$ and $z = -\lambda^2 z$, which forces $y = z = 0$, and then $x \neq 0$ and $\lambda = 0$. So 0 is the only real eigenvalue. But A^k is block diagonal with blocks (0) and R^k, and $R^k \neq 0$ since R is invertible. Over $\mathbb{C}$, A also has the eigenvalues $\pm i$, which is why (a) does not apply.

Solution to Exercise 7.5.37

Uses. Theorem 7.5.18 (Schur's theorem), Theorem 5.4.10 (Multiplicativity), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Schur's theorem gives a unitary U for each matrix separately. Does it give one U that works for both?
Hint 2. Try $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.
Sketch. The proof uses the same U for A and B, but Schur's theorem only provides a U depending on the matrix. The claim is false: $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ have only the eigenvalue 0, but $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has the eigenvalue 1.

Schur's theorem applied to A gives some unitary U, and applied to B gives some unitary U', which in general differs from U. The argument silently takes $U' = U$. A common triangularizing basis exists only in special situations, for instance for commuting matrices (Section 7.6).

The claim itself is false. Take $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$; both have 0 as their only eigenvalue, with multiplicity 2. But $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has the eigenvalues 1 and 0, while the claim predicts $0, 0$. What survives is the product of all of them: $\det(AB) = \det A \det B$ (Theorem 5.4.10), so by Theorem 7.5.22 the product of the eigenvalues of AB is $\prod \lambda_i \prod \mu_i$.

Solution to Exercise 7.5.38

Uses. Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Lemma 7.5.11 (Compressing an operator to a complement), Theorem 5.4.26 (Block-triangular matrices), Corollary 7.1.7 (No zero divisors, cancellation, and units), Proposition 4.4.16 (Complements exist), Theorem 6.4.24 (Orthogonal decomposition)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Follow the first proof of Theorem 7.5.13, but choose the first eigenvector for λ_1.
Hint 2. For the induction, you need the eigenvalues of the compression S. Show $p_T = (x - \lambda_1)p_S$ from the block form of Lemma 7.5.11(b).
Hint 3. For the orthonormal version, use $W = \operatorname{span}(\mathbf{b}_1)^\perp$ with $\mathbf{b}_1$ a unit vector, as in the second proof of Theorem 7.5.18.
Sketch. Induct on n. λ_1 is a root of p_T, hence an eigenvalue; take an eigenvector $\mathbf{b}_1$, a complement W, and the compression S. The block form $[T] = \begin{pmatrix} \lambda_1 & X \\ 0 & [S] \end{pmatrix}$ gives $p_T = (x - \lambda_1)p_S$, so $p_S = (x - \lambda_2)\cdots(x - \lambda_n)$. By induction W has a basis triangularizing S with diagonal $\lambda_2, \ldots, \lambda_n$, and the block form gives the diagonal $\lambda_1, \ldots, \lambda_n$ for T. With a unit $\mathbf{b}_1$ and $W = \operatorname{span}(\mathbf{b}_1)^\perp$ everything can be made orthonormal.

We induct on n; for $n = 1$ there is nothing to show. Let $n \geq 2$. Since $p_T = (x - \lambda_1)\cdots(x - \lambda_n)$, λ_1 is a root of p_T, hence an eigenvalue (Theorem 7.3.6). Let $\mathbf{b}_1$ be an eigenvector for λ_1, W a complement of $\operatorname{span}(\mathbf{b}_1)$ (Proposition 4.4.16), and S the compression of T to W. For a basis $\mathcal{D}$ of W and $\mathcal{B} = \mathbf{b}_1, \mathcal{D}$, Lemma 7.5.11(b) gives

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & X \\ 0 & [S]_{\mathcal{D}} \end{pmatrix},$$

so, taking $\det(xI_n - \cdot)$ with Theorem 5.4.26, $p_T = (x - \lambda_1)p_S$. Cancelling $x - \lambda_1$ from $(x - \lambda_1)p_S = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$ (Corollary 7.1.7) gives $p_S = (x - \lambda_2)\cdots(x - \lambda_n)$: the eigenvalues of S counted with multiplicity are $\lambda_2, \ldots, \lambda_n$. By the induction hypothesis, W has a basis $\mathcal{D} = \mathbf{b}_2, \ldots, \mathbf{b}_n$ with $[S]_{\mathcal{D}}$ upper triangular with diagonal $\lambda_2, \ldots, \lambda_n$ in this order. Then the displayed block matrix is upper triangular with diagonal $\lambda_1, \lambda_2, \ldots, \lambda_n$.

If V is an inner product space, choose $\mathbf{b}_1$ of norm 1 and $W = \operatorname{span}(\mathbf{b}_1)^\perp$, a complement by Theorem 6.4.24, and apply the inner product version of the induction hypothesis to W. Then $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is orthonormal.

Solution to Exercise 7.5.39

Uses. Theorem 6.4.38 (QR factorization), Corollary 7.5.14 (Triangularizing a matrix), Proposition 2.4.36 (Products of upper triangular matrices), Exercise 7.5.34, Definition 6.3.33 (Unitary and orthogonal matrices)
Hint 1. S is invertible, so its columns are independent and Theorem 6.4.38 applies, with Q square.
Hint 2. Substitute $S = QR'$ into $A = SRS^{-1}$. Inverses and products of invertible upper-triangular matrices are upper triangular (Exercise 7.5.34(b), Proposition 2.4.36).
Sketch. $A = QR'RR'^{-1}Q^{-1}$ and $Q^{-1} = Q^*$. The product $R'RR'^{-1}$ is upper triangular with diagonal $r'_{jj}r_{jj}r'^{-1}_{jj} = r_{jj}$. The columns of Q are Gram–Schmidt applied to the columns of S, the triangularizing basis, which is exactly the first proof.

The columns of S form a basis of $\mathbb{C}^n$, so by Theorem 6.4.38, $S = QR'$ with $Q \in \mathbb{C}^{n\times n}$ having orthonormal columns, hence unitary (Definition 6.3.33), and R' upper triangular with positive diagonal entries, hence invertible. Then

$$A = SRS^{-1} = QR'RR'^{-1}Q^{-1} = QR''Q^*.$$

By Exercise 7.5.34(b) (whose argument applies to any invertible upper-triangular matrix), R'^{-1} is upper triangular with diagonal entries $1/r'_{jj}$, and by Proposition 2.4.36, $R'' = R'RR'^{-1}$ is upper triangular with diagonal entries $r'_{jj}r_{jj}\left(1/r'_{jj}\right) = r_{jj}$.

By Theorem 6.4.38, the columns of Q are the output of Gram–Schmidt applied to the columns of S, which form a basis triangularizing T_A. So this is the first proof of Theorem 7.5.18: Gram–Schmidt applied to a triangularizing basis. The triangular factor R' records the change of basis between the two bases with the same flag, and conjugating by it keeps R triangular.

Solution to Exercise 7.5.40

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Exercise 7.5.34, Theorem 7.5.18 (Schur's theorem), Exercise 6.3.37, Proposition 6.3.27 (Properties of the conjugate transpose)
Hint 1. $R^{-1} = R^*$. One side is upper triangular and the other lower triangular.
Hint 2. Alternatively, look at the columns: the first column $(r_{11}, 0, \ldots, 0)$ is a unit vector, and every other column is orthogonal to it.
Hint 3. For (b), apply Schur's theorem and check that $R = U^*AU$ is unitary.
Sketch. (a) $R^* = R^{-1}$ is upper triangular (Exercise 7.5.34(b)) and lower triangular (as the conjugate transpose of an upper-triangular matrix), so it is diagonal, and so is R; then $R^*R = I_n$ gives $|r_{jj}|^2 = 1$. (b) Schur gives $A = URU^*$, and $R = U^*AU$ is a product of unitary matrices.

(a) Since R is unitary, $R^{-1} = R^*$. By (the argument of) Exercise 7.5.34(b), the inverse of an invertible upper-triangular matrix is upper triangular, so R^* is upper triangular. But the (i, j) entry of R^* is $\overline{r_{ji}}$, which is 0 for $i < j$, so R^* is also lower triangular. Hence R^* is diagonal, and so is $R = (R^*)^*$. Finally the (j, j) entry of $R^*R = I_n$ is $\overline{r_{jj}}r_{jj} = |r_{jj}|^2 = 1$.

(b) By Theorem 7.5.18, $A = URU^*$ with U unitary and R upper triangular. Then $R = U^*AU$ is a product of unitary matrices, so it is unitary (Exercise 6.3.37). By (a), R is diagonal; take $D = R$.

Solution to Exercise 7.5.41

Uses. Theorem 7.5.18 (Schur's theorem), Proposition 6.1.16 (Frobenius inner product), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.3.27 (Properties of the conjugate transpose), Proposition 7.5.8 (The diagonal of a triangular matrix)
Techniques. The trace of A^*A (Technique 6.1.18), Cyclicity of trace (Technique 2.4.29)
Hint 1. The right side is $\operatorname{tr}(A^*A)$ (Proposition 6.1.16). How does it change when A is replaced by U^*AU?
Hint 2. For a Schur form $R = U^*AU$, $\operatorname{tr}(A^*A) = \operatorname{tr}(R^*R) = \sum_{i,j}|r_{ij}|^2$, and the diagonal of R lists the eigenvalues.
Sketch. With $A = URU^*$, $A^*A = UR^*RU^*$, so $\operatorname{tr}(A^*A) = \operatorname{tr}(R^*RU^*U) = \operatorname{tr}(R^*R)$. Hence $\sum|a_{ij}|^2 = \sum|r_{ij}|^2 = \sum_j|\lambda_j|^2 + \sum_{i<j}|r_{ij}|^2$. Equality holds exactly when R is diagonal.

By Theorem 7.5.18, $A = URU^*$ with U unitary and R upper triangular with diagonal $\lambda_1, \ldots, \lambda_n$ (in some order). By Proposition 6.3.27, $A^* = UR^*U^*$, so $A^*A = UR^*U^*URU^* = UR^*RU^*$. By Proposition 6.1.16 and Proposition 2.4.28,

$$\sum_{i,j}|a_{ij}|^2 = \operatorname{tr}(A^*A) = \operatorname{tr}(U(R^*RU^*)) = \operatorname{tr}(R^*RU^*U) = \operatorname{tr}(R^*R) = \sum_{i,j}|r_{ij}|^2.$$

Since R is upper triangular,

$$\sum_{i,j}|r_{ij}|^2 = \sum_{j=1}^{n}|\lambda_j|^2 + \sum_{i<j}|r_{ij}|^2 \geq \sum_{j=1}^{n}|\lambda_j|^2,$$

which is the inequality.

If equality holds, then $r_{ij} = 0$ for all $i < j$, so R is diagonal and $A = URU^*$ has the required form. Conversely, if $A = UDU^*$ with D diagonal, then D is upper triangular, its diagonal lists the eigenvalues of A (Proposition 7.5.8, Proposition 7.5.9), and the computation above with $R = D$ gives equality. (In Section 8.2 such matrices are identified as the normal ones, $A^*A = AA^*$.)

Solution to Exercise 7.5.42

Uses. Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Lemma 7.5.11 (Compressing an operator to a complement), Theorem 5.4.26 (Block-triangular matrices), Corollary 7.1.7 (No zero divisors, cancellation, and units), Theorem 6.4.24 (Orthogonal decomposition), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Theorem 4.5.17 (Change of basis)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. The second proof of Theorem 7.5.18 only used $\mathbb{C}$ to find an eigenvalue. Over $\mathbb{R}$, with a split characteristic polynomial, an eigenvalue exists. What has to be checked for the compression?
Hint 2. Show that the compression S to $\operatorname{span}(\mathbf{u}_1)^\perp$ satisfies $p_A = (x-\lambda_1)p_S$, so p_S splits too.
Sketch. Prove the operator version on a real inner product space by induction on the dimension: a real root λ_1 of p_T gives a unit eigenvector $\mathbf{u}_1$; the compression S to $W = \operatorname{span}(\mathbf{u}_1)^\perp$ has $p_T = (x-\lambda_1)p_S$ by Lemma 7.5.11(b), so p_S splits; by induction W has an orthonormal basis triangularizing S, and $\mathbf{u}_1$ followed by it triangularizes T. Apply this to T_A on $\mathbb{R}^n$.

We prove: if V is a real inner product space with $1 \le \dim V = n < \infty$ and $T \in \mathrm{L}(V)$ has p_T splitting over $\mathbb{R}$, then V has an orthonormal basis triangularizing T. Applied to T_A on $\mathbb{R}^n$ with the dot product, with Q the matrix whose columns are that basis, this gives $Q^{\mathsf{T}}AQ = R$ upper triangular by Theorem 4.5.17, since $Q^{-1} = Q^{\mathsf{T}}$ for orthogonal Q.

Induct on n; for $n = 1$ any unit vector works. Let $n \ge 2$ and $p_T = (x-\mu_1)\cdots(x-\mu_n)$ with real μ_i. Then μ_1 is an eigenvalue (Theorem 7.3.6); let $\mathbf{u}_1$ be an eigenvector of norm 1. Let $W = \operatorname{span}(\mathbf{u}_1)^\perp$, a complement of $\operatorname{span}(\mathbf{u}_1)$ of dimension $n-1$ (Theorem 6.4.24), and let $S\mathbf{w} = T\mathbf{w} - \langle T\mathbf{w}, \mathbf{u}_1\rangle \mathbf{u}_1$, the compression of T to W (as in the second proof of Theorem 7.5.18). For any basis $\mathcal{D}$ of W, Lemma 7.5.11(b) and Theorem 5.4.26 give $p_T = (x-\mu_1)p_S$, and cancelling (Corollary 7.1.7) gives $p_S = (x-\mu_2)\cdots(x-\mu_n)$, which splits. By the induction hypothesis, W has an orthonormal basis $\mathbf{u}_2, \dots, \mathbf{u}_n$ with $S\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_2, \dots, \mathbf{u}_j)$. Exactly as in the second proof of Theorem 7.5.18, $\mathbf{u}_1, \dots, \mathbf{u}_n$ is an orthonormal basis of V and $T\mathbf{u}_j = \langle T\mathbf{u}_j, \mathbf{u}_1\rangle \mathbf{u}_1 + S\mathbf{u}_j \in \operatorname{span}(\mathbf{u}_1, \dots, \mathbf{u}_j)$, so T is upper triangular with respect to it (Proposition 7.5.2).

Solution to Exercise 7.5.43

Uses. Exercise 7.5.34, Exercise 7.5.36, Theorem 5.5.21 (The Vandermonde determinant), Corollary 3.5.25 (The invertible matrix theorem), Theorem 5.4.14 (The transpose)
Hint 1. Group the nonzero eigenvalues: let $\mu_1, \dots, \mu_r$ be the distinct nonzero ones, with algebraic multiplicities $m_1, \dots, m_r \ge 1$. By Exercise 7.5.36 it suffices to show $r = 0$.
Hint 2. By Exercise 7.5.34(a), $\sum_j m_j \mu_j^k = 0$ for $k = 1, \dots, r$. This is a square linear system for $(m_1, \dots, m_r)$.
Hint 3. Its matrix is a Vandermonde matrix times $\operatorname{diag}(\mu_1, \dots, \mu_r)$. For $\mathbb{F}_p$, try I_p.
Sketch. If there are distinct nonzero eigenvalues $\mu_1, \dots, \mu_r$ with multiplicities m_j, then $\sum_j m_j \mu_j^k = \operatorname{tr}\left(A^k\right) = 0$ for $k = 1, \dots, r$. The coefficient matrix $\left(\mu_j^k\right)_{k,j}$ has determinant $\mu_1 \cdots \mu_r \prod_{i<j}(\mu_j - \mu_i) \neq 0$, so all $m_j = 0$ in $\mathbb{C}$, which is absurd. So 0 is the only eigenvalue, and $A^n = 0$. Over $\mathbb{F}_p$, $\operatorname{tr}\left(I_p^k\right) = p * 1 = 0$ for all k, but $I_p^p \neq 0$.

Let $\mu_1, \dots, \mu_r$ be the distinct nonzero eigenvalues of A, with algebraic multiplicities $m_1, \dots, m_r \ge 1$, and suppose $r \ge 1$. The eigenvalues of A counted with multiplicity are these, each repeated m_j times, together with 0 repeated $n - \sum m_j$ times. By Exercise 7.5.34(a), for $k = 1, \dots, r$ (note $r \le n$),

$$0 = \operatorname{tr}\left(A^k\right) = m_1\mu_1^k + m_2\mu_2^k + \cdots + m_r\mu_r^k,$$

where m_j stands for $m_j * 1 \in \mathbb{C}$. So $M\mathbf{c} = \mathbf{0}$, where $\mathbf{c} = (m_1, \dots, m_r)$ and $M \in \mathbb{C}^{r\times r}$ has entries $M_{kj} = \mu_j^k$. Factoring μ_j out of column j, $M = V' \operatorname{diag}(\mu_1, \dots, \mu_r)$ with $V'_{kj} = \mu_j^{k-1}$, the transpose of the Vandermonde matrix of $\mu_1, \dots, \mu_r$. By Theorem 5.5.21 and Theorem 5.4.14,

$$\det M = \mu_1\mu_2\cdots\mu_r \prod_{i<j}\left(\mu_j - \mu_i\right) \neq 0,$$

since the μ_j are nonzero and distinct. So M is invertible (Corollary 3.5.25), $\mathbf{c} = \mathbf{0}$, and $m_1 = 0$ in $\mathbb{C}$, contradicting $m_1 \ge 1$. Hence $r = 0$: 0 is the only eigenvalue of A, and $A^n = 0$ by Exercise 7.5.36.

Over $\mathbb{F}_p$, take $A = I_p \in \mathbb{F}_p^{p\times p}$. Then $\operatorname{tr}\left(A^k\right) = \operatorname{tr} I_p = p * 1 = 0$ for every k, but $A^p = I_p \neq 0$. The proof breaks at its last step: in $\mathbb{F}_p$, the conclusion $m_1 * 1 = 0$ is no contradiction when p divides m_1, as here, where $m_1 = p$.

Solution to Exercise 7.5.44

Uses. Proposition 2.4.28 (Cyclicity of the trace), Proposition 2.4.27 (The trace is linear), Exercise 7.5.43
Techniques. Cyclicity of trace (Technique 2.4.29)
Hint 1. Compute $A^2B - BA^2$ in terms of A^2. Guess a formula for $A^kB - BA^k$.
Hint 2. Prove $A^kB - BA^k = kA^k$ by induction, then take the trace.
Sketch. By induction, $A^{k+1}B - BA^{k+1} = A\left(A^kB - BA^k\right) + (AB - BA)A^k = kA^{k+1} + A^{k+1}$, so $A^kB - BA^k = kA^k$. Taking traces, $k\operatorname{tr}\left(A^k\right) = 0$, so $\operatorname{tr}\left(A^k\right) = 0$ for all $k \geq 1$, and Exercise 7.5.43 gives $A^n = 0$.

We show $A^kB - BA^k = kA^k$ for $k \geq 1$ by induction. For $k = 1$ this is the hypothesis. If it holds for k, then

$$\begin{aligned} A^{k+1}B - BA^{k+1} &= A\left(A^kB - BA^k\right) + (AB - BA)A^k \\ &= A\left(kA^k\right) + AA^k = (k+1)A^{k+1}. \end{aligned}$$

Taking traces, using Proposition 2.4.27 and Proposition 2.4.28, $k\operatorname{tr}\left(A^k\right) = \operatorname{tr}\left(A^kB\right) - \operatorname{tr}\left(BA^k\right) = 0$, so $\operatorname{tr}\left(A^k\right) = 0$ for every $k \geq 1$, since $k \neq 0$ in $\mathbb{C}$. By Exercise 7.5.43, $A^n = 0$.

Solution to Exercise 7.5.44

Uses. Corollary 7.2.22 (At most $\dim V$ eigenvalues), Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29), Cyclicity of trace (Technique 2.4.29)
Hint 1. Consider the operator $L(X) = BX - XB$ on $\mathbb{C}^{n\times n}$. What does it do to A, to A^2, to A^k?
Hint 2. If no power of A were 0, L would have infinitely many eigenvalues.
Sketch. The identity $A^kB - BA^k = kA^k$ says $L\left(A^k\right) = -kA^k$ for $L(X) = BX - XB$. If $A^k \neq 0$ for every k, then $-1, -2, -3, \ldots$ are all eigenvalues of L, an operator on the n^2-dimensional space $\mathbb{C}^{n\times n}$, which has at most n^2 eigenvalues. So some $A^k = 0$, and then $A^n = 0$.

As in the first approach, $A^kB - BA^k = kA^k$ for all $k \geq 1$. Let $L \in \mathrm{L}(\mathbb{C}^{n\times n})$ be $L(X) = BX - XB$. Then $L\left(A^k\right) = -kA^k$. If $A^k \neq 0$ for every $k \geq 1$, then $-1, -2, \ldots, -\left(n^2+1\right)$ are n^2+1 distinct eigenvalues of L, which is impossible, since $\dim \mathbb{C}^{n\times n} = n^2$ (Corollary 7.2.22). So $A^k = 0$ for some k, and then $A^n = 0$ (Corollary 4.1.27).

Solution to Exercise 7.5.45

Uses. Lemma 2.4.24 (Matrix units), Proposition 7.5.2 (Upper-triangular matrices and invariant subspaces), Proposition 7.5.9 (Multiplicities on the diagonal), Corollary 7.5.14 (Triangularizing a matrix), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29), Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. Compute AU_{ij} (column i of A, moved to column j) and $U_{ij}A$ (row j of A, moved to row i), using that A is upper triangular.
Hint 2. Which matrix units appear in $L_A\left(U_{ij}\right)$ besides U_{ij}? Do they come before U_{ij} in the given order?
Hint 3. For (b), if $A = SRS^{-1}$, then L_A is conjugate to L_R by the invertible operator $X \mapsto SXS^{-1}$.
Sketch. (a) $AU_{ij} = \sum_{k\leq i} a_{ki}U_{kj}$ and $U_{ij}A = \sum_{l\geq j} a_{jl}U_{il}$, so $L_A\left(U_{ij}\right) = \left(a_{ii} - a_{jj}\right)U_{ij}$ plus units U_{kj} with $k < i$ (earlier rows) and U_{il} with $l > j$ (earlier in the same row). (b) With $A = SRS^{-1}$ and $C(X) = SXS^{-1}$, $L_A \circ C = C \circ L_R$, so L_A is upper triangular in the basis $C\left(U_{ij}\right)$ with diagonal $r_{ii} - r_{jj}$. (c) The sum of all $\lambda_i - \lambda_j$ is 0, and the terms with $i = j$ are 0, so $\det L_A = 0$.

(a) By Lemma 2.4.24, $AU_{ij} = (A\mathbf{e}_i)\mathbf{e}_j^{\mathsf{T}} = \sum_k a_{ki}U_{kj}$ and $U_{ij}A = \mathbf{e}_i\left(\mathbf{e}_j^{\mathsf{T}}A\right) = \sum_l a_{jl}U_{il}$. Since A is upper triangular, $a_{ki} = 0$ for $k > i$ and $a_{jl} = 0$ for $l < j$, so

$$L_A\left(U_{ij}\right) = \left(a_{ii} - a_{jj}\right)U_{ij} + \sum_{k<i} a_{ki}U_{kj} - \sum_{l>j} a_{jl}U_{il}.$$

In the given order, every U_{kj} with $k < i$ lies in an earlier row, and every U_{il} with $l > j$ lies earlier in the same row. So $L_A\left(U_{ij}\right)$ lies in the span of U_{ij} and the units before it, and by Proposition 7.5.2, L_A is upper triangular in this basis, with diagonal entry $a_{ii} - a_{jj}$ at U_{ij}.

(b) By Corollary 7.5.14, $A = SRS^{-1}$ with R upper triangular, $r_{ii} = \lambda_i$ after reordering. Let $C \in \mathrm{L}(\mathbb{C}^{n\times n})$ be $C(X) = SXS^{-1}$, which is invertible with inverse $X \mapsto S^{-1}XS$. For every X,

$$L_A(C(X)) = SRS^{-1}SXS^{-1} - SXS^{-1}SRS^{-1} = S(RX - XR)S^{-1} = C(L_R(X)).$$

So if $\mathcal{U}$ is the ordered basis of (a), then $C(\mathcal{U})$ is a basis (the image of a basis under an isomorphism), and L_A sends $C(U_{ij})$ to $C(L_R(U_{ij}))$, which has the same coordinates in $C(\mathcal{U})$ as $L_R(U_{ij})$ has in $\mathcal{U}$. Hence $[L_A]_{C(\mathcal{U})} = [L_R]_{\mathcal{U}}$ is upper triangular with diagonal entries $\lambda_i - \lambda_j$, and by Proposition 7.5.9 these are the eigenvalues of L_A counted with multiplicity.

(c) By Theorem 7.5.22, $\operatorname{tr} L_A = \sum_{i,j}(\lambda_i - \lambda_j) = n\sum_i \lambda_i - n\sum_j \lambda_j = 0$, and $\det L_A = \prod_{i,j}(\lambda_i - \lambda_j) = 0$, because the factors with $i = j$ vanish. So L_A is not invertible. (Directly: $L_A(I_n) = 0$.)

Solution to Exercise 7.5.46

Uses. Theorem 7.5.27 (Real block upper-triangular form), Lemma 7.5.11 (Compressing an operator to a complement)
Hint 1. Split $(1, 0, -i)$ into real and imaginary parts $\mathbf{x} + i\mathbf{y}$ and compute $A\mathbf{x}$ and $A\mathbf{y}$.
Hint 2. For (b), add a real eigenvector. For (c), the invariant plane already has an orthonormal basis; compress to its orthogonal complement.
Sketch. $A(1, 0, -i) = (i, 0, 1) = i(1, 0, -i)$. With $\mathbf{x} = \mathbf{e}_1$, $\mathbf{y} = -\mathbf{e}_3$: $A\mathbf{x} = -\mathbf{y}$, $A\mathbf{y} = \mathbf{x}$, so $\operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$ is invariant. (b) $(3, 2, 1)$ is an eigenvector for 1, and in the basis $\mathbf{x}, \mathbf{y}, (3, 2, 1)$ the matrix is $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. (c) Q with columns $\mathbf{e}_1, -\mathbf{e}_3, \mathbf{e}_2$ gives $Q^{\mathsf{T}}AQ = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

(a) $A(1, 0, -i) = (0 + 0 + i, 0, 1 + 0 + 0) = (i, 0, 1) = i(1, 0, -i)$, since $i(-i) = 1$. Write $(1, 0, -i) = \mathbf{x} + i\mathbf{y}$ with $\mathbf{x} = (1, 0, 0)$ and $\mathbf{y} = (0, 0, -1)$. Then

$$A\mathbf{x} = (0, 0, 1) = -\mathbf{y}, \quad A\mathbf{y} = (1, 0, 0) = \mathbf{x},$$

as in the complexification argument before Theorem 7.5.27, with $a = 0$ and $b = 1$. So $\operatorname{span}(\mathbf{x}, \mathbf{y}) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$ is invariant, and A acts on it by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ in the basis $\mathbf{x}, \mathbf{y}$, a block with eigenvalues $\pm i$.

(b) For the eigenvalue 1, $A - I_3 = \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix}$; adding the first and last rows gives $y - 2z = 0$, so $y = 2z$ and $x = y + z = 3z$, and $\mathbf{v} = (3, 2, 1)$ is an eigenvector. In the basis $\mathbf{x}, \mathbf{y}, \mathbf{v}$,

$$[T_A] = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

a 2×2 block without real eigenvalues followed by a 1×1 block.

(c) The plane $U = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_3)$ has the orthonormal basis $\mathbf{e}_1, -\mathbf{e}_3$, and $U^{\perp} = \operatorname{span}(\mathbf{e}_2)$. Let Q have columns $\mathbf{e}_1, -\mathbf{e}_3, \mathbf{e}_2$:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Its columns are orthonormal, so Q is orthogonal. The columns of $Q^{\mathsf{T}}AQ$ are the coordinates of $A\mathbf{e}_1 = -(-\mathbf{e}_3)$, $A(-\mathbf{e}_3) = \mathbf{e}_1$, and $A\mathbf{e}_2 = (2, 1, -1) = 2\mathbf{e}_1 + (-\mathbf{e}_3) + \mathbf{e}_2$ in this basis:

$$Q^{\mathsf{T}}AQ = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The compression of T_A to $U^{\perp}$ is multiplication by 1 on $\operatorname{span}(\mathbf{e}_2)$, which is the bottom right entry.

Section 7.6

Solution to Exercise 7.6.6

Uses. Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Proposition 2.4.23 (Transpose of a product), Remark 7.1.11 (Polynomials in a matrix), Definition 7.1.9 (Substitution)
Hint 1. Show that A and A^{T} have exactly the same annihilating polynomials. What does that say about their minimal polynomials?
Hint 2. For (b), check on $p = x^k$ that the polynomial $q(x) = p(x - c)$ satisfies $q(A + cI_n) = p(A)$, then extend by linearity.
Sketch. (a) $p\left(A^{\mathsf{T}}\right) = p(A)^{\mathsf{T}}$, so the annihilating polynomials of A and A^{T} coincide, and so do their monic generators. (b) Substituting $x - c$ is a bijection of $F[x]$ preserving degrees and monicity, and $q(A + cI_n) = p(A)$ for $q(x) = p(x - c)$; so it maps the annihilating polynomials of A onto those of $A + cI_n$. Hence $m_{\lambda I_n + N} = (x - \lambda)^n$.

(a) By Proposition 2.4.23, $\left(A^k\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^k$ for every $k \geq 0$, and transposition respects sums and scalar multiples, so $p\left(A^{\mathsf{T}}\right) = p(A)^{\mathsf{T}}$ for every $p \in F[x]$. A matrix is 0 if and only if its transpose is, so p annihilates A^{T} if and only if it annihilates A. By Theorem 7.6.2, m_A divides every polynomial that annihilates A^{T}, in particular $m_{A^{\mathsf{T}}}$, and likewise $m_{A^{\mathsf{T}}} \mid m_A$. Two monic polynomials that divide each other have the same degree, so the quotient is a constant, which is 1 by comparing leading coefficients. Hence $m_{A^{\mathsf{T}}} = m_A$.

(b) For $p \in F[x]$ let $q(x) = p(x - c)$, the result of substituting $x - c$ into p (Definition 7.1.9). Both $p \mapsto q(A + cI_n)$ and $p \mapsto p(A)$ respect sums and scalar multiples, and for $p = x^k$, using Remark 7.1.11,

$$q(A + cI_n) = \left((x - c)^k\right)(A + cI_n) = ((A + cI_n) - cI_n)^k = A^k.$$

So $q(A + cI_n) = p(A)$ for every p. Substituting $x - c$ is a bijection of $F[x]$ (its inverse substitutes $x + c$) that preserves degrees and leading coefficients. Hence it maps the monic annihilating polynomials of A bijectively onto those of $A + cI_n$, preserving degrees, and the one of least degree goes to the one of least degree: $m_{A+cI_n}(x) = m_A(x - c)$.

For N, Example 7.6.5(b) gives $m_N = x^n$, so $m_{\lambda I_n + N} = (x - \lambda)^n$.

Solution to Exercise 7.6.19

Uses. Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial)
Techniques. Use an annihilating polynomial (Technique 7.6.17)
Hint 1. Compute the squares.
Hint 2. None of the three is a scalar matrix, so each minimal polynomial has degree at least 2.
Sketch. $A^2 = 3A$, $B^2 = 0$, $C^2 = I_2$. None is scalar, so $m_A = x(x - 3)$, $m_B = x^2$, and $m_C = (x - 1)(x + 1)$. Hence A and C are diagonalizable and B is not.

Every entry of A^2 is $1 + 1 + 1 = 3$, so $A^2 = 3A$, and $x^2 - 3x = x(x - 3)$ annihilates A. Next, $B^2 = \begin{pmatrix} 1-1 & -1+1 \\ 1-1 & -1+1 \end{pmatrix} = 0$, and $C^2 = \begin{pmatrix} 9-8 & -6+6 \\ 12-12 & -8+9 \end{pmatrix} = I_2$.

None of A, B, C is a scalar multiple of the identity, so no monic polynomial of degree 1 annihilates them, and each minimal polynomial has degree 2. Since the minimal polynomial divides the quadratic found above (Theorem 7.6.2) and both are monic of degree 2, they are equal: $m_A = x(x - 3)$, $m_B = x^2$, $m_C = (x - 1)(x + 1)$. By Theorem 7.6.15, A (eigenvalues 0 and 3) and C (eigenvalues 1 and -1) are diagonalizable over $\mathbb{Q}$, and B is not, since x^2 has a repeated root.

Solution to Exercise 7.6.27

Uses. Theorem 7.6.20 (Cayley–Hamilton), Proposition 7.3.18 (The first and last coefficients), Corollary 7.6.9 (Inverses are polynomials)
Techniques. Use an annihilating polynomial (Technique 7.6.17)
Hint 1. $p_A = x^2 - (\operatorname{tr} A)x + \det A$.
Hint 2. From $A^2 = 5A + 2I_2$, factor out A to find A^{-1}, and multiply by A to find A^3.
Sketch. $p_A = x^2 - 5x - 2$, so $A^2 = 5A + 2I_2$. Then $A(A - 5I_2) = 2I_2$ gives $A^{-1} = \frac{1}{2}(A - 5I_2)$, and $A^3 = 5A^2 + 2A = 27A + 10I_2$.

Here $\operatorname{tr} A = 5$ and $\det A = 4 - 6 = -2$, so $p_A = x^2 - 5x - 2$ (Proposition 7.3.18), and Cayley–Hamilton gives $A^2 = 5A + 2I_2$. Then $A(A - 5I_2) = 2I_2$, and since $A - 5I_2$ commutes with A,

$$A^{-1} = \frac{1}{2}(A - 5I_2) = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}.$$

Next, $A^3 = A \cdot A^2 = 5A^2 + 2A = 5(5A + 2I_2) + 2A = 27A + 10I_2 = \begin{pmatrix} 37 & 54 \\ 81 & 118 \end{pmatrix}$.

Solution to Exercise 7.6.40

Uses. Corollary 7.6.24 (The minimal polynomial divides the characteristic polynomial), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Proposition 7.3.15 (Block-triangular matrices)
Hint 1. The minimal polynomial divides p and has the same roots. List the candidates and test them from the smallest degree up.
Hint 2. For C, the characteristic polynomial factors along the blocks.
Sketch. $p_A = (x-1)^3$, $A - I \neq 0 = (A-I)^2$, so $m_A = (x-1)^2$. $p_B = x^3$, $B \neq 0 = B^2$, so $m_B = x^2$. $p_C = x(x-2)^2$ and $C(C - 2I) = 0$, so $m_C = x(x-2)$. Only C is diagonalizable.

The matrix A**.** It is upper triangular with diagonal $1, 1, 1$, so $p_A = (x-1)^3$. By Corollary 7.6.24(c), m_A is $(x-1)^k$ for some $1 \leq k \leq 3$. Now $A - I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$ and $(A-I)^2 = 0$, so $m_A = (x-1)^2$. It has a repeated root, so A is not diagonalizable (Theorem 7.6.15).

The matrix B**.** It is strictly upper triangular, so $p_B = x^3$. Since $B \neq 0$ and $B^2 = 0$ (the only nonzero row of B is the first, and it picks out the third row of B, which is zero), $m_B = x^2$, and B is not diagonalizable.

The matrix C**.** It is block diagonal with blocks (2) and $K = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, so $p_C = (x-2)p_K = (x-2)\big(x^2 - 2x\big) = x(x-2)^2$ (Proposition 7.3.15). The candidates are $x(x-2)$ and $x(x-2)^2$. Since $K^2 = 2K$,

$$C(C - 2I) = \begin{pmatrix} 2 \cdot 0 & 0 \\ 0 & K(K - 2I_2) \end{pmatrix} = 0,$$

so $m_C = x(x-2)$, and C is diagonalizable, with eigenvalues $0, 2, 2$.

Solution to Exercise 7.6.41

Uses. Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Try small cases and small fields (Technique 1.2.38)
Hint 1. Every entry of J^2 is a sum of n ones. Is J a scalar matrix?
Hint 2. $R^2 = -I_2$ over every field. Factor $x^2 + 1$ over each $\mathbb{F}_p$ by trying all p values.
Sketch. (a) $J^2 = nJ$, and J is not scalar, so $m_J = x(x-n)$. If $n \neq 0$ in F, the roots $0, n$ are distinct and J is diagonalizable; if $n = 0$ in F, $m_J = x^2$ and it is not. (b) $m_R = x^2 + 1$ in each case. Over $\mathbb{F}_2$ it is $(x+1)^2$: not diagonalizable. Over $\mathbb{F}_3$ it has no roots: no eigenvalues. Over $\mathbb{F}_5$ it is $(x-2)(x-3)$: diagonalizable, with eigenvectors $(1,3)$ for 2 and $(1,2)$ for 3.

(a) The (i,j) entry of J^2 is $\sum_{l=1}^{n} 1 \cdot 1 = n$, so $J^2 = nJ$, and $x^2 - nx = x(x-n)$ annihilates J. Since $n \geq 2$, J has nonzero off-diagonal entries, so it is not a scalar matrix and $\deg m_J \geq 2$. Hence $m_J = x(x-n)$ (Theorem 7.6.2). If $n \neq 0$ in F, the roots 0 and n are distinct and J is diagonalizable, with eigenvalues 0 and n (Theorem 7.6.15). If $n = 0$ in F (for instance $F = \mathbb{F}_p$ with p dividing n), then $m_J = x^2$ and J is not diagonalizable; it is a nonzero matrix with $J^2 = 0$. For example, over $\mathbb{F}_2$, $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}^2 = 0$.

(b) Over every field $R^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2$, and R is not scalar, so $m_R = x^2 + 1$.

Over $\mathbb{F}_2$, $x^2 + 1 = x^2 - 1 = (x-1)^2$, with a repeated root, so R is not diagonalizable over $\mathbb{F}_2$.

Over $\mathbb{F}_3$, $x^2 + 1$ takes the values $1, 2, 2$ at $0, 1, 2$, so it has no roots, R has no eigenvalues (Theorem 7.6.7), and R is not diagonalizable.

Over $\mathbb{F}_5$, $2^2 + 1 = 5 = 0$ and $3^2 + 1 = 10 = 0$, so $x^2 + 1 = (x-2)(x-3)$ with distinct roots, and R is diagonalizable. If $R(1,a) = (-a, 1) = \lambda(1,a)$, then $\lambda = -a$ and $1 = \lambda a = -a^2$. For $a = 3$: $\lambda = -3 = 2$, and $R(1,3) =$

$(-3,1) = (2,1) = 2(1,3)$ in $\mathbb{F}_5$. For $a = 2$: $\lambda = -2 = 3$, and $R(1,2) = (-2,1) = (3,1) = 3(1,2)$. So $(1,3),(1,2)$ is an eigenbasis, with eigenvalues $2, 3$.

Solution to Exercise 7.6.42

Uses. Corollary 7.6.24 (The minimal polynomial divides the characteristic polynomial), Theorem 7.6.20 (Cayley–Hamilton), Corollary 7.6.9 (Inverses are polynomials), Exercise 7.6.6
Techniques. Use an annihilating polynomial (Technique 7.6.17)
Hint 1. A is triangular. For m_A, test the candidate $x(x-1)$.
Hint 2. For (b), Cayley–Hamilton gives $A^3 = 2A^2 - A$. Induct on k.
Hint 3. $m_B(x) = m_A(x-1)$. Expand it, and use Corollary 7.6.9.
Sketch. (a) $p_A = x(x-1)^2$ and $A(A-I) \neq 0$, so $m_A = p_A$. (b) $A^3 = 2A^2 - A$; if $A^k = (k-1)A^2 + (2-k)A$, then $A^{k+1} = (k-1)A^3 + (2-k)A^2 = kA^2 + (1-k)A$. (c) $m_B = (x-1)(x-2)^2 = x^3 - 5x^2 + 8x - 4$, so $B^{-1} = \frac{1}{4}\big(B^2 - 5B + 8I_3\big)$.

(a) A is upper triangular with diagonal $1, 1, 0$, so $p_A = x(x-1)^2 = x^3 - 2x^2 + x$. By Corollary 7.6.24(c), m_A is $x(x-1)$ or $x(x-1)^2$. Now

$$A(A - I_3) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0,$$

so $m_A = x(x-1)^2 = p_A$.

(b) By Cayley–Hamilton, $A^3 - 2A^2 + A = 0$. We induct on k. For $k = 1$ the formula reads $A = 0 \cdot A^2 + A$. If $A^k = (k-1)A^2 + (2-k)A$, then

$$A^{k+1} = (k-1)A^3 + (2-k)A^2 = (k-1)\big(2A^2 - A\big) + (2-k)A^2 = kA^2 + (1-k)A,$$

which is the formula for $k+1$. (Equivalently, $(k-1)x^2 + (2-k)x$ is the remainder of x^k on division by $x(x-1)^2$.)

(c) By Exercise 7.6.6(b), $m_B(x) = m_A(x-1) = (x-1)(x-2)^2 = x^3 - 5x^2 + 8x - 4$. Since $m_B(0) = -4 \neq 0$, B is invertible, and writing $m_B = x\big(x^2 - 5x + 8\big) - 4$, Corollary 7.6.9 gives

$$B^{-1} = \frac{1}{4}\big(B^2 - 5B + 8I_3\big).$$

With $B = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ and $B^2 = \begin{pmatrix} 4 & 4 & 4 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$,

$$B^{-1} = \frac{1}{4}\begin{pmatrix} 2 & -1 & -1 \\ 0 & 2 & -2 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/4 & -1/4 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiplying by B confirms this.

Solution to Exercise 7.6.43

Uses. Remark 7.6.11 (Neither polynomial determines the other), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Proposition 4.5.27 (Similarity invariants)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Rule out similarity with an invariant (Technique 4.5.29), Use an annihilating polynomial (Technique 7.6.17)
Hint 1. For (b), look at 4×4 matrices with $A^2 = 0$. Which similarity invariant can still tell them apart?
Hint 2. For (c), think of a nilpotent matrix.
Hint 3. For (d), if $m_{A^2} = (x-\mu_1)\cdots(x-\mu_m)$, which polynomial annihilates A? Are its roots distinct?
Sketch. (a) False: I_2 and $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. (b) False: with $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, the block matrices $\operatorname{diag}(N, 0)$ and $\operatorname{diag}(N, N)$ both have $p = x^4$ and $m = x^2$ but ranks 1 and 2. (c) False: $N^2 = 0$. (d) True: A is killed by $\prod(x^2 - \mu_i)$, whose $2m$ roots $\pm\sqrt{\mu_i}$ are distinct because the μ_i are distinct and nonzero.

(a) False. I_2 and $J = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ both have characteristic polynomial $(x-1)^2$, but $m_{I_2} = x - 1$ and $m_J = (x-1)^2$ (Remark 7.6.11).

(b) False. Let $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, and in $F^{4\times 4}$ let $A = \begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}$ in block form. Both are strictly upper triangular, so $p_A = p_B = x^4$. Both are nonzero with $A^2 = B^2 = 0$, so $m_A = m_B = x^2$. But $\operatorname{rank} A = 1$ and $\operatorname{rank} B = 2$, and similar matrices have the same rank (Proposition 4.5.27), so A and B are not similar.

(c) False. $N^2 = 0$ is diagonal, but N is not diagonalizable, since $m_N = x^2$.

(d) True. Since A^2 is diagonalizable, $m_{A^2} = (x-\mu_1)\cdots(x-\mu_m)$ with distinct $\mu_i \in \mathbb{C}$ (Theorem 7.6.15), and every μ_i is an eigenvalue of A^2. As A is invertible, so is A^2, so every $\mu_i \neq 0$. Choose $\nu_i \in \mathbb{C}$ with $\nu_i^2 = \mu_i$. Then

$$q = (x^2-\mu_1)\cdots(x^2-\mu_m) = \prod_{i=1}^{m}(x-\nu_i)(x+\nu_i)$$

satisfies $q(A) = m_{A^2}(A^2) = 0$. Its $2m$ roots $\pm\nu_i$ are distinct: $\nu_i \neq -\nu_i$ because $\nu_i \neq 0$, and for $i \neq j$ the roots $\pm\nu_i$ and $\pm\nu_j$ have different squares $\mu_i \neq \mu_j$. By Theorem 7.6.15(III), A is diagonalizable. (Invertibility was needed: in (c), $\mu_1 = 0$ and x^2 has a double root.)

Solution to Exercise 7.6.44

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 7.6.20 (Cayley–Hamilton)
Hint 1. Every step is correct except one. Do the eigenvectors of A always span $\mathbb{C}^n$?
Hint 2. Try $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.
Sketch. The eigenvectors of A span $\mathbb{C}^n$ only if A is diagonalizable; for $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ they span only $\operatorname{span}(\mathbf{e}_1)$, and the argument says nothing about $p_N(N)\mathbf{e}_2$. So the argument proves Cayley–Hamilton exactly for diagonalizable matrices. The perturbation proof extends the diagonalizable case to all matrices.

The computation $p_A(A)\mathbf{v} = p_A(\lambda)\mathbf{v} = \mathbf{0}$ is correct, and so is the principle that a linear map vanishing on a spanning set is zero. The error is the unstated claim that the eigenvectors of A span $\mathbb{C}^n$. They do if and only if $\mathbb{C}^n$ has a basis of eigenvectors, that is, if and only if A is diagonalizable (Theorem 7.4.2). For $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ the eigenvectors are the nonzero multiples of $\mathbf{e}_1$, and the argument gives no information about $p_N(N)\mathbf{e}_2$. (Here $p_N(N) = N^2 = 0$ anyway, but the argument does not show it.)

So the argument is a valid proof for diagonalizable A. The fourth proof of Theorem 7.6.20 starts from exactly this case and then reaches every matrix: it moves A to diagonalizable matrices A_t, and uses that the entries of $p_{A_t}(A_t)$ are polynomials in t that vanish for all but finitely many t, hence also at $t = 0$.

Solution to Exercise 7.6.45

Uses. Proposition 7.6.12 (Distinct linear factors force a diagonalization), Theorem 7.6.15 (Diagonalizability and the minimal polynomial)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Lagrange interpolation in T (Technique 7.6.13)
Hint 1. Factor $x^2 - x - 2$. When are its roots distinct?
Hint 2. The projection onto E_2 is $L(T)$, where L is the Lagrange polynomial with $L(2) = 1$ and $L(-1) = 0$.
Hint 3. Over $\mathbb{F}_3$, $x^2 - x - 2 = (x+1)^2$. Look for $A = -I_2 + N$ with $N^2 = 0$.
Sketch. (a) $x^2 - x - 2 = (x-2)(x+1)$, and $2 - (-1) = 3 \neq 0$, so the roots are distinct; Proposition 7.6.12 gives $V = E_2 \oplus E_{-1}$, and the projection is $L(T) = \frac{1}{3}(T + \mathrm{id}_V)$. (b) Over $\mathbb{F}_3$ the polynomial is $(x+1)^2$, and $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ works.

(a) The polynomial $q = x^2 - x - 2 = (x-2)(x+1)$ annihilates T. Its roots 2 and -1 differ by $3 \neq 0$, so they are distinct, and Proposition 7.6.12 gives $V = E_2 \oplus E_{-1}$; T is diagonalizable with eigenvalues in $\{2, -1\}$. The Lagrange polynomial of the node 2 (with the other node -1) is $L = (x+1)/3$, so $P = L(T) = \frac{1}{3}(T + \mathrm{id}_V)$. Indeed, P acts on E_2 as $(2+1)/3 = 1$ and on E_{-1} as $(-1+1)/3 = 0$, which is the projection onto E_2 along E_{-1}.

(b) In $\mathbb{F}_3$, $-1 = 2$ and $-2 = 1$, so $x^2 - x - 2 = x^2 + 2x + 1 = (x+1)^2$. Let $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = -I_2 + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then

$$A^2 = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A + 2I_2 = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

so $A^2 = A + 2I_2$. Since $A \neq -I_2$, $m_A = (x+1)^2$, and A is not diagonalizable (Theorem 7.6.15).

Solution to Exercise 7.6.46

Uses. Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 7.6.20 (Cayley–Hamilton), Proposition 7.3.18 (The first and last coefficients)
Techniques. Use an annihilating polynomial (Technique 7.6.17)
Hint 1. $q = x^2 + x + 1$ has no root in $\mathbb{F}_2$. What does that say about m_A, given that $m_A \mid q$?
Hint 2. So $m_A = q$, and Cayley–Hamilton forces $p_A = q$. Translate $p_A = x^2 + x + 1$ into conditions on $\operatorname{tr} A$ and $\det A$.
Sketch. (a) q has no roots in $\mathbb{F}_2$, so m_A, a divisor of q of degree at least 1, cannot be linear: $m_A = q$. By Cayley–Hamilton $q \mid p_A$, so $p_A = q$, that is, $\operatorname{tr} A = 1$ and $\det A = 1$. Conversely these give $A^2 + A + I_2 = 0$. The solutions are $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. (b) $A^2 = A + I_2$, $A(A + I_2) = I_2$, and $(A + I_2)^2 = A$.

(a) Let $q = x^2 + x + 1 \in \mathbb{F}_2[x]$; $q(0) = q(1) = 1$, so q has no roots in $\mathbb{F}_2$. Suppose $q(A) = 0$. Then $m_A \mid q$. If m_A had degree 1, say $m_A = x - c$, then c would be a root of q. So $m_A = q$. By Theorem 7.6.20, $m_A \mid p_A$, and both are monic of degree 2, so $p_A = q$. By Proposition 7.3.18, $p_A = x^2 - (\operatorname{tr} A)x + \det A$, so $\operatorname{tr} A = -1 = 1$ and $\det A = 1$. Conversely, if $\operatorname{tr} A = \det A = 1$, then $p_A = q$, and Cayley–Hamilton gives $q(A) = 0$.

So we need $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a + d = 1$ and $ad - bc = 1$. The first condition gives $(a, d) = (1, 0)$ or $(0, 1)$; in both cases $ad = 0$, so $bc = -1 = 1$, forcing $b = c = 1$. The solutions are

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

(b) From $A^2 + A + I_2 = 0$ and $-1 = 1$ in $\mathbb{F}_2$ we get $A^2 = A + I_2$. Sums: $A + A = 0$, $A + I_2$ is in the list, and $A + (A + I_2) = I_2$, $I_2 + (A + I_2) = A$; the rest are clear. Products: $A^2 = A + I_2$, $A(A + I_2) = A^2 + A = I_2$, and $(A + I_2)^2 = A^2 + I_2 = A$; products with 0 and I_2 are clear. So $A^{-1} = A + I_2$, $(A + I_2)^{-1} = A$, and $I_2^{-1} = I_2$. These four matrices form a field with four elements, $\mathbb{F}_4$ (Section 14.3).

Solution to Exercise 7.6.47

Uses. Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.6.20 (Cayley–Hamilton), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Proposition 7.1.22 (Counting roots), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
Techniques. Krylov sequences must become dependent (Technique 7.3.28)
Hint 1. For (a): if $\mathbf{v}$ is cyclic, $\mathbf{v}, T\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is a basis, so no nonzero polynomial of degree less than n kills $\mathbf{v}$.
Hint 2. For "$\Rightarrow$" in (b), combine (a) with Theorem 7.6.15.
Hint 3. For "$\Leftarrow$", try the sum $\mathbf{v}$ of an eigenbasis. What is $q(T)\mathbf{v}$ in that basis?
Sketch. (a) By Proposition 7.2.35, no nonzero polynomial of degree less than n kills a cyclic $\mathbf{v}$, so $\deg m_T \geq n$; with $m_T \mid p_T$, $m_T = p_T$. (b) If T is diagonalizable and cyclic, $m_T = \prod(x - \lambda_i)$ over the distinct eigenvalues has degree n. Conversely, with an eigenbasis $\mathbf{v}_i$ for distinct λ_i and $\mathbf{v} = \sum \mathbf{v}_i$, $q(T)\mathbf{v} = \sum q(\lambda_i)\mathbf{v}_i$ vanishes only if q has the n roots λ_i, so no nonzero q of degree less than n kills $\mathbf{v}$, and $\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is independent.

(a) Let $\mathbf{v}$ be cyclic. Then $\mathbf{v} \neq \mathbf{0}$ and $\dim K_T(\mathbf{v}) = n$, so by Proposition 7.2.35 the Krylov sequence first becomes dependent at $d = n$, and no nonzero polynomial of degree less than n sends $\mathbf{v}$ to $\mathbf{0}$. Since $m_T(T)\mathbf{v} = \mathbf{0}$, $\deg m_T \geq n$. By Theorem 7.6.20, $m_T \mid p_T$, which is monic of degree n, so $m_T = p_T$.

(b) Suppose T is diagonalizable with a cyclic vector. By (a), $\deg m_T = n$, and by Theorem 7.6.15, m_T is the product of $x - \lambda$ over the distinct eigenvalues λ. So T has n distinct eigenvalues.

Conversely, let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors for distinct eigenvalues $\lambda_1, \ldots, \lambda_n$; they form a basis (Corollary 7.4.5). Let $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_n$. For $q \in F[x]$, Proposition 7.2.28 gives

$$q(T)\mathbf{v} = q(\lambda_1)\mathbf{v}_1 + \cdots + q(\lambda_n)\mathbf{v}_n,$$

which is $\mathbf{0}$ only if $q(\lambda_i) = 0$ for every i. A nonzero polynomial of degree less than n has fewer than n roots (Proposition 7.1.22), so no such polynomial kills $\mathbf{v}$. A linear relation among $\mathbf{v}, T\mathbf{v}, \ldots, T^{n-1}\mathbf{v}$ is a polynomial of degree less than n killing $\mathbf{v}$, so these n vectors are independent, hence a basis of V, and $K_T(\mathbf{v}) = V$.

Solution to Exercise 7.6.48

Uses. Definition 7.2.33 (Krylov subspace), Proposition 7.2.35 (A basis of a Krylov subspace), Proposition 7.2.27 (Evaluating polynomials at an operator), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 7.6.4 (The minimal polynomial is the first dependence among the powers), Exercise 7.6.47
Techniques. Check it on a basis (Technique 2.2.7)
Hint 1. $S\mathbf{v}$ lies in $K_T(\mathbf{v}) = V$, so $S\mathbf{v} = q(T)\mathbf{v}$ for some q.
Hint 2. Compare S and $q(T)$ on the basis $\mathbf{v}, T\mathbf{v}, \dots, T^{n-1}\mathbf{v}$, using $ST^k = T^kS$.
Sketch. Polynomials in T commute with T. Conversely, write $S\mathbf{v} = q(T)\mathbf{v}$; then $ST^k\mathbf{v} = T^kS\mathbf{v} = q(T)T^k\mathbf{v}$, so S and $q(T)$ agree on a basis. The commutant is therefore $F[T]$, of dimension $\deg m_T = n$.

If $S = q(T)$, then $ST = TS$ by Proposition 7.2.27(b). Conversely, suppose $ST = TS$. Since $S\mathbf{v} \in V = K_T(\mathbf{v})$, there is $q \in F[x]$ with $S\mathbf{v} = q(T)\mathbf{v}$ (Definition 7.2.33). For every $k \geq 0$, $ST^k = T^kS$ (by induction on k), and $q(T)$ commutes with T^k, so

$$S\big(T^k\mathbf{v}\big) = T^k(S\mathbf{v}) = T^kq(T)\mathbf{v} = q(T)\big(T^k\mathbf{v}\big).$$

So S and $q(T)$ agree on the basis $\mathbf{v}, T\mathbf{v}, \dots, T^{n-1}\mathbf{v}$ of V (Proposition 7.2.35), and hence $S = q(T)$ (Theorem 2.2.1).

So the operators commuting with T are exactly the elements of $F[T]$. By Proposition 7.6.4, $\dim F[T] = \deg m_T$, and $m_T = p_T$ has degree n by Exercise 7.6.47(a).

Solution to Exercise 7.6.49

Uses. Corollary 7.6.16 (Restrictions of diagonalizable operators), Corollary 7.2.23 (Eigenspaces are independent), Proposition 4.4.16 (Complements exist), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$)
Techniques. Glue bases of the summands (Technique 4.4.14), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. For "$\Rightarrow$": $T{\restriction_U}$ is diagonalizable, so U is the direct sum of the pieces $U \cap E_\lambda$. Complete each piece inside E_λ.
Hint 2. For "$\Longleftarrow$": apply the hypothesis to $U = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$, the sum of all eigenspaces. What happens if its invariant complement is nonzero?
Sketch. ($\Rightarrow$) $U = \oplus_\lambda (U \cap E_\lambda)$ by Corollary 7.6.16. Choose a complement W_λ of $U \cap E_\lambda$ inside E_λ and let $W = \sum W_\lambda$; T acts on W_λ as λ, so W is invariant, and concatenated bases show $V = U \oplus W$. ($\Longleftarrow$) Let U be the sum of all eigenspaces and W an invariant complement. If $W \neq \{\mathbf{0}\}$, $T{\restriction_W}$ has an eigenvector, which lies in $U \cap W = \{\mathbf{0}\}$. So $U = V$. Over $\mathbb{R}$, a rotation of $\mathbb{R}^2$ is a counterexample.

($\Rightarrow$) Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T, so $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$, and let U be T-invariant. If $U = \{\mathbf{0}\}$, take $W = V$. Otherwise $T{\restriction_U}$ is diagonalizable (Corollary 7.6.16), and each of its eigenvectors lies in some $U_i = U \cap E_{\lambda_i}$, so $U = U_1 + \cdots + U_m$. For each i, choose a complement W_i of U_i inside E_{λ_i} (Proposition 4.4.16), and let $W = W_1 + \cdots + W_m$. Since T acts on $W_i \subseteq E_{\lambda_i}$ as multiplication by λ_i, W is T-invariant.

Choose bases $\mathcal{C}_i$ of U_i and $\mathcal{D}_i$ of W_i. Since $E_{\lambda_i} = U_i \oplus W_i$, the list $\mathcal{C}_i, \mathcal{D}_i$ is a basis of E_{λ_i}, and since V is the direct sum of the E_{λ_i}, all these lists together form a basis of V (Theorem 4.4.13(b)). The vectors from the $\mathcal{C}_i$ span U, and those from the $\mathcal{D}_i$ span W. A basis of V split into two parts, one spanning U and the other spanning W, shows that $U + W = V$ and that each vector has only one decomposition, so $V = U \oplus W$ (Theorem 4.4.13(b)).

($\Longleftarrow$) Let $U = E_{\lambda_1} + \cdots + E_{\lambda_m}$, the sum of all eigenspaces, which is direct (Corollary 7.2.23) and T-invariant. By hypothesis $V = U \oplus W$ with W invariant. If $W \neq \{\mathbf{0}\}$, then $T{\restriction_W}$ has an eigenvector $\mathbf{w}$ (Theorem 7.3.27), and $\mathbf{w}$ is an eigenvector of T, so $\mathbf{w} \in U \cap W = \{\mathbf{0}\}$, a contradiction. So $W = \{\mathbf{0}\}$, $V = U$, and concatenating bases of the eigenspaces gives an eigenbasis.

Over $\mathbb{R}$, "$\Longleftarrow$" fails: the rotation $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ of $\mathbb{R}^2$ has no invariant lines, so its only invariant subspaces are $\{\mathbf{0}\}$ and $\mathbb{R}^2$, which are complements of each other. Yet R is not diagonalizable over $\mathbb{R}$.

Solution to Exercise 7.6.50

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.5.22 (Trace and determinant from eigenvalues), Proposition 7.1.38 (Conjugate roots), Proposition B.1.8 (Properties of the modulus)

Techniques. Use an annihilating polynomial (Technique 7.6.17)
Hint 1. Every complex eigenvalue of A is a root of $f = x^3 - x - 1$. The determinant is the product of the complex eigenvalues with multiplicity.
Hint 2. Show that every real root of f is positive. What do the nonreal eigenvalues contribute to the product?
Sketch. The complex eigenvalues are roots of $f = x^3 - x - 1$. A real root r satisfies $r(r^2-1) = 1$; if $r < 0$ this needs $-1 < r < 0$, but then $r(r^2-1) = |r|(1-r^2) < 1$. So real eigenvalues are positive. Nonreal eigenvalues pair up with their conjugates, with equal multiplicity, contributing $|\lambda|^2 > 0$. So $\det A$, the product, is positive.

Regard A as a complex matrix, and let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues counted with multiplicity, so that $\det A = \lambda_1 \cdots \lambda_n$ (Theorem 7.5.22). Since $f = x^3 - x - 1$ annihilates A, every λ_j is a root of f (Proposition 7.2.28).

Real eigenvalues are positive. Let $r \in \mathbb{R}$ with $f(r) = 0$, that is, $r(r^2-1) = 1$. Clearly $r \neq 0$. If $r < 0$, then $r^2 - 1 < 0$ is needed for the product to be positive, so $-1 < r < 0$; but then $r(r^2-1) = |r|(1-r^2) \leq |r| < 1$, a contradiction. So $r > 0$.

Nonreal eigenvalues pair up. The polynomial p_A has real coefficients, so by Proposition 7.1.38 each nonreal λ and its conjugate $\overline{\lambda} \neq \lambda$ occur in the list $\lambda_1, \ldots, \lambda_n$ equally often. Grouping each such λ with a copy of $\overline{\lambda}$, the pair contributes $\lambda\overline{\lambda} = |\lambda|^2 > 0$ to the product (Proposition B.1.8).

So $\det A$ is a product of positive real eigenvalues and positive numbers $|\lambda|^2$, and $\det A > 0$.

Solution to Exercise 7.6.51

Uses. Corollary 7.6.33 (Eigenvalues of sums and products of commuting operators), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.6.20 (Cayley–Hamilton), Proposition 7.5.9 (Multiplicities on the diagonal), Corollary 7.1.35 (Factorization over $\mathbb{C}$)
Techniques. Prove it for triangular matrices first (Technique 7.5.23)
Hint 1. A nilpotent matrix has 0 as its only eigenvalue. Use Corollary 7.6.33.
Hint 2. A complex matrix whose only eigenvalue is 0 has $p = x^n$. Then apply Cayley–Hamilton.
Hint 3. For (c), compare the diagonals of AB and BA.
Sketch. (a) All eigenvalues of A and B are 0, so by Corollary 7.6.33 all eigenvalues of $A+B$ and AB are 0; then $p = x^n$ and Cayley–Hamilton gives $(A+B)^n = 0 = (AB)^n$. (b) $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and N^{T}. (c) AB and BA have the same diagonal $a_{ii}b_{ii}$, so $AB - BA$ is strictly upper triangular, with $p = x^n$. Example: $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

(a) If $A^k = 0$, every eigenvalue of A is a root of x^k (Proposition 7.2.28), so 0 is the only eigenvalue of A; likewise for B. By Corollary 7.6.33, every eigenvalue of $A+B$ is a sum $0+0$ and every eigenvalue of AB is a product $0 \cdot 0$. So p_{A+B} is a monic polynomial of degree n whose only complex root is 0, that is, $p_{A+B} = x^n$ (Corollary 7.1.35), and Cayley–Hamilton gives $(A+B)^n = 0$. In the same way $(AB)^n = 0$.

(A direct argument also works: since A and B commute, the binomial theorem gives $(A+B)^{2k} = \sum_j \binom{2k}{j} A^j B^{2k-j}$, and in each term $j \geq k$ or $2k - j \geq k$. And $(AB)^k = A^k B^k = 0$.)

(b) $N = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and N^{T} satisfy $N^2 = \left(N^{\mathsf{T}}\right)^2 = 0$, but $N + N^{\mathsf{T}} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ squares to I_2, and $NN^{\mathsf{T}} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is a nonzero projection; neither is nilpotent.

(c) As in the proof of Theorem 7.6.28, AB and BA are upper triangular with the same diagonal entries $a_{ii}b_{ii}$. So $C = AB - BA$ is upper triangular with zero diagonal, $p_C = x^n$ by Proposition 7.5.9, and $C^n = 0$ by Cayley–Hamilton. If $A = SRS^{-1}$ and $B = SR'S^{-1}$ with R, R' upper triangular, then $AB - BA = S(RR' - R'R)S^{-1}$, and $(AB - BA)^n = S(RR' - R'R)^n S^{-1} = 0$.

For $X = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $XY = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $YX = 0$. So matrices can be triangular in the same basis without commuting; the converse of Theorem 7.6.32 fails.

Solution to Exercise 7.6.52

Uses. Theorem 5.4.10 (Multiplicativity), Proposition 7.3.2 (Determinants over a commutative ring), Proposition B.1.8 (Properties of the modulus)
Hint 1. Factor $A^2 + B^2$ over $\mathbb{C}$, the way $a^2 + b^2 = (a+ib)(a-ib)$. Where is $AB = BA$ used?
Hint 2. How are $\det(A+iB)$ and $\det(A-iB)$ related when A, B are real?

Sketch. Since $AB = BA$, $(A+iB)(A-iB) = A^2+B^2$. As $A-iB$ is the entrywise conjugate of $A+iB$, $\det(A-iB) = \overline{\det(A+iB)}$, so $\det(A^2+B^2) = |\det(A+iB)|^2 \geq 0$. For $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $\det(A^2+B^2) = -1$.

Regard A and B as complex matrices. Then

$$(A+iB)(A-iB) = A^2 - iAB + iBA + B^2 = A^2 + B^2,$$

because $AB = BA$. Since A and B are real, $A-iB$ is the entrywise complex conjugate of $A+iB$. Complex conjugation $\mathbb{C} \to \mathbb{C}$ respects sums and products and fixes 1, so Proposition 7.3.2(g) gives $\det(A-iB) = \overline{\det(A+iB)}$. By Theorem 5.4.10 and Proposition B.1.8,

$$\det(A^2+B^2) = \det(A+iB)\overline{\det(A+iB)} = |\det(A+iB)|^2 \geq 0.$$

Without commuting this fails. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. Then $A^2 = A$ and $B^2 = B$, so $A^2 + B^2 = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$, with determinant -1. (Here $AB = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = BA$.)

Solution to Exercise 7.6.53

Uses. Theorem 7.6.34 (Simultaneous diagonalization), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization)
Hint 1. Find the eigenspaces of A first. B must preserve them.
Hint 2. On the plane $E_2(A)$, use the basis $\mathbf{e}_1, \mathbf{f} = \mathbf{e}_2 + \mathbf{e}_3$ and write down the 2×2 matrix of B.
Hint 3. For (c), what are the eigenvalues of B?
Sketch. (a) Both products equal $\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$. $E_2(A) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2+\mathbf{e}_3)$ and $E_0(A) = \operatorname{span}(\mathbf{e}_2 - \mathbf{e}_3)$; B has the three distinct eigenvalues $0, \pm\sqrt{2}$. (b) On $E_2(A)$, B has matrix $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, with eigenvectors $\pm\sqrt{2}\mathbf{e}_1 + \mathbf{f}$. The basis is $\left(\sqrt{2}, 1, 1\right), \left(-\sqrt{2}, 1, 1\right), (0, 1, -1)$. (c) No: B has irrational eigenvalues, so it is not diagonalizable over $\mathbb{Q}$.

(a) Computing,

$$AB = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} = BA.$$

For A: $A\mathbf{e}_1 = 2\mathbf{e}_1$, $A(0,1,1) = (0,2,2)$, and $A(0,1,-1) = \mathbf{0}$. These three vectors form a basis, so A is diagonalizable with $E_2(A) = \operatorname{span}(\mathbf{e}_1, \mathbf{e}_2+\mathbf{e}_3)$ and $E_0(A) = \operatorname{span}(\mathbf{e}_2 - \mathbf{e}_3)$. For B, part (b) produces three eigenvectors for the distinct eigenvalues $\sqrt{2}, -\sqrt{2}, 0$, so B is diagonalizable over $\mathbb{R}$ (Corollary 7.4.5).

(b) By Proposition 7.2.30(b), B preserves the eigenspaces of A. Let $\mathbf{f} = \mathbf{e}_2 + \mathbf{e}_3$. Then $B\mathbf{e}_1 = \mathbf{f}$ and $B\mathbf{f} = 2\mathbf{e}_1$, so on $E_2(A)$, in the basis $\mathbf{e}_1, \mathbf{f}$, B has matrix $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, with characteristic polynomial $x^2 - 2$. One checks

$$B\left(\sqrt{2}\mathbf{e}_1 + \mathbf{f}\right) = \sqrt{2}\mathbf{f} + 2\mathbf{e}_1 = \sqrt{2}\left(\sqrt{2}\mathbf{e}_1 + \mathbf{f}\right), \quad B\left(-\sqrt{2}\mathbf{e}_1 + \mathbf{f}\right) = -\sqrt{2}\left(-\sqrt{2}\mathbf{e}_1 + \mathbf{f}\right).$$

On $E_0(A)$, $B(\mathbf{e}_2 - \mathbf{e}_3) = \mathbf{0}$. So in the basis $\left(\sqrt{2}, 1, 1\right), \left(-\sqrt{2}, 1, 1\right), (0, 1, -1)$ of $\mathbb{R}^3$,

$$A \mapsto \operatorname{diag}(2, 2, 0), \quad B \mapsto \operatorname{diag}\left(\sqrt{2}, -\sqrt{2}, 0\right).$$

(c) No. If $\mathbf{v} \in \mathbb{Q}^3$ is an eigenvector of B with eigenvalue λ, then $\lambda = (B\mathbf{v})_i / v_i$ for any i with $v_i \neq 0$, so $\lambda \in \mathbb{Q}$. The eigenvalues of B are $\pm\sqrt{2}$ and 0, so the only rational eigenvectors of B lie in the line $E_0(B)$, and $\mathbb{Q}^3$ has no basis of eigenvectors of B. Simultaneous diagonalization needs each operator to be diagonalizable over the field at hand.

Solution to Exercise 7.6.54

Uses. Theorem 7.6.36 (Gershgorin's disk theorem), Theorem 7.5.22 (Trace and determinant from eigenvalues), Proposition 7.1.38 (Conjugate roots), Exercise 7.3.17
Techniques. Look at the largest coordinate (Technique 7.6.37)
Hint 1. Find the three disks. How far left does each reach?

Hint 2. The determinant is the product of the complex eigenvalues. What can a real eigenvalue be, and what does a conjugate pair contribute?
Hint 3. Apply Gershgorin to G^{T}, which has the same eigenvalues.
Sketch. (a) The disks have centers $6, 5, 4$ and radii $2, 3, 1$, so they lie in $\operatorname{Re} z \geq 4, 2, 3$. (b) Real eigenvalues are ≥ 2, nonreal ones pair with their conjugates, so $\det G$ is a product of positive numbers. (c) The column disks have radii $1, 2, 3$, so $|\lambda| \leq 6+1, 5+2, 4+3$; the row disks give $6+2, 5+3, 4+1$.

(a) The row disks are D_1: center 6, radius $1+1=2$; D_2: center 5, radius $1+2=3$; D_3: center 4, radius $0+1=1$. A point z of the disk with center $c \in \mathbb{R}$ and radius r has $\operatorname{Re} z \geq c - |z-c| \geq c - r$, which is 4, 2, 3 for the three disks. By Theorem 7.6.36 every eigenvalue lies in one of them, so $\operatorname{Re}\lambda \geq 2$. In particular 0 is not an eigenvalue, and G is invertible.

(b) By Theorem 7.5.22, $\det G = \lambda_1\lambda_2\lambda_3$, the product of the complex eigenvalues with multiplicity. A real eigenvalue is at least 2 by (a). A nonreal eigenvalue λ occurs as often as $\overline{\lambda}$ (Proposition 7.1.38, as p_G has real coefficients), and each such pair contributes $|\lambda|^2 > 0$. So $\det G > 0$. (In fact $\det G = 129$.)

(c) G^{T} has the same characteristic polynomial (Exercise 7.3.17), hence the same eigenvalues, and its row disks are the column disks of G: centers $6, 5, 4$ and radii $1+0=1$, $1+1=2$, $1+2=3$. A point of the disk with center c and radius r has $|z| \leq |c| + r$, which here is 7 in all three cases. So $|\lambda| \leq 7$. The row disks give $6+2=8$, $5+3=8$, and $4+1=5$, so only $|\lambda| \leq 8$.

Solution to Exercise 7.6.55

Uses. Theorem 7.6.36 (Gershgorin's disk theorem), Theorem B.1.11 (Triangle inequality in $\mathbb{C}$)
Techniques. Look at the largest coordinate (Technique 7.6.37)
Hint 1. The disk D_i has center a_{ii} and radius $1 - a_{ii}$. Which vector does A fix?
Hint 2. If $|\lambda| = 1$ and $\lambda \in D_i$, then $1 = |\lambda| \leq |\lambda - a_{ii}| + a_{ii} \leq 1$. When is the triangle inequality an equality?
Sketch. (a) D_i has center $a_{ii} \geq 0$ and radius $1 - a_{ii}$, so it lies in the closed unit disk; $A(1,\dots,1) = (1,\dots,1)$. (b) If $|\lambda| = 1$, equality holds in $|\lambda| \leq |\lambda - a_{ii}| + a_{ii}$, so $\lambda - a_{ii}$ is a nonnegative multiple of the positive number a_{ii}, and λ is a positive real with $|\lambda| = 1$. (c) $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has eigenvalue -1.

(a) Since the entries are nonnegative and row i sums to 1, $R_i = \sum_{j \neq i} a_{ij} = 1 - a_{ii}$. By Theorem 7.6.36, $\lambda \in D_i$ for some i, and by the triangle inequality

$$|\lambda| \leq |\lambda - a_{ii}| + |a_{ii}| \leq (1 - a_{ii}) + a_{ii} = 1.$$

The row sums say $A\mathbf{1} = \mathbf{1}$ for $\mathbf{1} = (1,\dots,1)$, so 1 is an eigenvalue.

(b) Suppose $|\lambda| = 1$, and let i be as in (a). Then both inequalities in (a) are equalities. Equality in the triangle inequality for $z_1 = \lambda - a_{ii}$ and $z_2 = a_{ii}$ means that $z_1 = |z_1|u$ and $z_2 = |z_2|u$ for some $u \in \mathbb{C}$ with $|u| = 1$ (Theorem B.1.11(b)). Since $z_2 = a_{ii} > 0$, $u = 1$, so $\lambda - a_{ii}$ is a nonnegative real number. Then λ is a real number with $\lambda \geq a_{ii} > 0$ and $|\lambda| = 1$, so $\lambda = 1$.

(c) $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ has nonnegative entries and row sums 1, and $A(1,-1) = -(1,-1)$, so -1 is an eigenvalue of absolute value 1.

Solution to Exercise 7.6.56

Uses. Proposition 2.4.28 (Cyclicity of the trace), Corollary 7.6.30 (Trace, determinant, and invertibility of $p(T)$), Theorem 5.5.21 (The Vandermonde determinant), Theorem 7.6.20 (Cayley–Hamilton), Corollary 7.1.35 (Factorization over $\mathbb{C}$)
Techniques. Cyclicity of trace (Technique 2.4.29), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Show that $\operatorname{tr}\left(C^k\right) = 0$ for every $k \geq 1$. Write $C^k = C^{k-1}AB - C^{k-1}BA$ and move A past C^{k-1}.
Hint 2. If $\lambda_1, \dots, \lambda_n$ are the eigenvalues of C, then $\operatorname{tr}\left(C^k\right) = \sum_j \lambda_j^k$. Suppose some are nonzero, group the distinct nonzero ones, and get a Vandermonde system.
Hint 3. Over $\mathbb{F}_p$, think of the $p \times p$ matrices of x and d/dx on polynomials of degree less than p.
Sketch. For $k \geq 1$, $\operatorname{tr}\left(C^{k-1}AB\right) = \operatorname{tr}\left(AC^{k-1}B\right) = \operatorname{tr}\left(C^{k-1}BA\right)$, using $CA = AC$ and cyclicity, so $\operatorname{tr}\left(C^k\right) = 0$. With distinct nonzero eigenvalues $\mu_1, \dots, \mu_r$ of multiplicities m_l, this gives $\sum_l m_l \mu_l^k = 0$ for $k = 1, \dots, r$, an invertible Vandermonde-type system forcing $m_l = 0$. So 0 is the only eigenvalue, $p_C = x^n$, and $C^n = 0$. Over $\mathbb{F}_p$, $AB - BA = I_p$ is possible.

Traces of powers vanish. Let $k \geq 1$. Since A commutes with C, it commutes with C^{k-1}, so

$$\operatorname{tr}\big(C^{k-1}AB\big) = \operatorname{tr}\big(AC^{k-1}B\big) = \operatorname{tr}\big(\big(C^{k-1}B\big)A\big)$$

by cyclicity of the trace (Proposition 2.4.28). Hence $\operatorname{tr}\big(C^k\big) = \operatorname{tr}\big(C^{k-1}AB\big) - \operatorname{tr}\big(C^{k-1}BA\big) = 0$.

All eigenvalues are 0**.** Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of C counted with multiplicity. By Corollary 7.6.30 with $p = x^k$, $\operatorname{tr}\big(C^k\big) = \lambda_1^k + \cdots + \lambda_n^k$. Suppose some $\lambda_j \neq 0$, let $\mu_1, \dots, \mu_r$ be the distinct nonzero values among them ($r \geq 1$), and let $m_l \geq 1$ be the number of j with $\lambda_j = \mu_l$. Then for $k = 1, \dots, r$,

$$0 = \operatorname{tr}\big(C^k\big) = m_1\mu_1^k + \cdots + m_r\mu_r^k.$$

This is $M\mathbf{m} = \mathbf{0}$ with $\mathbf{m} = (m_1, \dots, m_r)$ (as elements of $\mathbb{C}$) and $M = \big(\mu_l^k\big)_{k,l}$. Factoring μ_l out of column l leaves the Vandermonde matrix $\big(\mu_l^{k-1}\big)$, so $\det M = \mu_1 \cdots \mu_r \prod_{l<l'}(\mu_{l'} - \mu_l) \neq 0$ (Theorem 5.5.21), as the μ_l are distinct and nonzero. So $\mathbf{m} = \mathbf{0}$, contradicting $m_1 \geq 1$. Hence every $\lambda_j = 0$, $p_C = x^n$ (Corollary 7.1.35), and $C^n = 0$ by Cayley–Hamilton.

Over $\mathbb{F}_p$**.** On $V = \mathbb{F}_p[x]_{<p}$, with basis $1, x, \dots, x^{p-1}$, let D be differentiation and X the operator $f \mapsto xf$ followed by deleting the x^p term. For $j < p-1$, $(DX - XD)x^j = (j+1)x^j - jx^j = x^j$. For $j = p-1$, $Xx^{p-1} = \mathbf{0}$ and $DXx^{p-1} = \mathbf{0}$, while $XDx^{p-1} = (p-1)x^{p-1} = -x^{p-1}$; so $(DX - XD)x^{p-1} = x^{p-1}$ as well. With $A = D$, $B = X$, we get $C = AB - BA = \mathrm{id}_V$, which commutes with A and is not nilpotent. The proof breaks at the last step: over $\mathbb{F}_p$ the multiplicities m_l are elements of $\mathbb{F}_p$, and $m_l = p$ is 0 there. (Here $C = \mathrm{id}_V$ has $\operatorname{tr}\big(C^k\big) = p = 0$ for all k.)

Section 7.7

Solution to Exercise 7.7.4

Uses. Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Lemma 7.4.13 (Geometric multiplicity through a rank), Theorem 7.4.16 (Criteria for diagonalizability), Proposition 7.3.30 (Odd dimension), Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (a) Factor $x^3 - x$ over $\mathbb{Q}$ and over $\mathbb{F}_2$. (b) Compute $\mathrm{gm}(1)$ from a rank. (d) Diagonalize and apply the polynomial to the diagonal.
Hint 2. (a) Over $\mathbb{F}_2$, $x^3 - x = x(x+1)^2$; try $\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ in a block. (e) $p_A = x^2 + \det A$.
Sketch. (a) True over $\mathbb{Q}$: $x(x-1)(x+1)$ has distinct roots. False over $\mathbb{F}_2$: $A = \left(\begin{smallmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{smallmatrix}\right)$ has $A^3 = A$ and $\mathrm{gm}(1) = 1 < 2 = \mathrm{am}(1)$. (b) False: $\mathrm{gm}(1) = 4 - 3 = 1 < 2$. (c) True: odd dimension. (d) True: $A = S\,\mathrm{diag}(1,2,3)S^{-1}$, so $A^2 - 3A + 2I_3 = S\,\mathrm{diag}(0,0,2)S^{-1}$. (e) True: $p_A = x^2 + \det A$ has the distinct roots $\pm\sqrt{-\det A}$.

(a) True over $\mathbb{Q}$. The polynomial $x^3 - x = x(x-1)(x+1)$ annihilates A and has the distinct roots $0, 1, -1$, so A is diagonalizable by Theorem 7.6.15. Over $\mathbb{F}_2$ it is false, because $-1 = 1$ there and $x^3 - x = x(x+1)^2$ has a repeated root. For

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{F}_2^{3\times 3},$$

the block $J = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ has $J^2 = \left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right) = I_2$ over $\mathbb{F}_2$, so $J^3 = J$ and $A^3 = A$. But $p_A = x(x-1)^2$ and $A - I_3 = \left(\begin{smallmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{smallmatrix}\right)$ has rank 2, so $\mathrm{gm}(1) = 1 < 2 = \mathrm{am}(1)$ (Lemma 7.4.13), and A is not diagonalizable (Theorem 7.4.16).

(b) False. By Lemma 7.4.13, $\mathrm{gm}(1) = 4 - \operatorname{rank}(A - I_4) = 1$, while $\mathrm{am}(1) = 2$, so A is not diagonalizable. Such matrices exist: the block diagonal matrix with blocks $0_{2\times 2}$ and $\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$ has $p_A = x^2(x-1)^2$, rank 2, and $\operatorname{rank}(A - I_4) = 2 + 1 = 3$.

(c) True by Proposition 7.3.30: p_A is a real polynomial of odd degree 5, so it has a real root.

(d) True. Three distinct eigenvalues make A diagonalizable (Corollary 7.4.5): $A = SDS^{-1}$ with $D = \operatorname{diag}(1,2,3)$. With $q = x^2 - 3x + 2 = (x-1)(x-2)$, Proposition 7.4.23 gives $q(A) = Sq(D)S^{-1} = S\operatorname{diag}(0,0,2)S^{-1}$, which has rank 1.

(e) True. By Proposition 7.3.18, $p_A = x^2 - (\operatorname{tr} A)x + \det A = x^2 - c$ with $c = -\det A > 0$. Its roots $\pm\sqrt{c}$ are real and distinct, so A is diagonalizable over $\mathbb{R}$ (Corollary 7.4.5).

Solution to Exercise 7.7.5

Uses. Definition 7.4.12 (Geometric multiplicity), Definition 7.4.11 (Algebraic multiplicity), Theorem 7.4.15 (Geometric multiplicity is at most algebraic multiplicity), Proposition 7.1.34 (Roots of a split polynomial)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Which multiplicity is $\dim E_0$, and which one does "counted with multiplicity" refer to?
Hint 2. Try the shift N with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$. For the correct version, use $\mathrm{gm} \le \mathrm{am}$.
Sketch. $\dim E_0 = \mathrm{gm}(0)$, but "counted with multiplicity" means $\mathrm{am}(0)$. The claim is false: the $n \times n$ shift ($n \ge 2$) has rank $n-1$ and no nonzero eigenvalue. Correct: the number of nonzero eigenvalues with multiplicity is $n - \mathrm{am}(0) \le n - \mathrm{gm}(0) = r$, with equality when A is diagonalizable.

The error. $\dim\ker A = \dim E_0$ is the geometric multiplicity $\mathrm{gm}(0)$. The eigenvalues "counted with multiplicity" are the roots of p_A counted with algebraic multiplicity (Proposition 7.1.34), and $\mathrm{am}(0)$ can be larger than $\mathrm{gm}(0)$. The claim is false: for $n \ge 2$, the shift $N \in \mathbb{C}^{n\times n}$ with $N\mathbf{e}_1 = \mathbf{0}$ and $N\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \ge 2$ has rank $n-1$, but $N^n = 0$, so its only eigenvalue is 0 (Proposition 7.2.28) and no eigenvalue is nonzero.

A correct version. At most r eigenvalues, counted with multiplicity, are nonzero. Indeed p_A splits over $\mathbb{C}$, its roots with multiplicity number n, and 0 occurs $\mathrm{am}(0)$ times, so the number of nonzero ones is $n - \mathrm{am}(0) \le n - \mathrm{gm}(0) = n - (n-r) = r$ by Theorem 7.4.15. Equality holds exactly when $\mathrm{am}(0) = \mathrm{gm}(0)$, for instance when A is diagonalizable.

Solution to Exercise 7.7.6

Uses. Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Theorem 7.6.34 (Simultaneous diagonalization), Theorem 7.6.31 (Commuting operators have a common eigenvector)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Two steps need a hypothesis that is not given. Which operator must be diagonalizable for each of them?
Hint 2. Try $A = I_2$ with B a nonzero nilpotent matrix, and the other way round.
Sketch. The basis of $E_\lambda(A)$ made of eigenvectors of B exists only if $B|_{E_\lambda(A)}$ is diagonalizable, and the eigenspaces of A add up to $\mathbb{C}^n$ only if A is diagonalizable. $A = I_2$, $B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ breaks the first step; $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $B = I_2$ breaks the second. With both diagonalizable the argument is correct (Theorem 7.6.34).

The first step is right: $E_\lambda(A)$ is B-invariant (Proposition 7.2.30(b)). There are two errors after it.

First, $E_\lambda(A)$ need not have a basis of eigenvectors of B: that requires the restriction of B to $E_\lambda(A)$ to be diagonalizable, which follows if B is (Corollary 7.6.16) and can fail otherwise. With $A = I_2$ and $B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, which commute, $E_1(A) = \mathbb{C}^2$ and the only eigenvectors of B are the nonzero multiples of $\mathbf{e}_1$. So no basis of common eigenvectors exists, and the claim is false.

Second, the eigenspaces of A need not add up to $\mathbb{C}^n$; that is the statement that A is diagonalizable. With $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $B = I_2$, the bases chosen span only $E_0(A) = \operatorname{span}(\mathbf{e}_1)$.

When A and B are both diagonalizable, both gaps close and the argument is the proof of Theorem 7.6.34. Without that, commuting still gives one common eigenvector (Theorem 7.6.31), and a common triangular form.

Solution to Exercise 7.7.7

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Example 7.7.2 (An idempotent is diagonalizable, three ways)

Techniques. Test an identity on eigenvectors (Technique 7.4.25), Use an annihilating polynomial (Technique 7.6.17)
Hint 1. The last step needs a basis of eigenvectors. Look for a counterexample with only one eigenvector line.
Hint 2. (a) Compute A^4. (b) If A is diagonalizable, the argument above is correct. Conversely, $A^2 = A$ is an annihilating polynomial with distinct roots.
Sketch. Agreement on eigenvectors gives agreement everywhere only when the eigenvectors span. $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has $N^3 = N^2 = 0$ and $N^2 \neq N$. (a) $A^4 = AA^3 = AA^2 = A^3 = A^2$. (b) If A is diagonalizable, the argument works on an eigenbasis; if $A^2 = A$, then $x(x-1)$ annihilates A.

The error. Two linear maps that agree on a spanning list are equal, but the eigenvectors of A need not span $\mathbb{C}^n$; the last step assumes that A is diagonalizable. The claim is false: $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has $N^2 = 0$, so $N^3 = N^2$, but $N^2 = 0 \neq N$. Its eigenvectors are the nonzero multiples of $\mathbf{e}_1$, on which N^2 and N do agree.

(a) $A^4 = AA^3 = AA^2 = A^3 = A^2$, so $\left(A^2\right)^2 = A^2$.

(b) If A is diagonalizable, it has an eigenbasis, and the argument above shows that A^2 and A agree on it. So $A^2 = A$ (Technique 7.4.25). Conversely, if $A^2 = A$, then $x(x-1)$ annihilates A, and A is diagonalizable (Theorem 7.6.15, or Example 7.7.2).

In terms of polynomials: $x^2(x-1)$ annihilates A, so m_A divides it, and A is diagonalizable exactly when m_A divides $x(x-1)$. The "proof" tested the identity on eigenvectors, which sees only the roots of m_A, never their multiplicities.

Solution to Exercise 7.7.8

Uses. Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues), Proposition 7.3.30 (Odd dimension), Proposition 7.4.7 (One eigenvalue), Theorem 7.1.39 (Factorization over $\mathbb{R}$)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18), Use an annihilating polynomial (Technique 7.6.17)
Hint 1. The roots of m_A in $\mathbb{R}$ are the real eigenvalues. Can m_A have a factor without real roots?
Hint 2. Over $\mathbb{R}$, $x^3 - 1 = (x-1)\left(x^2 + x + 1\right)$. Look for a 2×2 integer matrix with $p = x^2 + x + 1$, and add a block 1.
Sketch. Knowing the real roots of m_A does not determine m_A: it can be $(x-1)\left(x^2 + x + 1\right)$. Counterexample: A with diagonal blocks 1 and $C = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$, where $C^3 = I_2$. (a) p_A has odd degree, so there is a real eigenvalue λ, and $\lambda^3 = 1$ forces $\lambda = 1$. (b) If A is diagonalizable over $\mathbb{R}$ with only eigenvalue 1, then $A = I_3$ by Proposition 7.4.7.

The error. By Theorem 7.6.7, the roots of m_A in $\mathbb{R}$ are exactly the real eigenvalues of A, so m_A has 1 as its only real root. That does not make $m_A = x - 1$: over $\mathbb{R}$, $x^3 - 1 = (x-1)\left(x^2 + x + 1\right)$, and $x^2 + x + 1$ has no real root, so m_A could be $(x-1)\left(x^2 + x + 1\right)$. The step "its roots are the eigenvalues, so it is $x - 1$" assumed that m_A splits over $\mathbb{R}$.

A counterexample. Let $C = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$. Then $C^2 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$ and $C^3 = CC^2 = I_2$. The block diagonal matrix A with blocks (1) and C has $A^3 = I_3$ and $A \neq I_3$. Over $\mathbb{C}$ its eigenvalues are 1 and the two primitive cube roots of unity, the roots of $p_C = x^2 + x + 1$; geometrically, C is similar to a rotation of the plane by $2\pi/3$.

(a) A has a real eigenvalue λ by Proposition 7.3.30, and $\lambda^3 = 1$ with $\lambda \in \mathbb{R}$ forces $\lambda = 1$.

(b) If $A = I_3$, it is diagonal. Conversely, if A is diagonalizable over $\mathbb{R}$, its eigenvalues are real roots of $x^3 - 1$, so 1 is its only eigenvalue, and $A = I_3$ by Proposition 7.4.7. Over $\mathbb{C}$ the situation is different: $x^3 - 1$ has three distinct complex roots, so every A with $A^3 = I_3$ is diagonalizable over $\mathbb{C}$ (Theorem 7.6.15), and $A \neq I_3$ exactly when the eigenvalues are $1, \omega, \overline{\omega}$ with $\omega = e^{2\pi i/3}$.

Solution to Exercise 7.7.9

Uses. Corollary 4.1.27 (A nilpotent operator dies within $\dim V$ steps), Proposition 4.1.26 (Nilpotent chains are independent), Theorem 4.2.11 (Rank–nullity), Exercise 4.2.51
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Dimension counting (Technique 4.2.24)
Hint 1. If $B^k = N$, then B is nilpotent. How fast must a nilpotent $n \times n$ matrix die?
Hint 2. Powers: $N^j = B^{kj}$, and $kj \geq n$ already for $j = \lceil n/k \rceil$. Kernels: $\dim \ker N = 1$, and $\ker B \subseteq \ker B^2 \subseteq \ker N$.

Sketch. Powers: $B^{kn} = N^n = 0$, so $B^n = 0$ (Corollary 4.1.27). With $j = \lceil n/k \rceil \le n-1$, $N^j = B^{kj} = 0$, so $N^{n-1} = 0$, a contradiction. Kernels: a chain of length n shows $\dim\ker N = 1$. B is nilpotent, so $\{\mathbf{0}\} \neq \ker B \subseteq \ker B^2 \subseteq \ker B^k = \ker N$, forcing $\ker B = \ker B^2$; then all $\ker B^j$ equal $\ker B$, which is not F^n, but $B^{kn} = 0$.

Suppose $B^k = N$ with $k \ge 2$. Then $B^{kn} = N^n = 0$, so B is nilpotent.

Powers. By Corollary 4.1.27, $B^n = 0$. Let $j = \lceil n/k \rceil$, so that $kj \ge n$ and $N^j = B^{kj} = B^{kj-n}B^n = 0$. Since $k \ge 2$ and $n \ge 2$, $j \le \lceil n/2 \rceil \le n-1$, so $N^{n-1} = N^{n-1-j}N^j = 0$, a contradiction.

Kernels. Choose $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$. By Proposition 4.1.26, $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is independent, so it is a basis of F^n, and its image under N spans $\operatorname{span}\left(N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}\right)$, of dimension $n-1$. By Theorem 4.2.11, $\dim\ker N = 1$. Since B is nilpotent and $n \ge 1$, B is not injective, so $\ker B \neq \{\mathbf{0}\}$. Also $\ker B \subseteq \ker B^2 \subseteq \ker B^k = \ker N$, because $k \ge 2$. A nonzero subspace of a line is the line, so $\ker B = \ker B^2 = \ker N$. By Exercise 4.2.51(b), $\ker B^m = \ker B$ for every $m \ge 1$. But $B^{kn} = 0$, so $\ker B^{kn} = F^n$, while $\ker B$ is a line and $n \ge 2$. This is a contradiction.

The first proof needs only the bound $B^n = 0$; the second shows where a root would have to fail: $\ker B^2$ would have to lie in the line $\ker N$ and also strictly contain the nonzero $\ker B$, because B is nilpotent and $n \ge 2$. By contrast $I_n + N$ does have a square root over $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$, namely the truncated binomial series $\sum_{j=0}^{n-1}\binom{1/2}{j}N^j$, since the series for $(1+t)^{1/2}$ squares to $1+t$ and $N^n = 0$.

Solution to Exercise 7.7.10

Uses. Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Exercise 7.3.40, Exercise 7.4.40
Techniques. Compare AB and BA (Technique 7.3.21), Use an annihilating polynomial (Technique 7.6.17)
Hint 1. $B(AB)^jA = (BA)^{j+1}$. Apply this with $f = m_{AB}$.
Hint 2. (c) Map $E_\lambda(AB)$ to $E_\lambda(BA)$ by $\mathbf{v} \mapsto B\mathbf{v}$. (d) If 0 is not a root of m_{AB}, then xm_{AB} still has distinct roots.
Sketch. (a) $B(AB)^jA = (BA)^{j+1}$, sum. With $f = m_{AB}$ the left side is 0. (b) A nonzero root of m_{BA} is a root of m_{AB}, and symmetrically. $A = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$: $AB = 0$, $BA = A$. (c) $\mathbf{v} \mapsto B\mathbf{v}$ is injective on $E_\lambda(AB)$ and lands in $E_\lambda(BA)$; symmetric. (d) xm_{AB} is a product of distinct linear factors and kills BA. The example in (b) has $AB = 0$ diagonalizable and BA not.

(a) By associativity, $B(AB)^jA = (BA)^{j+1}$ for every $j \ge 0$. If $f = \sum_j c_j x^j$, summing gives $Bf(AB)A = \sum_j c_j(BA)^{j+1} = (xf)(BA)$. With $f = m_{AB}$ the left side is $B0A = 0$, so xm_{AB} annihilates BA, and m_{BA} divides it (Theorem 7.6.2).

(b) By symmetry, m_{AB} also divides xm_{BA}. If $\lambda \neq 0$ is an eigenvalue of BA, it is a root of m_{BA} (Theorem 7.6.7), hence of xm_{AB}, hence of m_{AB}, so it is an eigenvalue of AB. The converse is symmetric. For $A = \begin{pmatrix}0&1\\0&0\end{pmatrix}$ and $B = \begin{pmatrix}1&0\\0&0\end{pmatrix}$, $AB = 0$ and $BA = A$, so $m_{AB} = x$ while $m_{BA} = x^2$.

(c) Let $\lambda \neq 0$ and $\mathbf{v} \in E_\lambda(AB)$. Then $BA(B\mathbf{v}) = B(AB\mathbf{v}) = \lambda B\mathbf{v}$, so $B\mathbf{v} \in E_\lambda(BA)$. If $B\mathbf{v} = \mathbf{0}$, then $\lambda\mathbf{v} = AB\mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$. So $\mathbf{v} \mapsto B\mathbf{v}$ is an injective linear map $E_\lambda(AB) \to E_\lambda(BA)$, and $\operatorname{gm}_{AB}(\lambda) \le \operatorname{gm}_{BA}(\lambda)$. Exchanging the roles of A and B gives the reverse inequality.

(d) By Theorem 7.6.15, $m_{AB} = (x-\lambda_1)\cdots(x-\lambda_r)$ with distinct λ_i, and none of them is 0 since AB is invertible. So xm_{AB} is a product of distinct linear factors, and by (a) it annihilates BA. Hence BA is diagonalizable. In the example of (b), $AB = 0$ is diagonalizable but not invertible, and $BA = \begin{pmatrix}0&1\\0&0\end{pmatrix}$ is not diagonalizable.

For $m = 1$, write $A = \mathbf{v}^{\mathsf{T}}$ and $B = \mathbf{u}$. Then $AB = \mathbf{v}^{\mathsf{T}}\mathbf{u}$ is a 1×1 matrix and $BA = \mathbf{u}\mathbf{v}^{\mathsf{T}}$, and (d) says that $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ is diagonalizable when $\mathbf{v}^{\mathsf{T}}\mathbf{u} \neq 0$, one half of Exercise 7.4.40. Compare the three proofs of "same nonzero eigenvalues" in Exercise 7.3.40: characteristic polynomials see multiplicities, eigenvectors see geometric multiplicities as in (c), and minimal polynomials see diagonalizability.

Solution to Exercise 7.7.11

Uses. Theorem 7.4.16 (Criteria for diagonalizability), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22), Study a linear map on a space of matrices (Technique 4.2.29)
Hint 1. Write $A = SDS^{-1}$ and $B = RER^{-1}$, and substitute $X = SYR^{-1}$.
Hint 2. $DY = YE$ says $(d_i - e_j)y_{ij} = 0$ for all i, j. Count the pairs with $d_i = e_j$.
Sketch. $X \mapsto S^{-1}XR$ is an isomorphism of $F^{m\times n}$ taking W to $\{Y \mid DY = YE\}$, whose elements have $y_{ij} = 0$ unless $d_i = e_j$. The number of such pairs is $\sum_\lambda \operatorname{gm}_A(\lambda)\operatorname{gm}_B(\lambda)$. (b) is the case $\dim W = 0$. (c) With $B = A$: $\sum \operatorname{gm}(\lambda)^2 \geq \sum \operatorname{gm}(\lambda) = m$, with equality iff every $\operatorname{gm}(\lambda) \leq 1$.

(a) Write $A = SDS^{-1}$ and $B = RER^{-1}$ with $D = \operatorname{diag}(d_1, \ldots, d_m)$ and $E = \operatorname{diag}(e_1, \ldots, e_n)$. The map $X \mapsto S^{-1}XR$ is linear on $F^{m\times n}$ with inverse $Y \mapsto SYR^{-1}$, so it is an isomorphism. For $Y = S^{-1}XR$,

$$AX = XB \iff SDS^{-1}X = XRER^{-1} \iff DY = YE,$$

multiplying on the left by S^{-1} and on the right by R. The (i, j) entry of $DY - YE$ is $(d_i - e_j)y_{ij}$. So W corresponds to the space of Y with $y_{ij} = 0$ whenever $d_i \neq e_j$, which has as a basis the matrix units at the positions with $d_i = e_j$. By Theorem 7.4.16, each λ occurs exactly $\operatorname{gm}_A(\lambda)$ times among the d_i and $\operatorname{gm}_B(\lambda)$ times among the e_j, so the number of positions is $\sum_\lambda \operatorname{gm}_A(\lambda)\operatorname{gm}_B(\lambda)$.

(b) $W = \{0\}$ exactly when every term of the sum vanishes, that is, when no λ is an eigenvalue of both.

(c) Take $B = A$. Then $\dim W = \sum_\lambda \operatorname{gm}(\lambda)^2$, where $\sum_\lambda \operatorname{gm}(\lambda) = m$ since A is diagonalizable. As each $\operatorname{gm}(\lambda)$ is a nonnegative integer, $\operatorname{gm}(\lambda)^2 \geq \operatorname{gm}(\lambda)$, with equality iff $\operatorname{gm}(\lambda) \in \{0, 1\}$. So $\dim W \geq m$, with equality iff every eigenvalue has $\operatorname{gm} = 1$, that is, iff there are m distinct eigenvalues.

The "if" direction of (b) holds for all $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$, diagonalizable or not. If $AX = XB$, then $A^jX = XB^j$ for all j, so $p_B(A)X = Xp_B(B) = 0$ by Cayley–Hamilton (Theorem 7.6.20). The eigenvalues of $p_B(A)$ are the numbers $p_B(\lambda)$ for eigenvalues λ of A (Theorem 7.6.28), and these are nonzero when no eigenvalue is shared, so $p_B(A)$ is invertible and $X = 0$.

Solution to Exercise 7.7.12

Uses. Theorem 7.6.31 (Commuting operators have a common eigenvector), Proposition 7.3.34 (Eigenvalues of a complexification), Example 7.2.18 (Rotations)
Techniques. Complexify, then take real and imaginary parts (Technique 7.3.35), Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31)
Hint 1. Regard A, B as commuting complex matrices and take a common eigenvector $\mathbf{z} = \mathbf{u} + i\mathbf{w}$ with $\mathbf{u}, \mathbf{w} \in \mathbb{R}^n$.
Hint 2. Compare real and imaginary parts of $A\mathbf{z} = \lambda\mathbf{z}$. For (b), a rotation has no invariant line.
Sketch. By Theorem 7.6.31 there is $\mathbf{z} = \mathbf{u} + i\mathbf{w} \neq \mathbf{0}$ with $A\mathbf{z} = \lambda\mathbf{z}$ and $B\mathbf{z} = \mu\mathbf{z}$. With $\lambda = a + bi$, real and imaginary parts give $A\mathbf{u} = a\mathbf{u} - b\mathbf{w}$ and $A\mathbf{w} = b\mathbf{u} + a\mathbf{w}$; the same for B. So $U = \operatorname{span}(\mathbf{u}, \mathbf{w})$ works. (b) $A = B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

(a) As complex matrices, A and B commute, so by Theorem 7.6.31 there is a nonzero $\mathbf{z} \in \mathbb{C}^n$ with $A\mathbf{z} = \lambda\mathbf{z}$ and $B\mathbf{z} = \mu\mathbf{z}$. Write $\mathbf{z} = \mathbf{u} + i\mathbf{w}$ with $\mathbf{u}, \mathbf{w} \in \mathbb{R}^n$, and $\lambda = a + bi$, $\mu = c + di$ with $a, b, c, d \in \mathbb{R}$. Since A has real entries,

$$A\mathbf{u} + iA\mathbf{w} = A\mathbf{z} = (a + bi)(\mathbf{u} + i\mathbf{w}) = (a\mathbf{u} - b\mathbf{w}) + i(b\mathbf{u} + a\mathbf{w}),$$

and comparing real and imaginary parts, $A\mathbf{u} = a\mathbf{u} - b\mathbf{w}$ and $A\mathbf{w} = b\mathbf{u} + a\mathbf{w}$. In the same way $B\mathbf{u} = c\mathbf{u} - d\mathbf{w}$ and $B\mathbf{w} = d\mathbf{u} + c\mathbf{w}$. So $U = \operatorname{span}(\mathbf{u}, \mathbf{w})$ is invariant under A and B (Proposition 7.2.4). It is nonzero because $\mathbf{z} \neq \mathbf{0}$, and it has dimension at most 2.

(b) Let $A = B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the rotation by $\pi/2$. An invariant line of A would be spanned by a real eigenvector (Proposition 7.2.7), and A has no real eigenvalue (Example 7.2.18). So the only common invariant subspace of dimension 1 or 2 is $\mathbb{R}^2$.

This is the real counterpart of Theorem 7.6.31, and the proof is Technique 7.3.35: solve over $\mathbb{C}$, then take real and imaginary parts. The same move with one matrix is Theorem 7.3.31.

Solution to Exercise 7.7.13

Uses. Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Lemma 7.5.11 (Compressing an operator to a complement), Lemma 7.6.21 (A triangular operator is killed by its diagonal)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. (a) $\ker A$ is nonzero and B-invariant. What is B on it?
Hint 2. (b) Induct on n. Take $U = \operatorname{span}(\mathbf{v})$ from (a) and compress A and B to a complement W. Check that $T\mathbf{w} - T'\mathbf{w} \in U$ (Lemma 7.5.11(a)) gives $(AB)' = A'B'$ and $\left(A^k\right)' = (A')^k$.
Sketch. (a) $\ker A \neq \{\mathbf{0}\}$ is B-invariant and $B|_{\ker A}$ is nilpotent, hence not injective. (b) Induct on n. Compress to a complement W of $U = \operatorname{span}(\mathbf{v})$; since U is invariant under both, compressions respect products, so A', B' commute and are nilpotent. By induction they are strictly upper triangular in a basis of W; prepend $\mathbf{v}$. (c) Sums and products of strictly upper triangular matrices are strictly upper triangular, hence nilpotent.

(a) If A were injective, so would be every power of A, but $A^k = 0$ on $F^n \neq \{\mathbf{0}\}$. So $K = \ker A \neq \{\mathbf{0}\}$. By Proposition 7.2.30, K is B-invariant, and the restriction $B|_K$ satisfies $(B|_K)^k = 0$ whenever $B^k = 0$, so it is not injective either: some $\mathbf{v} \in K$, $\mathbf{v} \neq \mathbf{0}$, has $B\mathbf{v} = \mathbf{0}$. Then $A\mathbf{v} = B\mathbf{v} = \mathbf{0}$.

(b) We induct on n. For $n = 1$, a nilpotent 1×1 matrix is 0. Let $n \geq 2$, take $\mathbf{v}$ from (a), let $U = \operatorname{span}(\mathbf{v})$, and choose a complement W of U, with projection $P = \mathrm{Pr}_{W,U}$. For $T \in \{A, B\}$, let $T' = P \circ T|_W \in \mathrm{L}(W)$ be the compression (Lemma 7.5.11).

We claim that $(T_1T_2)' = T_1'T_2'$ for all T_1, T_2 that map U into U. Let $\mathbf{w} \in W$. By Lemma 7.5.11(a), $T_2\mathbf{w} = T_2'\mathbf{w} + \mathbf{u}$ with $\mathbf{u} \in U$, and $T_1T_2'\mathbf{w} = T_1'T_2'\mathbf{w} + \mathbf{u}'$ with $\mathbf{u}' \in U$. So $T_1T_2\mathbf{w} = T_1'T_2'\mathbf{w} + \mathbf{u}' + T_1\mathbf{u}$, where $\mathbf{u}' + T_1\mathbf{u} \in U$; applying P gives the claim. Here U is killed by A and B, so the claim applies to all products of A and B. Hence $A'B' = (AB)' = (BA)' = B'A'$, and $(A')^k = \left(A^k\right)' = 0$, and likewise for B'. By the induction hypothesis, W has a basis $\mathbf{w}_2, \ldots, \mathbf{w}_n$ in which A' and B' are strictly upper triangular. In the basis $\mathbf{v}, \mathbf{w}_2, \ldots, \mathbf{w}_n$ of F^n, Lemma 7.5.11(b) gives

$$[T_A] = \begin{pmatrix} 0 & * \\ 0 & [A'] \end{pmatrix},$$

with 0 in the corner because $A\mathbf{v} = \mathbf{0}$, and $[A']$ strictly upper triangular. So $[T_A]$ is strictly upper triangular, and the same holds for B. Let S have columns $\mathbf{v}, \mathbf{w}_2, \ldots, \mathbf{w}_n$.

(c) With S from (b), $S^{-1}(A+B)S$ and $S^{-1}ABS = \left(S^{-1}AS\right)\left(S^{-1}BS\right)$ are strictly upper triangular. A strictly upper triangular $n \times n$ matrix M has $M^n = 0$ (Lemma 7.6.21 with all diagonal entries 0), so $A + B$ and AB are nilpotent.

Nothing in this proof uses eigenvalues, which is why it works over every field. Over $\mathbb{C}$, (b) is a special case of Theorem 7.6.32, and (c) also follows from the binomial theorem (Exercise 7.6.51).

Solution to Exercise 7.7.14

Uses. Theorem 7.6.20 (Cayley–Hamilton), Proposition 7.3.18 (The first and last coefficients), Theorem 5.5.9 (The adjugate and the inverse), Corollary 7.1.23 (Polynomials and their values), Aside (Rational functions) (Rational functions)
Techniques. Perturbation and density (Technique 7.6.23), Polynomial identity arguments (Technique 5.5.24)
Hint 1. For invertible A, Cayley–Hamilton gives $AQ = -c_0 I_n$ with Q the bracket. What is c_0?
Hint 2. For general A, replace A by $A - tI_n$. Both sides are matrices of polynomials in t, and they agree whenever t is not an eigenvalue of A.
Sketch. Let Q be the bracket. Cayley–Hamilton: $AQ = -c_0 I_n = (-1)^{n+1}(\det A)I_n$. Multiply by $\operatorname{adj}(A)$ on the left: $(\det A)Q = (-1)^{n+1}(\det A)\operatorname{adj}(A)$, and cancel $\det A \neq 0$. For general A, both sides for $A - tI_n$ are polynomial in t entrywise (the c_j of $p_{A-tI_n}(x) = p_A(x+t)$ are polynomials in t), and they agree for all t outside the at most n eigenvalues of A; so they agree at $t = 0$.

Write $Q(A) = A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I_n$, so that $p_A(A) = AQ(A) + c_0 I_n$.

Invertible A. By Cayley–Hamilton (Theorem 7.6.20), $AQ(A) = -c_0 I_n$, and $c_0 = (-1)^n \det A$ by Proposition 7.3.18, so $AQ(A) = (-1)^{n+1}(\det A)I_n$. Multiplying on the left by $\operatorname{adj}(A)$ and using $\operatorname{adj}(A)A = (\det A)I_n$ (Theorem 5.5.9),

$$(\det A)Q(A) = (-1)^{n+1}(\det A)\operatorname{adj}(A).$$

Since $\det A \neq 0$, $\operatorname{adj}(A) = (-1)^{n+1}Q(A)$.

All A, for infinite F. For $t \in F$, let $A_t = A - tI_n$. Its characteristic polynomial is $\det(xI_n - A + tI_n) = p_A(x + t)$, so each coefficient $c_j(t)$ of p_{A_t} is a polynomial in t, and every entry of

$$Q_t = A_t^{n-1} + c_{n-1}(t)A_t^{n-2} + \cdots + c_1(t)I_n$$

is a polynomial in t. So is every entry of $\operatorname{adj}(A_t)$, being a cofactor of A_t. By the first part, $\operatorname{adj}(A_t) = (-1)^{n+1}Q_t$ whenever A_t is invertible, that is, for every t that is not one of the at most n eigenvalues of A. Each entry of the difference is then a polynomial in t with infinitely many roots, hence $\mathbf{0}$ (Corollary 7.1.23). At $t = 0$ this is the claim. (For a finite field, run the same argument over the field of rational functions $F(t)$ of Aside (Rational functions), where $A - tI_n$ is invertible because its determinant $(-1)^n p_A(t)$ is nonzero, and then set $t = 0$.)

$n = 2$. Here $p_A = x^2 - (\operatorname{tr} A)x + \det A$, so $c_1 = -\operatorname{tr} A$ and the formula reads $\operatorname{adj}(A) = -(A - (\operatorname{tr} A)I_2) = (\operatorname{tr} A)I_2 - A$. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ this is $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, the adjugate of Example 5.5.7.

In particular $\operatorname{adj}(A)$ commutes with every matrix that commutes with A, even when A is singular. The proof is Technique 7.6.23 in its algebraic form: prove the identity where it is easy (invertible matrices), then use that both sides are polynomials.

Problems for Chapter 7

Solution to Problem 7.1

Uses. Theorem 7.2.21 (Eigenvectors for distinct eigenvalues are independent), Corollary 7.2.23 (Eigenspaces are independent), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. (b) Apply $\prod_{j \neq i}(T - \lambda_j \operatorname{id})$ to $\sum c_j \mathbf{v}_j = \mathbf{0}$. For the second proof, apply a single $T - \lambda_j \operatorname{id}$ to a shortest relation.
Hint 2. (c) $D^2 \cos kx = -k^2 \cos kx$. For the full list, group $a_k \cos kx + b_k \sin kx$ into one vector of $E_{-k^2}(D^2)$ and use Corollary 7.2.23.
Sketch. (b) The product kills every term except the ith, which becomes $c_i \prod_{j \neq i}(\lambda_i - \lambda_j)\mathbf{v}_i$. Or: in a shortest relation, applying $T - \lambda_j$ id for an index j in it gives a shorter nontrivial relation. (c) The eigenvalues $0, -1, -4, \ldots, -n^2$ of D^2 are distinct. For the full list, $f_k = a_k \cos kx + b_k \sin kx \in E_{-k^2}$ and $\sum f_k = \mathbf{0}$, so each $f_k = \mathbf{0}$; evaluating at 0 and $\pi/(2k)$ gives $a_k = b_k = 0$.

(a) Let $T \in \mathrm{L}(V)$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be eigenvectors of T with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is linearly independent (Theorem 7.2.21). No hypothesis on $\dim V$ is needed.

(b) *Killing terms.* Suppose $c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$, and fix i. Apply $P_i = \prod_{j \neq i}(T - \lambda_j \operatorname{id})$. By Proposition 7.2.28, $P_i\mathbf{v}_k = \prod_{j \neq i}(\lambda_k - \lambda_j)\mathbf{v}_k$, which is $\mathbf{0}$ for $k \neq i$ (the factor $j = k$ vanishes). So $c_i \prod_{j \neq i}(\lambda_i - \lambda_j)\mathbf{v}_i = \mathbf{0}$. The product is nonzero because the λ's are distinct, and $\mathbf{v}_i \neq \mathbf{0}$, so $c_i = 0$.

A shortest relation. Suppose the list is dependent, and among all nontrivial relations choose one, $\sum_{i \in S} c_i\mathbf{v}_i = \mathbf{0}$ with every $c_i \neq 0$, for which $|S|$ is smallest. Since eigenvectors are nonzero, $|S| \geq 2$. Pick $j \in S$ and apply $T - \lambda_j$ id:

$$\sum_{i \in S, i \neq j} c_i(\lambda_i - \lambda_j)\mathbf{v}_i = \mathbf{0}.$$

The coefficients are nonzero and the index set $S \setminus \{j\}$ is nonempty, so this is a nontrivial relation that is shorter, a contradiction.

(c) V is a subspace of $\mathbb{R}^{\mathbb{R}}$, and D^2 is linear. Since $D^2 \cos kx = -k^2 \cos kx$, $D^2 \sin kx = -k^2 \sin kx$, and $D^2 1 = 0$, the functions $1, \cos x, \ldots, \cos nx$ are eigenvectors of D^2 with the distinct eigenvalues $0, -1, -4, \ldots, -n^2$, so they are independent by (a). The same holds for $\sin x, \ldots, \sin nx$.

For the full list, suppose $a_0 + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx) = 0$ for all x. Let $f_0 = a_0$ and $f_k = a_k \cos kx + b_k \sin kx$. Each f_k lies in the eigenspace $E_{-k^2}(D^2)$, and $f_0 + \cdots + f_n = \mathbf{0}$. By Corollary 7.2.23 the eigenspaces are independent, so every $f_k = \mathbf{0}$. So $a_0 = 0$, and for $k \geq 1$, evaluating f_k at $x = 0$ gives $a_k = 0$, and then at $x = \pi/(2k)$ gives $b_k = 0$.

The theorem separates eigenvalues and nothing more. Inside one eigenspace, here the plane spanned by $\cos kx$ and $\sin kx$, independence needs a separate argument, which the two evaluations supply.

Solution to Problem 7.2

> *Uses.* Theorem 7.4.16 (Criteria for diagonalizability), Corollary 7.2.23 (Eigenspaces are independent), Theorem 4.4.13 (Bases and dimension of a sum), Theorem 7.4.15 (Geometric multiplicity is at most algebraic multiplicity), Proposition 7.1.27 (Multiplicities add up to at most the degree), Proposition 7.1.34 (Roots of a split polynomial), Proposition 7.3.15 (Block-triangular matrices)
> *Techniques.* Dimension counting (Technique 4.2.24), Block decomposition (Technique 2.4.35)
> *Hint 1.* (b) The eigenspaces are independent, so their sum has dimension $\sum \operatorname{gm}(\lambda_i)$. Then chain $\sum \operatorname{gm} \leq \sum \operatorname{am} \leq n$.
> *Hint 2.* (c) Use criterion (IV) for M. For an eigenvalue of D, map $E_\lambda(D)$ into $E_\lambda(M)$ by $\mathbf{y} \mapsto \left(-(B - \lambda I_k)^{-1} C\mathbf{y}, \mathbf{y}\right)$.
> *Sketch.* (b) $\sum E_{\lambda_i}$ is direct, of dimension $\sum \operatorname{gm}$; it is V iff $\sum \operatorname{gm} = n$, and then glued bases are an eigenbasis; conversely an eigenbasis lies in the sum. Also $\sum \operatorname{gm} \leq \sum \operatorname{am} \leq n$ with $\sum \operatorname{am} = n$ iff p_T splits, so $\sum \operatorname{gm} = n$ iff both inequalities are equalities. (c) $p_M = p_B p_D$ splits. For an eigenvalue λ of B, $\mathbf{x} \mapsto (\mathbf{x}, \mathbf{0})$ embeds $E_\lambda(B)$ in $E_\lambda(M)$; for one of D, $\mathbf{y} \mapsto \left(-(B - \lambda I_k)^{-1} C\mathbf{y}, \mathbf{y}\right)$ embeds $E_\lambda(D)$. So $\operatorname{gm}_M \geq \operatorname{am}_M$ everywhere. Counterexample: $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

(a) Let $\dim V = n$, $T \in \mathrm{L}(V)$, with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Then T is diagonalizable if and only if $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}$, if and only if $\sum_i \operatorname{gm}(\lambda_i) = n$, if and only if p_T splits over F and $\operatorname{gm}(\lambda_i) = \operatorname{am}(\lambda_i)$ for all i (Theorem 7.4.16).

(b) The sum $E_{\lambda_1} + \cdots + E_{\lambda_m}$ is direct (Corollary 7.2.23), so its dimension is $\sum_i \operatorname{gm}(\lambda_i)$ (Theorem 4.4.13). If $\sum_i \operatorname{gm}(\lambda_i) = n$, the sum is all of V, and a basis of each E_{λ_i}, glued together, is a basis of V (Theorem 4.4.13) consisting of eigenvectors, so T is diagonalizable. Conversely, if $\mathcal{B}$ is an eigenbasis, every vector of $\mathcal{B}$ lies in some E_{λ_i}, so $V = \operatorname{span} \mathcal{B} \subseteq \sum_i E_{\lambda_i}$ and $\sum_i \operatorname{gm}(\lambda_i) = n$.

For the second equivalence, $\operatorname{gm}(\lambda_i) \leq \operatorname{am}(\lambda_i)$ for each i (Theorem 7.4.15), and $\sum_i \operatorname{am}(\lambda_i) \leq \deg p_T = n$ (Proposition 7.1.27), with equality if and only if p_T splits: if it splits, the multiplicities add up to n (Proposition 7.1.34); if they add up to n, then $\prod (x - \lambda_i)^{\operatorname{am}(\lambda_i)}$ divides p_T and has the same degree, and both are monic, so they are equal. Hence

$$\sum_i \operatorname{gm}(\lambda_i) \leq \sum_i \operatorname{am}(\lambda_i) \leq n,$$

and $\sum \operatorname{gm} = n$ exactly when both inequalities are equalities, that is, when p_T splits and every $\operatorname{gm}(\lambda_i) = \operatorname{am}(\lambda_i)$.

(c) By Proposition 7.3.15, $p_M = p_B p_D$, which splits because p_B and p_D do (Theorem 7.4.16). Let λ be an eigenvalue of M; it is a root of exactly one of p_B, p_D, so $\operatorname{am}_M(\lambda)$ is $\operatorname{am}_B(\lambda)$ or $\operatorname{am}_D(\lambda)$.

If $p_B(\lambda) = 0$, the map $\mathbf{x} \mapsto (\mathbf{x}, \mathbf{0})$ is injective and sends $E_\lambda(B)$ into $E_\lambda(M)$, because $M(\mathbf{x}, \mathbf{0}) = (B\mathbf{x}, \mathbf{0})$. So $\operatorname{gm}_M(\lambda) \geq \operatorname{gm}_B(\lambda) = \operatorname{am}_B(\lambda) = \operatorname{am}_M(\lambda)$.

If $p_D(\lambda) = 0$, then λ is not an eigenvalue of B, so $B - \lambda I_k$ is invertible. For $\mathbf{y} \in E_\lambda(D)$ let $\mathbf{x} = -(B - \lambda I_k)^{-1} C\mathbf{y}$. Then $B\mathbf{x} + C\mathbf{y} = (B - \lambda I_k)\mathbf{x} + \lambda\mathbf{x} + C\mathbf{y} = \lambda\mathbf{x}$ and $D\mathbf{y} = \lambda\mathbf{y}$, so $(\mathbf{x}, \mathbf{y}) \in E_\lambda(M)$. The map $\mathbf{y} \mapsto (\mathbf{x}, \mathbf{y})$ is linear and injective, so $\operatorname{gm}_M(\lambda) \geq \operatorname{gm}_D(\lambda) = \operatorname{am}_M(\lambda)$.

With Theorem 7.4.15, $\operatorname{gm}_M = \operatorname{am}_M$ at every eigenvalue, and M is diagonalizable by (b). Without the hypothesis, $M = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ($B = D = (1)$, $C = (1)$) is not diagonalizable (Example 7.4.8).

Another route: if X solves $BX - XD = -C$, then $\begin{pmatrix} I_k & X \\ 0 & I_l \end{pmatrix}^{-1} M \begin{pmatrix} I_k & X \\ 0 & I_l \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$, and with no common eigenvalue such an X always exists (Exercise 7.7.11(b) and rank–nullity on $X \mapsto BX - XD$).

Solution to Problem 7.3

Uses. Proposition 7.2.37 (Chains are independent), Corollary 7.2.38 (The span of a chain), Theorem 7.5.13 (Upper-triangular form), Proposition 7.5.9 (Multiplicities on the diagonal), Lemma 7.6.21 (A triangular operator is killed by its diagonal), Theorem 4.2.11 (Rank–nullity)
Techniques. Apply an operator to a relation to kill terms (Technique 4.1.25), Dimension counting (Technique 4.2.24)
Hint 1. (b) Apply N^k to a relation $\sum c_j N^j \mathbf{v} = \mathbf{0}$, then N^{k-1}, and so on.
Hint 2. (c) First $N^n = 0$, from a triangular form. Then show $\dim\ker N^{j+1} \le \dim\ker N^j + 1$ by restricting N to $\ker N^{j+1}$.
Sketch. (b) Applying N^k to $\sum_{j=0}^{k} c_j N^j \mathbf{v} = \mathbf{0}$ leaves $c_0 N^k \mathbf{v} = \mathbf{0}$, so $c_0 = 0$; then N^{k-1} gives $c_1 = 0$, and so on. $N\left(N^k \mathbf{v}\right) = \mathbf{0}$ with $N^k \mathbf{v} \neq \mathbf{0}$. (c) In a triangular basis the diagonal is all λ, so $N^n = 0$. N maps $\ker N^{j+1}$ into $\ker N^j$ with kernel $\ker N$, so $\dim\ker N^{j+1} \le \dim\ker N^j + 1$ and $\dim\ker N^{n-1} \le n - 1$. A $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$ gives a chain of length n, a basis, and Corollary 7.2.38(c) gives the matrix.

(a) Let $T \in \mathrm{L}(V)$, $\lambda \in F$, $N = T - \lambda\,\mathrm{id}$, and suppose $N^k \mathbf{v} \neq \mathbf{0}$ and $N^{k+1}\mathbf{v} = \mathbf{0}$. Then $\mathbf{v}, N\mathbf{v}, \ldots, N^k\mathbf{v}$ is linearly independent (Proposition 7.2.37). Moreover $N^k\mathbf{v}$ is an eigenvector of T with eigenvalue λ, and the span of the chain is the Krylov subspace $K_T(\mathbf{v})$, of dimension $k+1$ (Corollary 7.2.38).

(b) Suppose $\sum_{j=0}^{k} c_j N^j \mathbf{v} = \mathbf{0}$, and suppose some c_j is nonzero; let i be the smallest such index. Apply N^{k-i}: every term with $j > i$ becomes $c_j N^{j+k-i}\mathbf{v} = \mathbf{0}$, since $j + k - i \ge k + 1$. What remains is $c_i N^k \mathbf{v} = \mathbf{0}$, so $c_i = 0$, a contradiction. For the last vector, $N\left(N^k\mathbf{v}\right) = N^{k+1}\mathbf{v} = \mathbf{0}$ and $N^k\mathbf{v} \neq \mathbf{0}$, so $TN^k\mathbf{v} = \lambda N^k\mathbf{v}$.

(c) Let $N = T - \lambda\,\mathrm{id}$. By Theorem 7.5.13, T has an upper-triangular matrix in some basis, and by Proposition 7.5.9 its diagonal entries are the roots of p_T, all equal to λ. So the matrix of N is strictly upper triangular, and $N^n = 0$ (Lemma 7.6.21).

Next, for $j \ge 0$, N maps $\ker N^{j+1}$ into $\ker N^j$, and the kernel of this restriction is $\ker N \cap \ker N^{j+1} = \ker N$. By Theorem 4.2.11,

$$\dim\ker N^{j+1} \le \dim\ker N + \dim\ker N^j = 1 + \dim\ker N^j,$$

using $\dim\ker N = \mathrm{gm}(\lambda) = 1$. Starting from $\dim\ker N^0 = 0$, induction gives $\dim\ker N^j \le j$. In particular $\dim\ker N^{n-1} \le n - 1 < n$, so $N^{n-1} \neq 0$.

Choose $\mathbf{v}$ with $N^{n-1}\mathbf{v} \neq \mathbf{0}$; also $N^n\mathbf{v} = \mathbf{0}$. By (b), the chain $\mathbf{v}, N\mathbf{v}, \ldots, N^{n-1}\mathbf{v}$ is independent, and it has n vectors, so it is a basis of V. By Corollary 7.2.38(c), in the basis $N^{n-1}\mathbf{v}, \ldots, N\mathbf{v}, \mathbf{v}$ the matrix of T is as claimed.

Conversely, that matrix has $\mathrm{gm}(\lambda) = n - \mathrm{rank}(N) = 1$. So for an operator with a single eigenvalue, "one eigenvector line" and "one chain" are the same condition. Over any field the same proof works once $N^n = 0$ is known, which Cayley–Hamilton supplies (Theorem 7.6.20). This is the simplest case of the Jordan form (Section 10.3).

Solution to Problem 7.4

Uses. Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Theorem 7.1.45 (Bézout's identity)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Apply an operator to a relation to kill terms (Technique 4.1.25)
Hint 1. Find polynomials with $(x-2)^2 - (x-3)q(x) = 1$, and apply the identity in A to $\mathbf{u} \in U$.
Hint 2. $G = \mathrm{span}(\mathbf{e}_1, \mathbf{e}_2)$ and $E = \mathrm{span}(\mathbf{e}_3)$. On G, A acts as $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, which has one invariant line.
Sketch. $(x-2)^2 - (x-3)(x-1) = 1$, so $\mathbf{u} = (A-2I)^2\mathbf{u} - (A-3I)(A-I)\mathbf{u}$. The first term lies in E and the second in G, because $(A-3I)(A-2I)^2 = 0$, and both lie in U. So $U = (U \cap G) \oplus (U \cap E)$. $U \cap E$ is $\{\mathbf{0}\}$ or E, and $U \cap G$ is $\{\mathbf{0}\}$, $\mathrm{span}(\mathbf{e}_1)$, or G. Six subspaces.

We have $A - 2I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $(A-2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $A - 3I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. So $G = \mathrm{span}(\mathbf{e}_1, \mathbf{e}_2)$, $E = \mathrm{span}(\mathbf{e}_3)$, and $(A-3I_3)(A-2I_3)^2 = 0$, since the third column of $A - 3I_3$ is zero.

The splitting. In $\mathbb{Q}[x]$, $(x-2)^2-(x-3)(x-1)=\left(x^2-4x+4\right)-\left(x^2-4x+3\right)=1$. Substituting A (Proposition 7.2.27), for every $\mathbf{u}\in\mathbb{Q}^3$,

$$\mathbf{u}=(A-2I_3)^2\mathbf{u}-(A-3I_3)(A-I_3)\mathbf{u}.$$

The first term is killed by $A-3I_3$, so it lies in E; the second is killed by $(A-2I_3)^2$, since $(A-2I_3)^2(A-3I_3)(A-I_3)=(A-I_3)(A-3I_3)(A-2I_3)^2=0$, so it lies in G. If U is invariant, it is invariant under every polynomial in A, so for $\mathbf{u}\in U$ both terms lie in U. Hence $U=(U\cap G)+(U\cap E)$, and the sum is direct because $G\cap E=\{\mathbf{0}\}$.

The pieces. $U\cap E$ is a subspace of the line E, so it is $\{\mathbf{0}\}$ or E. $U\cap G$ is an invariant subspace of G, on which A acts by $A\mathbf{e}_1=2\mathbf{e}_1$, $A\mathbf{e}_2=\mathbf{e}_1+2\mathbf{e}_2$. An invariant line is spanned by an eigenvector (Proposition 7.2.7), and the eigenvectors in G are the nonzero multiples of $\mathbf{e}_1$, since $(A-2I_3)(a\mathbf{e}_1+b\mathbf{e}_2)=b\mathbf{e}_1$. So $U\cap G$ is $\{\mathbf{0}\}$, $\operatorname{span}(\mathbf{e}_1)$, or G.

Conversely, each of the $2\cdot 3$ combinations is invariant, being a sum of invariant subspaces. The six invariant subspaces are

$$\{\mathbf{0}\},\quad \operatorname{span}(\mathbf{e}_1),\quad \operatorname{span}(\mathbf{e}_1,\mathbf{e}_2),\quad \operatorname{span}(\mathbf{e}_3),\quad \operatorname{span}(\mathbf{e}_1,\mathbf{e}_3),\quad \mathbb{Q}^3.$$

The same argument shows, for any T with $p(T)q(T)=0$ and p,q coprime, that invariant subspaces split along $\ker p(T)\oplus\ker q(T)$ (compare Exercise 7.2.55). With three distinct eigenvalues there would be $2^3=8$ invariant subspaces (Exercise 7.2.48); the chain in G is what removes two of them.

Solution to Problem 7.5

> *Uses.* Corollary 7.5.14 (Triangularizing a matrix), Proposition 7.5.8 (The diagonal of a triangular matrix), Proposition 7.3.11 (Similar matrices have the same characteristic polynomial), Proposition 7.4.7 (One eigenvalue), Proposition 7.2.28 (Polynomials act on eigenvectors by their values)
> *Techniques.* Finding counterexamples and minimal examples (Technique 1.4.18), Prove it for triangular matrices first (Technique 7.5.23)
> *Hint 1.* (a) The eigenvalues of A^2 are the squares of those of A. Follow $\lambda,\lambda^2,\lambda^4,\dots$.
> *Hint 2.* (b) Try $A=I_n$. (c) Try $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ and compute $A+A^{-1}$. (d) What are the eigenvalues of I_n+N?
> *Sketch.* (a) True: the set S of eigenvalues satisfies $\lambda\in S\Rightarrow\lambda^2\in S$, so $\lambda^{2^i}=\lambda^{2^j}$ for some $i<j$, giving $\lambda=0$ or $\lambda^{2^j-2^i}=1$. (b) False: $A=I_n$, $B=-I_n$. (c) False: $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ has $A+A^{-1}=2I_2$. (d) True: I_n+N has only the eigenvalue 1, and a diagonalizable matrix with one eigenvalue is scalar.

(a) True. Write $A=SRS^{-1}$ with R upper triangular (Corollary 7.5.14). Then $A^2=SR^2S^{-1}$, and R^2 is upper triangular with diagonal entries r_{ii}^2. By Proposition 7.5.8, the eigenvalues of A are the r_{ii} and those of A^2 are the r_{ii}^2. Similar matrices have the same eigenvalues (Proposition 7.3.11), so the finite set S of eigenvalues of A equals $\{\lambda^2\mid\lambda\in S\}$. In particular $\lambda\in S$ implies $\lambda^2,\lambda^4,\lambda^8,\dots\in S$. Since S is finite, $\lambda^{2^i}=\lambda^{2^j}$ for some $i<j$, that is, $\lambda^{2^i}\left(\lambda^{2^j-2^i}-1\right)=0$. So $\lambda=0$ or $\lambda^{2^j-2^i}=1$.

(b) False. With $A=I_n$ and $B=-I_n$, $A^2=B^2=I_n$, but a matrix similar to I_n is $SI_nS^{-1}=I_n$, and $B\neq I_n$.

(c) False. Let $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$. Then $A^{-1}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$ and $A+A^{-1}=2I_2$, which is diagonal, while A is not diagonalizable (Example 7.4.8). The map $\lambda\mapsto\lambda+\lambda^{-1}$ has derivative 0 at $\lambda=1$, and that is where the shear disappears.

(d) True. If $N^k=0$, the only eigenvalue of N is 0 (Proposition 7.2.28). If $(I_n+N)\mathbf{v}=\mu\mathbf{v}$ with $\mathbf{v}\neq\mathbf{0}$, then $N\mathbf{v}=(\mu-1)\mathbf{v}$, so $\mu=1$. A diagonalizable matrix with the single eigenvalue 1 is I_n (Proposition 7.4.7), so $I_n+N=I_n$ and $N=0$. The same proof works over every field.

Solution to Problem 7.6

> *Uses.* Proposition 7.3.2 (Determinants over a commutative ring), Theorem 2.4.34 (Block multiplication), Lemma 7.4.13 (Geometric multiplicity through a rank), Theorem 7.4.16 (Criteria for diagonalizability), Proposition 7.4.7 (One eigenvalue)

Techniques. Block decomposition (Technique 2.4.35), Choose a convenient basis (Technique 4.5.33)
Hint 1. (a) In $xI_{2n} - M$, add the second block column to the first, then subtract the first block row from the second.
Hint 2. (b) Conjugate by $S = \begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}$, which satisfies $S^2 = 2I_{2n}$. Then compare gm and am of $\begin{pmatrix} 2A & 0 \\ 0 & 0 \end{pmatrix}$ with those of $2A$.
Sketch. (a) $(xI - M)\begin{pmatrix} I & 0 \\ I & I \end{pmatrix} = \begin{pmatrix} xI-2A & -A \\ xI-2A & xI-A \end{pmatrix}$, and $\begin{pmatrix} I & 0 \\ -I & I \end{pmatrix}$ times that is $\begin{pmatrix} xI-2A & -A \\ 0 & xI \end{pmatrix}$, with determinant $p_{2A}x^n$. (b) $S^{-1}MS = \begin{pmatrix} 2A & 0 \\ 0 & 0 \end{pmatrix}$. For $\mu \neq 0$ the multiplicities of μ for this block matrix and for $2A$ agree, and for $\mu = 0$ both increase by n; so criterion (IV) holds for one iff for the other, and $2A$ is diagonalizable iff A is. (c) M^2 has blocks $2A^2 = 0$. A diagonalizable nilpotent matrix is 0.

(a) Work in $F[x]^{2n\times 2n}$, where determinants are multiplicative and block triangular matrices have determinant the product of the diagonal blocks (Proposition 7.3.2). By Theorem 2.4.34,

$$(xI_{2n} - M)\begin{pmatrix} I_n & 0 \\ I_n & I_n \end{pmatrix} = \begin{pmatrix} xI_n - 2A & -A \\ xI_n - 2A & xI_n - A \end{pmatrix}, \quad \begin{pmatrix} I_n & 0 \\ -I_n & I_n \end{pmatrix}\begin{pmatrix} xI_n - 2A & -A \\ xI_n - 2A & xI_n - A \end{pmatrix} = \begin{pmatrix} xI_n - 2A & -A \\ 0 & xI_n \end{pmatrix}.$$

The two outer factors have determinant 1, so $p_M = \det(xI_n - 2A)\det(xI_n) = x^n p_{2A}$.

(b) Let $S = \begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix}$. Then $S^2 = 2I_{2n}$, so $S^{-1} = (1/2)S$, and

$$MS = \begin{pmatrix} 2A & 0 \\ 2A & 0 \end{pmatrix}, \quad S^{-1}MS = \frac{1}{2}\begin{pmatrix} 4A & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2A & 0 \\ 0 & 0 \end{pmatrix} =: M'.$$

For $\mu \in F$, $\ker(M' - \mu I_{2n}) = \ker(2A - \mu I_n) \times \ker(-\mu I_n)$, so $\mathrm{gm}_{M'}(\mu) = \mathrm{gm}_{2A}(\mu)$ for $\mu \neq 0$ and $\mathrm{gm}_{M'}(0) = \mathrm{gm}_{2A}(0) + n$. By (a) applied to M', or directly, $p_{M'} = x^n p_{2A}$, so $\mathrm{am}_{M'}(\mu) = \mathrm{am}_{2A}(\mu)$ for $\mu \neq 0$, $\mathrm{am}_{M'}(0) = \mathrm{am}_{2A}(0) + n$, and $p_{M'}$ splits iff p_{2A} does. By Theorem 7.4.16(IV), M' is diagonalizable iff $2A$ is. Finally $2A = S_0 D S_0^{-1}$ iff $A = S_0(D/2)S_0^{-1}$, so $2A$ is diagonalizable iff A is, and M is similar to M'.

(c) By Theorem 2.4.34 every block of M^2 is $A^2 + A^2 = 2A^2 = 0$ in characteristic 2. So M is nilpotent, and its only eigenvalue is 0. If M is diagonalizable, then $M = 0$ (Proposition 7.4.7), that is, $A = 0$; conversely $A = 0$ gives $M = 0$. This matches (a): in characteristic 2, $2A = 0$ and $p_M = x^{2n}$. The similarity of (b) fails because S is singular when $2 = 0$.

Solution to Problem 7.7

Uses. Proposition 2.4.23 (Transpose of a product), Theorem 4.3.21 (Column–row factorization), Theorem 2.4.34 (Block multiplication)
Techniques. Transpose to trade rows for columns (Technique 3.2.27), Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. (a) Compute $\mathbf{w}^{\mathsf{T}}A\mathbf{v}$ in two ways.
Hint 2. (b) Transpose $S^{-1}A = DS^{-1}$. The products $\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j$ are the entries of $S^{-1}S$. For the sum, multiply SD by S^{-1} column by row.
Sketch. (a) $\lambda\mathbf{w}^{\mathsf{T}}\mathbf{v} = \mathbf{w}^{\mathsf{T}}A\mathbf{v} = \left(A^{\mathsf{T}}\mathbf{w}\right)^{\mathsf{T}}\mathbf{v} = \mu\mathbf{w}^{\mathsf{T}}\mathbf{v}$. (b) $A^{\mathsf{T}}\left(S^{-1}\right)^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}}D$; $\left(S^{-1}S\right)_{ij} = \mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j$; $SDS^{-1} = \sum_i(\lambda_i\mathbf{v}_i)\mathbf{w}_i^{\mathsf{T}}$. (c) $P_iP_j = \mathbf{v}_i\left(\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j\right)\mathbf{w}_j^{\mathsf{T}}$, $\sum P_i = SS^{-1}$, and $A^k = SD^kS^{-1}$.

(a) $\lambda\mathbf{w}^{\mathsf{T}}\mathbf{v} = \mathbf{w}^{\mathsf{T}}(A\mathbf{v}) = \left(A^{\mathsf{T}}\mathbf{w}\right)^{\mathsf{T}}\mathbf{v} = \mu\mathbf{w}^{\mathsf{T}}\mathbf{v}$, using $\left(A^{\mathsf{T}}\mathbf{w}\right)^{\mathsf{T}} = \mathbf{w}^{\mathsf{T}}A$ (Proposition 2.4.23). So $(\lambda - \mu)\mathbf{w}^{\mathsf{T}}\mathbf{v} = 0$, and $\lambda \neq \mu$ gives $\mathbf{w}^{\mathsf{T}}\mathbf{v} = 0$.

(b) From $AS = SD$ we get $S^{-1}A = DS^{-1}$, and transposing, $A^{\mathsf{T}}\left(S^{-1}\right)^{\mathsf{T}} = \left(S^{-1}\right)^{\mathsf{T}}D$. The columns of $\left(S^{-1}\right)^{\mathsf{T}}$ are $\mathbf{w}_1, \ldots, \mathbf{w}_n$, so column i of this equation reads $A^{\mathsf{T}}\mathbf{w}_i = \lambda_i\mathbf{w}_i$. The (i, j) entry of $S^{-1}S = I_n$ is row i of S^{-1} times column j of S, which is $\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j$. Finally, SD has columns $\lambda_i\mathbf{v}_i$, and a product of a matrix with columns $\mathbf{a}_i$ and a matrix with rows $\mathbf{b}_i^{\mathsf{T}}$ is $\sum_i \mathbf{a}_i\mathbf{b}_i^{\mathsf{T}}$ (Theorem 2.4.34, or the column–row expansion of Theorem 4.3.21). So $A = (SD)S^{-1} = \sum_i \lambda_i\mathbf{v}_i\mathbf{w}_i^{\mathsf{T}}$.

(c) By (b), $P_iP_j = \mathbf{v}_i\left(\mathbf{w}_i^{\mathsf{T}}\mathbf{v}_j\right)\mathbf{w}_j^{\mathsf{T}}$, which is 0 for $i \neq j$ and P_i for $i = j$. By the column–row expansion again, $\sum_i P_i = SS^{-1} = I_n$. And $A^k = SD^kS^{-1} = \sum_i \lambda_i^k P_i$ by the same expansion applied to SD^k.

So a diagonalizable A is a combination of rank-one idempotents P_i, the projection onto $\mathrm{span}(\mathbf{v}_i)$ along the span of the other $\mathbf{v}_j$; the left eigenvectors $\mathbf{w}_i$ measure the $\mathbf{v}_i$-coordinate. When the λ_i are distinct, P_i is the Lagrange polynomial $L_i(A)$ of Technique 7.6.13. Over $\mathbb{R}$ the $\mathbf{w}_i$ equal the $\mathbf{v}_i$ exactly when $S^{-1} = S^{\mathsf{T}}$, which makes $A = SDS^{\mathsf{T}}$ symmetric; conversely every real symmetric matrix arises this way (Section 8.2).

Solution to Problem 7.8

Uses. Corollary 7.4.5 (Distinct eigenvalues give a diagonalization), Corollary 7.2.22 (At most $\dim V$ eigenvalues), Theorem 7.5.22 (Trace and determinant from eigenvalues), Proposition 5.5.18 (Tridiagonal determinants)
Techniques. Plug in clever test vectors (Technique 6.1.40), Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. (a) $(A_n\mathbf{v})_j = v_{j-1} + v_{j+1}$, and $\sin((j-1)\theta) + \sin((j+1)\theta) = 2\cos\theta\sin j\theta$.
Hint 2. (b) The error term vanishes when $(n+1)\theta \in \pi\mathbb{Z}$. (c) Expand along the last row: $d_m = -d_{m-2}$.
Sketch. (a) The sum formula gives the middle rows; row 1 uses $\sin 0 = 0$, and row n is short of $\sin((n+1)\theta)$. (b) $\theta_k = k\pi/(n+1)$, $1 \le k \le n$, gives eigenvectors ($v_1 = \sin\theta_k \neq 0$) with n distinct eigenvalues, since cos is injective on $(0,\pi)$. (c) $d_1 = 0$, $d_2 = -1$, $d_m = -d_{m-2}$; so $\det A_n = 0$ for odd n and $(-1)^{n/2}$ for even n, and this is the product.

(a) Let $v_j = \sin j\theta$, and set $v_0 = \sin 0 = 0$. Row j of A_n has ones in columns $j-1$ and $j+1$ (when they exist), so $(A_n\mathbf{v})_j = v_{j-1} + v_{j+1}$ for $j \le n-1$ and $(A_n\mathbf{v})_n = v_{n-1}$. By the sum formula $\sin(a-b) + \sin(a+b) = 2\sin a\cos b$,

$$v_{j-1} + v_{j+1} = \sin((j-1)\theta) + \sin((j+1)\theta) = 2\cos\theta\sin j\theta,$$

which gives the rows $1 \le j \le n-1$ (row 1 because $v_0 = 0$). For row n, $v_{n-1} = 2\cos\theta\sin n\theta - \sin((n+1)\theta)$ by the same formula. This is the claim.

(b) Let $\theta_k = k\pi/(n+1)$ for $k = 1, \dots, n$. Then $\sin((n+1)\theta_k) = \sin k\pi = 0$, so $A_n\mathbf{v}_k = 2\cos\theta_k\mathbf{v}_k$, and $\mathbf{v}_k \neq \mathbf{0}$ because its first entry is $\sin\theta_k > 0$. So each $2\cos\theta_k$ is an eigenvalue. The θ_k lie in $(0,\pi)$, where cos is strictly decreasing, so these n eigenvalues are distinct. An $n \times n$ matrix has at most n eigenvalues (Corollary 7.2.22), so these are all of them, and A_n is diagonalizable (Corollary 7.4.5).

(c) Let $d_m = \det A_m$. Directly, $d_1 = 0$ and $d_2 = \det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1$. By Proposition 5.5.18 with diagonal entries 0 and off-diagonal entries 1, $d_m = 0 \cdot d_{m-1} - 1 \cdot 1 \cdot d_{m-2} = -d_{m-2}$ for $m \ge 3$. Hence $\det A_n = 0$ for odd n and $\det A_n = (-1)^{n/2}$ for even n. Since A_n is diagonalizable with the eigenvalues of (b), its determinant is their product (Theorem 7.5.22), so

$$\prod_{k=1}^{n} 2\cos\frac{k\pi}{n+1} = \begin{cases} 0 & \text{if } n \text{ is odd} \\ (-1)^{n/2} & \text{if } n \text{ is even.} \end{cases}$$

For odd n this is visible directly, since the factor with $k = (n+1)/2$ is $2\cos(\pi/2) = 0$. In the same way, $\operatorname{tr} A_n = 0$ says $\sum_k \cos(k\pi/(n+1)) = 0$, which also follows from $\cos(\pi - x) = -\cos x$.

A_n is the adjacency matrix of a path with n vertices; Section 16.5 studies such matrices. The eigenvector was guessed from the recurrence $v_{j-1} + v_{j+1} = cv_j$, whose solutions are combinations of $\sin j\theta$ and $\cos j\theta$, and the boundary rows chose sin and θ.

Solution to Problem 7.9

Uses. Theorem 7.1.14 (Division with remainder), Proposition 7.2.35 (A basis of a Krylov subspace), Theorem 7.6.2 (Existence and uniqueness of the minimal polynomial), Theorem 2.2.1 (Linear maps are determined by a basis), Theorem 7.6.15 (Diagonalizability and the minimal polynomial)
Techniques. Krylov sequences must become dependent (Technique 7.3.28), Check it on a basis (Technique 2.2.7)
Hint 1. (a) Divide f by $p_{\mathbf{v}}$ and apply the remainder to $\mathbf{v}$.
Hint 2. (b) Two operators agree if they agree on a basis. (c) Compute $A\mathbf{e}_j$ and $A^2\mathbf{e}_j$ and find the first dependence for each j.
Sketch. (a) $f = qp_{\mathbf{v}} + r$ with $\deg r < \deg p_{\mathbf{v}}$, and $r(T)\mathbf{v} = f(T)\mathbf{v}$; minimality forces $r = \mathbf{0}$ when this is $\mathbf{0}$. (b) $f(T) = 0$ iff $f(T)\mathbf{b}_i = \mathbf{0}$ for all i, iff each $p_{\mathbf{b}_i}$ divides f by (a). m_T is the monic annihilating polynomial of least degree. (c) Each $\mathbf{e}_j$ has $p_{\mathbf{e}_j} = (x-2)^2$, so $m_A = (x-2)^2$ and A is not diagonalizable.

(a) If $p_{\mathbf{v}}$ divides f, say $f = qp_{\mathbf{v}}$, then $f(T)\mathbf{v} = q(T)p_{\mathbf{v}}(T)\mathbf{v} = \mathbf{0}$ (Proposition 7.2.27). Conversely, suppose $f(T)\mathbf{v} = \mathbf{0}$, and divide: $f = qp_{\mathbf{v}} + r$ with $r = \mathbf{0}$ or $\deg r < \deg p_{\mathbf{v}}$ (Theorem 7.1.14). Then $r(T)\mathbf{v} = f(T)\mathbf{v} - q(T)p_{\mathbf{v}}(T)\mathbf{v} = \mathbf{0}$. If $r \neq \mathbf{0}$, dividing r by its leading coefficient gives a monic polynomial of smaller degree than $p_{\mathbf{v}}$ killing $\mathbf{v}$, which contradicts the choice of $p_{\mathbf{v}}$. So $r = \mathbf{0}$.

(b) $f(T) = 0$ if and only if $f(T)\mathbf{b}_i = \mathbf{0}$ for every i, since a linear map is determined by its values on a basis (Theorem 2.2.1). By (a), this holds if and only if every $p_{\mathbf{b}_i}$ divides f. By definition, m_T is the monic polynomial of least degree with $m_T(T) = 0$ (Theorem 7.6.2), so it is the monic polynomial of least degree divisible by all the $p_{\mathbf{b}_i}$.

(c) For $\mathbf{e}_1$: $A\mathbf{e}_1 = (3, 0, 1)$ and $A^2\mathbf{e}_1 = (8, 0, 4)$. The first two are independent, and $A^2\mathbf{e}_1 - 4A\mathbf{e}_1 + 4\mathbf{e}_1 = (8 - 12 + 4, 0, 4 - 4) = \mathbf{0}$, so $p_{\mathbf{e}_1} = x^2 - 4x + 4 = (x-2)^2$. For $\mathbf{e}_2$: $A\mathbf{e}_2 = (1, 2, 1)$ and $A^2\mathbf{e}_2 = (4, 4, 4) = 4A\mathbf{e}_2 - 4\mathbf{e}_2$, so again $p_{\mathbf{e}_2} = (x-2)^2$. For $\mathbf{e}_3$: $A\mathbf{e}_3 = (-1, 0, 1)$ and $A^2\mathbf{e}_3 = (-4, 0, 0) = 4A\mathbf{e}_3 - 4\mathbf{e}_3$, so $p_{\mathbf{e}_3} = (x-2)^2$. By (b), $m_A = (x-2)^2$. It has a repeated root, so A is not diagonalizable (Theorem 7.6.15).

This costs a few matrix–vector products per basis vector, and no determinant. Here $p_A = (x-2)^3$ (its only root is 2, and $\operatorname{tr} A = 6$), so m_A is a proper divisor of p_A, and in fact $(A - 2I_3)^2 = 0$.

Solution to Problem 7.10

Uses. Corollary 7.5.14 (Triangularizing a matrix), Proposition 7.5.8 (The diagonal of a triangular matrix), Corollary 7.4.5 (Distinct eigenvalues give a diagonalization)
Techniques. Prove it for triangular matrices first (Technique 7.5.23), Choose a convenient basis (Technique 4.5.33)
Hint 1. Triangularize, $A = SRS^{-1}$. Which diagonal matrix can you subtract from R so that the rest has distinct diagonal entries?
Hint 2. A triangular matrix with distinct diagonal entries has distinct eigenvalues. For (c), choose the diagonal entries to avoid 0 as well.
Sketch. (a) $A = SRS^{-1}$ with R triangular. Let $E = \operatorname{diag}(r_{11} - 1, r_{22} - 2, \ldots, r_{nn} - n)$. Then $R - E$ is triangular with diagonal $1, 2, \ldots, n$, so diagonalizable, and $A = S(R-E)S^{-1} + SES^{-1}$. (b) $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. (c) Choose E diagonal with $e_{ii} \neq 0$ and $e_{ii} \neq r_{ii}$.

(a) By Corollary 7.5.14, $A = SRS^{-1}$ with S invertible and R upper triangular. Let $E = \operatorname{diag}(r_{11} - 1, r_{22} - 2, \ldots, r_{nn} - n)$. Then $R - E$ is upper triangular with diagonal entries $1, 2, \ldots, n$, which are its eigenvalues (Proposition 7.5.8); they are distinct, so $R - E$ is diagonalizable (Corollary 7.4.5). So are $S(R-E)S^{-1}$ and SES^{-1}, being similar to a diagonalizable and to a diagonal matrix, and

$$A = S(R-E)S^{-1} + SES^{-1}.$$

(b) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The first summand is triangular with distinct diagonal entries $1, 0$, and the second is diagonal.

(c) With $A = SRS^{-1}$ as in (a), choose $e_i \in \mathbb{C}$ with $e_i \neq 0$ and $e_i \neq r_{ii}$, and let $E = \operatorname{diag}(e_1, \ldots, e_n)$. Then E and $R - E$ are triangular with nonzero diagonal entries, hence invertible, and $A = S(R-E)S^{-1} + SES^{-1}$.

Diagonalizable matrices are not closed under addition, so (a) says that they generate everything additively. The same construction with $E = -\varepsilon \operatorname{diag}(1, 2, \ldots, n)$ shows more: $A + \varepsilon S \operatorname{diag}(1, \ldots, n) S^{-1}$ is diagonalizable whenever $r_{ii} + i\varepsilon$ are distinct, which holds for all but finitely many ε, so every complex matrix is a limit of diagonalizable ones (Technique 7.6.23).

Solution to Problem 7.11

Uses. Proposition 7.2.4 (Invariance is checked on a spanning list), Proposition 1.5.14 (Appending a vector), Theorem 4.1.10 (Independent lists are short), Exercise 7.3.43
Techniques. Extend a basis of a subspace (Technique 4.1.18), Choose a convenient basis (Technique 4.5.33)
Hint 1. Build the list greedily: if $W_k = \operatorname{span}(\mathbf{v}_1, A\mathbf{v}_1, \ldots, \mathbf{v}_k, A\mathbf{v}_k)$ is not all of $\mathbb{R}^n$, pick $\mathbf{v}_{k+1} \notin W_k$. W_k is A-invariant.
Hint 2. To see the new list is independent, show $a\mathbf{v} + bA\mathbf{v} \in W_k$ forces $a = b = 0$: apply $a\,\mathrm{id} - bA$ and use $(a - bA)(a + bA) = (a^2 + b^2)\,\mathrm{id}$.
Sketch. Inductively build $\mathbf{v}_1, \ldots, \mathbf{v}_k$ with $\mathbf{v}_i, A\mathbf{v}_i$ independent; W_k is A-invariant since $A(A\mathbf{v}_i) = -\mathbf{v}_i$. For $\mathbf{v} \notin W_k$, if $(a + bA)\mathbf{v} \in W_k$ with $(a, b) \neq (0, 0)$, then $(a^2 + b^2)\mathbf{v} = (a - bA)(a + bA)\mathbf{v} \in W_k$, a contradiction. So the list grows by two independent vectors until it spans; $n = 2m$. In this basis $A\mathbf{v}_i = A\mathbf{v}_i$ and $A(A\mathbf{v}_i) = -\mathbf{v}_i$, which is the matrix J.

(a) We construct $\mathbf{v}_1, \mathbf{v}_2, \ldots$ such that for each k the list $\mathcal{L}_k = \mathbf{v}_1, \ldots, \mathbf{v}_k, A\mathbf{v}_1, \ldots, A\mathbf{v}_k$ is independent. The empty list ($k = 0$) is independent. Suppose $\mathcal{L}_k$ is independent and $W_k = \operatorname{span} \mathcal{L}_k \neq \mathbb{R}^n$. Since $A(\mathbf{v}_i) = A\mathbf{v}_i$ and $A(A\mathbf{v}_i) = -\mathbf{v}_i$, A maps the list into W_k, so W_k is A-invariant (Proposition 7.2.4). Choose $\mathbf{v} = \mathbf{v}_{k+1} \notin W_k$.

We claim that if $a, b \in \mathbb{R}$ and $a\mathbf{v} + bA\mathbf{v} \in W_k$, then $a = b = 0$. Since $A^2 = -I_n$,

$$(aI_n - bA)(aI_n + bA) = a^2 I_n - b^2 A^2 = \left(a^2 + b^2\right) I_n.$$

So $\left(a^2 + b^2\right)\mathbf{v} = (aI_n - bA)(a\mathbf{v} + bA\mathbf{v}) \in W_k$, because W_k is invariant under A and hence under $aI_n - bA$. As $\mathbf{v} \notin W_k$, $a^2 + b^2 = 0$, so $a = b = 0$.

Now suppose $\mathbf{w} + a\mathbf{v} + bA\mathbf{v} = \mathbf{0}$ with $\mathbf{w} \in W_k$. Then $a\mathbf{v} + bA\mathbf{v} = -\mathbf{w} \in W_k$, so $a = b = 0$ by the claim, and then $\mathbf{w} = \mathbf{0}$. So $\mathcal{L}_{k+1}$ is independent. An independent list in $\mathbb{R}^n$ has at most n vectors (Theorem 4.1.10), so the process stops, at some m with $W_m = \mathbb{R}^n$. Then $\mathcal{L}_m$ is a basis with $2m$ vectors, and $n = 2m$.

(b) Let S have columns $\mathbf{v}_1, \ldots, \mathbf{v}_m, A\mathbf{v}_1, \ldots, A\mathbf{v}_m$. For $i \le m$, A sends column i to column $m + i$, and column $m + i$ to $-\mathbf{v}_i$, which is -1 times column i. So $AS = SJ$, with J as given, and $A = SJS^{-1}$. Two real $n \times n$ matrices with square $-I_n$ are both similar to J, hence to each other.

The same argument shows that a real operator with $T^2 = -\mathrm{id}$ turns V into a complex vector space, with $i\mathbf{v} \coloneqq T\mathbf{v}$, and then (a) is the statement that a complex basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ gives the real basis $\mathbf{v}_j, i\mathbf{v}_j$. Exercise 7.3.43 proved only that n is even; the basis is what gives the similarity class.

Solution to Problem 7.12

> *Uses.* Proposition 4.1.26 (Nilpotent chains are independent), Corollary 7.2.38 (The span of a chain), Proposition 7.4.7 (One eigenvalue), Exercise A.2.8, Exercise 7.1.56, Theorem 4.2.11 (Rank–nullity)
> *Techniques.* Apply an operator to a relation to kill terms (Technique 4.1.25), Dimension counting (Technique 4.2.24), Compare degrees and leading coefficients (Technique 7.1.6)
> *Hint 1.* (a) $(x+1)^m - x^m = mx^{m-1} + \cdots$. Over $\mathbb{Q}$, $m \neq 0$. So the chain starting at x^n ends at the nonzero constant $n!$.
> *Hint 2.* (b) $T^p f = f(x + p)$. For Δ^p, use the binomial theorem and $p \mid \binom{p}{j}$ for $0 < j < p$. For $\mathrm{gm}(1)$, compare $\dim\ker\Delta^{j+1}$ with $\dim\ker\Delta^j$, or find a nonconstant f with $f(x+1) = f$.
> *Sketch.* (a) The leading term of $\Delta(ax^m + \cdots)$ is max^{m-1}, so $\Delta^n x^n = n! \neq 0$ and $\Delta^{n+1}x^n = 0$; the chain is independent of length $n + 1 = \dim V$. $\ker\Delta$ is the constants, T has only eigenvalue 1, and $T \neq \mathrm{id}$, so not diagonalizable. (b) $T^p f = f(x + p) = f$. $(T - \mathrm{id})^p = T^p - \mathrm{id}$ in characteristic p. $x^p - x$ and 1 lie in $\ker\Delta$, so $\mathrm{gm}(1) \ge 2$.

(a) For $m \ge 1$, $(x+1)^m - x^m = mx^{m-1} + (\text{terms of degree} < m-1)$ by the binomial theorem. So if $f = ax^m + (\text{lower terms})$ with $a \neq 0$ and $m \ge 1$, then $\Delta f = max^{m-1} + (\text{lower terms})$, and $ma \neq 0$ in $\mathbb{Q}$: $\deg\Delta f = m - 1$. Constants are killed by Δ. Hence $\Delta^j x^n$ has degree $n - j$ for $j \le n$, with $\Delta^n x^n = n!$, and $\Delta^{n+1}x^n = 0$. By Proposition 4.1.26 (with T there replaced by Δ), $x^n, \Delta x^n, \ldots, \Delta^n x^n$ is independent; it has $n + 1 = \dim V$ vectors, so it is a basis.

If $Tf = \lambda f$ with $f \neq \mathbf{0}$, compare leading coefficients: $f(x+1)$ has the same leading coefficient as f, so $\lambda = 1$. And 1 is an eigenvalue, with $E_1 = \ker\Delta$ = the constants, since Δ lowers the degree of every nonconstant polynomial. So $\mathrm{gm}(1) = 1$. In the chain basis, read from the end, $T = \mathrm{id} + \Delta$ has the matrix of Corollary 7.2.38(c) with $\lambda = 1$. T is not diagonalizable: its only eigenvalue is 1 and $T \neq \mathrm{id}_V$ (Proposition 7.4.7).

(b) For $f \in V$, $T^p f = f(x + 1 + \cdots + 1) = f(x + p) = f$, since $p = 0$ in $\mathbb{F}_p$. So $T^p = \mathrm{id}_V$. T and id_V commute, so by the binomial theorem (Exercise A.2.8),

$$\Delta^p = (T - \mathrm{id}_V)^p = \sum_{j=0}^{p} \binom{p}{j} (-1)^{p-j} T^j = T^p + (-1)^p \mathrm{id}_V = T^p - \mathrm{id}_V = 0,$$

because p divides $\binom{p}{j}$ for $0 < j < p$, and $(-1)^p = -1$ in $\mathbb{F}_p$ (for $p = 2$, $1 = -1$). For $\mathrm{gm}(1)$: in $\mathbb{F}_p[x]$, $(x+1)^p = x^p + 1$ (the same binomial fact), so $g = x^p - x$ satisfies $g(x+1) = x^p + 1 - x - 1 = g$. Since $n \ge p$, $g \in V$, and $1, g$ are independent elements of $\ker\Delta = E_1$. So $\mathrm{gm}(1) \ge 2$. (Compare Exercise 7.1.56.)

Over $\mathbb{Q}$, T is one long chain; over $\mathbb{F}_p$, every chain has length at most p because $\Delta^p = 0$, so V, of dimension $n + 1 > p$, needs at least two chains, and each chain contributes one eigenvector. Counting with Theorem 4.2.11 gives more: $N = \Delta$ maps $\ker N^{j+1}$ to $\ker N^j$ with kernel $\ker N$, so $n + 1 = \dim\ker N^p \le p\,\mathrm{gm}(1)$, that is, $\mathrm{gm}(1) \ge (n+1)/p$.

Solution to Problem 7.13

Uses. Lemma 2.4.24 (Matrix units), Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem 5.4.10 (Multiplicativity), Proposition 5.3.17 (Block-diagonal matrices)
Techniques. Study a linear map on a space of matrices (Technique 4.2.29), Diagonalize, solve, conjugate back (Technique 7.4.22), Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. (b) If S and R have the eigenvectors of A and of B^{T} as columns, then $X \mapsto SXR^{\mathsf{T}}$ sends the matrix units U_{ij} to $\mathbf{v}_i\mathbf{w}_j^{\mathsf{T}}$.
Hint 2. (c) L is the composite of $X \mapsto AX$ and $X \mapsto XB$. The first acts on each column of X separately.
Sketch. (a) $A\mathbf{v}\mathbf{w}^{\mathsf{T}}B = \lambda\mathbf{v}\left(B^{\mathsf{T}}\mathbf{w}\right)^{\mathsf{T}} = \lambda\mu\mathbf{v}\mathbf{w}^{\mathsf{T}}$, nonzero. (b) $B^{\mathsf{T}} = \left(R_0^{-1}\right)^{\mathsf{T}}ER_0^{\mathsf{T}}$ is diagonalizable with eigenbasis $\mathbf{w}_j$; the n^2 matrices $\mathbf{v}_i\mathbf{w}_j^{\mathsf{T}} = SU_{ij}R^{\mathsf{T}}$ form a basis, an eigenbasis of L. Sum and product of $\lambda_i\mu_j$. (c) Ordering the matrix units column by column, $X \mapsto AX$ is block diagonal with n blocks A; ordering row by row, $X \mapsto XB$ is block diagonal with n blocks B^{T}. Multiply.

(a) $L\left(\mathbf{v}\mathbf{w}^{\mathsf{T}}\right) = (A\mathbf{v})\left(\mathbf{w}^{\mathsf{T}}B\right) = \lambda\mathbf{v}\left(B^{\mathsf{T}}\mathbf{w}\right)^{\mathsf{T}} = \lambda\mu\mathbf{v}\mathbf{w}^{\mathsf{T}}$, and $\mathbf{v}\mathbf{w}^{\mathsf{T}} \neq 0$ since its (i,j) entry v_iw_j is nonzero when $v_i \neq 0 \neq w_j$.

(b) Let S be invertible with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, eigenvectors of A for $\lambda_1, \ldots, \lambda_n$. If $B = R_0ER_0^{-1}$ with $E = \operatorname{diag}(\mu_1, \ldots, \mu_n)$, then $B^{\mathsf{T}} = \left(R_0^{-1}\right)^{\mathsf{T}}ER_0^{\mathsf{T}}$, so the columns $\mathbf{w}_1, \ldots, \mathbf{w}_n$ of $R = \left(R_0^{-1}\right)^{\mathsf{T}}$ are eigenvectors of B^{T} for $\mu_1, \ldots, \mu_n$, and R is invertible. The map $X \mapsto SXR^{\mathsf{T}}$ is an isomorphism of $F^{n\times n}$, and it sends the matrix unit $U_{ij} = \mathbf{e}_i\mathbf{e}_j^{\mathsf{T}}$ (Lemma 2.4.24) to $S\mathbf{e}_i\left(R\mathbf{e}_j\right)^{\mathsf{T}} = \mathbf{v}_i\mathbf{w}_j^{\mathsf{T}}$. So the n^2 matrices $\mathbf{v}_i\mathbf{w}_j^{\mathsf{T}}$ form a basis of $F^{n\times n}$, and by (a) it is an eigenbasis of L with eigenvalues $\lambda_i\mu_j$. So L is diagonalizable (Theorem 7.4.2), and its trace and determinant are the sum and product of these eigenvalues (Theorem 7.5.22):

$$\operatorname{tr} L = \sum_{i,j}\lambda_i\mu_j = \left(\sum_i \lambda_i\right)\left(\sum_j \mu_j\right) = \operatorname{tr} A \operatorname{tr} B, \quad \det L = \prod_{i,j}\lambda_i\mu_j = \left(\prod_i \lambda_i\right)^n\left(\prod_j \mu_j\right)^n = (\det A)^n(\det B)^n.$$

(c) Write $L = L_1 \circ L_2$ with $L_1(X) = AX$ and $L_2(X) = XB$, so $\det L = \det L_1 \det L_2$ (Theorem 5.4.10). Order the matrix units column by column: $U_{11}, \ldots, U_{n1}, U_{12}, \ldots, U_{n2}, \ldots$. Column j of AX is A times column j of X, so in this basis L_1 has the block diagonal matrix with n diagonal blocks A, and $\det L_1 = (\det A)^n$ (Proposition 5.3.17). Order them row by row instead: row i of XB is $\left(B^{\mathsf{T}}(\text{row } i \text{ of } X)^{\mathsf{T}}\right)^{\mathsf{T}}$, so L_2 has the block diagonal matrix with n blocks B^{T}, and $\det L_2 = \left(\det B^{\mathsf{T}}\right)^n = (\det B)^n$. The determinant of an operator does not depend on the basis, so $\det L = (\det A)^n(\det B)^n$.

The trace formula also holds for all A, B (it is linear in each of A and B and can be checked on matrix units). Over $\mathbb{C}$, the eigenvalues of L are always the $\lambda_i\mu_j$ counted with multiplicity: triangularize A and B^{T} and order the basis $\mathbf{v}_i\mathbf{w}_j^{\mathsf{T}}$ suitably, or use Technique 7.6.23. The map $X \mapsto AX - XA$ of Exercise 7.5.45 works the same way, with $\lambda_i - \lambda_j$ in place of $\lambda_i\mu_j$.

Solution to Problem 7.14

Uses. Lemma 7.5.11 (Compressing an operator to a complement), Corollary 7.6.16 (Restrictions of diagonalizable operators), Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Theorem 7.6.7 (The roots of the minimal polynomial are the eigenvalues)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Is W invariant? Try $T = \operatorname{diag}(1,2)$ on F^2, $U = \operatorname{span}(\mathbf{e}_1)$, and $W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$.
Hint 2. Show $T^k\mathbf{w} - S^k\mathbf{w} \in U$ for every $\mathbf{w} \in W$ and $k \geq 0$, so that $p(S) = \Pr_{W,U} \circ p(T)\restriction_W$ for every polynomial p. Then use m_T.
Sketch. W need not be invariant, so $S \neq T\restriction_W$ in general, and an eigenvector of S need not be one of T: with $T = \operatorname{diag}(1,2)$, $U = \operatorname{span}(\mathbf{e}_1)$, $W = \operatorname{span}(\mathbf{e}_1 + \mathbf{e}_2)$, $S = 2\mathrm{id}_W$ but $\mathbf{e}_1 + \mathbf{e}_2$ is not an eigenvector of T. Correct proof: induction gives $T^k\mathbf{w} - S^k\mathbf{w} \in U$, so $p(S)\mathbf{w} = \Pr_{W,U} p(T)\mathbf{w}$. With $p = m_T$, which has distinct linear factors, $m_T(S) = 0$, so S is diagonalizable with eigenvalues among the roots of m_T.

The error. Corollary 7.6.16 applies to invariant subspaces, and W need not be invariant; then $T\restriction_W$ is a map $W \to V$, not an operator on W, and S is a different map. The second sentence fails too. Let $T = \operatorname{diag}(1,2)$ on

F^2, $U = \operatorname{span}(\mathbf{e}_1)$, and $W = \operatorname{span}(\mathbf{w})$ with $\mathbf{w} = \mathbf{e}_1 + \mathbf{e}_2$. Then $T\mathbf{w} = \mathbf{e}_1 + 2\mathbf{e}_2 = -\mathbf{e}_1 + 2\mathbf{w}$, so $S\mathbf{w} = 2\mathbf{w}$, and $\mathbf{w}$ is an eigenvector of S but not of T.

A correct proof. Write $P = \Pr_{W,U}$, so that P fixes W and kills U. We first show that $T^k\mathbf{w} - S^k\mathbf{w} \in U$ for all $\mathbf{w} \in W$ and $k \geq 0$, by induction on k. For $k = 0$ the difference is $\mathbf{0}$. If $T^k\mathbf{w} = S^k\mathbf{w} + \mathbf{u}$ with $\mathbf{u} \in U$, then

$$T^{k+1}\mathbf{w} = T\big(S^k\mathbf{w}\big) + T\mathbf{u} = S^{k+1}\mathbf{w} + \mathbf{u}' + T\mathbf{u},$$

where $\mathbf{u}' = T\big(S^k\mathbf{w}\big) - S\big(S^k\mathbf{w}\big) \in U$ by Lemma 7.5.11(a) applied to $S^k\mathbf{w} \in W$, and $T\mathbf{u} \in U$ by invariance. This completes the induction. Taking linear combinations, $p(T)\mathbf{w} - p(S)\mathbf{w} \in U$ for every $p \in F[x]$, and applying P gives $p(S)\mathbf{w} = Pp(T)\mathbf{w}$.

Since T is diagonalizable, $m_T = (x - \lambda_1)\cdots(x - \lambda_m)$ with distinct eigenvalues λ_i of T (Theorem 7.6.15). Then $m_T(S)\mathbf{w} = Pm_T(T)\mathbf{w} = \mathbf{0}$ for all $\mathbf{w}$, so m_T annihilates S. A product of distinct linear factors annihilates S, so S is diagonalizable (Theorem 7.6.15), and its eigenvalues are roots of m_T (Proposition 7.2.28), which are eigenvalues of T (Theorem 7.6.7).

The compression is the quotient operator T/U on V/U in disguise (Remark 7.5.16), and the claim says that quotients of diagonalizable operators are diagonalizable. The eigenvalues are preserved, the eigenvectors are not: they are eigenvectors only modulo U.

Solution to Problem 7.15

> *Uses.* Problem 7.9, Proposition 7.2.35 (A basis of a Krylov subspace), Corollary 7.1.48 (Coprime polynomials), Definition 7.1.24 (Multiplicity of a root), Theorem 7.6.20 (Cayley–Hamilton), Exercise 7.6.47
> *Techniques.* Krylov sequences must become dependent (Technique 7.3.28), Apply an operator to a relation to kill terms (Technique 4.1.25)
> *Hint 1.* Let $q_i = m_T/(x - \lambda_i)^{k_i}$. Since $(x - \lambda_i)^{k_i - 1}q_i$ has smaller degree than m_T, some $\mathbf{u}_i$ is not killed by it. Put $\mathbf{v}_i = q_i(T)\mathbf{u}_i$ and $\mathbf{v} = \sum \mathbf{v}_i$.
> *Hint 2.* If $f(T)\mathbf{v} = \mathbf{0}$, apply $q_j(T)$: the terms $i \neq j$ die. Then use Bézout for q_j and $\big(x - \lambda_j\big)^{k_j}$ to get $\big(x - \lambda_j\big)^{k_j} \mid f$.
> *Sketch.* $\mathbf{v}_i = q_i(T)\mathbf{u}_i$ is killed by $(x - \lambda_i)^{k_i}$ but not by $(x - \lambda_i)^{k_i - 1}$. If $f(T)\mathbf{v} = \mathbf{0}$, applying $q_j(T)$ kills $\mathbf{v}_i$ for $i \neq j$, so $q_j(T)f(T)\mathbf{v}_j = \mathbf{0}$; Bézout $(aq_j + b\big(x - \lambda_j\big)^{k_j} = 1)$ gives $f(T)\mathbf{v}_j = \mathbf{0}$, so $p_{\mathbf{v}_j} = \big(x - \lambda_j\big)^{k_j}$ divides f. With $f = p_{\mathbf{v}}$, each $\big(x - \lambda_j\big)^{k_j}$ divides $p_{\mathbf{v}}$, which divides m_T; compare multiplicities. Then $\dim K_T(\mathbf{v}) = \deg p_{\mathbf{v}}$; and $\deg m_T = \deg p_T = \dim V$ iff $m_T = p_T$ by Cayley–Hamilton.

For each i let $q_i = \prod_{j \neq i} \big(x - \lambda_j\big)^{k_j}$, so $m_T = (x - \lambda_i)^{k_i}q_i$. The polynomial $(x - \lambda_i)^{k_i - 1}q_i$ has degree less than $\deg m_T$, so it does not annihilate T, and some $\mathbf{u}_i$ has $(T - \lambda_i\,\mathrm{id})^{k_i - 1}q_i(T)\mathbf{u}_i \neq \mathbf{0}$. Let $\mathbf{v}_i = q_i(T)\mathbf{u}_i$. Then $(T - \lambda_i\,\mathrm{id})^{k_i}\mathbf{v}_i = m_T(T)\mathbf{u}_i = \mathbf{0}$ and $(T - \lambda_i\,\mathrm{id})^{k_i - 1}\mathbf{v}_i \neq \mathbf{0}$. By Problem 7.9(a), $p_{\mathbf{v}_i}$ divides $(x - \lambda_i)^{k_i}$ but not $(x - \lambda_i)^{k_i - 1}$; a monic divisor of $(x - \lambda_i)^{k_i}$ has only the root λ_i, so it is $(x - \lambda_i)^e$ (Corollary 7.1.35), and $e = k_i$. So $p_{\mathbf{v}_i} = (x - \lambda_i)^{k_i}$.

Let $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_r$ and suppose $f(T)\mathbf{v} = \mathbf{0}$. Fix j. For $i \neq j$, $(x - \lambda_i)^{k_i}$ divides q_j, so $q_j(T)\mathbf{v}_i = \mathbf{0}$. Hence

$$\mathbf{0} = q_j(T)f(T)\mathbf{v} = q_j(T)f(T)\mathbf{v}_j.$$

The polynomials q_j and $\big(x - \lambda_j\big)^{k_j}$ have no common root, so they are coprime, and $aq_j + b\big(x - \lambda_j\big)^{k_j} = 1$ for some $a, b \in \mathbb{C}[x]$ (Corollary 7.1.48). The vector $\mathbf{y} = f(T)\mathbf{v}_j$ satisfies $q_j(T)\mathbf{y} = \mathbf{0}$ and $\big(T - \lambda_j\,\mathrm{id}\big)^{k_j}\mathbf{y} = f(T)\big(T - \lambda_j\,\mathrm{id}\big)^{k_j}\mathbf{v}_j = \mathbf{0}$, so $\mathbf{y} = a(T)q_j(T)\mathbf{y} + b(T)\big(T - \lambda_j\,\mathrm{id}\big)^{k_j}\mathbf{y} = \mathbf{0}$. By Problem 7.9(a), $\big(x - \lambda_j\big)^{k_j} = p_{\mathbf{v}_j}$ divides f.

Apply this to $f = p_{\mathbf{v}}$: each $\big(x - \lambda_j\big)^{k_j}$ divides $p_{\mathbf{v}}$, so λ_j is a root of $p_{\mathbf{v}}$ of multiplicity at least k_j (Definition 7.1.24). On the other hand $m_T(T)\mathbf{v} = \mathbf{0}$, so $p_{\mathbf{v}}$ divides m_T, and its roots are among the λ_j with multiplicity at most k_j. Over $\mathbb{C}$ a monic polynomial is the product of $(x - \mu)$ over its roots with multiplicity (Corollary 7.1.35), so $p_{\mathbf{v}} = m_T$.

By Proposition 7.2.35, $\dim K_T(\mathbf{v}) = \deg p_{\mathbf{v}} = \deg m_T$. If $m_T = p_T$, this is $\dim V$, so $\mathbf{v}$ is cyclic. Conversely, if $\mathbf{w}$ is cyclic, then $\dim V = \dim K_T(\mathbf{w}) = \deg p_{\mathbf{w}} \leq \deg m_T$, since $p_{\mathbf{w}}$ divides m_T; as m_T divides p_T

(Theorem 7.6.20) and $\deg p_T = \dim V$, the monic polynomials m_T and p_T are equal. This is the converse of Exercise 7.6.47(a). The result holds over every field, with a longer proof using the factorization of m_T into irreducibles.

Solution to Problem 7.16

Uses. Theorem 7.6.15 (Diagonalizability and the minimal polynomial), Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Theorem 7.5.22 (Trace and determinant from eigenvalues), Proposition 2.4.23 (Transpose of a product)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. Transpose the equation to eliminate A^{T}, and find a polynomial that annihilates A.
Hint 2. $x^4 - 2x^2 + x = x(x-1)\left(x^2+x-1\right)$ has four distinct real roots. Take traces of A and of A^2.
Sketch. $A^{\mathsf{T}} = I - A^2$, so $A = I - \left(I - A^2\right)^2 = 2A^2 - A^4$, and $x(x-1)\left(x^2+x-1\right)$ annihilates A. So A is diagonalizable over $\mathbb{R}$ with eigenvalues in $\{0, 1, \alpha, \beta\}$, $\alpha, \beta = \left(-1 \pm \sqrt{5}\right)/2$. With multiplicities a, b, c, d: $\operatorname{tr} A = 0$ and irrationality of $\sqrt{5}$ give $b = c = d$; $\operatorname{tr} A^2 = \operatorname{tr} I - \operatorname{tr} A^{\mathsf{T}} = 3$ and $\alpha^2 = 1 - \alpha$, $\beta^2 = 1 - \beta$ give $2b + c + d = 3$. So $4b = 3$, impossible. No.

No. Suppose A is such a matrix. Then $A^{\mathsf{T}} = I_3 - A^2$, and transposing (Proposition 2.4.23), $A = I_3 - \left(A^{\mathsf{T}}\right)^2 = I_3 - \left(I_3 - A^2\right)^2 = 2A^2 - A^4$. So $q(A) = 0$ with

$$q = x^4 - 2x^2 + x = x(x-1)\left(x^2+x-1\right),$$

whose roots $0, 1, \alpha = \left(-1+\sqrt{5}\right)/2, \beta = \left(-1-\sqrt{5}\right)/2$ are real and distinct. By Theorem 7.6.15, A is diagonalizable over $\mathbb{R}$, and its eigenvalues lie in $\{0, 1, \alpha, \beta\}$ (Proposition 7.2.28). Let a, b, c, d be the multiplicities of $0, 1, \alpha, \beta$, so $a + b + c + d = 3$, and the eigenvalues of A^2 are their squares (diagonalize).

First trace. $0 = \operatorname{tr} A = b + c\alpha + d\beta = b - (c+d)/2 + (c-d)\sqrt{5}/2$ (Theorem 7.5.22). Since $\sqrt{5}$ is irrational and b, c, d are integers, $c = d$ and then $b = c$.

Second trace. Taking the trace of $A^2 = I_3 - A^{\mathsf{T}}$ gives $\operatorname{tr} A^2 = 3 - \operatorname{tr} A = 3$. Since $\alpha^2 = 1 - \alpha$ and $\beta^2 = 1 - \beta$,

$$\operatorname{tr} A^2 = b + c(1-\alpha) + d(1-\beta) = b + c + d - (c\alpha + d\beta) = b + c + d + b,$$

using $c\alpha + d\beta = -b$ from the first trace. So $2b + c + d = 3$, and with $b = c = d$ this gives $4b = 3$, which has no integer solution. Hence no such A exists.

The equation mixes A and A^{T}, which have the same eigenvalues but different eigenvectors; transposing it once eliminates A^{T} and leaves a polynomial identity in A alone. After that the problem is counting with multiplicities. This is Problem 2 of the 2011 International Mathematics Competition for University Students (IMC).

Solution to Problem 7.17

Uses. Proposition 7.2.28 (Polynomials act on eigenvectors by their values), Proposition 7.1.38 (Conjugate roots), Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem 5.4.10 (Multiplicativity)
Techniques. Use an annihilating polynomial (Technique 7.6.17), Complexify, then take real and imaginary parts (Technique 7.3.35)
Hint 1. Multiply $(A+B)\left(A^{-1}+B^{-1}\right) = I_n$ out and set $X = AB^{-1}$.
Hint 2. $X^2 + X + I_n = 0$. What are the complex eigenvalues of a real X with this property, and with which multiplicities?
Sketch. $(A+B)\left(A^{-1}+B^{-1}\right) = I$ gives $X + X^{-1} + I = 0$, so $X^2 + X + I = 0$ with $X = AB^{-1}$. The complex eigenvalues of X are roots of $x^2 + x + 1$, i.e. $\omega, \overline{\omega}$; p_X is real, so they occur equally often, $n = 2k$, and $\det X = (\omega\overline{\omega})^k = 1$. So $\det A = \det B$. Over $\mathbb{C}$: no; $B = I_n$, $A = \omega I_n$ (if $3 \nmid n$) or $A = \operatorname{diag}(\omega, \ldots, \omega, \overline{\omega})$ (if $3 \mid n$) satisfy the equation with $\det A \neq 1$.

Multiplying the hypothesis on the left by $A + B$ gives $I_n = (A+B)\left(A^{-1}+B^{-1}\right) = 2I_n + AB^{-1} + BA^{-1}$. With $X = AB^{-1}$, so that $X^{-1} = BA^{-1}$, this is $X + X^{-1} + I_n = 0$, and multiplying by X,

$$X^2 + X + I_n = 0.$$

By Proposition 7.2.28, every complex eigenvalue of X is a root of $x^2 + x + 1$, that is, $\omega = e^{2\pi i/3}$ or $\overline{\omega}$. Since X is real, p_X has real coefficients, so ω and $\overline{\omega}$ are roots of p_X of the same multiplicity k (Proposition 7.1.38),

and $n = 2k$. By Theorem 7.5.22 over $\mathbb{C}$, $\det X = \omega^k \overline{\omega}^k = |\omega|^{2k} = 1$. As $\det X = \det A / \det B$ (Theorem 5.4.10), $\det A = \det B$.

Over $\mathbb{C}$ the conclusion fails for every n. The computation above can be reversed: if X is invertible with $X^2 + X + I_n = 0$ and $X + I_n$ is invertible, then $A = X$ and $B = I_n$ satisfy the equation, since $(X + I_n)\left(X^{-1} + I_n\right) = X + X^{-1} + 2I_n = I_n$. Take $X = \omega I_n$ if 3 does not divide n, with $\det X = \omega^n \neq 1$, and $X = \operatorname{diag}(\omega, \ldots, \omega, \overline{\omega})$ if 3 divides n, with $\det X = \omega^{n-2} = \omega^{-2} = \omega \neq 1$. In both cases $X + I_n$ has eigenvalues $1 + \omega = -\overline{\omega}$ and $1 + \overline{\omega} = -\omega$, so it is invertible, and $\det A \neq 1 = \det B$.

The real hypothesis enters only through the pairing of conjugate eigenvalues, which makes $\det X$ a product of factors $\omega\overline{\omega} = 1$; Exercise 7.6.50 uses the pairing in the same way. This is Problem 1 of the 2015 International Mathematics Competition for University Students (IMC).

Solution to Problem 7.18

Uses. Theorem 5.4.10 (Multiplicativity), Proposition 7.3.2 (Determinants over a commutative ring), Exercise 7.3.17
Techniques. Factor the matrix (Technique 5.6.1), Conjugate the equation (Technique B.1.3)
Hint 1. With $\omega = e^{2\pi i/3}$, expand $(A + \omega B)(A + \overline{\omega} B)$ and use $A^2 + B^2 = AB$ and $1 + \omega + \overline{\omega} = 0$.
Hint 2. The product is $-\omega(AB - BA)$. Take determinants; the left side is $|\det(A + \omega B)|^2$ because A, B are real.
Sketch. $(A + \omega B)(A + \overline{\omega} B) = A^2 + B^2 + \overline{\omega} AB + \omega BA = (1 + \overline{\omega})AB + \omega BA = -\omega(AB - BA)$. Determinants: $|\det(A + \omega B)|^2 = (-\omega)^n \det(AB - BA)$. The right side has a nonzero real factor, so $(-\omega)^n = e^{-\pi i n/3}$ is real, which happens iff $3 \mid n$.

Let $\omega = e^{2\pi i/3}$, so $\omega\overline{\omega} = 1$ and $1 + \omega + \overline{\omega} = 0$. Expanding (the scalars commute with all matrices) and using $A^2 + B^2 = AB$,

$$(A + \omega B)(A + \overline{\omega} B) = A^2 + B^2 + \overline{\omega} AB + \omega BA = (1 + \overline{\omega})AB + \omega BA = -\omega(AB - BA),$$

because $1 + \overline{\omega} = -\omega$. Let $Z = A + \omega B$. Since A and B are real, $A + \overline{\omega} B$ is the entrywise conjugate $\overline{Z}$, and $\det \overline{Z} = \overline{\det Z}$ (Exercise 7.3.17, or Proposition 7.3.2(g) with complex conjugation). Taking determinants (Theorem 5.4.10),

$$|\det Z|^2 = \det Z \cdot \overline{\det Z} = (-\omega)^n \det(AB - BA).$$

Here $d = \det(AB - BA)$ is a nonzero real number, so $(-\omega)^n = |\det Z|^2 / d$ is real. Now $-\omega = e^{i\pi} e^{2\pi i/3} = e^{-\pi i/3}$, so $(-\omega)^n = e^{-\pi i n/3}$, which is real exactly when $\sin(\pi n/3) = 0$, that is, when 3 divides n.

The proof factors a real quadratic expression over $\mathbb{C}$ as $Z\overline{Z}$, and the conjugate pair forces a determinant to be a nonnegative real number; this is the device of Exercise 7.6.52, with ω in place of i because the quadratic $a^2 - ab + b^2$ has roots $-\omega, -\overline{\omega}$ instead of $\pm i$. It also gives the sign: $\det(AB - BA)$ has the sign of $(-1)^{n/3}$. This is a classical olympiad problem.

Section 8.1

Solution to Exercise 8.1.6

Uses. Definition 8.1.5 (Skew-adjoint operators, skew-Hermitian matrices), Definition 6.6.8 (Adjoint), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm)
Techniques. Move an operator across the inner product (Technique 6.6.20), Show a vector is zero via its norm (Technique 6.1.36)
Hint 1. Expand the left side. The cross terms are $-c\langle K\mathbf{v}, \mathbf{v}\rangle - c\langle \mathbf{v}, K\mathbf{v}\rangle$, since c is real.
Hint 2. Move K across in the second cross term: $\langle \mathbf{v}, K\mathbf{v}\rangle = \langle K^*\mathbf{v}, \mathbf{v}\rangle = -\langle K\mathbf{v}, \mathbf{v}\rangle$.
Sketch. The cross terms are $-c(\langle K\mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{v}, K\mathbf{v}\rangle) = -c(\langle K\mathbf{v}, \mathbf{v}\rangle - \langle K\mathbf{v}, \mathbf{v}\rangle) = 0$. If $K\mathbf{v} = c\mathbf{v}$ with $c \neq 0$ real and $\mathbf{v} \neq \mathbf{0}$, the identity gives $0 \geq c^2\|\mathbf{v}\|^2 > 0$. For $\mathrm{id} + K$, take $c = -1$.

Since c is real, $\overline{c} = c$, and expanding gives

$$\|K\mathbf{v} - c\mathbf{v}\|^2 = \|K\mathbf{v}\|^2 - c\langle K\mathbf{v}, \mathbf{v}\rangle - c\langle \mathbf{v}, K\mathbf{v}\rangle + c^2\|\mathbf{v}\|^2.$$

By the conjugated form of the defining identity, $\langle \mathbf{v}, K\mathbf{v}\rangle = \langle K^*\mathbf{v}, \mathbf{v}\rangle = -\langle K\mathbf{v}, \mathbf{v}\rangle$, so the two middle terms cancel, which proves the identity.

If $K\mathbf{v} = c\mathbf{v}$ with $c \in \mathbb{R}$, $c \neq 0$, and $\mathbf{v} \neq \mathbf{0}$, the identity gives $0 = \|K\mathbf{v}\|^2 + c^2\|\mathbf{v}\|^2 \geq c^2\|\mathbf{v}\|^2 > 0$, a contradiction. So no nonzero real number is an eigenvalue of K. Finally, with $c = -1$, $\|(\mathrm{id} + K)\mathbf{v}\|^2 = \|K\mathbf{v}\|^2 + \|\mathbf{v}\|^2$, which is positive for $\mathbf{v} \neq \mathbf{0}$. So $\ker(\mathrm{id} + K) = \{\mathbf{0}\}$, and $\mathrm{id} + K$ is injective.

Solution to Exercise 8.1.12

Uses. Definition 8.1.8 (Normal operator, normal matrix), Proposition 6.6.18 (Properties of the adjoint)
Hint 1. Use Proposition 6.6.18 to write down each adjoint: $(T^*)^* = T$, $(cT)^* = \bar{c}T^*$, $\left(T^k\right)^* = (T^*)^k$, $\left(T^{-1}\right)^* = (T^*)^{-1}$.
Hint 2. For T^k: since T and T^* commute, any product of copies of T and T^* can be reordered. For T^{-1}, invert both sides of $T^*T = TT^*$.
Sketch. $(T^*)^*T^* = TT^* = T^*T = T^*(T^*)^*$. $(cT)^*(cT) = |c|^2T^*T = |c|^2TT^* = (cT)(cT)^*$. $(T^*)^kT^k = T^k(T^*)^k$ because T and T^* commute. Inverting $T^*T = TT^*$ gives $T^{-1}(T^*)^{-1} = (T^*)^{-1}T^{-1}$, and $(T^*)^{-1} = \left(T^{-1}\right)^*$.

By Proposition 6.6.18(c), $(T^*)^* = T$, so $(T^*)^*T^* = TT^* = T^*T = T^*(T^*)^*$, and T^* is normal. By (b) of the same proposition, $(cT)^* = \bar{c}T^*$, so

$$(cT)^*(cT) = \bar{c}cT^*T = |c|^2TT^* = (cT)(cT)^*.$$

By (d), applied $k-1$ times, $\left(T^k\right)^* = (T^*)^k$. Since $TT^* = T^*T$, we can move each factor T^* in $(T^*)^kT^k$ past each factor T, one swap at a time, so $(T^*)^kT^k = T^k(T^*)^k$, and T^k is normal.

Finally, suppose T is invertible and T^{-1} has an adjoint. By Proposition 6.6.18(f), T^* is invertible with $(T^*)^{-1} = \left(T^{-1}\right)^*$. Taking inverses of both sides of $T^*T = TT^*$ gives $T^{-1}(T^*)^{-1} = (T^*)^{-1}T^{-1}$, that is, $T^{-1}\left(T^{-1}\right)^* = \left(T^{-1}\right)^*T^{-1}$.

Solution to Exercise 8.1.19

Uses. Theorem 8.1.18 (Eigenspaces of a normal operator are orthogonal), Theorem 8.1.16 (Eigenvectors of a normal operator and of its adjoint), Theorem 6.3.3 (Pythagorean theorem), Proposition 6.1.4 (Sesquilinearity)
Hint 1. The eigenvalues 2 and i are distinct. What does that say about $\langle \mathbf{v}, \mathbf{w}\rangle$?
Hint 2. $\mathbf{v} \perp \mathbf{w}$, and $T^*\mathbf{v} = 2\mathbf{v}$, $T^*\mathbf{w} = -i\mathbf{w}$. Then use Pythagoras.
Sketch. $\mathbf{v} \perp \mathbf{w}$. $T(\mathbf{v} + \mathbf{w}) = 2\mathbf{v} + i\mathbf{w}$ has squared norm $4 + 9 = 13$. $T^*(\mathbf{v} - \mathbf{w}) = 2\mathbf{v} + i\mathbf{w}$ too, with norm $\sqrt{13}$. Finally $\langle 2\mathbf{v} + i\mathbf{w}, \mathbf{v} - \mathbf{w}\rangle = 2 - 9i$.

Since $2 \neq i$, $\langle \mathbf{v}, \mathbf{w}\rangle = 0$ by Theorem 8.1.18, and by Theorem 8.1.16, $T^*\mathbf{v} = 2\mathbf{v}$ and $T^*\mathbf{w} = \bar{i}\mathbf{w} = -i\mathbf{w}$.

First, $T(\mathbf{v} + \mathbf{w}) = 2\mathbf{v} + i\mathbf{w}$, a sum of orthogonal vectors, so by Theorem 6.3.3 its squared norm is $4\|\mathbf{v}\|^2 + |i|^2\|\mathbf{w}\|^2 = 4 + 9 = 13$, and $\|T(\mathbf{v} + \mathbf{w})\| = \sqrt{13}$.

Next, $T^*(\mathbf{v} - \mathbf{w}) = 2\mathbf{v} - (-i)\mathbf{w} = 2\mathbf{v} + i\mathbf{w}$, so $\|T^*(\mathbf{v} - \mathbf{w})\| = \sqrt{13}$ as well.

Finally, using $\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{v}\rangle = 0$ and Proposition 6.1.4,

$$\langle 2\mathbf{v} + i\mathbf{w}, \mathbf{v} - \mathbf{w}\rangle = 2\langle \mathbf{v}, \mathbf{v}\rangle - i\langle \mathbf{w}, \mathbf{w}\rangle = 2 - 9i.$$

As a check, this also equals $\langle \mathbf{v} + \mathbf{w}, T^*(\mathbf{v} - \mathbf{w})\rangle = \langle \mathbf{v} + \mathbf{w}, 2\mathbf{v} + i\mathbf{w}\rangle = 2 + \bar{i} \cdot 9 = 2 - 9i$.

Solution to Exercise 8.1.22

Uses. Definition 6.3.26 (Conjugate transpose), Definition 8.1.5 (Skew-adjoint operators, skew-Hermitian matrices), Definition 6.3.33 (Unitary and orthogonal matrices), Definition 8.1.8 (Normal operator, normal matrix), Example 8.1.9 (Normal operators)
Hint 1. Compute A^* first: it decides Hermitian and skew-Hermitian at once. A Hermitian or skew-Hermitian matrix is automatically normal.
Hint 2. For the rest, compute A^*A and AA^*. Unitary means $A^*A = I$.
Sketch. (a) Normal ($A^{\mathsf{T}}A = AA^{\mathsf{T}} = 5I$) and nothing else. (b) $A^* = -A$: skew-Hermitian, hence normal; not unitary. (c) $A^*A = AA^* = 5I$: normal only. (d) A permutation matrix: orthogonal, hence unitary and normal. (e) $A^*A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix} = AA^*$: none.

(a) $A^{\mathsf{T}} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ is neither A nor $-A$. This is $aI + bJ$ with $a = 1$, $b = -2$, so by Example 8.1.9(c), $A^{\mathsf{T}}A = AA^{\mathsf{T}} = 5I_2$. So A is normal, and not orthogonal since $5I_2 \neq I_2$.

(b) $A^* = \begin{pmatrix} -i & -1 \\ 1 & -2i \end{pmatrix} = -A$, so A is skew-Hermitian, hence normal. It is not Hermitian, since $A \neq 0$. It is not unitary: $A^*A = -A^2 = -\begin{pmatrix} -2 & 3i \\ -3i & -5 \end{pmatrix} = \begin{pmatrix} 2 & -3i \\ 3i & 5 \end{pmatrix} \neq I_2$.

(c) $A^* = \begin{pmatrix} 2 & -i \\ -i & 2 \end{pmatrix}$, which is neither A nor $-A$. Now

$$A^*A = \begin{pmatrix} 2 & -i \\ -i & 2 \end{pmatrix}\begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}, \quad AA^* = \begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}\begin{pmatrix} 2 & -i \\ -i & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix},$$

so A is normal but not unitary. (Its parts in Proposition 8.1.21 are $S = 2I_2$ and $K = iB$ with $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and S commutes with B, as that proposition predicts.)

(d) The columns are $\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1$, an orthonormal basis, so A is orthogonal (Definition 6.3.33), hence unitary and normal. It is not symmetric, since $a_{13} = 1 \neq 0 = a_{31}$, and not skew-symmetric, since $a_{13} = 1 \neq 0 = -a_{31}$.

(e) $A^* = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}$, and

$$A^*A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}, \quad AA^* = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix}.$$

These differ, so A is not normal, and therefore none of the other three.

Solution to Exercise 8.1.23

Uses. Definition 7.2.13 (Eigenspace), Corollary 7.2.23 (Eigenspaces are independent), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Algorithm 6.4.14 (Gram–Schmidt)
Hint 1. $A\mathbf{x} = \mathbf{x} + (x_1 + x_2 + x_3)\mathbf{1}$. What happens for $\mathbf{x} = \mathbf{1}$, and for $\mathbf{x}$ with $x_1 + x_2 + x_3 = 0$?
Hint 2. $E_4 = \operatorname{span}(\mathbf{1})$ and $E_1 = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 0\}$. Their dimensions add up to 3, so there are no other eigenvalues.
Hint 3. For (c), run Gram–Schmidt inside E_1.
Sketch. $A\mathbf{1} = 4\mathbf{1}$, and $A\mathbf{x} = \mathbf{x}$ when $x_1 + x_2 + x_3 = 0$. So $E_4 = \operatorname{span}(\mathbf{1})$ and E_1 is the plane $x_1 + x_2 + x_3 = 0$; $E_1 \perp E_4$ since that equation says $\mathbf{x} \cdot \mathbf{1} = 0$. In E_1, $(1, -1, 0) \cdot (1, 0, -1) = 1$. Orthonormal eigenbasis: $(1, 1, 1)/\sqrt{3}$, $(1, -1, 0)/\sqrt{2}$, $(1, 1, -2)/\sqrt{6}$.

(a) For $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{1}\mathbf{1}^{\mathsf{T}}\mathbf{x} = (x_1 + x_2 + x_3)\mathbf{1}$, so $A\mathbf{x} = \mathbf{x} + (x_1 + x_2 + x_3)\mathbf{1}$. Hence $(A - I_3)\mathbf{x} = \mathbf{0}$ iff $x_1 + x_2 + x_3 = 0$, so $E_1 = \{\mathbf{x} \mid x_1 + x_2 + x_3 = 0\}$, a plane. And $(A - 4I_3)\mathbf{x} = (x_1 + x_2 + x_3)\mathbf{1} - 3\mathbf{x}$ vanishes iff $\mathbf{x}$ is a multiple of $\mathbf{1}$, say $\mathbf{x} = c\mathbf{1}$, and then indeed $3c\mathbf{1} - 3c\mathbf{1} = \mathbf{0}$; so $E_4 = \operatorname{span}(\mathbf{1})$. The eigenspaces of distinct eigenvalues form a direct sum (Corollary 7.2.23), and $\dim E_1 + \dim E_4 = 3$ already, so there is no room for another eigenspace: the eigenvalues are 1 and 4.

(b) For $\mathbf{x} \in E_1$, $\mathbf{x} \cdot \mathbf{1} = x_1 + x_2 + x_3 = 0$, so $E_1 \perp E_4$, as Proposition 8.1.1(b) predicts. Inside E_1, the eigenvectors $(1, -1, 0)$ and $(1, 0, -1)$ have dot product $1 \neq 0$.

(c) Apply Gram–Schmidt (Algorithm 6.4.14) to the basis $(1, -1, 0), (1, 0, -1)$ of E_1: the second vector minus its projection onto the first is

$$(1, 0, -1) - \frac{1}{2}(1, -1, 0) = \frac{1}{2}(1, 1, -2).$$

Normalizing, and adding the normalized $\mathbf{1}$ from E_4, which is orthogonal to E_1 by (b),

$$\frac{1}{\sqrt{3}}(1, 1, 1), \quad \frac{1}{\sqrt{2}}(1, -1, 0), \quad \frac{1}{\sqrt{6}}(1, 1, -2)$$

is an orthonormal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.

Solution to Exercise 8.1.24

Uses. Example 8.1.9 (Normal operators), Non-example 8.1.10 (Operators that are not normal), Proposition 6.6.18 (Properties of the adjoint), Example 6.1.39 (A rotation)
Hint 1. The nilpotent $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is the source of counterexamples. Write it as a sum of a symmetric and a skew-symmetric matrix; compute N^2.
Hint 2. For (c), look for an invertible operator that is not normal.
Hint 3. For (d), compute $(T + cT^*)^* = T^* + \overline{c}T$ and expand both products.
Sketch. (a) False: $N = \left(N + N^{\mathsf{T}}\right)/2 + \left(N - N^{\mathsf{T}}\right)/2$. (b) False: $N^2 = 0$ is normal. (c) False: the shear is invertible, so both kernels are $\{\mathbf{0}\}$. (d) True: the two products differ by $\left(1 - |c|^2\right)(T^*T - TT^*)$. (e) False: the rotation J.

Let $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, acting on $\mathbb{F}^2$ with the dot product. It is not normal (Non-example 8.1.10).

(a) False. $N = S + K$ with $S = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ symmetric and $K = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ skew-symmetric. Both are normal (Example 8.1.9(a)), and their sum N is not.

(b) False. $N^2 = 0$, which is normal, but N is not.

(c) False. The shear $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is invertible, and so is $S^* = S^{\mathsf{T}}$, so $\ker S = \ker S^* = \{\mathbf{0}\}$; but S is not normal (Non-example 8.1.10).

(d) True. By Proposition 6.6.18, $(T + cT^*)^* = T^* + \overline{c}T$. Expanding,

$$(T^* + \overline{c}T)(T + cT^*) = T^*T + c(T^*)^2 + \overline{c}T^2 + |c|^2 TT^*,$$
$$(T + cT^*)(T^* + \overline{c}T) = TT^* + \overline{c}T^2 + c(T^*)^2 + |c|^2 T^*T.$$

The difference is $(T^*T - TT^*) - |c|^2(T^*T - TT^*) = \left(1 - |c|^2\right)(T^*T - TT^*)$, which is 0 when T is normal.

(e) False. The rotation $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is orthogonal, hence normal, and $\langle J\mathbf{x}, \mathbf{x}\rangle = 0$ for every $\mathbf{x} \in \mathbb{R}^2$ (Example 6.1.39), but $J \neq 0$. (Over $\mathbb{C}$ the statement is true for every operator, by Proposition 6.1.41.)

Solution to Exercise 8.1.25

Uses. Definition 8.1.8 (Normal operator, normal matrix), Example 8.1.9 (Normal operators), Example 7.3.5 (Small cases)
Hint 1. Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and compare $A^{\mathsf{T}}A$ with AA^{T} entry by entry.
Hint 2. The diagonal entries give $b^2 = c^2$. Treat $c = b$ and $c = -b \neq b$ separately.
Sketch. $A^{\mathsf{T}}A = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}$, $AA^{\mathsf{T}} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$. Normal iff $b^2 = c^2$ and $ab + cd = ac + bd$. If $c = b$, A is symmetric. If $c = -b \neq 0$, the second equation gives $b(a - d) = 0$, so $a = d$. The second family is a rotation followed by a scaling, with characteristic polynomial $(x - a)^2 + b^2$.

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$A^{\mathsf{T}}A = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}, \quad AA^{\mathsf{T}} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}.$$

So A is normal if and only if $c^2 = b^2$ and $ab + cd = ac + bd$. The first condition says $c = b$ or $c = -b$.

If $c = b$, then A is symmetric, and the second condition holds automatically. If $c = -b$ and $b \neq 0$, the second condition reads $ab - bd = -ab + bd$, that is, $2b(a - d) = 0$, so $a = d$, and $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$; renaming $-b$ as b gives the stated form. Conversely, symmetric matrices are normal, and so are the matrices $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ by Example 8.1.9(c).

Writing $a + bi = r(\cos\theta + i\sin\theta)$ with $r = \sqrt{a^2 + b^2}$, the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is a rotation by θ followed by scaling by r; it is multiplication by $a + bi$ on $\mathbb{C} = \mathbb{R}^2$. Its characteristic polynomial is $(x - a)^2 + b^2$ (Example 7.3.5), which has no real root when $b \neq 0$. So a real normal 2×2 matrix either is symmetric (and then, by Section 8.2, has an orthonormal eigenbasis), or is a scaled rotation, with no real eigenvalues at all (unless $b = 0$, when it is aI_2).

Solution to Exercise 8.1.26

Uses. Proposition 6.3.27 (Properties of the conjugate transpose), Definition 8.1.8 (Normal operator, normal matrix), Proposition 2.4.14 (Algebra of matrices)
Techniques. Write a rank-one matrix as $\mathbf{u}\mathbf{v}^{\mathsf{T}}$ (Technique 4.3.17)
Hint 1. $A^* = \mathbf{y}\mathbf{x}^*$. Compute A^*A and AA^*, remembering that $\mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2$ is a scalar.
Hint 2. $A^*A = \|\mathbf{x}\|^2\mathbf{y}\mathbf{y}^*$ and $AA^* = \|\mathbf{y}\|^2\mathbf{x}\mathbf{x}^*$. If these are equal, apply both to $\mathbf{y}$.
Sketch. $A^*A = \|\mathbf{x}\|^2\mathbf{y}\mathbf{y}^*$ and $AA^* = \|\mathbf{y}\|^2\mathbf{x}\mathbf{x}^*$. If $\mathbf{y} = c\mathbf{x}$, both are $|c|^2\|\mathbf{x}\|^2\mathbf{x}\mathbf{x}^*$. Conversely, applying both to $\mathbf{y}$ gives $\|\mathbf{x}\|^2\|\mathbf{y}\|^2\mathbf{y} = \|\mathbf{y}\|^2(\mathbf{x}^*\mathbf{y})\mathbf{x}$, so $\mathbf{y}$ is a multiple of $\mathbf{x}$. Then $A = \bar{c}\mathbf{x}\mathbf{x}^*$ and $A^* = c\mathbf{x}\mathbf{x}^*$.

By Proposition 6.3.27, $A^* = (\mathbf{y}^*)^*\mathbf{x}^* = \mathbf{y}\mathbf{x}^*$. Since $\mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2$ and $\mathbf{y}^*\mathbf{y} = \|\mathbf{y}\|^2$ are 1×1 matrices, that is, scalars,

$$A^*A = \mathbf{y}(\mathbf{x}^*\mathbf{x})\mathbf{y}^* = \|\mathbf{x}\|^2\mathbf{y}\mathbf{y}^*, \quad AA^* = \mathbf{x}(\mathbf{y}^*\mathbf{y})\mathbf{x}^* = \|\mathbf{y}\|^2\mathbf{x}\mathbf{x}^*.$$

If $\mathbf{y} = c\mathbf{x}$, then $\mathbf{y}\mathbf{y}^* = |c|^2\mathbf{x}\mathbf{x}^*$ and $\|\mathbf{y}\|^2 = |c|^2\|\mathbf{x}\|^2$, so both products equal $|c|^2\|\mathbf{x}\|^2\mathbf{x}\mathbf{x}^*$, and A is normal.

Conversely, suppose $A^*A = AA^*$, and apply both sides to $\mathbf{y}$:

$$\|\mathbf{x}\|^2\|\mathbf{y}\|^2\mathbf{y} = \|\mathbf{y}\|^2(\mathbf{x}^*\mathbf{y})\mathbf{x}.$$

Dividing by $\|\mathbf{x}\|^2\|\mathbf{y}\|^2 \neq 0$ gives $\mathbf{y} = c\mathbf{x}$ with $c = (\mathbf{x}^*\mathbf{y})/\|\mathbf{x}\|^2$.

In that case $A = \mathbf{x}(c\mathbf{x})^* = \bar{c}\mathbf{x}\mathbf{x}^*$ and $A^* = c\mathbf{x}\mathbf{x}^*$. Since $\mathbf{x}\mathbf{x}^* \neq 0$ (its (j,j) entry is $|x_j|^2$, which is nonzero for some j), $A^* = A$ if and only if $c = \bar{c}$, that is, $c \in \mathbb{R}$.

Solution to Exercise 8.1.27

Uses. Proposition 6.1.37 (Inner products distinguish vectors), Non-example 8.1.10 (Operators that are not normal), Theorem 8.1.16 (Eigenvectors of a normal operator and of its adjoint)
Hint 1. Two vectors are equal if they have the same inner product with *every* vector (Proposition 6.1.37). How many vectors does the argument test against?
Hint 2. The argument shows only that $T^*\mathbf{v} - \overline{\lambda}\mathbf{v} \perp \mathbf{v}$. Try $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ with $\mathbf{v} = \mathbf{e}_1$.
Sketch. The computation is correct but shows only $\langle T^*\mathbf{v} - \overline{\lambda}\mathbf{v}, \mathbf{v}\rangle = 0$; testing against the single vector $\mathbf{v}$ does not determine $T^*\mathbf{v}$. For N, $N\mathbf{e}_1 = 0\mathbf{e}_1$ but $N^*\mathbf{e}_1 = \mathbf{e}_2 \neq \mathbf{0}$, and indeed $\mathbf{e}_2 \perp \mathbf{e}_1$. The claim holds for normal T (Theorem 8.1.16).

The error. The displayed computation is correct, and it shows that $\langle T^*\mathbf{v} - \overline{\lambda}\mathbf{v}, \mathbf{v}\rangle = 0$, that is, $T^*\mathbf{v} - \overline{\lambda}\mathbf{v}$ is orthogonal to $\mathbf{v}$. The last sentence is wrong. To conclude that two vectors are equal, they must have the same inner product with every vector of V (Proposition 6.1.37), and the argument tests only against the single vector $\mathbf{v}$.

A counterexample. Let $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on $\mathbb{F}^2$ with the dot product, $\mathbf{v} = \mathbf{e}_1$, and $\lambda = 0$. Then $N\mathbf{e}_1 = \mathbf{0} = 0\mathbf{e}_1$, but $N^*\mathbf{e}_1 = N^{\mathsf{T}}\mathbf{e}_1 = \mathbf{e}_2 \neq \mathbf{0}$. Consistently with what the argument does prove, $\mathbf{e}_2 \perp \mathbf{e}_1$.

A correct version. If T is normal, the conclusion holds by Theorem 8.1.16. The missing ingredient is the equality of lengths $\|(T - \lambda\,\mathrm{id})\mathbf{v}\| = \|(T^* - \overline{\lambda}\,\mathrm{id})\mathbf{v}\|$, which controls all of $T^*\mathbf{v} - \overline{\lambda}\mathbf{v}$ and not just its component along $\mathbf{v}$.

Solution to Exercise 8.1.28

Uses. Definition 6.6.26 (Self-adjoint operator), Proposition 6.1.4 (Sesquilinearity), Proposition 6.1.27 (Basic properties of the norm), Corollary 4.2.18 (Injective if and only if surjective), Theorem 2.6.11 (Injectivity is decided at $\mathbf{0}$)
Techniques. Move an operator across the inner product (Technique 6.6.20), Shift by a scalar (Technique 8.1.15)
Hint 1. Let $R = T - a\,\mathrm{id}$, which is self-adjoint since a is real. Then $T - \lambda\,\mathrm{id} = R - ib\,\mathrm{id}$; expand $\|R\mathbf{v} - ib\mathbf{v}\|^2$.
Hint 2. The cross terms are $ib\langle R\mathbf{v}, \mathbf{v}\rangle - ib\langle \mathbf{v}, R\mathbf{v}\rangle$, and $\langle \mathbf{v}, R\mathbf{v}\rangle = \langle R\mathbf{v}, \mathbf{v}\rangle$.
Hint 3. For (c), (a) gives $\|(T - \lambda\,\mathrm{id})\mathbf{v}\| \geq |b|\|\mathbf{v}\|$. Apply it to $\mathbf{v} = (T - \lambda\,\mathrm{id})^{-1}\mathbf{w}$.
Sketch. With $R = T - a\,\mathrm{id}$ self-adjoint, $\|R\mathbf{v} - ib\mathbf{v}\|^2 = \|R\mathbf{v}\|^2 + ib\langle R\mathbf{v}, \mathbf{v}\rangle - ib\langle \mathbf{v}, R\mathbf{v}\rangle + b^2\|\mathbf{v}\|^2$, and the middle terms cancel. If $b \neq 0$ the right side is at least $b^2\|\mathbf{v}\|^2 > 0$ for $\mathbf{v} \neq \mathbf{0}$, so $T - \lambda\,\mathrm{id}$ is injective, hence invertible in finite dimensions, and $\|\mathbf{w}\| \geq |b|\|(T - \lambda\,\mathrm{id})^{-1}\mathbf{w}\|$.

(a) Let $R = T - a\,\mathrm{id}$. Since a is real, $R^* = T^* - a\,\mathrm{id} = R$, so R is self-adjoint. Now $T - \lambda\,\mathrm{id} = R - ib\,\mathrm{id}$, and expanding with Proposition 6.1.4,

$$\|R\mathbf{v} - ib\mathbf{v}\|^2 = \|R\mathbf{v}\|^2 + \overline{-ib}\langle R\mathbf{v}, \mathbf{v}\rangle + (-ib)\langle \mathbf{v}, R\mathbf{v}\rangle + |ib|^2\|\mathbf{v}\|^2.$$

Since $\overline{-ib} = ib$ and $\langle \mathbf{v}, R\mathbf{v}\rangle = \langle R\mathbf{v}, \mathbf{v}\rangle$ by self-adjointness, the two middle terms are $ib\langle R\mathbf{v}, \mathbf{v}\rangle - ib\langle R\mathbf{v}, \mathbf{v}\rangle = 0$. This leaves $\|R\mathbf{v}\|^2 + b^2\|\mathbf{v}\|^2$.

(b) If $T\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, then (a) gives $0 = \|(T - a\,\mathrm{id})\mathbf{v}\|^2 + b^2\|\mathbf{v}\|^2 \geq b^2\|\mathbf{v}\|^2$, so $b = 0$ and $\lambda = a \in \mathbb{R}$.

(c) By (a), $\|(T - \lambda\,\mathrm{id})\mathbf{v}\| \geq |b|\|\mathbf{v}\|$ for every $\mathbf{v}$. If $b \neq 0$, this is positive for $\mathbf{v} \neq \mathbf{0}$, so $\ker(T - \lambda\,\mathrm{id}) = \{\mathbf{0}\}$ and $T - \lambda\,\mathrm{id}$ is injective (Theorem 2.6.11). An injective operator on a finite-dimensional space is surjective (Corollary 4.2.18), hence invertible. For $\mathbf{w} \in V$, put $\mathbf{v} = (T - \lambda\,\mathrm{id})^{-1}\mathbf{w}$; then $\|\mathbf{w}\| \geq |b|\|\mathbf{v}\|$, which is the claimed bound.

Solution to Exercise 8.1.29

Uses. Proposition 8.1.21 (Normal means commuting parts), Proposition 6.6.18 (Properties of the adjoint), Proposition 8.1.4 (Self-adjoint operators are determined by their diagonal values)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. $\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle$ and $T^* = S - K$. Expand $(S - K)(S + K)$.
Hint 2. For (b), the operator $SK - KS$ is self-adjoint. Then use Proposition 8.1.4 and Proposition 8.1.21.
Sketch. $T^*T = (S - K)(S + K) = S^2 - K^2 + (SK - KS)$, and $\langle S^2\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2$, $-\langle K^2\mathbf{v}, \mathbf{v}\rangle = \|K\mathbf{v}\|^2$. For (b), $(SK - KS)^* = K^*S^* - S^*K^* = SK - KS$, so its diagonal values all vanish iff $SK = KS$, iff T is normal.

(a) By Proposition 8.1.21, $T^* = S - K$, so

$$T^*T = (S - K)(S + K) = S^2 - K^2 + (SK - KS).$$

Now $\langle S^2\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, S^*\mathbf{v}\rangle = \|S\mathbf{v}\|^2$ and $-\langle K^2\mathbf{v}, \mathbf{v}\rangle = -\langle K\mathbf{v}, K^*\mathbf{v}\rangle = \langle K\mathbf{v}, K\mathbf{v}\rangle = \|K\mathbf{v}\|^2$. Hence

$$\|T\mathbf{v}\|^2 = \langle T^*T\mathbf{v}, \mathbf{v}\rangle = \|S\mathbf{v}\|^2 + \|K\mathbf{v}\|^2 + \langle (SK - KS)\mathbf{v}, \mathbf{v}\rangle.$$

(b) By Proposition 6.6.18, $(SK - KS)^* = K^*S^* - S^*K^* = (-K)S - S(-K) = SK - KS$, so $SK - KS$ is self-adjoint. By (a), the identity $\|T\mathbf{v}\|^2 = \|S\mathbf{v}\|^2 + \|K\mathbf{v}\|^2$ holds for every $\mathbf{v}$ if and only if $\langle (SK - KS)\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$, which by Proposition 8.1.4 holds if and only if $SK = KS$. By Proposition 8.1.21 this says that T is normal.

Solution to Exercise 8.1.30

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Definition 6.3.26 (Conjugate transpose), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 6.4.29 (The orthogonal projection as a linear map), Theorem 6.4.31 (The matrix of an orthogonal projection)
Hint 1. $U^2 = U^*U$.
Hint 2. $P^2 = \big(I + 2U + U^2\big)/4$. For the geometry, what does U do to $\operatorname{im} P$ and to $\ker P = (\operatorname{im} P)^\perp$?
Sketch. (a) $U^2 = U^*U = I$. (b) $P^* = P$ and $P^2 = (2I + 2U)/4 = P$. $U = 2P - I$ fixes $\operatorname{im} P$ and negates $(\operatorname{im} P)^\perp$: it is the reflection across $\operatorname{im} P$. (c) $(2P - I)^* = 2P - I$ and $(2P - I)^2 = 4P^2 - 4P + I = I$. (d) $P = \frac{1}{2}\left(\begin{smallmatrix}1 & 1\\ 1 & 1\end{smallmatrix}\right)$, $U = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$.

(a) $U^2 = U^*U = I_n$, since $U^* = U$ and U is unitary.

(b) $P^* = (I_n + U^*)/2 = P$, and by (a),

$$P^2 = \frac{I_n + 2U + U^2}{4} = \frac{2I_n + 2U}{4} = P.$$

By Corollary 6.6.29, applied to $\mathbf{x} \mapsto P\mathbf{x}$ on $\mathbb{F}^n$ with the dot product (Example 6.6.10), this map is the orthogonal projection onto $W = \operatorname{im} P$, and clearly $U = 2P - I_n$. If $\mathbf{x} \in W$, then $P\mathbf{x} = \mathbf{x}$ and $U\mathbf{x} = \mathbf{x}$; if $\mathbf{x} \in W^\perp$, then $P\mathbf{x} = \mathbf{0}$ (Proposition 6.4.29) and $U\mathbf{x} = -\mathbf{x}$. Since $\mathbb{F}^n = W \oplus W^\perp$, U is the reflection across W: it keeps the component in W and reverses the component in $W^\perp$.

(c) If $P^2 = P = P^*$, then $(2P - I_n)^* = 2P - I_n$, and $(2P - I_n)^*(2P - I_n) = 4P^2 - 4P + I_n = I_n$. So $2P - I_n$ is Hermitian and unitary.

(d) With the unit vector $\mathbf{u} = (1,1)/\sqrt{2}$, the projection matrix is $P = \mathbf{u}\mathbf{u}^{\mathsf{T}} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (Theorem 6.4.31), so $U = 2P - I_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, the reflection across the line $x_1 = x_2$, which swaps the coordinates.

Solution to Exercise 8.1.31

> *Uses.* Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial), Example 7.3.5 (Small cases), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators)
> *Hint 1.* The eigenvalues are the roots of $x^2 - (\operatorname{tr} A)x + \det A$ (Example 7.3.5). Choose A so that this polynomial has no root in the field.
> *Hint 2.* Over $\mathbb{Q}$, try to get $x^2 - 2$. Over $\mathbb{F}_2$ there are only two candidates, 0 and 1; try to get $x^2 + x + 1$.
> *Hint 3.* For (c), the proof needs a field containing the complex roots of p_A, and an inner product to show that those roots lie in the original field.
> *Sketch.* (a) False: $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ has $p = x^2 - 2$. (b) False: $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ has $p = x^2 + x + 1$, with $p(0) = p(1) = 1$. (c) Over $\mathbb{Q}$ the proof still produces a real eigenvalue, $\pm\sqrt{2}$, but the argument only shows that it is real, not rational. Over $\mathbb{F}_2$ there is no inner product at all: the dot product has $(1,1)\cdot(1,1) = 0$, and the step $\lambda\|\mathbf{v}\|^2 = \overline{\lambda}\|\mathbf{v}\|^2 \Rightarrow \lambda = \overline{\lambda}$ has no meaning.

(a) False. $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is symmetric with $p_A = (x-1)(x+1) - 1 = x^2 - 2$ (Example 7.3.5). Its roots $\pm\sqrt{2}$ are irrational, so by Theorem 7.3.6 A has no eigenvalue in $\mathbb{Q}$.

(b) False. $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{2\times 2}$ is symmetric with $p_A = x(x-1) - 1 = x^2 + x + 1$ over $\mathbb{F}_2$. Since $p_A(0) = 1$ and $p_A(1) = 1 + 1 + 1 = 1$, it has no root in $\mathbb{F}_2$, and A has no eigenvalue there.

(c) Over $\mathbb{Q}$, we may still regard A as a complex Hermitian matrix, and the proof shows that its eigenvalues are real. The step that fails is the last one: the proof concludes that a real root λ of p_T lies in the field and is therefore an eigenvalue (Corollary 7.3.13), and $\mathbb{R} \subseteq \mathbb{Q}$ is false. Here the real roots $\pm\sqrt{2}$ are not rational. The theorem holds over $\mathbb{R}$ because $\mathbb{R}$ is the field in which the proof places the eigenvalue.

Over $\mathbb{F}_2$ the proof breaks at the start. There is no inner product on $\mathbb{F}_2^2$ in our sense, since positivity makes no sense in $\mathbb{F}_2$, and the dot product is degenerate: $(1,1)\cdot(1,1) = 1 + 1 = 0$. The key step of Proposition 8.1.1(a), dividing by $\|\mathbf{v}\|^2 \neq 0$ and comparing λ with $\overline{\lambda}$, has nothing to stand on. The roots of $x^2 + x + 1$ do exist in a larger field with four elements, but nothing forces them back into $\mathbb{F}_2$.

Solution to Exercise 8.1.32

> *Uses.* Proposition 8.1.14 (Kernels and images of normal operators), Theorem 6.6.21 (Kernels and images of adjoints), Theorem 6.4.6 (The projection is characterized by orthogonality)
> *Techniques.* Recognize a projection by its two conditions (Technique 6.4.7)
> *Hint 1.* First show $T^2 = T$. The operator $R = T^2 - T$ satisfies $TR = 0$. Where does $\operatorname{im} R$ lie?
> *Hint 2.* $\operatorname{im} R \subseteq \ker T$ because $TR = 0$, and $\operatorname{im} R \subseteq \operatorname{im} T$ because $R = T(T - \mathrm{id})$. Use Proposition 8.1.14(c).
> *Hint 3.* Once $T^2 = T$, check the two conditions characterizing $\operatorname{Pr}_{\operatorname{im} T}(\mathbf{v})$.
> *Sketch.* $R = T^2 - T = T(T - \mathrm{id})$ has $\operatorname{im} R \subseteq \operatorname{im} T$, and $TR = T^3 - T^2 = 0$ gives $\operatorname{im} R \subseteq \ker T$. Since $\ker T \cap \operatorname{im} T = \{0\}$ for normal T, $R = 0$. Then $T\mathbf{v} \in \operatorname{im} T$ and $T(\mathbf{v} - T\mathbf{v}) = \mathbf{0}$, so $\mathbf{v} - T\mathbf{v} \in \ker T = (\operatorname{im} T)^{\perp}$, and $T\mathbf{v} = \operatorname{Pr}_{\operatorname{im} T}(\mathbf{v})$.

Let $R = T^2 - T$. Since $R = T(T - \mathrm{id})$, every vector $R\mathbf{v} = T((T - \mathrm{id})\mathbf{v})$ lies in $\operatorname{im} T$. Since $TR = T^3 - T^2 = 0$, every vector $R\mathbf{v}$ lies in $\ker T$. By Proposition 8.1.14(c), $\ker T$ and $\operatorname{im} T$ are orthogonal, so their intersection is $\{\mathbf{0}\}$ (a vector orthogonal to itself is $\mathbf{0}$). Hence $R = 0$, that is, $T^2 = T$.

Let $W = \operatorname{im} T$ and $\mathbf{v} \in V$. Then $T\mathbf{v} \in W$, and $T(\mathbf{v} - T\mathbf{v}) = T\mathbf{v} - T^2\mathbf{v} = \mathbf{0}$, so $\mathbf{v} - T\mathbf{v} \in \ker T = \ker T^* = (\operatorname{im} T)^{\perp} = W^{\perp}$ by Proposition 8.1.14(a) and Theorem 6.6.21(a). By Theorem 6.4.6, $T\mathbf{v} = \operatorname{Pr}_W(\mathbf{v})$. So $T = \operatorname{Pr}_W$.

Solution to Exercise 8.1.33

> *Uses.* Proposition 2.4.28 (Cyclicity of the trace), Theorem 6.4.20 (Orthonormal bases exist), Proposition 6.3.34 (Characterizations of unitary matrices), Proposition 8.1.4 (Self-adjoint operators are determined by their diagonal values), Example 6.6.10 (The adjoint of a matrix)
> *Techniques.* Cyclicity of trace (Technique 2.4.29)
> *Hint 1.* Let $S = A^*A - AA^*$. Then S is Hermitian and $\mathbf{x}^*S\mathbf{x} = \|A\mathbf{x}\|^2 - \|A^*\mathbf{x}\|^2 \geq 0$. What is $\operatorname{tr} S$?
> *Hint 2.* $\operatorname{tr} S = 0$ by cyclicity of the trace. Fix a unit vector $\mathbf{u}$, extend it to an orthonormal basis, and let Q be the unitary matrix with these columns. Compute $\operatorname{tr}(Q^*SQ)$ in two ways.
> *Hint 3.* The diagonal entries of Q^*SQ are the nonnegative numbers $\mathbf{u}_j^*S\mathbf{u}_j$, and they add up to 0.

Sketch. $S = A^*A - AA^*$ is Hermitian with $\mathbf{x}^*S\mathbf{x} = \|A\mathbf{x}\|^2 - \|A^*\mathbf{x}\|^2 \geq 0$ and $\operatorname{tr} S = 0$. For a unit vector $\mathbf{u} = \mathbf{u}_1$ completed to an orthonormal basis (columns of a unitary Q), $0 = \operatorname{tr} S = \operatorname{tr}(Q^*SQ) = \sum_j \mathbf{u}_j^* S\mathbf{u}_j$, a sum of nonnegative terms, so $\mathbf{u}^*S\mathbf{u} = 0$. By scaling, $\mathbf{x}^*S\mathbf{x} = 0$ for all $\mathbf{x}$, and $S = 0$ by Proposition 8.1.4.

Let $S = A^*A - AA^*$. Then $S^* = S$, and for $\mathbf{x} \in \mathbb{F}^n$,

$$\mathbf{x}^*S\mathbf{x} = (A\mathbf{x})^*(A\mathbf{x}) - (A^*\mathbf{x})^*(A^*\mathbf{x}) = \|A\mathbf{x}\|^2 - \|A^*\mathbf{x}\|^2 \geq 0.$$

By cyclicity of the trace (Proposition 2.4.28), $\operatorname{tr} S = \operatorname{tr}(A^*A) - \operatorname{tr}(AA^*) = 0$.

Let $\mathbf{u} \in \mathbb{F}^n$ be a unit vector. By Theorem 6.4.20(b), it extends to an orthonormal basis $\mathbf{u} = \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ of $\mathbb{F}^n$; let Q be the unitary matrix with these columns. Then $QQ^* = I_n$ (Proposition 6.3.34), so by cyclicity

$$0 = \operatorname{tr} S = \operatorname{tr}(SQQ^*) = \operatorname{tr}(Q^*SQ) = \sum_{j=1}^{n} \mathbf{u}_j^* S\mathbf{u}_j.$$

Each term is ≥ 0, so each is 0; in particular $\mathbf{u}^*S\mathbf{u} = 0$. Every nonzero $\mathbf{x}$ is a positive multiple of a unit vector, so $\mathbf{x}^*S\mathbf{x} = 0$ for all $\mathbf{x}$. Now $\mathbf{x}^*S\mathbf{x} = \langle S\mathbf{x}, \mathbf{x}\rangle$ for the dot product, and $\mathbf{x} \mapsto S\mathbf{x}$ is self-adjoint since $S^* = S$ (Example 6.6.10). By Proposition 8.1.4, $S = 0$, so A is normal.

Solution to Exercise 8.1.34

Uses. Proposition 2.4.23 (Transpose of a product), Definition 8.1.8 (Normal operator, normal matrix), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Exercise 8.1.25, Theorem 6.4.31 (The matrix of an orthogonal projection)
Hint 1. Transpose the equation $A^{\mathsf{T}} = A^2$. For normality, compute $A^{\mathsf{T}}A$ and AA^{T}.
Hint 2. $A^3 = A^{\mathsf{T}}A$ is symmetric. Compute $\left(A^3\right)^2$ using $A^4 = A$.
Hint 3. For (c), use Exercise 8.1.25: A is symmetric or a scaled rotation. In the symmetric case $A = A^2$. In the other case compare A^{T} and A^2 entry by entry.
Sketch. (a) $A = \left(A^2\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^2 = A^4$, and $A^{\mathsf{T}}A = A^3 = AA^{\mathsf{T}}$. (b) $A^3 = A^{\mathsf{T}}A$ is symmetric, and $\left(A^3\right)^2 = A^4A^2 = A^3$. (c) Symmetric: $A = A^2$, so A is an orthogonal projection: 0, I_2, or $\mathbf{u}\mathbf{u}^{\mathsf{T}}$ with $\|\mathbf{u}\| = 1$. Scaled rotation $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$, $b \neq 0$: $A^{\mathsf{T}} = A^2$ gives $a = a^2 - b^2$ and $b = -2ab$, so $a = -1/2$, $b = \pm\sqrt{3}/2$: the rotations by $\pm 2\pi/3$.

(a) Transposing $A^{\mathsf{T}} = A^2$ and using Proposition 2.4.23, $A = \left(A^2\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^2 = \left(A^2\right)^2 = A^4$. Moreover $A^{\mathsf{T}}A = A^2A = A^3$ and $AA^{\mathsf{T}} = AA^2 = A^3$, so $A^{\mathsf{T}}A = AA^{\mathsf{T}}$, and A is normal.

(b) $A^3 = A^{\mathsf{T}}A$ is symmetric, since $\left(A^{\mathsf{T}}A\right)^{\mathsf{T}} = A^{\mathsf{T}}A$. And by (a), $\left(A^3\right)^2 = A^6 = A^4A^2 = AA^2 = A^3$. So $P = A^3$ satisfies $P^2 = P = P^{\mathsf{T}}$, and $\mathbf{x} \mapsto P\mathbf{x}$ is an orthogonal projection by Corollary 6.6.29.

(c) By (a) and Exercise 8.1.25, either A is symmetric or $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b \neq 0$.

If A is symmetric, then $A = A^{\mathsf{T}} = A^2$, so $A^2 = A = A^{\mathsf{T}}$ and A is the matrix of an orthogonal projection Pr_W (Corollary 6.6.29). For $W = \{\mathbf{0}\}$ or $W = \mathbb{R}^2$ this gives $A = 0$ or $A = I_2$. For a line $W = \operatorname{span}(\mathbf{u})$ with $\|\mathbf{u}\| = 1$, it gives $A = \mathbf{u}\mathbf{u}^{\mathsf{T}}$ (Theorem 6.4.31). All of these satisfy $A^{\mathsf{T}} = A = A^2$.

If $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b \neq 0$, then

$$A^{\mathsf{T}} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad A^2 = \begin{pmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{pmatrix}.$$

Equality requires $a = a^2 - b^2$ and $b = -2ab$. As $b \neq 0$, the second gives $a = -1/2$, and then the first gives $b^2 = a^2 - a = 3/4$, so $b = \pm\sqrt{3}/2$. Conversely these values satisfy both equations. They give

$$A = \begin{pmatrix} -1/2 & \mp\sqrt{3}/2 \\ \pm\sqrt{3}/2 & -1/2 \end{pmatrix},$$

the rotations by $\pm 2\pi/3$. This is consistent with (a): for a rotation R by $\pm 2\pi/3$, $R^3 = I_2$, so $R^{\mathsf{T}} = R^{-1} = R^2$.

So for $n = 2$ the solutions are 0, I_2, the matrices $\mathbf{u}\mathbf{u}^{\mathsf{T}}$ with $\mathbf{u}$ a unit vector, and the two rotations by $\pm 2\pi/3$.

Solution to Exercise 8.1.35

Uses. Exercise 8.1.6, Example 6.6.10 (The adjoint of a matrix), Corollary 3.5.25 (The invertible matrix theorem), Proposition 6.3.27 (Properties of the conjugate transpose), Definition 6.3.33 (Unitary and orthogonal matrices)
Techniques. Guess the inverse, then check both products (Technique 2.5.16)
Hint 1. For (a), apply Exercise 8.1.6 to $\mathbf{x} \mapsto K\mathbf{x}$ and to $\mathbf{x} \mapsto -K\mathbf{x}$, both skew-adjoint for the dot product.
Hint 2. For (b), first show that $I_n - K$ and $(I_n + K)^{-1}$ commute, using $(I_n - K)(I_n + K) = I_n - K^2 = (I_n + K)(I_n - K)$. Then $Q^* = (I_n - K)^{-1}(I_n + K)$.
Hint 3. If $Q\mathbf{x} = -\mathbf{x}$, write $\mathbf{x} = (I_n + K)\mathbf{y}$.
Sketch. (a) $\|(I+K)\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|K\mathbf{x}\|^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$, and the same for $-K$. (b) $Q^* = \left((I+K)^{-1}\right)^*(I-K)^* = (I-K)^{-1}(I+K)$, and $Q^*Q = (I-K)^{-1}(I+K)(I-K)(I+K)^{-1} = I$ because $I+K$ and $I-K$ commute. If $Q\mathbf{x} = -\mathbf{x}$ with $\mathbf{x} = (I+K)\mathbf{y}$, then $(I-K)\mathbf{y} = -(I+K)\mathbf{y}$, so $\mathbf{y} = \mathbf{0}$. (c) $Q = \frac{1}{1+t^2}\begin{pmatrix} 1-t^2 & 2t \\ -2t & 1-t^2 \end{pmatrix}$, the rotation by $-2\arctan t$.

(a) By Example 6.6.10, $\mathbf{x} \mapsto K\mathbf{x}$ on $\mathbb{F}^n$ with the dot product has adjoint $\mathbf{x} \mapsto K^*\mathbf{x} = -K\mathbf{x}$, so it is skew-adjoint, and so is $\mathbf{x} \mapsto -K\mathbf{x}$. By Exercise 8.1.6, $\|(I_n + K)\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|K\mathbf{x}\|^2$, which is positive for $\mathbf{x} \neq \mathbf{0}$, and likewise with $-K$ in place of K. So $I_n + K$ and $I_n - K$ have trivial null space and are invertible (Corollary 3.5.25).

(b) Since $(I_n - K)(I_n + K) = I_n - K^2 = (I_n + K)(I_n - K)$, multiplying on both sides by $(I_n + K)^{-1}$ gives $(I_n + K)^{-1}(I_n - K) = (I_n - K)(I_n + K)^{-1}$. By Proposition 6.3.27, $(I_n - K)^* = I_n + K$ and $\left((I_n + K)^{-1}\right)^* = \left((I_n + K)^*\right)^{-1} = (I_n - K)^{-1}$, so

$$Q^* = \left((I_n + K)^{-1}\right)^*(I_n - K)^* = (I_n - K)^{-1}(I_n + K).$$

Hence, using that $I_n + K$ and $I_n - K$ commute,

$$Q^*Q = (I_n - K)^{-1}(I_n + K)(I_n - K)(I_n + K)^{-1} = (I_n - K)^{-1}(I_n - K)(I_n + K)(I_n + K)^{-1} = I_n,$$

so Q is unitary (Definition 6.3.33). Suppose $Q\mathbf{x} = -\mathbf{x}$, and let $\mathbf{y} = (I_n + K)^{-1}\mathbf{x}$, so that $\mathbf{x} = (I_n + K)\mathbf{y}$ and $Q\mathbf{x} = (I_n - K)\mathbf{y}$. Then $(I_n - K)\mathbf{y} = -(I_n + K)\mathbf{y}$, so $2\mathbf{y} = \mathbf{0}$, hence $\mathbf{y} = \mathbf{0}$ and $\mathbf{x} = \mathbf{0}$. So -1 is not an eigenvalue of Q.

(c) Here $I_2 + K = \begin{pmatrix} 1 & -t \\ t & 1 \end{pmatrix}$ with inverse $\frac{1}{1+t^2}\begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix}$, and $I_2 - K = \begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix}$. So

$$Q = \frac{1}{1+t^2}\begin{pmatrix} 1 & t \\ -t & 1 \end{pmatrix}^2 = \frac{1}{1+t^2}\begin{pmatrix} 1-t^2 & 2t \\ -2t & 1-t^2 \end{pmatrix}.$$

With $t = \tan(\varphi/2)$, the double-angle formulas give $\cos\varphi = \left(1-t^2\right)/\left(1+t^2\right)$ and $\sin\varphi = 2t/\left(1+t^2\right)$, so $Q = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, the rotation by $-\varphi = -2\arctan t$. As t runs over $\mathbb{R}$, φ runs over $(-\pi, \pi)$, so every rotation except the rotation by π, which is $-I_2$, arises, in line with (b).

Solution to Exercise 8.1.36

Uses. Theorem 7.4.2 (Diagonalizable means an eigenbasis), Theorem 6.4.20 (Orthonormal bases exist), Theorem 8.1.18 (Eigenspaces of a normal operator are orthogonal), Proposition 6.3.13 (Orthogonal lists are independent), Proposition 8.1.7 (What an orthonormal eigenbasis forces)
Techniques. Glue bases of the summands (Technique 4.4.14)
Hint 1. Choose an orthonormal basis of each eigenspace and put them together. Why is the result orthonormal?
Hint 2. Vectors from different eigenspaces are orthogonal by Theorem 8.1.18. Why does the combined list span V?
Sketch. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues. Concatenate orthonormal bases of $E_{\lambda_1}, \ldots, E_{\lambda_k}$. Within each piece the list is orthonormal, and across pieces orthogonal by Theorem 8.1.18. Its span contains every eigenspace, hence every vector of an eigenbasis, hence V. The converse is Proposition 8.1.7.

Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. By Theorem 6.4.20, each eigenspace E_{λ_j} has an orthonormal basis $\mathcal{O}_j$. Let $\mathcal{O}$ be the concatenation of $\mathcal{O}_1, \ldots, \mathcal{O}_k$. Two vectors of $\mathcal{O}$ from the same $\mathcal{O}_j$ are orthogonal unit vectors, and two vectors from different eigenspaces are orthogonal by Theorem 8.1.18. So $\mathcal{O}$ is orthonormal, hence independent (Proposition 6.3.13), and it consists of eigenvectors.

It remains to show that $\mathcal{O}$ spans V. Since T is diagonalizable, V has a basis of eigenvectors (Theorem 7.4.2). Each of its vectors lies in some $E_{\lambda_j} = \operatorname{span}\left(\mathcal{O}_j\right) \subseteq \operatorname{span}(\mathcal{O})$, so $V \subseteq \operatorname{span}(\mathcal{O})$. Thus $\mathcal{O}$ is an orthonormal basis of eigenvectors.

For the last sentence: a normal diagonalizable operator has an orthonormal eigenbasis by what we just proved, and an operator with an orthonormal eigenbasis is normal by Proposition 8.1.7. (Section 8.2 shows that over $\mathbb{C}$ the hypothesis "diagonalizable" is automatic for normal operators.)

Solution to Exercise 8.1.37

> *Uses.* Theorem 7.5.18 (Schur's theorem), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.1.16 (Frobenius inner product), Proposition 6.3.34 (Characterizations of unitary matrices), Proposition 8.1.7 (What an orthonormal eigenbasis forces)
> *Techniques.* The trace of A^*A (Technique 6.1.18), Cyclicity of trace (Technique 2.4.29)
> *Hint 1.* Triangularize A unitarily with Schur's theorem: $A = URU^*$. How do $\operatorname{tr}(A^*A)$ and $\operatorname{tr}(R^*R)$ compare?
> *Hint 2.* $\operatorname{tr}(A^*A) = \operatorname{tr}(R^*R) = \sum_{i \le j} |r_{ij}|^2$, and the diagonal of R lists the eigenvalues.
> *Hint 3.* Equality forces every entry of R above the diagonal to vanish.
> *Sketch.* $A = URU^*$ with R upper triangular and diagonal $\lambda_1, \dots, \lambda_n$ (Theorem 7.5.18). Then $\operatorname{tr}(A^*A) = \operatorname{tr}(UR^*RU^*) = \operatorname{tr}(R^*R) = \sum_{i \le j} |r_{ij}|^2 = \sum_j |\lambda_j|^2 + \sum_{i<j} |r_{ij}|^2$. Equality means R is diagonal, so $A^*A = UR^*RU^* = URR^*U^* = AA^*$.

(a) By Theorem 7.5.18(b), $A = URU^*$ with U unitary and $R = (r_{ij})$ upper triangular with diagonal entries $\lambda_1, \dots, \lambda_n$. Then $A^*A = UR^*U^*URU^* = UR^*RU^*$, and by cyclicity of the trace (Proposition 2.4.28) and $U^*U = I_n$,

$$\operatorname{tr}(A^*A) = \operatorname{tr}(R^*RU^*U) = \operatorname{tr}(R^*R).$$

By Proposition 6.1.16, $\operatorname{tr}(R^*R) = \sum_{i,j} |r_{ij}|^2$, and since $r_{ij} = 0$ for $i > j$,

$$\operatorname{tr}(A^*A) = \sum_{j=1}^{n} |\lambda_j|^2 + \sum_{i<j} |r_{ij}|^2 \ge \sum_{j=1}^{n} |\lambda_j|^2.$$

The formula $\operatorname{tr}(A^*A) = \sum_{i,j} |a_{ij}|^2$ is Proposition 6.1.16 again.

(b) If equality holds, then $\sum_{i<j} |r_{ij}|^2 = 0$, so $r_{ij} = 0$ for all $i < j$ and R is diagonal. Diagonal matrices commute with their conjugate transposes, so

$$A^*A = UR^*RU^* = URR^*U^* = AA^*.$$

(Equivalently, the columns of U are an orthonormal eigenbasis of A, and Proposition 8.1.7 applies.) The converse, that normal matrices give equality, follows from the complex spectral theorem in Section 8.2.

Solution to Exercise 8.1.38

> *Uses.* Proposition 6.4.29 (The orthogonal projection as a linear map), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents), Proposition 2.4.28 (Cyclicity of the trace), Proposition 6.1.16 (Frobenius inner product), Exercise 6.6.38
> *Techniques.* The trace of A^*A (Technique 6.1.18), Cyclicity of trace (Technique 2.4.29)
> *Hint 1.* $P^2 = P = P^*$, and $AW \subseteq W$ means $PAP = AP$. Taking conjugate transposes, $PA^*P = PA^*$.
> *Hint 2.* Write $PAA^*P = PA(P + (I - P))A^*P$, so $\operatorname{tr}(PAA^*P) = \operatorname{tr}(PAPA^*P) + \operatorname{tr}(XX^*)$. Compute $\operatorname{tr}(PAA^*P)$ another way, using $AA^* = A^*A$ and $AP = PAP$.
> *Hint 3.* You should find $\operatorname{tr}(XX^*) = 0$, so $X = 0$. What does $PA(I - P) = 0$ say about A^*?
> *Sketch.* With $X = PA(I - P)$, $\operatorname{tr}(PAA^*P) = \operatorname{tr}(PAPA^*P) + \operatorname{tr}(XX^*)$. On the other hand $\operatorname{tr}(PAA^*P) = \operatorname{tr}(PA^*AP) = \operatorname{tr}\big((AP)^*(AP)\big) = \operatorname{tr}\big((PAP)^*(PAP)\big) = \operatorname{tr}(PA^*PAP) = \operatorname{tr}(PAPA^*P)$ by cyclicity. So $\operatorname{tr}(XX^*) = 0$, $X = 0$, and $X^* = (I - P)A^*P = 0$ says $A^*P = PA^*P$, that is, $A^*W \subseteq W$. The last claim is Exercise 6.6.38 applied to A^*.

Since P is the matrix of an orthogonal projection, $P^2 = P = P^*$ (Corollary 6.6.29), and $\operatorname{im} P = W$, with $P\mathbf{w} = \mathbf{w}$ for $\mathbf{w} \in W$ (Proposition 6.4.29). The hypothesis $AW \subseteq W$ says that $AP\mathbf{x} \in W$ for every $\mathbf{x}$, that is, $PAP = AP$.

Let $X = PA(I_n - P)$. Inserting $I_n = P + (I_n - P)$ and using $(I_n - P)^2 = I_n - P = (I_n - P)^*$,

$$PAA^*P = PAPA^*P + PA(I_n - P)(I_n - P)A^*P = (PAP)(PAP)^* + XX^*.$$

On the other hand, by normality, then $AP = PAP$, then cyclicity of the trace (Proposition 2.4.28),

$$\operatorname{tr}(PAA^*P) = \operatorname{tr}(PA^*AP) = \operatorname{tr}\big((AP)^*(AP)\big) = \operatorname{tr}\big((PAP)^*(PAP)\big) = \operatorname{tr}\big((PAP)(PAP)^*\big).$$

Comparing the traces of the two expressions for PAA^*P gives $\operatorname{tr}(XX^*) = 0$. By Proposition 6.1.16, $\operatorname{tr}(XX^*) = \operatorname{tr}\big((X^*)^*X^*\big)$ is the sum of the squared moduli of the entries of X^*, so $X^* = 0$, that is, $(I_n - P)A^*P = 0$. So $A^*P = PA^*P$, which says that $A^*\mathbf{w} = PA^*\mathbf{w} \in W$ for every $\mathbf{w} \in W$: $A^*W \subseteq W$.

Finally, by Exercise 6.6.38 applied to A^* (whose adjoint is A), $A^*W \subseteq W$ implies $AW^\perp \subseteq W^\perp$.

Section 8.2

Solution to Exercise 8.2.9

Uses. Theorem 8.2.5 (Complex spectral theorem), Proposition 8.2.2 (Necessary conditions)
Hint 1. By Theorem 8.2.5, check $A^*A = AA^*$ for the first question. For the second, an orthogonally diagonalizable real matrix must be symmetric (Proposition 8.2.2).
Hint 2. In (e), the matrix is $I_3 + P$ for a permutation matrix P with $P^{\mathsf{T}} = P^{-1}$.
Sketch. (a) Not normal: neither. (b) Normal, not symmetric: unitarily only. (c) $A^*A = 2I_2 = AA^*$: unitarily. (d) Both: $(1,1)$ and $(1,-1)$ are orthogonal eigenvectors. (e) $I + P$ with $P^{\mathsf{T}}P = PP^{\mathsf{T}} = I$ is normal but not symmetric: unitarily only.

By Theorem 8.2.5 a complex matrix is unitarily diagonalizable if and only if it is normal. A real orthogonally diagonalizable matrix is symmetric by Proposition 8.2.2(b), since its eigenvalues are real. (The converse is Theorem 8.2.12 below.)

(a) With $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, $(A^*A)_{11} = 1$ but $(AA^*)_{11} = 1 + 4 = 5$. Not normal, so neither.

(b) $A = 2J$ with $J^{\mathsf{T}}J = JJ^{\mathsf{T}} = I_2$, so $A^{\mathsf{T}}A = 4I_2 = AA^{\mathsf{T}}$ and A is normal: unitarily diagonalizable. It is not symmetric, so not orthogonally diagonalizable.

(c) $A^* = \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}$, and $A^*A = \begin{pmatrix} 1+1 & i-i \\ -i+i & 1+1 \end{pmatrix} = 2I_2$; similarly $AA^* = 2I_2$. Normal, so unitarily diagonalizable (although $A^* \neq A$). It is not real.

(d) $A(1,1) = 3(1,1)$ and $A(1,-1) = -(1,-1)$, and these eigenvectors are orthogonal. So $A = Q\,\operatorname{diag}(3,-1)Q^{\mathsf{T}}$ with $Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$: orthogonally, and in particular unitarily, diagonalizable.

(e) $A = I_3 + P$ where P is the permutation matrix with $P\mathbf{e}_1 = \mathbf{e}_3$, $P\mathbf{e}_2 = \mathbf{e}_1$, $P\mathbf{e}_3 = \mathbf{e}_2$. Its columns are orthonormal, so $P^{\mathsf{T}}P = I_3$, hence $P^{\mathsf{T}} = P^{-1}$ and $PP^{\mathsf{T}} = I_3$ as well. Then $A^{\mathsf{T}}A = I + P + P^{\mathsf{T}} + I = AA^{\mathsf{T}}$, so A is normal: unitarily diagonalizable. It is not symmetric ($a_{12} = 1 \neq 0 = a_{21}$), so not orthogonally diagonalizable.

Solution to Exercise 8.2.16

Uses. Algorithm 8.2.14 (Orthogonal diagonalization)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. The trace is 11 and the determinant is 24.
Hint 2. The eigenvalues are 3 and 8. Find one eigenvector for each and normalize; they are automatically orthogonal.
Sketch. $p = x^2 - 11x + 24 = (x-3)(x-8)$. $E_8 = \operatorname{span}((2,-1))$ and $E_3 = \operatorname{span}((1,2))$. $Q = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$, $D = \operatorname{diag}(8,3)$.

$p_A = (x-7)(x-4) - 4 = x^2 - 11x + 24 = (x-3)(x-8)$. From $A - 8I_2 = \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix}$ we get $E_8 = \operatorname{span}((2,-1))$, and from $A - 3I_2 = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$ we get $E_3 = \operatorname{span}((1,2))$. These are orthogonal, as they must be. Normalizing,

$$Q = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 8 & 0 \\ 0 & 3 \end{pmatrix},$$

and one checks $AQ = \frac{1}{\sqrt{5}}\begin{pmatrix} 16 & 3 \\ -8 & 6 \end{pmatrix} = QD$.

Solution to Exercise 8.2.22

Uses. Proposition 8.2.20 (Spectral projections are polynomials in T), Theorem 8.2.18 (Spectral decomposition), Theorem 6.4.31 (The matrix of an orthogonal projection)
Hint 1. $p_H = (x-3)^2 - 1$.
Hint 2. With eigenvalues 2 and 4, $P_4 = (H - 2I_2)/2$ and $P_2 = (H - 4I_2)/(-2)$.
Sketch. Eigenvalues $2, 4$. $P_4 = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = \mathbf{u}\mathbf{u}^*$ with $\mathbf{u} = \frac{1}{\sqrt{2}}(i, 1)$, and $P_2 = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$. $P_4 P_2 = \frac{1}{4}\begin{pmatrix} 1+i^2 & -i+i \\ -i+i & i^2+1 \end{pmatrix} = 0$.

$p_H = (x-3)^2 - i(-i) = (x-3)^2 - 1 = (x-2)(x-4)$. By Proposition 8.2.20,

$$P_4 = \frac{H - 2I_2}{4-2} = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, \quad P_2 = \frac{H - 4I_2}{2-4} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

From eigenvectors: $H(i, 1) = (3i + i, 1 + 3) = 4(i, 1)$ and $H(-i, 1) = (-3i + i, -1 + 3) = 2(-i, 1)$. With $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(i, 1)$ and $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(-i, 1)$, Theorem 6.4.31 gives

$$\mathbf{u}_1\mathbf{u}_1^* = \frac{1}{2}\begin{pmatrix} i \\ 1 \end{pmatrix}(-i \;\; 1) = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = P_4, \quad \mathbf{u}_2\mathbf{u}_2^* = \frac{1}{2}\begin{pmatrix} -i \\ 1 \end{pmatrix}(i \;\; 1) = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} = P_2.$$

So $H = 4P_4 + 2P_2$. Finally

$$P_4 P_2 = \frac{1}{4}\begin{pmatrix} 1\cdot 1 + i \cdot i & 1\cdot(-i) + i \cdot 1 \\ -i \cdot 1 + 1 \cdot i & (-i)(-i) + 1 \cdot 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Solution to Exercise 8.2.32

Uses. Algorithm 8.2.14 (Orthogonal diagonalization), Theorem 6.4.15 (Gram–Schmidt is correct), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Techniques. Diagonalize, solve, conjugate back (Technique 7.4.22)
Hint 1. One eigenvalue is -2, with eigenvector $(2, 1, -2)$. Use the trace to find the others.
Hint 2. The other eigenvalue is 7, with a two-dimensional eigenspace: the plane orthogonal to $(2, 1, -2)$. Find a basis of it and apply Gram–Schmidt.
Sketch. $A(2, 1, -2) = -2(2, 1, -2)$. Since $\operatorname{tr} A = 12$ and p_A splits with real roots, and $A - 7I_3$ has rank 1, the eigenvalues are $-2, 7, 7$. $E_7 = (2, 1, -2)^\perp$ has basis $(1, 0, 1), (1, -2, 0)$; Gram–Schmidt gives $\frac{1}{\sqrt{2}}(1, 0, 1)$ and $\frac{1}{3\sqrt{2}}(1, -4, -1)$.

Let A be the matrix. We have

$$A - 7I_3 = \begin{pmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{pmatrix},$$

whose rows are all multiples of $(2, 1, -2)$. So $\operatorname{rank}(A - 7I_3) = 1$ and $E_7 = \ker(A - 7I_3) = \{\mathbf{x} \mid 2x_1 + x_2 - 2x_3 = 0\}$ is a plane. Since $\operatorname{tr} A = 12$ and A is diagonalizable with eigenvalue 7 of multiplicity 2 (Theorem 8.2.12), the third eigenvalue is $12 - 14 = -2$ (Theorem 7.5.22). Indeed $A(2, 1, -2) = (6 - 2 - 8, -4 + 6 - 4, 8 + 2 - 6) = (-4, -2, 4) = -2(2, 1, -2)$.

For E_7, take the basis $(1, 0, 1), (1, -2, 0)$ (both satisfy $2x_1 + x_2 - 2x_3 = 0$). Gram–Schmidt: $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, 0, 1)$, and

$$(1, -2, 0) - \frac{1 \cdot 1 + 0 + 0}{2}(1, 0, 1) = \left(\frac{1}{2}, -2, -\frac{1}{2}\right) = \frac{1}{2}(1, -4, -1),$$

which normalizes to $\mathbf{u}_3 = \frac{1}{3\sqrt{2}}(1, -4, -1)$. With $\mathbf{u}_1 = \frac{1}{3}(2, 1, -2)$,

$$Q = \begin{pmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{3} & 0 & -\frac{4}{3\sqrt{2}} \\ -\frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \end{pmatrix}, \quad D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

The columns are orthonormal: $\mathbf{u}_1 \perp E_7$ automatically, and $\mathbf{u}_2 \perp \mathbf{u}_3$ by construction.

Solution to Exercise 8.2.33

Uses. Theorem 8.2.18 (Spectral decomposition), Proposition 8.2.19 (Uniqueness of the spectral decomposition), Proposition 8.2.24 (Properties of $f(T)$), Corollary 6.6.29 (Orthogonal projections are the self-adjoint idempotents)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25), Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. Write $T = \sum_j \lambda_j P_j$. Then $T^* = \sum_j \overline{\lambda_j} P_j$, $T^2 = \sum_j \lambda_j^2 P_j$, and $T^*T = \sum_j |\lambda_j|^2 P_j$.
Hint 2. Two operators $\sum a_j P_j$ and $\sum b_j P_j$ are equal if and only if $a_j = b_j$ for all j: apply both to a vector of E_{λ_j}.
Sketch. With $T = \sum \lambda_j P_j$: $T^* = T$ iff $\overline{\lambda_j} = \lambda_j$; $T^* = -T$ iff $\overline{\lambda_j} = -\lambda_j$; $T^2 = T = T^*$ iff $\lambda_j^2 = \lambda_j = \overline{\lambda_j}$, i.e. $\lambda_j \in \{0, 1\}$; $T^*T = \mathrm{id}$ iff $|\lambda_j|^2 = 1$. Counterexample to (a): $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Write $T = \sum_{j=1}^{k} \lambda_j P_j$ as in Theorem 8.2.18. By Proposition 8.2.24, $T^* = \sum_j \overline{\lambda_j} P_j$, $T^2 = \sum_j \lambda_j^2 P_j$, and $T^*T = \sum_j |\lambda_j|^2 P_j$, while $\mathrm{id}_V = \sum_j P_j$. Two operators $\sum_j a_j P_j$ and $\sum_j b_j P_j$ are equal if and only if $a_j = b_j$ for every j: apply both to a nonzero $\mathbf{v} \in E_{\lambda_j}$, on which $\sum_i a_i P_i$ acts as a_j (Proposition 8.2.24(a)).

(a) $T^* = T$ if and only if $\overline{\lambda_j} = \lambda_j$ for all j, that is, every λ_j is real.

(b) $T^* = -T$ if and only if $\overline{\lambda_j} = -\lambda_j$ for all j, that is, $\operatorname{Re} \lambda_j = 0$.

(c) By Corollary 6.6.29, T is an orthogonal projection if and only if $T^2 = T = T^*$, that is, $\lambda_j^2 = \lambda_j = \overline{\lambda_j}$ for all j. The first equation says $\lambda_j \in \{0, 1\}$, and then the second holds automatically.

(d) $T^*T = \mathrm{id}_V$ if and only if $|\lambda_j|^2 = 1$ for all j.

For the example, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has only the real eigenvalue 1, but $A^* = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \neq A$.

Solution to Exercise 8.2.34

Uses. Theorem 8.2.5 (Complex spectral theorem), Theorem 8.2.12 (Real spectral theorem), Proposition 8.2.24 (Properties of $f(T)$), Theorem 7.3.24 (Every monic polynomial is a characteristic polynomial), Proposition 8.2.2 (Necessary conditions)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25), Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a) and (b), apply the relation to an eigenvector: what equation does each eigenvalue satisfy?
Hint 2. For (c), look for a real 2×2 matrix whose characteristic polynomial is the other factor of $x^3 + x - 2 = (x - 1)(x^2 + x + 2)$.
Hint 3. For (d), think of repeated eigenvalues. For (e), choose an orthonormal basis inside each eigenspace.
Sketch. (a) True: $\lambda^m = 0$ forces every eigenvalue to be 0, and $T = \sum \lambda_j P_j = 0$. (b) True: real eigenvalues of T satisfy $\lambda^3 + \lambda - 2 = (\lambda - 1)(\lambda^2 + \lambda + 2) = 0$, so $\lambda = 1$ and $T = P_1 = \mathrm{id}$. (c) False: the companion matrix C of $x^2 + x + 2$ satisfies $C^2 = -C - 2I$, hence $C^3 + C = 2I$. (d) False: I_2 with the basis $(1, 0), (1, 1)$. (e) True: orthonormal bases of the eigenspaces combine into an orthonormal eigenbasis.

(a) True. By Theorem 8.2.5, $T = \sum_j \lambda_j P_j$. If $\mathbf{v} \in E_{\lambda_j}$ is nonzero, then $\mathbf{0} = T^m \mathbf{v} = \lambda_j^m \mathbf{v}$, so $\lambda_j = 0$. Hence the only eigenvalue is 0 and $T = 0 \cdot P_1 = 0$.

(b) True. By Theorem 8.2.12, V has an orthonormal eigenbasis $\mathbf{u}_1, \ldots, \mathbf{u}_n$, with $T\mathbf{u}_i = \mu_i \mathbf{u}_i$ and $\mu_i \in \mathbb{R}$. Applying $T^3 + T = 2\mathrm{id}_V$ to $\mathbf{u}_i$ gives $\mu_i^3 + \mu_i - 2 = 0$. Since $x^3 + x - 2 = (x - 1)(x^2 + x + 2)$ and $x^2 + x + 2$ has discriminant $1 - 8 < 0$, the only real root is 1. So $T\mathbf{u}_i = \mathbf{u}_i$ for all i, and $T = \mathrm{id}_V$.

(c) False. Let $C = \begin{pmatrix} 0 & -2 \\ 1 & -1 \end{pmatrix}$, the companion matrix of $x^2 + x + 2$ (Theorem 7.3.24); directly, $C^2 = \begin{pmatrix} -2 & 2 \\ -1 & -1 \end{pmatrix} = -C - 2I_2$. Then $C^3 = -C^2 - 2C = C + 2I_2 - 2C = 2I_2 - C$, so $C^3 + C = 2I_2$, but $C \neq I_2$. (Here C is not symmetric, and its eigenvalues are not real.)

(d) False. $A = I_2$ is symmetric, and every nonzero vector is an eigenvector, so $(1, 0), (1, 1)$ is an eigenbasis. It is not orthogonal. (What is true: eigenvectors for *different* eigenvalues are orthogonal.)

(e) True. Choose an orthonormal basis of each eigenspace (Theorem 6.4.20). Since A is diagonalizable, the eigenspaces together contain a basis, so their dimensions add up to n (Theorem 7.4.16), and the concatenation is an orthonormal list of n eigenvectors, because vectors from different eigenspaces are orthogonal. It is an orthonormal eigenbasis, so A is symmetric by Proposition 8.2.2(b) (or Theorem 8.2.12).

Solution to Exercise 8.2.35

Uses. Example 8.2.3 (Which matrices have orthonormal eigenbases), Lemma 8.2.10 (Self-adjoint operators have eigenvalues), Theorem 8.2.12 (Real spectral theorem)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Every step after the first one is the second proof of Theorem 8.2.5, which works over $\mathbb{R}$ as well. Look at the first step.
Hint 2. Does every operator on $\mathbb{R}^2$ have an eigenvector?
Sketch. The proof assumes that T has an eigenvector. Over $\mathbb{R}$ it need not: the rotation J is normal ($J^{\mathsf{T}}J = I = JJ^{\mathsf{T}}$) and has no real eigenvalue. If T is self-adjoint, Lemma 8.2.10 supplies the eigenvector, and the proof becomes that of Theorem 8.2.12.

The error is in the second sentence: "let $\mathbf{u}$ be a unit eigenvector of T". Over $\mathbb{R}$ an operator need not have any eigenvector. The rest of the argument is correct over $\mathbb{R}$: the computation $\|(T-\lambda\,\mathrm{id})\mathbf{u}\| = \|(T^*-\lambda\,\mathrm{id})\mathbf{u}\|$ for a real λ, the invariance of W, and the normality of $T{\restriction_W}$ go through as in the second proof of Theorem 8.2.5.

Counterexample: $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ satisfies $J^{\mathsf{T}}J = I_2 = JJ^{\mathsf{T}}$, so T_J is normal on $\mathbb{R}^2$, but $p_J = x^2+1$ has no real root, so T_J has no eigenvector, let alone an orthonormal eigenbasis (Example 8.2.3).

Repair: assume T is self-adjoint. Then Lemma 8.2.10 provides the eigenvector at every stage (restrictions of self-adjoint operators to invariant subspaces are self-adjoint), and the argument becomes the first proof of Theorem 8.2.12.

Solution to Exercise 8.2.36

Uses. Theorem 8.2.18 (Spectral decomposition), Proposition 8.2.20 (Spectral projections are polynomials in T), Theorem 8.2.12 (Real spectral theorem), Corollary 8.2.7 (Hermitian matrices), Definition 6.6.26 (Self-adjoint operator)
Techniques. Look for invariant subspaces among kernels and images of commuting operators (Technique 7.2.31), Glue bases of the summands (Technique 4.4.14)
Hint 1. The easy direction: in a common orthonormal eigenbasis both matrices are diagonal.
Hint 2. For the other, S maps each eigenspace E_{λ_j} of T into itself (Proposition 8.2.20(a)). Is $S{\restriction_{E_{\lambda_j}}}$ self-adjoint?
Hint 3. Apply the spectral theorem to $S{\restriction_{E_{\lambda_j}}}$ on each E_{λ_j}, and glue the bases.
Sketch. If both are diagonal in an orthonormal basis $\mathcal{O}$, their matrices commute, so $ST = TS$. Conversely, $V = \oplus E_{\lambda_j}$ for the eigenvalues of T, and S maps each E_{λ_j} into itself; $S{\restriction_{E_{\lambda_j}}}$ is self-adjoint, so E_{λ_j} has an orthonormal eigenbasis for it. These vectors are eigenvectors of T (with eigenvalue λ_j) and of S; the eigenspaces are orthogonal, so the concatenation is orthonormal.

If $\mathcal{O}$ is an orthonormal basis of common eigenvectors, then $[S]_{\mathcal{O}}$ and $[T]_{\mathcal{O}}$ are diagonal, so they commute, and $[ST]_{\mathcal{O}} = [S]_{\mathcal{O}}[T]_{\mathcal{O}} = [T]_{\mathcal{O}}[S]_{\mathcal{O}} = [TS]_{\mathcal{O}}$. Hence $ST = TS$.

Conversely, suppose $ST = TS$. A self-adjoint operator satisfies the spectral hypothesis (over $\mathbb{C}$ it is normal), so by Theorem 8.2.18, $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ with pairwise orthogonal eigenspaces of T. By Proposition 8.2.20(a), S maps each E_{λ_j} into itself. The restriction $S_j = S{\restriction_{E_{\lambda_j}}} \in \mathrm{L}\left(E_{\lambda_j}\right)$ is self-adjoint, because $\langle S_j\mathbf{v},\mathbf{w}\rangle = \langle S\mathbf{v},\mathbf{w}\rangle = \langle \mathbf{v}, S\mathbf{w}\rangle = \langle \mathbf{v}, S_j\mathbf{w}\rangle$ for $\mathbf{v},\mathbf{w} \in E_{\lambda_j}$. By Theorem 8.2.12 or Corollary 8.2.7, E_{λ_j} has an orthonormal basis $\mathcal{O}_j$ of eigenvectors of S_j, which are eigenvectors of S; they lie in E_{λ_j}, so they are eigenvectors of T too. The concatenation $\mathcal{O}_1, \ldots, \mathcal{O}_k$ is orthonormal, since the eigenspaces are pairwise orthogonal, and it has $\sum_j \dim E_{\lambda_j} = \dim V$ vectors. So it is an orthonormal basis of common eigenvectors.

Solution to Exercise 8.2.37

Uses. Proposition 8.2.20 (Spectral projections are polynomials in T), Proposition 7.2.27 (Evaluating polynomials at an operator), Proposition 6.6.18 (Properties of the adjoint), Definition 8.1.8 (Normal operator, normal matrix)
Hint 1. By Proposition 8.2.20(b), T^* is a polynomial in T.
Hint 2. For (b), try $T = S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.
Hint 3. For (c), apply (a) twice, and take adjoints of $ST = TS$.
Sketch. (a) $T^* = p(T)$, and S commutes with every polynomial in T. (b) $T = S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$: $ST^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = T^*S$. (c) $ST^* = T^*S$ and (by symmetry) $TS^* = S^*T$; taking adjoints of $ST = TS$ gives $S^*T^* = T^*S^*$; normality gives the rest. Then $(ST)(ST)^* = STT^*S^*$ and $(ST)^*(ST) = T^*S^*ST$ are products of the same commuting operators.

(a) By Proposition 8.2.20(b), $T^* = p(T)$ for some $p \in \mathbb{C}[x]$. Since S commutes with T, it commutes with T^m for every m (induction on m), hence with $p(T) = T^*$.

(b) Let $T = S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, which commute. Then $T^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and

$$ST^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = T^*S.$$

(This T is not normal, since $TT^* \neq T^*T$ by the same computation.)

(c) By (a), S commutes with T^*; by (a) with the roles of S and T exchanged (now S is normal and T commutes with it), T commutes with S^*. Taking adjoints of $ST = TS$ gives $T^*S^* = S^*T^*$ (Proposition 6.6.18(d)). Finally $SS^* = S^*S$ and $TT^* = T^*T$ by normality. So the four operators pairwise commute. Then

$$(ST)(ST)^* = STT^*S^* \quad \text{and} \quad (ST)^*(ST) = T^*S^*ST$$

are products of the same four pairwise commuting operators, so they are equal, and ST is normal. Similarly $(S+T)(S+T)^* = SS^* + ST^* + TS^* + TT^*$ and $(S+T)^*(S+T) = S^*S + S^*T + T^*S + T^*T$ agree term by term.

Solution to Exercise 8.2.38

Uses. Theorem 8.2.18 (Spectral decomposition), Theorem 6.4.31 (The matrix of an orthogonal projection), Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. For a symmetric A, the eigenspace for 1 is forced to be $(1,2)^\perp$. So the spectral decomposition is determined.
Hint 2. For (b), drop the orthogonality: choose any other eigenvector for 1, such as $(1,0)$, and compute SDS^{-1}.
Sketch. (a) $E_3 = \operatorname{span}((1,2))$ and $E_1 = E_3^\perp$ are forced, so $A = 3P + (I - P)$ with $P = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$: $A = \frac{1}{5}\begin{pmatrix} 7 & 4 \\ 4 & 13 \end{pmatrix}$. (b) $S = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$, $D = \operatorname{diag}(3,1)$: $B = SDS^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$.

(a) Let A be such a matrix. By Theorem 8.2.18, $\mathbb{R}^2 = E_3 \oplus E_1$ with $E_3 \perp E_1$. Each eigenspace is at least a line, so both are lines; $E_3 = \operatorname{span}((1,2))$ and $E_1 = E_3^\perp = \operatorname{span}((2,-1))$. So $A = 3P_3 + 1P_1$ with $P_3 = \operatorname{Pr}_{E_3}$ and $P_1 = I_2 - P_3$. With $\mathbf{u} = \frac{1}{\sqrt{5}}(1,2)$, Theorem 6.4.31 gives $P_3 = \mathbf{u}\mathbf{u}^{\mathsf{T}} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, so

$$A = I_2 + 2P_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{2}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 7 & 4 \\ 4 & 13 \end{pmatrix}.$$

It is symmetric, and $A(1,2) = \frac{1}{5}(15,30) = 3(1,2)$, $A(2,-1) = \frac{1}{5}(10,-5) = (2,-1)$. The argument shows it is the only one.

(b) Take $(1,0)$ as eigenvector for 1, and let $S = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$ (columns: the eigenvectors for 3 and 1) and $D = \operatorname{diag}(3,1)$. Then $S^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}$ and

$$B = SDS^{-1} = \begin{pmatrix} 3 & 1 \\ 6 & 0 \end{pmatrix}\begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}.$$

Then $B(1,2) = (3,6)$ and $B(1,0) = (1,0)$, and B is not symmetric. Uniqueness in (a) came from the orthogonality of eigenspaces, which only symmetric matrices guarantee.

Solution to Exercise 8.2.39

Uses. Proposition 8.2.24 (Properties of $f(T)$), Proposition 7.2.30 (Kernels and images of commuting operators are invariant), Proposition 7.2.7 (Invariant lines are spanned by eigenvectors), Theorem 8.2.18 (Spectral decomposition)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. A has eigenvalues 16 and 4, with eigenvectors $(1,1)$ and $(1,-1)$.
Hint 2. If $B^2 = A$, then B commutes with A. What does B do to the eigenspaces of A?

Hint 3. Each eigenspace of A is a line mapped into itself by B, so it consists of eigenvectors of B. For (c), every reflection squares to I_2.
Sketch. (a) $P = \frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}$, $Q = \frac{1}{2}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}$, $A = 16P + 4Q$, $\sqrt{A} = 4P + 2Q = \begin{pmatrix}3&1\\1&3\end{pmatrix}$. (b) B commutes with A, so preserves the lines E_{16} and E_4, so $B = \beta_1 P + \beta_2 Q$ with $\beta_1^2 = 16$ and $\beta_2^2 = 4$: four matrices $\pm\begin{pmatrix}3&1\\1&3\end{pmatrix}$, $\pm\begin{pmatrix}1&3\\3&1\end{pmatrix}$. (c) For I_2 the eigenspace is all of $\mathbb{R}^2$, which imposes no constraint; any reflection is a square root.

(a) $p_A = (x-10)^2 - 36 = (x-16)(x-4)$, with $E_{16} = \operatorname{span}((1,1))$ and $E_4 = \operatorname{span}((1,-1))$. With $P = \operatorname{Pr}_{E_{16}} = \frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}$ and $Q = \operatorname{Pr}_{E_4} = \frac{1}{2}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}$, we have $A = 16P + 4Q$, and

$$B = \sqrt{A} = 4P + 2Q = \begin{pmatrix}3&1\\1&3\end{pmatrix}.$$

Check: $\begin{pmatrix}3&1\\1&3\end{pmatrix}^2 = \begin{pmatrix}10&6\\6&10\end{pmatrix}$.

(b) Suppose $B^2 = A$. Then $BA = B^3 = AB$, so B maps $E_{16} = \ker(A - 16I_2)$ and $E_4 = \ker(A - 4I_2)$ into themselves (Proposition 7.2.30). These are lines, so by Proposition 7.2.7, $B(1,1) = \beta_1(1,1)$ and $B(1,-1) = \beta_2(1,-1)$ for some $\beta_1, \beta_2 \in \mathbb{R}$. Then $A(1,1) = B^2(1,1) = \beta_1^2(1,1)$, so $\beta_1^2 = 16$, and similarly $\beta_2^2 = 4$. Since $(1,1), (1,-1)$ is a basis, $B = \beta_1 P + \beta_2 Q$, with $\beta_1 = \pm 4$ and $\beta_2 = \pm 2$. Conversely each of these four matrices squares to $16P + 4Q = A$ by Proposition 8.2.24. They are

$$\begin{pmatrix}3&1\\1&3\end{pmatrix}, \quad \begin{pmatrix}-3&-1\\-1&-3\end{pmatrix}, \quad \begin{pmatrix}1&3\\3&1\end{pmatrix}, \quad \begin{pmatrix}-1&-3\\-3&-1\end{pmatrix}.$$

(c) For I_2 there is a single eigenspace, $\mathbb{R}^2$ itself, and the argument of (b) gives no information. Every reflection $R_\theta = \begin{pmatrix}\cos\theta&\sin\theta\\\sin\theta&-\cos\theta\end{pmatrix}$ satisfies $R_\theta^2 = I_2$ (it is $\operatorname{Pr}_\ell - \operatorname{Pr}_{\ell^\perp}$ for a line ℓ), and these are infinitely many distinct square roots. In (b), the distinct eigenvalues of A forced every square root to have the same eigenvectors as A.

Solution to Exercise 8.2.40

Uses. Definition 8.2.23 (Functions of an operator), Example 8.2.27 (The exponential), Example 8.2.3 (Which matrices have orthonormal eigenbases), Theorem 6.4.31 (The matrix of an orthogonal projection)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. A has eigenvalues ± 1 with eigenvectors $(1, \pm 1)$. Write down P_1 and P_{-1}.
Hint 2. J has eigenvalues $\pm i$ with eigenvectors $(1, \mp i)$. Use $e^{it} + e^{-it} = 2\cos t$ and $e^{it} - e^{-it} = 2i\sin t$.
Sketch. $e^{tA} = e^t P_1 + e^{-t} P_{-1} = \begin{pmatrix}\cosh t&\sinh t\\\sinh t&\cosh t\end{pmatrix}$. $e^{tJ} = e^{it}P_i + e^{-it}P_{-i}$ with $P_{\pm i} = \frac{1}{2}\begin{pmatrix}1&\pm i\\\mp i&1\end{pmatrix}$, giving $\begin{pmatrix}\cos t&-\sin t\\\sin t&\cos t\end{pmatrix}$: rotation by t.

For A: $E_1 = \operatorname{span}((1,1))$ and $E_{-1} = \operatorname{span}((1,-1))$, so $P_1 = \frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}$ and $P_{-1} = \frac{1}{2}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}$, and

$$e^{tA} = e^t P_1 + e^{-t} P_{-1} = \frac{1}{2}\begin{pmatrix}e^t + e^{-t}&e^t - e^{-t}\\e^t - e^{-t}&e^t + e^{-t}\end{pmatrix} = \begin{pmatrix}\cosh t&\sinh t\\\sinh t&\cosh t\end{pmatrix}.$$

For J: by Example 8.2.3, $E_i = \operatorname{span}((1,-i))$ and $E_{-i} = \operatorname{span}((1,i))$. With $\mathbf{u} = \frac{1}{\sqrt{2}}(1,-i)$, $P_i = \mathbf{u}\mathbf{u}^* = \frac{1}{2}\begin{pmatrix}1\\-i\end{pmatrix}(1\ \ i) = \frac{1}{2}\begin{pmatrix}1&i\\-i&1\end{pmatrix}$, and similarly $P_{-i} = \frac{1}{2}\begin{pmatrix}1&-i\\i&1\end{pmatrix}$. So

$$e^{tJ} = e^{it}P_i + e^{-it}P_{-i} = \frac{1}{2}\begin{pmatrix}e^{it} + e^{-it}&i\left(e^{it} - e^{-it}\right)\\-i\left(e^{it} - e^{-it}\right)&e^{it} + e^{-it}\end{pmatrix} = \begin{pmatrix}\cos t&-\sin t\\\sin t&\cos t\end{pmatrix},$$

using $e^{it} + e^{-it} = 2\cos t$ and $i\left(e^{it} - e^{-it}\right) = i \cdot 2i\sin t = -2\sin t$. So e^{tJ} is the rotation by the angle t: the solutions of $\mathbf{x}' = J\mathbf{x}$, that is $x_1' = -x_2$ and $x_2' = x_1$, run around circles at unit angular speed. The complex projections were only a tool; the answer is real.

Solution to Exercise 8.2.41

Uses. Corollary 8.2.28 (Principal axes), Remark 8.2.30 (Classifying conics)
Hint 1. Write each left side as $\mathbf{x}^{\mathsf{T}} A\mathbf{x}$, splitting the cross term evenly.
Hint 2. All three matrices have eigenvectors $(1,1)$ and $(1,-1)$.

Sketch. (a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, eigenvalues $3, -1$: $3y_1^2 - y_2^2 = 1$, a hyperbola with vertices on the line $y = x$. (b) $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, eigenvalues $4, 2$: $y_1^2/2 + y_2^2/4 = 1$, an ellipse with semi-axes $\sqrt{2}$ along $(1,1)$ and 2 along $(1,-1)$. (c) $A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, eigenvalues $4, 0$: $4y_1^2 = 1$, the lines $x + y = \pm 1/\sqrt{2}$.

In each case let $\mathbf{q}_1 = \frac{1}{\sqrt{2}}(1,1)$ and $\mathbf{q}_2 = \frac{1}{\sqrt{2}}(1,-1)$, and $y_1 = (x+y)/\sqrt{2}$, $y_2 = (x-y)/\sqrt{2}$. Every symmetric matrix $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$ has $A\mathbf{q}_1 = (a+b)\mathbf{q}_1$ and $A\mathbf{q}_2 = (a-b)\mathbf{q}_2$, so Corollary 8.2.28 applies with these coordinates. The principal axes are the lines $y = x$ and $y = -x$ in all three cases.

(a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, eigenvalues 3 and -1. The curve is $3y_1^2 - y_2^2 = 1$: a hyperbola, with vertices at $y_1 = \pm 1/\sqrt{3}$, $y_2 = 0$, that is, at $\pm\frac{1}{\sqrt{6}}(1,1)$ on the line $y = x$.

(b) $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, eigenvalues 4 and 2. The curve is $4y_1^2 + 2y_2^2 = 8$, that is, $y_1^2/2 + y_2^2/4 = 1$: an ellipse with semi-axis $\sqrt{2}$ along $y = x$ and semi-axis 2 along $y = -x$.

(c) $A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, eigenvalues 4 and 0. The curve is $4y_1^2 = 1$, that is, $y_1 = \pm 1/2$: the two parallel lines $x + y = \pm 1/\sqrt{2}$. (Directly: $2(x+y)^2 = 1$.)

Solution to Exercise 8.2.42

Uses. Theorem 8.2.12 (Real spectral theorem), Corollary 8.2.7 (Hermitian matrices), Proposition 6.3.18 (Coordinates in an orthonormal basis), Theorem 6.3.22 (Parseval's identity)
Techniques. Pass to orthonormal coordinates (Technique 6.7.2)
Hint 1. Expand $\mathbf{v} = \sum c_i \mathbf{u}_i$ in an orthonormal eigenbasis. Then $\langle T\mathbf{v}, \mathbf{v}\rangle = \sum \mu_i |c_i|^2$ and $\|\mathbf{v}\|^2 = \sum |c_i|^2$.
Hint 2. For (b), take $\mathbf{v} = \mathbf{e}_i$. For (c), find a vector $\mathbf{v}$ with $\langle A\mathbf{v}, \mathbf{v}\rangle > 3\|\mathbf{v}\|^2$.
Sketch. (a) $\langle T\mathbf{v}, \mathbf{v}\rangle = \sum_i \mu_i |c_i|^2$ is a weighted average of the eigenvalues times $\sum |c_i|^2 = \|\mathbf{v}\|^2$; equality on the right iff $c_i = 0$ whenever $\mu_i < \lambda_{\max}$. (b) $\mathbf{e}_i^{\mathsf{T}} A \mathbf{e}_i = a_{ii}$. (c) $\mathbf{v} = (1,1,1)$: $\mathbf{v}^{\mathsf{T}} A \mathbf{v} = 3 + 2 \cdot (2+2) = 11 > 9$.

(a) Let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be an orthonormal eigenbasis with $T\mathbf{u}_i = \mu_i \mathbf{u}_i$ (Theorem 8.2.12 or Corollary 8.2.7), and write $\mathbf{v} = \sum_i c_i \mathbf{u}_i$. Then $T\mathbf{v} = \sum_i \mu_i c_i \mathbf{u}_i$, and by orthonormality

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \sum_i \mu_i c_i \overline{c_i} = \sum_i \mu_i |c_i|^2, \quad \|\mathbf{v}\|^2 = \sum_i |c_i|^2$$

(Theorem 6.3.22). Since $\lambda_{\min} \le \mu_i \le \lambda_{\max}$ and $|c_i|^2 \ge 0$,

$$\lambda_{\min} \sum_i |c_i|^2 \le \sum_i \mu_i |c_i|^2 \le \lambda_{\max} \sum_i |c_i|^2.$$

Equality on the right means $\sum_i (\lambda_{\max} - \mu_i)|c_i|^2 = 0$, a sum of nonnegative terms, so $c_i = 0$ whenever $\mu_i \ne \lambda_{\max}$; that is, $\mathbf{v} \in \operatorname{span}\{\mathbf{u}_i \mid \mu_i = \lambda_{\max}\} = E_{\lambda_{\max}}$ (as in the proof of Theorem 8.2.18). Conversely, $\mathbf{v} \in E_{\lambda_{\max}}$ gives $\langle T\mathbf{v}, \mathbf{v}\rangle = \lambda_{\max}\|\mathbf{v}\|^2$.

(b) Apply (a) to T_A on $\mathbb{R}^n$ with the dot product and $\mathbf{v} = \mathbf{e}_i$: $\langle A\mathbf{e}_i, \mathbf{e}_i\rangle = a_{ii}$ and $\|\mathbf{e}_i\| = 1$.

(c) With $\mathbf{v} = (1,1,1)$, $A\mathbf{v} = (3,5,3)$, so $\langle A\mathbf{v}, \mathbf{v}\rangle = 11$ while $\|\mathbf{v}\|^2 = 3$. By (a), $\lambda_{\max} \ge 11/3 > 3$. (Part (b) alone gives only $\lambda_{\max} \ge 1$. In fact $\lambda_{\max} = 1 + 2\sqrt{2} \approx 3.83$.)

Solution to Exercise 8.2.43

Uses. Theorem 8.2.12 (Real spectral theorem), Proposition 4.5.27 (Similarity invariants), Proposition 6.2.22 (The square of a sum), Proposition 7.4.23 (Powers and polynomials of a diagonalizable matrix)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Diagonalize: all three quantities can be read off from the eigenvalues $\mu_1, \dots, \mu_n$.
Hint 2. $\operatorname{rank} A$ is the number of nonzero μ_i, $\operatorname{tr} A = \sum \mu_i$, and $\operatorname{tr}(A^2) = \sum \mu_i^2 > 0$. Apply Cauchy–Schwarz to the nonzero μ_i.
Sketch. $A = QDQ^{\mathsf{T}}$; rank and trace are similarity invariants, and $A^2 = QD^2Q^{\mathsf{T}}$. With r nonzero eigenvalues, $(\sum \mu_i)^2 \le r \sum \mu_i^2$ by Cauchy–Schwarz (Proposition 6.2.22), with equality iff the nonzero μ_i are equal. Since $A \ne 0$, $\operatorname{tr}(A^2) > 0$.

By Theorem 8.2.12, $A = QDQ^{\mathsf{T}}$ with Q orthogonal and $D = \operatorname{diag}(\mu_1, \dots, \mu_n)$, and $A^2 = QD^2Q^{\mathsf{T}}$ (Proposition 7.4.23). Rank and trace are invariant under similarity (Proposition 4.5.27), so with r the number of nonzero μ_i,

$$\operatorname{rank} A = \operatorname{rank} D = r, \quad \operatorname{tr} A = \sum_i \mu_i, \quad \operatorname{tr}\left(A^2\right) = \sum_i \mu_i^2.$$

Since $A \neq 0$, $D \neq 0$, so $r \geq 1$ and $\operatorname{tr}\left(A^2\right) > 0$. The sum $\sum_i \mu_i$ only involves the r nonzero eigenvalues, and by Proposition 6.2.22 (Cauchy–Schwarz for the vectors $(1, \ldots, 1)$ and $(\mu_i)_{\mu_i \neq 0}$ in $\mathbb{R}^r$),

$$\left(\sum_i \mu_i\right)^2 \leq r \sum_{\mu_i \neq 0} \mu_i^2 = r \operatorname{tr}\left(A^2\right).$$

Dividing by $\operatorname{tr}\left(A^2\right) > 0$ gives the inequality. Equality holds in Cauchy–Schwarz if and only if $(\mu_i)_{\mu_i \neq 0}$ is a multiple of $(1, \ldots, 1)$, that is, if and only if all nonzero eigenvalues of A are equal. Equivalently, $A = cP$ for a scalar $c \neq 0$ and an orthogonal projection P (by Theorem 8.2.18).

Solution to Exercise 8.2.44

Uses. Theorem 8.2.5 (Complex spectral theorem), Exercise 6.6.38, Theorem 7.3.27 (Existence of eigenvalues over $\mathbb{C}$), Corollary 6.4.27 (Testing whether a subspace is everything)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9), Move an operator across the inner product (Technique 6.6.20)
Hint 1. One direction was shown inside the second proof of Theorem 8.2.5. For the other, imitate the third proof.
Hint 2. If every vector of E_λ is an eigenvector of T^*, then $T^*\restriction_{E_\lambda}$ is an operator all of whose nonzero vectors are eigenvectors. What kind of operator is that?
Hint 3. Show that $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$ for $\mathbf{u} \in E_\lambda$ by comparing $\langle T\mathbf{u}, \mathbf{u}\rangle$ and $\langle \mathbf{u}, T^*\mathbf{u}\rangle$. Then the third proof of Theorem 8.2.5 runs unchanged.
Sketch. ($\Rightarrow$) shown in the second proof of Theorem 8.2.5. ($\Leftarrow$) For a nonzero $\mathbf{u} \in E_\lambda$, $T^*\mathbf{u} = \nu\mathbf{u}$, and $\lambda\|\mathbf{u}\|^2 = \langle T\mathbf{u}, \mathbf{u}\rangle = \langle \mathbf{u}, T^*\mathbf{u}\rangle = \overline{\nu}\|\mathbf{u}\|^2$, so $\nu = \overline{\lambda}$. Now the third proof applies: $E = \sum E_\lambda$ is T^*-invariant, $E^\perp$ is T-invariant and contains no eigenvector, so $E = V$; eigenspaces are orthogonal. So T has an orthonormal eigenbasis and is normal.

($\Rightarrow$) If T is normal and $T\mathbf{u} = \lambda\mathbf{u}$, then $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$, as shown in the second proof of Theorem 8.2.5.

($\Leftarrow$) Let $\mathbf{u} \neq \mathbf{0}$ with $T\mathbf{u} = \lambda\mathbf{u}$. By hypothesis $T^*\mathbf{u} = \nu\mathbf{u}$ for some $\nu \in \mathbb{C}$, and

$$\lambda\|\mathbf{u}\|^2 = \langle T\mathbf{u}, \mathbf{u}\rangle = \langle \mathbf{u}, T^*\mathbf{u}\rangle = \langle \mathbf{u}, \nu\mathbf{u}\rangle = \overline{\nu}\|\mathbf{u}\|^2,$$

so $\nu = \overline{\lambda}$. Thus $T^*\mathbf{u} = \overline{\lambda}\mathbf{u}$ for every $\mathbf{u} \in E_\lambda$ (including $\mathbf{0}$). This is the only property of normal operators used in the third proof of Theorem 8.2.5, so that proof applies verbatim: $E = \sum_\lambda E_\lambda$ is T^*-invariant, so $E^\perp$ is T-invariant (Exercise 6.6.38); an eigenvector of $T\restriction_{E^\perp}$ would lie in $E \cap E^\perp = \{\mathbf{0}\}$, so $E^\perp = \{\mathbf{0}\}$ and $E = V$ (Theorem 7.3.27, Corollary 6.4.27); and $\lambda\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, T^*\mathbf{v}\rangle = \mu\langle \mathbf{u}, \mathbf{v}\rangle$ for $\mathbf{u} \in E_\lambda$, $\mathbf{v} \in E_\mu$ shows that the eigenspaces are orthogonal. Concatenating orthonormal bases of the eigenspaces gives an orthonormal eigenbasis, so T is normal by Theorem 8.2.5.

Solution to Exercise 8.2.45

Uses. Definition 8.1.8 (Normal operator, normal matrix), Theorem 7.3.6 (Eigenvalues are the roots of the characteristic polynomial)
Techniques. Compute entry by entry (Technique 2.4.15)
Hint 1. Write $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and compare the entries of $A^{\mathsf{T}}A$ and AA^{T}.
Hint 2. The diagonal entries give $b^2 = c^2$. Split into the cases $c = b$ and $c = -b \neq 0$.
Sketch. $A^{\mathsf{T}}A = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}$ and $AA^{\mathsf{T}} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$. Equality means $b^2 = c^2$ and $ab + cd = ac + bd$. If $c = b$, A is symmetric. If $c = -b \neq 0$, the second equation gives $2b(a - d) = 0$, so $d = a$. (b) $p = (x - a)^2 + b^2$ has no real root for $b \neq 0$; the matrix is $\sqrt{a^2 + b^2}$ times a rotation.

(a) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

$$A^{\mathsf{T}}A = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}, \quad AA^{\mathsf{T}} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}.$$

So A is normal if and only if $b^2 = c^2$ and $ab + cd = ac + bd$. If $c = b$, the second equation holds and A is symmetric. Otherwise $c = -b$ with $b \neq 0$, and the second equation becomes $ab - bd = -ab + bd$, that is,

$2b(a-d)=0$, so $d=a$ and $A=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, which is of the stated form with b replaced by $-b$. Conversely, symmetric matrices are normal, and for $A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ both products equal $\left(a^2+b^2\right)I_2$.

(b) If A is normal and not symmetric, then $A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b\neq 0$, and $p_A=(x-a)^2+b^2>0$ for every real x. So A has no real eigenvalue (Theorem 7.3.6). Writing $r=\sqrt{a^2+b^2}>0$ and $(a,b)=r(\cos\theta,\sin\theta)$, we get $A=r\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$: a rotation by θ followed by scaling by r. It is multiplication by $a+bi$ on $\mathbb{C}=\mathbb{R}^2$.

Solution to Exercise 8.2.46

Uses. Lemma 8.2.10 (Self-adjoint operators have eigenvalues), Exercise 6.6.38, Corollary 6.4.27 (Testing whether a subspace is everything), Theorem 6.4.20 (Orthonormal bases exist), Definition 6.6.26 (Self-adjoint operator)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Let E be the sum of the eigenspaces of T. Why is $E^\perp$ invariant? (This is easier than in the complex case.)
Hint 2. To find an eigenvector in $E^\perp$, you need $T{\restriction}_{E^\perp}$ to be an operator that has one. Which lemma provides it?
Sketch. $E=\sum E_\lambda$ is T-invariant, and $T=T^*$, so $E^\perp$ is T-invariant (Exercise 6.6.38). The restriction $T{\restriction}_{E^\perp}$ is self-adjoint, so if $E^\perp\neq\{\mathbf{0}\}$ it has an eigenvector (Lemma 8.2.10), which would lie in $E\cap E^\perp$. So $E=V$. Orthogonality: $\lambda\langle\mathbf{u},\mathbf{v}\rangle=\langle T\mathbf{u},\mathbf{v}\rangle=\langle\mathbf{u},T\mathbf{v}\rangle=\mu\langle\mathbf{u},\mathbf{v}\rangle$.

Let T be self-adjoint, let $\lambda_1,\ldots,\lambda_k$ be its distinct eigenvalues ($k\geq 1$ by Lemma 8.2.10), and let $E=E_{\lambda_1}+\cdots+E_{\lambda_k}$. Each E_{λ_j} is T-invariant, so E is T-invariant. Since $T^*=T$, Exercise 6.6.38 shows that $E^\perp$ is T^*-invariant, that is, T-invariant. (This is the first use of self-adjointness; in the complex case we needed the fact that eigenvectors of T are eigenvectors of T^*.)

Suppose $E^\perp\neq\{\mathbf{0}\}$. The restriction $S=T{\restriction}_{E^\perp}$ is self-adjoint, since $\langle S\mathbf{v},\mathbf{w}\rangle=\langle T\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{v},T\mathbf{w}\rangle=\langle\mathbf{v},S\mathbf{w}\rangle$ for $\mathbf{v},\mathbf{w}\in E^\perp$. By Lemma 8.2.10 (the second use of self-adjointness, replacing the existence of eigenvalues over $\mathbb{C}$), S has an eigenvector $\mathbf{w}\in E^\perp$. It is an eigenvector of T, so $\mathbf{w}\in E\cap E^\perp=\{\mathbf{0}\}$, a contradiction. So $E^\perp=\{\mathbf{0}\}$ and $E=V$ by Corollary 6.4.27.

For $\mathbf{u}\in E_\lambda$ and $\mathbf{v}\in E_\mu$ with $\lambda\neq\mu$, $\lambda\langle\mathbf{u},\mathbf{v}\rangle=\langle T\mathbf{u},\mathbf{v}\rangle=\langle\mathbf{u},T\mathbf{v}\rangle=\mu\langle\mathbf{u},\mathbf{v}\rangle$ (the third use; the eigenvalues are real), so $\langle\mathbf{u},\mathbf{v}\rangle=0$. Concatenating orthonormal bases of the eigenspaces (Theorem 6.4.20) gives an orthonormal list spanning $E=V$, which is an orthonormal eigenbasis.

Solution to Exercise 8.2.47

Uses. Corollary 8.2.7 (Hermitian matrices), Theorem 8.2.5 (Complex spectral theorem), Proposition 8.2.24 (Properties of $f(T)$), Exercise 8.2.33
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. Write $H=UDU^*$ with D real diagonal. What are the eigenvalues of $I+iH$?
Hint 2. $C=f(H)$ for $f(\lambda)=(1-i\lambda)/(1+i\lambda)$. What is $|f(\lambda)|$ for real λ?
Hint 3. For (c), solve $z=(1-i\lambda)/(1+i\lambda)$ for λ, and apply that function to C, which is normal.
Sketch. (a) The eigenvalues of $I+iH$ are $1+i\lambda_j\neq 0$. (b) $C=Uf(D)U^*$ with $f(\lambda)=(1-i\lambda)/(1+i\lambda)$; for real λ, $|f(\lambda)|=1$ and $f(\lambda)\neq -1$; so $C^*C=I$. (c) C is normal with eigenvalues z on the unit circle, $z\neq -1$; $g(z)=i(z-1)/(z+1)$ is real there and inverts f, so $H=g(C)$ is Hermitian with $f(H)=C$. Uniqueness: $H=g(f(H))$.

By Corollary 8.2.7, $H=UDU^*$ with U unitary and $D=\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$ real.

(a) $I_n+iH=U(I_n+iD)U^*$, and I_n+iD is diagonal with entries $1+i\lambda_j$, which are nonzero because λ_j is real. So I_n+iH is a product of invertible matrices.

(b) Let $f(\lambda)=(1-i\lambda)/(1+i\lambda)$ for $\lambda\in\mathbb{R}$. With the matrix form of Definition 8.2.23, $(I_n-iH)(I_n+iH)^{-1}=U(I_n-iD)(I_n+iD)^{-1}U^*=Uf(D)U^*=f(H)$. For real λ, $1-i\lambda=\overline{1+i\lambda}$, so $|f(\lambda)|=1$; and $f(\lambda)=-1$ would mean $1-i\lambda=-1-i\lambda$, which is impossible. So $C=Uf(D)U^*$ with $f(D)$ diagonal, unimodular entries, and $C^*C=U\overline{f(D)}f(D)U^*=I_n$: C is unitary, and its eigenvalues $f\left(\lambda_j\right)$ are not -1.

(c) Let C be unitary with no eigenvalue -1. Then $C^*C=I_n=CC^*$, so C is normal, and by Theorem 8.2.5, $C=VZV^*$ with V unitary and $Z=\operatorname{diag}(z_1,\ldots,z_n)$; $\left|z_j\right|=1$ by Exercise 8.2.33(d), and $z_j\neq -1$. Let $g(z)=i(z-1)/(z+1)$. For $|z|=1$, $\overline{z}=1/z$, so

$$\overline{g(z)} = -i\frac{1/z - 1}{1/z + 1} = -i\frac{1 - z}{1 + z} = g(z),$$

and $g(z)$ is real. Solving $z = (1 - i\lambda)/(1 + i\lambda)$ for λ gives $z + i\lambda z = 1 - i\lambda$, so $\lambda = i(z - 1)/(z + 1) = g(z)$: thus $f(g(z)) = z$ and $g(f(\lambda)) = \lambda$. Let $H = g(C) = Vg(Z)V^*$. It is Hermitian since $g(Z)$ is real diagonal, and $f(H) = Vf(g(Z))V^* = VZV^* = C$. For uniqueness, if $f(H') = C$ with H' Hermitian, $H' = U'D'U'^*$, then $C = U'f(D')U'^*$, and $g(C) = U'g(f(D'))U'^* = H'$ (the matrix $f(D')$ is a diagonalization of C, and $g(C)$ does not depend on the choice). So $H' = g(C) = H$.

Solution to Exercise 8.2.48

Uses. Theorem 8.2.5 (Complex spectral theorem), Exercise 8.2.45, Proposition 7.3.34 (Eigenvalues of a complexification), Exercise 6.6.38, Theorem 6.4.20 (Orthonormal bases exist), Theorem 4.5.17 (Change of basis)
Techniques. Complexify, then take real and imaginary parts (Technique 7.3.35), Induction on dimension via an invariant subspace (Technique 7.5.12)
Hint 1. Induct on n. If A has a real eigenvalue, split off a line as in the second proof of Theorem 8.2.5; it works over $\mathbb{R}$ for a real eigenvector.
Hint 2. Otherwise take a complex eigenvector $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ of A for $\lambda = a + bi$, $b \neq 0$. Since A is normal as a complex matrix, $A^{\mathsf{T}}\mathbf{z} = A^*\mathbf{z} = \overline{\lambda}\mathbf{z}$. Take real and imaginary parts.
Hint 3. $U = \operatorname{span}(\mathbf{x}, \mathbf{y})$ is a plane invariant under A and A^{T}, so $U^\perp$ is too. The restriction of T_A to U, in an orthonormal basis of U, is a real normal 2×2 matrix without real eigenvalues. Use Exercise 8.2.45.
Sketch. Induction on n, splitting off an invariant subspace U of dimension 1 or 2 that is invariant under both A and A^{T}; then $U^\perp$ is invariant under both, and both restrictions are normal. A real eigenvector gives U of dimension 1. Otherwise a complex eigenvector $\mathbf{x} + i\mathbf{y}$ for $a + bi$ ($b \neq 0$) is also an eigenvector of $A^* = A^{\mathsf{T}}$ for $a - bi$, and comparing real and imaginary parts shows $U = \operatorname{span}(\mathbf{x}, \mathbf{y})$ is invariant under A and A^{T}. In an orthonormal basis of U, the restriction is normal with no real eigenvalue, hence of the form $\left(\begin{smallmatrix} a' & -b' \\ b' & a' \end{smallmatrix}\right)$ (Exercise 8.2.45); replacing the second basis vector by its negative makes $b' > 0$.

We prove the statement for normal operators T on finite-dimensional real inner product spaces V (with "orthonormal basis" in place of Q), by strong induction on $\dim V$; for A apply it to T_A on $\mathbb{R}^n$ and let Q have the basis as columns (Theorem 4.5.17). Dimension 0 is trivial. Suppose $\dim V = n \geq 1$. It suffices to find a subspace $U \neq \{\mathbf{0}\}$ of dimension 1 or 2 that is invariant under T and T^*, and an orthonormal basis of U in which $T{\restriction}_U$ has a block of the required form. Indeed, then $W = U^\perp$ is invariant under T and T^* (Exercise 6.6.38 applied to T and to T^*), $T{\restriction}_W$ is normal with adjoint $T^*{\restriction}_W$ (as in the second proof of Theorem 8.2.5), and the induction hypothesis applies to it; the two orthonormal bases together give the required basis of $V = U \oplus W$.

Case 1: T has a real eigenvalue λ. Let $\mathbf{u}$ be a unit eigenvector. The computation $\|(T - \lambda\,\mathrm{id})\mathbf{v}\| = \|(T^* - \lambda\,\mathrm{id})\mathbf{v}\|$ from the second proof of Theorem 8.2.5 works over $\mathbb{R}$, so $T^*\mathbf{u} = \lambda\mathbf{u}$, and $U = \operatorname{span}(\mathbf{u})$ is invariant under T and T^*, with 1×1 block (λ).

Case 2: T has no real eigenvalue. Choose an orthonormal basis $\mathcal{O}$ of V and let $M = [T]_{\mathcal{O}} \in \mathbb{R}^{n \times n}$; then $M^{\mathsf{T}} = [T^*]_{\mathcal{O}}$ and M is normal. As a complex matrix, $M^* = M^{\mathsf{T}}$, so M is normal in $\mathbb{C}^{n \times n}$, and it has an eigenvalue $\lambda = a + bi$ with eigenvector $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$; $b \neq 0$ by assumption. By the second proof of Theorem 8.2.5, $M^{\mathsf{T}}\mathbf{z} = M^*\mathbf{z} = \overline{\lambda}\mathbf{z}$. Comparing real and imaginary parts of $M\mathbf{z} = \lambda\mathbf{z}$ and $M^{\mathsf{T}}\mathbf{z} = \overline{\lambda}\mathbf{z}$ (as in Proposition 7.3.34),

$$M\mathbf{x} = a\mathbf{x} - b\mathbf{y}, \quad M\mathbf{y} = b\mathbf{x} + a\mathbf{y}, \quad M^{\mathsf{T}}\mathbf{x} = a\mathbf{x} + b\mathbf{y}, \quad M^{\mathsf{T}}\mathbf{y} = -b\mathbf{x} + a\mathbf{y}.$$

So $U_0 = \operatorname{span}(\mathbf{x}, \mathbf{y})$ is invariant under M and M^{T}, and it has dimension 2 by Proposition 7.3.34(c). Let $U \subseteq V$ be the corresponding subspace (the vectors whose $\mathcal{O}$-coordinates lie in U_0); it is invariant under T and T^*. Let $\mathbf{f}_1, \mathbf{f}_2$ be an orthonormal basis of U and $N = [T{\restriction}_U]_{\mathbf{f}_1, \mathbf{f}_2} \in \mathbb{R}^{2\times 2}$. Since $T{\restriction}_U$ is normal with adjoint $T^*{\restriction}_U$, N is normal (Theorem 6.6.15). It has no real eigenvalue, since T has none, so by Exercise 8.2.45, $N = \left(\begin{smallmatrix} a' & -b' \\ b' & a' \end{smallmatrix}\right)$ with $b' \neq 0$. If $b' < 0$, replace $\mathbf{f}_2$ by $-\mathbf{f}_2$, which changes the signs of the off-diagonal entries. This gives the required block.

Section 8.3

Solution to Exercise 8.3.7

Uses. Theorem 8.3.5 (Characterizations of positive operators), Definition 8.3.1 (Positive operator)
Hint 1. First check self-adjointness. For the symmetric ones, compute the eigenvalues from the trace and the determinant.
Hint 2. For (c), try to write $\mathbf{x}^* A\mathbf{x}$ as $|\cdots|^2$, or A as B^*B with B a 1×2 matrix.
Sketch. (a) Positive: eigenvalues $4, 2$. (b) Not positive: eigenvalues $4, -2$, and $\mathbf{x} = (1,-1)$ gives -4. (c) Positive: $A = B^*B$ for $B = \begin{pmatrix}1 & i\end{pmatrix}$, so $\mathbf{x}^*A\mathbf{x} = |x_1 + ix_2|^2$. (d) and (e) are not symmetric, so not positive, although $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = 2\|\mathbf{x}\|^2 \geq 0$ in (e).

(a) The matrix is symmetric with trace 6 and determinant 8, so its eigenvalues are 4 and 2 (the roots of $x^2 - 6x + 8$). Both are nonnegative, so it is positive by Theorem 8.3.5(II).

(b) The matrix is symmetric, but its eigenvalues are the roots of $x^2 - 2x - 8$, namely 4 and -2. So it is not positive; concretely, $\mathbf{x} = (1,-1)$ gives $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = 1 - 3 - 3 + 1 = -4 < 0$.

(c) Let $B = \begin{pmatrix}1 & i\end{pmatrix} \in \mathbb{C}^{1\times 2}$. Then

$$B^*B = \begin{pmatrix}1\\ -i\end{pmatrix}\begin{pmatrix}1 & i\end{pmatrix} = \begin{pmatrix}1 & i\\ -i & -i\cdot i\end{pmatrix} = \begin{pmatrix}1 & i\\ -i & 1\end{pmatrix},$$

so the matrix is positive by Theorem 8.3.5(VI). Explicitly, $\mathbf{x}^*A\mathbf{x} = |B\mathbf{x}|^2 = |x_1 + ix_2|^2 \geq 0$. Its eigenvalues are 0 and 2.

(d) The matrix is not symmetric, so it is not positive. (Also $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = x_1x_2$, which is -1 at $\mathbf{x} = (1,-1)$.)

(e) Here $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = 2x_1^2 + x_1x_2 - x_2x_1 + 2x_2^2 = 2\|\mathbf{x}\|^2 \geq 0$ for every $\mathbf{x} \in \mathbb{R}^2$, but the matrix is not symmetric, so it is not positive (Remark 8.3.4).

Solution to Exercise 8.3.12

Uses. Theorem 8.3.9 (The positive square root), Proposition 8.2.20 (Spectral projections are polynomials in T)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. The eigenvalues of A are 9 and 1, with eigenvectors $(1,-1)$ and $(1,1)$.
Hint 2. For (b), A is a positive square root of A^2, while B has eigenvalues 3 and -1.
Sketch. (a) $P_9 = \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}$, $P_1 = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$, $\sqrt{A} = 3P_9 + P_1 = \begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}$; $B^2 = \begin{pmatrix}5 & -4\\ -4 & 5\end{pmatrix}$. (b) A is positive and squares to A^2, so $\sqrt{A^2} = A$ by uniqueness. $B^2 = A$, so $\sqrt{B^2} = \sqrt{A} = \begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix} \neq B$; B is not positive.

(a) $p_A = (x-5)^2 - 16 = (x-9)(x-1)$. By Proposition 8.2.20,

$$P_9 = \frac{A - I_2}{8} = \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}, \quad P_1 = \frac{A - 9I_2}{-8} = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix},$$

so by Theorem 8.3.9,

$$\sqrt{A} = 3P_9 + 1P_1 = \frac{1}{2}\begin{pmatrix}3+1 & -3+1\\ -3+1 & 3+1\end{pmatrix} = \begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}.$$

Check: $\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}^2 = \begin{pmatrix}4+1 & -2-2\\ -2-2 & 1+4\end{pmatrix} = A$. Also $B^2 = \begin{pmatrix}1+4 & -2-2\\ -2-2 & 4+1\end{pmatrix} = A$.

(b) A is positive (its eigenvalues are 9 and 1) and $A^2 = A^2$, so A is a positive square root of A^2, and by uniqueness $\sqrt{A^2} = A$. On the other hand $B^2 = A$, so $\sqrt{B^2} = \sqrt{A} = \begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix} \neq B$. The difference is that B is not positive: $B = 3P_9 - P_1$ has the eigenvalue -1. In general $\sqrt{S^2} = S$ for positive S, and for a self-adjoint S the operator $\sqrt{S^2}$ replaces each eigenvalue of S by its absolute value (Exercise 8.3.37).

Solution to Exercise 8.3.18

Uses. Theorem 8.3.5 (Characterizations of positive operators), Proposition 8.3.15 (Characterizations of positive definite operators)
Hint 1. The eigenvalues of A_t are real. Their sum is 5 and their product is $4 - t^2$.
Hint 2. Alternatively, complete the square in $x^2 + 2txy + 4y^2$.

Sketch. Both eigenvalues are ≥ 0 if and only if their sum 5 and product $4 - t^2$ are ≥ 0: positive for $|t| \leq 2$. Positive definite iff both > 0, iff $|t| < 2$. Or: $x^2 + 2txy + 4y^2 = (x + ty)^2 + (4 - t^2)y^2$.

A_t is symmetric, so its eigenvalues λ_1, λ_2 are real, with $\lambda_1 + \lambda_2 = \operatorname{tr} A_t = 5$ and $\lambda_1\lambda_2 = \det A_t = 4 - t^2$. Two real numbers with a positive sum are both nonnegative exactly when their product is nonnegative, and both positive exactly when their product is positive. So by Theorem 8.3.5 and Proposition 8.3.15, A_t is positive if and only if $|t| \leq 2$, and positive definite if and only if $|t| < 2$.

The same answer comes from completing the square:

$$(x, y)A_t\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 2txy + 4y^2 = (x + ty)^2 + (4 - t^2)y^2.$$

If $|t| \leq 2$ this is ≥ 0, and if $|t| < 2$ it is 0 only when $y = 0$ and $x + ty = 0$, that is, at $(0, 0)$. If $|t| > 2$, then $(x, y) = (-t, 1)$ gives $4 - t^2 < 0$. At $t = \pm 2$, the vector $(-t, 1) \neq \mathbf{0}$ gives 0.

Solution to Exercise 8.3.28

Uses. Definition 8.3.1 (Positive operator), Proposition 6.6.18 (Properties of the adjoint)
Techniques. Move an operator across the inner product (Technique 6.6.20)
Hint 1. Check the two conditions of Definition 8.3.1 directly. For (b), move X across the inner product.
Hint 2. For (c), write $T^{2k} = \left(T^k\right)^* \mathrm{id}_V T^k$ and $T^{2k+1} = \left(T^k\right)^* TT^k$, and use (b).
Hint 3. For (d), substitute $\mathbf{v} = T\mathbf{w}$ in $\left\langle T^{-1}\mathbf{v}, \mathbf{v}\right\rangle$.
Sketch. (a) Sums and real multiples of self-adjoint operators are self-adjoint, and the forms add and scale. (b) $\langle X^*TX\mathbf{u}, \mathbf{u}\rangle = \langle TX\mathbf{u}, X\mathbf{u}\rangle \geq 0$. (c) T^k is self-adjoint, and $T^{2k} = \left(T^k\right)^* \mathrm{id}_V T^k$, $T^{2k+1} = \left(T^k\right)^* TT^k$ are positive by (b). (d) $\left(T^{-1}\right)^* = (T^*)^{-1} = T^{-1}$ and $\left\langle T^{-1}\mathbf{v}, \mathbf{v}\right\rangle = \langle \mathbf{w}, T\mathbf{w}\rangle \geq 0$ for $\mathbf{w} = T^{-1}\mathbf{v}$.

(a) By Proposition 6.6.18, $(S + T)^* = S + T$ and $(cT)^* = \bar{c}T = cT$, as c is real. For every $\mathbf{v}$, $\langle (S + T)\mathbf{v}, \mathbf{v}\rangle = \langle S\mathbf{v}, \mathbf{v}\rangle + \langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$ and $\langle cT\mathbf{v}, \mathbf{v}\rangle = c\langle T\mathbf{v}, \mathbf{v}\rangle \geq 0$.

(b) $(X^*TX)^* = X^*T^*X^{**} = X^*TX$ by Proposition 6.6.18, and for $\mathbf{u} \in U$, $\langle X^*TX\mathbf{u}, \mathbf{u}\rangle = \langle TX\mathbf{u}, X\mathbf{u}\rangle \geq 0$.

(c) For $m = 0$, $T^0 = \mathrm{id}_V$ is positive. By Proposition 6.6.18(d) and induction, $\left(T^k\right)^* = (T^*)^k = T^k$. If $m = 2k$, then $T^m = \left(T^k\right)^* \mathrm{id}_V T^k$, which is positive by (b) applied to the positive operator id_V and $X = T^k$. If $m = 2k + 1$, then $T^m = \left(T^k\right)^* TT^k$, positive by (b) with $X = T^k$. (Alternatively, $T^m = \sum_j \lambda_j^m P_j$ has the nonnegative eigenvalues λ_j^m, by Proposition 8.2.24.)

(d) By Proposition 6.6.18(f), $\left(T^{-1}\right)^* = (T^*)^{-1} = T^{-1}$. For $\mathbf{v} \in V$, let $\mathbf{w} = T^{-1}\mathbf{v}$. Then

$$\left\langle T^{-1}\mathbf{v}, \mathbf{v}\right\rangle = \langle \mathbf{w}, T\mathbf{w}\rangle = \overline{\langle T\mathbf{w}, \mathbf{w}\rangle} = \langle T\mathbf{w}, \mathbf{w}\rangle \geq 0,$$

since $\langle T\mathbf{w}, \mathbf{w}\rangle$ is a nonnegative real number.

Solution to Exercise 8.3.29

Uses. Theorem 8.3.5 (Characterizations of positive operators), Definition 8.3.1 (Positive operator), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. For (a), the diagonal entries are nonnegative in orthonormal bases. Try $T = \mathrm{diag}(0, 1)$ on $\mathbb{R}^2$ and a basis that is far from orthogonal.
Hint 2. For (b) and (c), compare the eigenvalues of T with those of T^3 and T^2.
Hint 3. For (d), try a projection and a matrix of ones in $\mathbb{R}^{2\times 2}$. For (e), real test vectors cannot see an imaginary part: try $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$.
Sketch. (a) False: $T = \mathrm{diag}(0, 1)$ and $\mathcal{B} = (1, 1), (1, 2)$ give $[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix}$. (b) True: an eigenvalue λ of T gives the eigenvalue $\lambda^3 \geq 0$ of T^3, so $\lambda \geq 0$. (c) False: $T = -\mathrm{id}_V$. (d) False: $S = \mathrm{diag}(1, 0)$, $T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ give $\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$, with determinant -1. (e) False: $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ vanishes on real vectors but has the eigenvalue -1.

(a) False. Let $T = T_A$ on $\mathbb{R}^2$ with $A = \mathrm{diag}(0, 1)$, which is positive, and let $\mathcal{B} = \mathbf{b}_1, \mathbf{b}_2$ with $\mathbf{b}_1 = (1, 1)$, $\mathbf{b}_2 = (1, 2)$. Then $T\mathbf{b}_1 = (0, 1) = -\mathbf{b}_1 + \mathbf{b}_2$ and $T\mathbf{b}_2 = (0, 2) = -2\mathbf{b}_1 + 2\mathbf{b}_2$, so

$$[T]_{\mathcal{B}\mathcal{B}} = \begin{pmatrix} -1 & -2 \\ 1 & 2 \end{pmatrix},$$

whose $(1,1)$ entry is negative. For an orthonormal basis the statement is true, since the diagonal entries are $\langle T\mathbf{e}_i, \mathbf{e}_i\rangle \geq 0$.

(b) True. Let λ be an eigenvalue of T, with $T\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$; λ is real by Proposition 8.1.1. Then $T^3\mathbf{v} = \lambda^3\mathbf{v}$, so λ^3 is an eigenvalue of the positive operator T^3, and $\lambda^3 \geq 0$ by Theorem 8.3.5. For a real number this forces $\lambda \geq 0$. So T is self-adjoint with nonnegative eigenvalues, hence positive by Theorem 8.3.5(II).

(c) False. If $\dim V \geq 1$, then $T = -\mathrm{id}_V$ is self-adjoint and $T^2 = \mathrm{id}_V$ is positive, but $\langle T\mathbf{v}, \mathbf{v}\rangle = -\|\mathbf{v}\|^2 < 0$ for $\mathbf{v} \neq \mathbf{0}$. Here the argument of (b) breaks down because $\lambda^2 \geq 0$ does not force $\lambda \geq 0$.

(d) False. Let $S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, both positive (a projection, and $T = \mathbf{u}\mathbf{u}^{\mathsf{T}}$ with $\mathbf{u} = (1,1)$). Then

$$ST + TS = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix},$$

and $\mathbf{x} = (1,-2)$ gives $\mathbf{x}^{\mathsf{T}}(ST+TS)\mathbf{x} = 2 - 2 - 2 + 0 = -2 < 0$.

(e) False. Let $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$. Then $A^* = A$, and for real $\mathbf{x}$, $\mathbf{x}^*A\mathbf{x} = ix_1x_2 - ix_2x_1 = 0$. But $p_A = x^2 - 1$, so A has the eigenvalue -1; for instance $\mathbf{x} = (1,i)$ gives $A\mathbf{x} = (-1,-i)$ and $\mathbf{x}^*A\mathbf{x} = -1 + (-i)(-i) = -2 < 0$. Positivity of a complex matrix must be tested on complex vectors.

Solution to Exercise 8.3.30

Uses. Proposition 6.6.30 (Self-adjoint means real values), Remark 8.3.4 (Why self-adjointness is part of the definition), Corollary 6.6.28 (Self-adjointness in an orthonormal basis)
Hint 1. For (ii) $\Rightarrow$ (i), over $\mathbb{C}$ nonnegative values force self-adjointness (Proposition 6.6.30). A real Hermitian matrix is symmetric.
Hint 2. For (i) $\Rightarrow$ (ii), write $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and expand $\mathbf{x}^*A\mathbf{x}$. Where does symmetry make the imaginary part vanish?
Hint 3. For the last part, use $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ of Non-example 8.3.3 and $\mathbf{x} = (1,i)$.
Sketch. (ii) $\Rightarrow$ (i): the values $\mathbf{x}^*A\mathbf{x}$ are real, so A is Hermitian, hence symmetric as it is real; restrict to real $\mathbf{x}$. (i) $\Rightarrow$ (ii): for $\mathbf{x} = \mathbf{u} + i\mathbf{v}$, $\mathbf{x}^*A\mathbf{x} = \mathbf{u}^{\mathsf{T}}A\mathbf{u} + \mathbf{v}^{\mathsf{T}}A\mathbf{v} + i\big(\mathbf{u}^{\mathsf{T}}A\mathbf{v} - \mathbf{v}^{\mathsf{T}}A\mathbf{u}\big)$, and the imaginary part is 0 by symmetry. For $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{x}^{\mathsf{T}}B\mathbf{x} = \|\mathbf{x}\|^2$ on $\mathbb{R}^2$ but $(1,i)^*B(1,i) = 2 - 2i$.

(ii) $\Rightarrow$ (i). All values $\mathbf{x}^*A\mathbf{x} = \langle T_A\mathbf{x}, \mathbf{x}\rangle$ on $\mathbb{C}^n$ are real, so T_A is self-adjoint on $\mathbb{C}^n$ by Proposition 6.6.30, and $A^* = A$ (Corollary 6.6.28). Since A is real, $A^* = A^{\mathsf{T}}$, so $A^{\mathsf{T}} = A$. Taking $\mathbf{x} \in \mathbb{R}^n$ in (ii) gives $\mathbf{x}^{\mathsf{T}}A\mathbf{x} \geq 0$.

(i) $\Rightarrow$ (ii). Let $\mathbf{x} = \mathbf{u} + i\mathbf{v}$ with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{x}^* = \mathbf{u}^{\mathsf{T}} - i\mathbf{v}^{\mathsf{T}}$, and since A is real,

$$\mathbf{x}^*A\mathbf{x} = \mathbf{u}^{\mathsf{T}}A\mathbf{u} + \mathbf{v}^{\mathsf{T}}A\mathbf{v} + i\big(\mathbf{u}^{\mathsf{T}}A\mathbf{v} - \mathbf{v}^{\mathsf{T}}A\mathbf{u}\big).$$

The number $\mathbf{v}^{\mathsf{T}}A\mathbf{u}$ is a 1×1 matrix, so it equals its transpose $\mathbf{u}^{\mathsf{T}}A^{\mathsf{T}}\mathbf{v} = \mathbf{u}^{\mathsf{T}}A\mathbf{v}$. So the imaginary part vanishes, and $\mathbf{x}^*A\mathbf{x} = \mathbf{u}^{\mathsf{T}}A\mathbf{u} + \mathbf{v}^{\mathsf{T}}A\mathbf{v} \geq 0$.

For the last part, let $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Then $\mathbf{x}^{\mathsf{T}}B\mathbf{x} = \|\mathbf{x}\|^2 \geq 0$ for $\mathbf{x} \in \mathbb{R}^2$ (Non-example 8.3.3(b)). For $\mathbf{x} = (1,i)$, $B\mathbf{x} = (1-i, 1+i)$ and $\mathbf{x}^*B\mathbf{x} = (1-i) - i(1+i) = 2 - 2i$, which is not even real. The computation above shows exactly what was lost: without symmetry, the imaginary part $\mathbf{u}^{\mathsf{T}}A\mathbf{v} - \mathbf{v}^{\mathsf{T}}A\mathbf{u}$ need not vanish.

Solution to Exercise 8.3.31

Uses. Theorem 8.3.5 (Characterizations of positive operators), Proposition 8.3.15 (Characterizations of positive definite operators), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Hint 1. A is Hermitian, so its eigenvalues λ_1, λ_2 are real, with sum $a + c$ and product $ac - |b|^2$.
Hint 2. Two real numbers are both ≥ 0 if and only if their sum and their product are ≥ 0. Then translate "$a + c \geq 0$ and $ac \geq |b|^2$" into the stated conditions.

Sketch. (a) Positive iff $\lambda_1+\lambda_2 = a+c \geq 0$ and $\lambda_1\lambda_2 = ac-|b|^2 \geq 0$. If $ac \geq |b|^2 \geq 0$, then a and c do not have opposite strict signs, so with $a+c \geq 0$ both are ≥ 0. (b) Both eigenvalues > 0 iff $ac-|b|^2 > 0$ and $a+c>0$; given the first, $a>0$ iff $c>0$ iff $a+c>0$. (c) $\operatorname{diag}(0,-1)$.

A is Hermitian, so it has real eigenvalues λ_1, λ_2 (counted with multiplicity), and by Theorem 7.5.22, $\lambda_1+\lambda_2 = a+c$ and $\lambda_1\lambda_2 = ac-|b|^2$.

(a) By Theorem 8.3.5, A is positive if and only if $\lambda_1, \lambda_2 \geq 0$. For real numbers this holds if and only if $\lambda_1+\lambda_2 \geq 0$ and $\lambda_1\lambda_2 \geq 0$: if the product is ≥ 0, the two numbers do not have opposite strict signs, and then a nonnegative sum rules out both being negative. So A is positive if and only if $a+c \geq 0$ and $ac-|b|^2 \geq 0$. These two conditions are equivalent to the three in (a). If $a, c \geq 0$, then $a+c \geq 0$. Conversely, if $ac \geq |b|^2 \geq 0$ and, say, $a<0$, then $c \leq 0$ and $a+c<0$; so $a+c \geq 0$ forces $a \geq 0$, and likewise $c \geq 0$.

(b) By Proposition 8.3.15, A is positive definite if and only if $\lambda_1, \lambda_2 > 0$, if and only if $\lambda_1\lambda_2 > 0$ and $\lambda_1+\lambda_2 > 0$, that is, $ac-|b|^2 > 0$ and $a+c>0$. If $ac > |b|^2 \geq 0$, then a and c are nonzero with the same sign, so $a+c>0$ if and only if $a>0$.

(c) $A = \operatorname{diag}(0,-1)$ has $a = 0 \geq 0$ and $\det A = 0 \geq 0$, but $\mathbf{e}_2^{\mathsf{T}} A \mathbf{e}_2 = -1 < 0$. The condition $c \geq 0$ in (a) cannot be dropped. In (b), by contrast, $a>0$ and $\det A > 0$ suffice, which is the 2×2 case of Sylvester's criterion (Section 12.3).

Solution to Exercise 8.3.31

Uses. Definition 8.3.1 (Positive operator), Definition 8.3.14 (Positive definite)
Hint 1. If $a>0$, write $\mathbf{x}^* A\mathbf{x}$ as $a|x_1+(b/a)x_2|^2 + (\det A/a)|x_2|^2$.
Hint 2. If $a=0$, positivity forces $b=0$: choose $x_2=1$ and $x_1=-tb$ with $t>0$ large.
Sketch. For $a>0$, $\mathbf{x}^*A\mathbf{x} = a|x_1+(b/a)x_2|^2 + \big((ac-|b|^2)/a\big)|x_2|^2$, which is ≥ 0 for all $\mathbf{x}$ iff $ac-|b|^2 \geq 0$, and >0 for $\mathbf{x}\neq\mathbf{0}$ iff $ac-|b|^2>0$ (test $\mathbf{x}=(-b/a,1)$). For $a=0$: $\mathbf{x}^*A\mathbf{x} = 2\operatorname{Re}(\overline{x_1}bx_2) + c|x_2|^2$ is ≥ 0 for all $\mathbf{x}$ iff $b=0$ and $c\geq 0$, and never positive definite. For $a<0$, $\mathbf{e}_1$ gives $a<0$.

Expanding, $\mathbf{x}^*A\mathbf{x} = a|x_1|^2 + 2\operatorname{Re}(\overline{x_1}bx_2) + c|x_2|^2$. We split into cases by the sign of a.

If $a<0$, then $\mathbf{e}_1^* A\mathbf{e}_1 = a<0$, so A is not positive, and the conditions in (a) and (b) fail.

If $a>0$, then expanding $a|x_1+(b/a)x_2|^2 = a|x_1|^2 + 2\operatorname{Re}(\overline{x_1}bx_2) + \big(|b|^2/a\big)|x_2|^2$ gives

$$\mathbf{x}^*A\mathbf{x} = a\left|x_1+\frac{b}{a}x_2\right|^2 + \frac{ac-|b|^2}{a}|x_2|^2.$$

If $ac-|b|^2 \geq 0$, this is ≥ 0; and if $ac-|b|^2>0$, it is 0 only when $x_2=0$ and $x_1+(b/a)x_2=0$, that is, $\mathbf{x}=\mathbf{0}$. Conversely, $\mathbf{x}=(-b/a,1)\neq\mathbf{0}$ gives the value $\big(ac-|b|^2\big)/a$, which must be ≥ 0 if A is positive and >0 if A is positive definite. Also $c\geq 0$ follows from $ac\geq|b|^2\geq 0$.

If $a=0$, then A is not positive definite ($\mathbf{e}_1^*A\mathbf{e}_1=0$), and the condition in (b) fails. If $b\neq 0$, take $x_2=1$ and $x_1=-tb$ with $t>0$: the value is $-2t|b|^2+c$, which is negative for large t; and indeed $ac-|b|^2 = -|b|^2 < 0$. If $b=0$, then $\mathbf{x}^*A\mathbf{x} = c|x_2|^2$, which is ≥ 0 for all $\mathbf{x}$ if and only if $c\geq 0$. In every case A is positive exactly when $a\geq 0$, $c\geq 0$, and $ac-|b|^2\geq 0$, which proves (a) and (b). For (c), see the first solution.

Solution to Exercise 8.3.32

Uses. Theorem 8.3.5 (Characterizations of positive operators), Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 8.3.9 (The positive square root)
Techniques. Gram matrices (Technique 6.3.31), Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. By the proof of Theorem 8.3.5, A is the Gram matrix of $\sqrt{T}\mathbf{e}_1, \ldots, \sqrt{T}\mathbf{e}_n$.
Hint 2. Apply Cauchy–Schwarz to $a_{ij} = \langle \sqrt{T}\mathbf{e}_j, \sqrt{T}\mathbf{e}_i\rangle$.
Sketch. $a_{ij} = \langle \mathbf{w}_j, \mathbf{w}_i\rangle$ with $\mathbf{w}_i = \sqrt{T}\mathbf{e}_i$, so $a_{ii} = \|\mathbf{w}_i\|^2$ and Cauchy–Schwarz gives (a). (b) $|a_{ij}|^2 \leq 0$, and $a_{ji} = \overline{a_{ij}}$. (c) $|a_{ij}| \leq \sqrt{a_{ii}a_{jj}} \leq \max(a_{ii}, a_{jj})$.

(a) Let $\mathbf{w}_i = \sqrt{T}\mathbf{e}_i$. Since $\sqrt{T}$ is self-adjoint, $a_{ij} = \langle T\mathbf{e}_j, \mathbf{e}_i\rangle = \langle \sqrt{T}\mathbf{e}_j, \sqrt{T}\mathbf{e}_i\rangle = \langle \mathbf{w}_j, \mathbf{w}_i\rangle$, and in particular $a_{ii} = \|\mathbf{w}_i\|^2$. By Cauchy–Schwarz (Theorem 6.2.1),

$$|a_{ij}|^2 = |\langle \mathbf{w}_j, \mathbf{w}_i\rangle|^2 \le \|\mathbf{w}_j\|^2\|\mathbf{w}_i\|^2 = a_{jj}a_{ii}.$$

(b) If $a_{ii} = 0$, then $|a_{ij}|^2 \le 0$, so $a_{ij} = 0$ for every j: row i is zero. Since A is Hermitian, $a_{ji} = \overline{a_{ij}} = 0$, so column i is zero too.

(c) The diagonal entries are nonnegative reals. For all i, j, by (a),

$$|a_{ij}| \le \sqrt{a_{ii}a_{jj}} \le \max(a_{ii}, a_{jj}) \le \max_k a_{kk},$$

and $\max_k a_{kk}$ is itself one of the numbers $|a_{ij}|$.

Solution to Exercise 8.3.32

Uses. Theorem 8.3.5 (Characterizations of positive operators), Exercise 8.3.31, Proposition 6.3.18 (Coordinates in an orthonormal basis)
Techniques. Gram matrices (Technique 6.3.31), Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Restrict the form $\langle T\mathbf{v}, \mathbf{v}\rangle$ to vectors $\mathbf{v} = x\mathbf{e}_i + y\mathbf{e}_j$ with $i \ne j$.
Hint 2. The matrix $\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj}\end{pmatrix}$ is then positive, and Exercise 8.3.31 applies.
Sketch. For $i \ne j$ and $\mathbf{v} = x\mathbf{e}_i + y\mathbf{e}_j$, $\langle T\mathbf{v}, \mathbf{v}\rangle = (x, y)^* A_{ij}(x, y)$ with $A_{ij} = \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj}\end{pmatrix}$, which is therefore positive (and Hermitian). By Exercise 8.3.31, $a_{ii}a_{jj} - |a_{ij}|^2 \ge 0$. For $i = j$ the inequality is $a_{ii}^2 \le a_{ii}^2$.

(a) For $i = j$ there is nothing to prove. Let $i \ne j$, and for $(x, y) \in \mathbb{F}^2$ let $\mathbf{v} = x\mathbf{e}_i + y\mathbf{e}_j$. Since $a_{kl} = \langle T\mathbf{e}_l, \mathbf{e}_k\rangle$,

$$\langle T\mathbf{v}, \mathbf{v}\rangle = \overline{x}a_{ii}x + \overline{x}a_{ij}y + \overline{y}a_{ji}x + \overline{y}a_{jj}y = \begin{pmatrix} x \\ y\end{pmatrix}^* \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj}\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix}.$$

The 2×2 matrix is Hermitian (as A is) and the left side is ≥ 0, so it is positive, and Exercise 8.3.31(a) gives $a_{ii}a_{jj} - |a_{ij}|^2 \ge 0$. Parts (b) and (c) follow as in the first solution.

Solution to Exercise 8.3.33

Uses. Theorem 8.3.9 (The positive square root), Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 4.2.11 (Rank–nullity), Lemma 8.3.8 (Where the form vanishes)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)
Hint 1. Write $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \sqrt{T}\mathbf{v}, \sqrt{T}\mathbf{w}\rangle$ and apply Cauchy–Schwarz.
Hint 2. For (b), take $\mathbf{w} = T\mathbf{v}$ in (a).
Hint 3. For (c), $\langle T\mathbf{v}, \mathbf{v}\rangle = \|\sqrt{T}\mathbf{v}\|^2$. For the images, one inclusion is clear and rank–nullity does the rest.
Sketch. (a) $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle \sqrt{T}\mathbf{v}, \sqrt{T}\mathbf{w}\rangle$, and $\|\sqrt{T}\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle$. (b) With $\mathbf{w} = T\mathbf{v}$: $\|T\mathbf{v}\|^4 \le \langle T\mathbf{v}, \mathbf{v}\rangle\langle T^2\mathbf{v}, T\mathbf{v}\rangle$, which is 0 if $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$. (c) $\sqrt{T}\mathbf{v} = \mathbf{0} \Rightarrow T\mathbf{v} = \mathbf{0}$, and $T\mathbf{v} = \mathbf{0} \Rightarrow \|\sqrt{T}\mathbf{v}\|^2 = 0$. $\operatorname{im} T = \operatorname{im}(\sqrt{T}\sqrt{T}) \subseteq \operatorname{im}\sqrt{T}$, with equal dimensions.

Let $R = \sqrt{T}$, which is self-adjoint with $R^2 = T$ (Theorem 8.3.9). For all $\mathbf{v}, \mathbf{w}$, $\langle T\mathbf{v}, \mathbf{w}\rangle = \langle R\mathbf{v}, R\mathbf{w}\rangle$, and in particular $\langle T\mathbf{v}, \mathbf{v}\rangle = \|R\mathbf{v}\|^2$.

(a) By Cauchy–Schwarz (Theorem 6.2.1),

$$|\langle T\mathbf{v}, \mathbf{w}\rangle|^2 = |\langle R\mathbf{v}, R\mathbf{w}\rangle|^2 \le \|R\mathbf{v}\|^2\|R\mathbf{w}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle\langle T\mathbf{w}, \mathbf{w}\rangle.$$

(b) Suppose $\langle T\mathbf{v}, \mathbf{v}\rangle = 0$, and take $\mathbf{w} = T\mathbf{v}$ in (a):

$$\|T\mathbf{v}\|^4 = |\langle T\mathbf{v}, T\mathbf{v}\rangle|^2 \le \langle T\mathbf{v}, \mathbf{v}\rangle\langle T^2\mathbf{v}, T\mathbf{v}\rangle = 0.$$

So $T\mathbf{v} = \mathbf{0}$. (This proof uses $\sqrt{T}$ through (a); one can also prove (a) without it, by applying the argument of Theorem 6.2.1 to the form $(\mathbf{v}, \mathbf{w}) \mapsto \langle T\mathbf{v}, \mathbf{w}\rangle$.)

(c) If $R\mathbf{v} = \mathbf{0}$, then $T\mathbf{v} = R(R\mathbf{v}) = \mathbf{0}$. If $T\mathbf{v} = \mathbf{0}$, then $\|R\mathbf{v}\|^2 = \langle T\mathbf{v}, \mathbf{v}\rangle = 0$, so $R\mathbf{v} = \mathbf{0}$. Hence $\ker R = \ker T$. Next, $\operatorname{im} T = \operatorname{im}(RR) \subseteq \operatorname{im} R$, and by rank–nullity (Theorem 4.2.11) $\dim\operatorname{im} T = \dim V - \dim\ker T = \dim V - \dim\ker R = \dim\operatorname{im} R$. A subspace of $\operatorname{im} R$ with the same dimension is all of $\operatorname{im} R$ (Proposition 4.1.21), so $\operatorname{im} T = \operatorname{im} R$.

Solution to Exercise 8.3.34

Uses. Theorem 8.3.5 (Characterizations of positive operators), Proposition 7.5.25 (Cyclicity of the trace for linear maps), Theorem 8.3.9 (The positive square root), Exercise 8.3.28, Proposition 6.6.18 (Properties of the adjoint)
Techniques. Sandwich between square roots (Technique 8.3.19), Cyclicity of trace (Technique 2.4.29)
Hint 1. For (a), compute tr and det in an orthonormal eigenbasis.
Hint 2. For (b), $\operatorname{tr}(ST) = \operatorname{tr}\big(\sqrt{S}\big(\sqrt{S}T\big)\big) = \operatorname{tr}\big(\sqrt{S}T\sqrt{S}\big)$, and $\sqrt{S}T\sqrt{S}$ is positive.
Hint 3. For (c), if $\sqrt{S}T\sqrt{S} = 0$, then $\big(\sqrt{T}\sqrt{S}\big)^*\big(\sqrt{T}\sqrt{S}\big) = 0$. Deduce $TS = 0$ and take adjoints.
Sketch. (a) $[T]_{\mathcal{O}} = \operatorname{diag}(d_i)$ with $d_i \geq 0$, so $\operatorname{tr} T = \sum d_i \geq 0$, $\det T = \prod d_i \geq 0$, and $\operatorname{tr} T = 0$ forces all $d_i = 0$. (b) By cyclicity $\operatorname{tr}(ST) = \operatorname{tr}(C)$ with $C = \sqrt{S}T\sqrt{S}$ positive. (c) If $\operatorname{tr} C = 0$, then $C = 0$ by (a). Since $C = M^*M$ with $M = \sqrt{T}\sqrt{S}$, $\|M\mathbf{v}\|^2 = \langle C\mathbf{v}, \mathbf{v}\rangle = 0$, so $M = 0$, $TS = \sqrt{T}M\sqrt{S} = 0$, and $ST = (TS)^* = 0$.

(a) By Theorem 8.3.5(III), $[T]_{\mathcal{O}} = \operatorname{diag}(d_1, \ldots, d_n)$ with every $d_i \geq 0$ for some orthonormal basis $\mathcal{O}$. So $\operatorname{tr} T = \sum_i d_i \geq 0$ and $\det T = \prod_i d_i \geq 0$. If $\operatorname{tr} T = 0$, then every $d_i = 0$, so $[T]_{\mathcal{O}} = 0$ and $T = 0$.

(b) By Proposition 7.5.25, $\operatorname{tr}\big(\sqrt{S}\big(\sqrt{S}T\big)\big) = \operatorname{tr}\big(\big(\sqrt{S}T\big)\sqrt{S}\big)$, so $\operatorname{tr}(ST) = \operatorname{tr}(C)$ with $C = \sqrt{S}T\sqrt{S}$. Since $\sqrt{S}$ is self-adjoint, $C = \big(\sqrt{S}\big)^* T\sqrt{S}$ is positive by Exercise 8.3.28(b). By (a), $\operatorname{tr}(ST) = \operatorname{tr} C \geq 0$.

(c) If $ST = 0$, then $\operatorname{tr}(ST) = 0$. Conversely, suppose $\operatorname{tr}(ST) = 0$. Then $\operatorname{tr} C = 0$, so $C = 0$ by (a). Let $M = \sqrt{T}\sqrt{S}$. Since $\sqrt{S}$ and $\sqrt{T}$ are self-adjoint, $M^*M = \sqrt{S}\sqrt{T}\sqrt{T}\sqrt{S} = C = 0$, so $\|M\mathbf{v}\|^2 = \langle M^*M\mathbf{v}, \mathbf{v}\rangle = 0$ for every $\mathbf{v}$, and $M = 0$. Then $TS = \sqrt{T}\big(\sqrt{T}\sqrt{S}\big)\sqrt{S} = \sqrt{T}M\sqrt{S} = 0$, and $ST = S^*T^* = (TS)^* = 0$ by Proposition 6.6.18.

Solution to Exercise 8.3.35

Uses. Proposition 8.3.15 (Characterizations of positive definite operators), Proposition 6.1.13 (Integral inner products), Definition 6.3.28 (Gram matrix), Proposition 7.1.22 (Counting roots), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Techniques. Gram matrices (Technique 6.3.31)
Hint 1. Look for functions $f_1, \ldots, f_n$ on $[0,1]$ with $\int_0^1 f_i(t) f_j(t)\,\mathrm{d}t = 1/(i+j-1)$.
Hint 2. $f_i(t) = t^{i-1}$ works for H_n. For K_n you need $\int_0^1 f_i f_j = 1/(i+j)$; try $f_i(t) = t^{i-1/2}$.
Hint 3. A Gram matrix of a linearly independent list is positive definite (Proposition 8.3.15). Independence comes down to a polynomial with infinitely many roots.
Sketch. (a) H_n is the Gram matrix of $1, t, \ldots, t^{n-1}$ in $C[0,1]$ with $\langle f, g\rangle = \int_0^1 fg$, an independent list, so it is positive definite and $\det H_n$ is a product of positive eigenvalues. (b) K_n is the Gram matrix of $t^{1/2}, t^{3/2}, \ldots, t^{n-1/2}$; if $\sum c_i t^{i-1/2} = 0$ on $[0,1]$, multiplying by $t^{1/2}$ gives a polynomial vanishing on $[0,1]$.

Give $C[0,1]$ the inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,\mathrm{d}t$ (Proposition 6.1.13).

(a) Let $f_i(t) = t^{i-1}$ for $i = 1, \ldots, n$. Then

$$\langle f_j, f_i\rangle = \int_0^1 t^{i+j-2}\,\mathrm{d}t = \frac{1}{i+j-1} = h_{ij},$$

so H_n is the Gram matrix of $f_1, \ldots, f_n$ (Definition 6.3.28). This list is linearly independent: if $\sum_i c_i t^{i-1} = 0$ for every $t \in [0,1]$, the polynomial $\sum_i c_i x^{i-1}$ has infinitely many roots, so it is zero (Proposition 7.1.22) and every $c_i = 0$. By Proposition 8.3.15, H_n is positive definite. Its eigenvalues are positive and $\det H_n$ is their product (Theorem 7.5.22), so $\det H_n > 0$.

(b) Let $g_i(t) = t^{i-1/2}$ for $i = 1, \ldots, n$, continuous on $[0,1]$ since $i - 1/2 > 0$. Then $\langle g_j, g_i\rangle = \int_0^1 t^{i+j-1}\,\mathrm{d}t = 1/(i+j) = k_{ij}$. If $\sum_i c_i t^{i-1/2} = 0$ for every $t \in [0,1]$, then multiplying by $t^{1/2}$ gives $\sum_i c_i t^i = 0$ on $[0,1]$, so as in (a) every $c_i = 0$. Hence K_n is the Gram matrix of an independent list and is positive definite.

No eigenvalue had to be computed. The eigenvalues of H_n are in fact tiny (the smallest is below 10^{-5} for $n = 5$), so a numerical test of positivity would be delicate, while the Gram description settles the question exactly.

Solution to Exercise 8.3.36

Uses. Theorem 8.3.9 (The positive square root), Theorem 8.3.5 (Characterizations of positive operators), Theorem 7.5.22 (Trace and determinant from eigenvalues), Exercise 8.3.31
Hint 1. Every $R \in \mathbb{F}^{2\times 2}$ satisfies $R^2 - (\operatorname{tr} R)R + (\det R)I_2 = 0$. Check it by multiplying out, and apply it to $R = \sqrt{A}$.

Hint 2. If R has eigenvalues $\mu_1, \mu_2 \geq 0$, then $A = R^2$ has eigenvalues μ_1^2, μ_2^2. Express $\det R$ and $(\operatorname{tr} R)^2$ through $\det A$ and $\operatorname{tr} A$.
Sketch. With $R = \sqrt{A}$, eigenvalues $\mu_1, \mu_2 \geq 0$: $\det R = \mu_1\mu_2 = \delta$ and $(\operatorname{tr} R)^2 = \mu_1^2 + \mu_2^2 + 2\mu_1\mu_2 = \operatorname{tr} A + 2\delta$, so $\operatorname{tr} R = s = \sqrt{\operatorname{tr} A + 2\delta} > 0$ (as $R \neq 0$). Then $A - sR + \delta I_2 = 0$ gives $R = (A + \delta I_2)/s$. (b) $s = 1 + \sqrt{3}$ and $\sqrt{A} = \frac{1}{2}\begin{pmatrix} 1+\sqrt{3} & \sqrt{3}-1 \\ \sqrt{3}-1 & 1+\sqrt{3} \end{pmatrix}$; then $\delta = 3$, $s = 4$, $\sqrt{A} = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$.

(a) For $R = \begin{pmatrix} p & q \\ r & s' \end{pmatrix}$, multiplying out gives

$$R^2 - (p + s')R = \begin{pmatrix} p^2 + qr - p^2 - ps' & pq + qs' - pq - qs' \\ rp + s'r - pr - s'r & rq + s'^2 - ps' - s'^2 \end{pmatrix} = -(ps' - qr)I_2,$$

that is, $R^2 - (\operatorname{tr} R)R + (\det R)I_2 = 0$. Let $R = \sqrt{A}$. By Theorem 8.3.5(III), $[R]_{\mathcal{O}} = \operatorname{diag}(\mu_1, \mu_2)$ for an orthonormal basis $\mathcal{O}$ with $\mu_1, \mu_2 \geq 0$, and then $[A]_{\mathcal{O}} = \operatorname{diag}\left(\mu_1^2, \mu_2^2\right)$. By Theorem 7.5.22,

$$\det R = \mu_1\mu_2 = \sqrt{\mu_1^2\mu_2^2} = \sqrt{\det A} = \delta, \quad (\operatorname{tr} R)^2 = (\mu_1 + \mu_2)^2 = \mu_1^2 + \mu_2^2 + 2\mu_1\mu_2 = \operatorname{tr} A + 2\delta.$$

Since $\operatorname{tr} R = \mu_1 + \mu_2 \geq 0$, $\operatorname{tr} R = s$ with $s = \sqrt{\operatorname{tr} A + 2\delta}$. If $s = 0$, then $\mu_1 = \mu_2 = 0$, so $R = 0$ and $A = 0$, which is excluded. So $s > 0$, and the identity $R^2 - sR + \delta I_2 = 0$ with $R^2 = A$ gives $R = (A + \delta I_2)/s$.

(b) For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$: $\det A = 3$, $\operatorname{tr} A = 4$, and A is positive (Exercise 8.3.31). So $\delta = \sqrt{3}$ and $s = \sqrt{4 + 2\sqrt{3}} = 1 + \sqrt{3}$, since $\left(1 + \sqrt{3}\right)^2 = 4 + 2\sqrt{3}$. Using $\left(2 + \sqrt{3}\right)/\left(1 + \sqrt{3}\right) = \left(1 + \sqrt{3}\right)/2$ and $1/\left(1 + \sqrt{3}\right) = \left(\sqrt{3} - 1\right)/2$,

$$\sqrt{A} = \frac{1}{1 + \sqrt{3}}\begin{pmatrix} 2 + \sqrt{3} & 1 \\ 1 & 2 + \sqrt{3} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 + \sqrt{3} & \sqrt{3} - 1 \\ \sqrt{3} - 1 & 1 + \sqrt{3} \end{pmatrix}.$$

Check: the square of the right side is $\frac{1}{4}\begin{pmatrix} \left(1+\sqrt{3}\right)^2 + \left(\sqrt{3}-1\right)^2 & 2\left(1+\sqrt{3}\right)\left(\sqrt{3}-1\right) \\ \cdots & \cdots \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix} = A$.

For $A = \begin{pmatrix} 5 & 4i \\ -4i & 5 \end{pmatrix}$: A is Hermitian with $a = 5 > 0$ and $\det A = 25 - (4i)(-4i) = 25 - 16 = 9 > 0$, so it is positive (even positive definite) by Exercise 8.3.31. Here $\delta = 3$, $s = \sqrt{10 + 6} = 4$, and

$$\sqrt{A} = \frac{1}{4}\begin{pmatrix} 8 & 4i \\ -4i & 8 \end{pmatrix} = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}.$$

Check: $\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}^2 = \begin{pmatrix} 4+i(-i) & 2i+2i \\ -2i-2i & (-i)i+4 \end{pmatrix} = \begin{pmatrix} 5 & 4i \\ -4i & 5 \end{pmatrix}$.

Solution to Exercise 8.3.37

Uses. Proposition 8.2.24 (Properties of $f(T)$), Theorem 8.3.9 (The positive square root), Theorem 8.3.5 (Characterizations of positive operators)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. Let $f(\lambda) = |\lambda|$. Show that $f(T)$ is positive and squares to T^2, then use uniqueness of the positive square root.
Hint 2. For (b), if T is positive, then T itself is a positive square root of T^2.
Hint 3. For (c), $T_+ = \sum_j \max(\lambda_j, 0)P_j$ and $T_- = \sum_j \max(-\lambda_j, 0)P_j$.
Sketch. (a) $f(T) = \sum|\lambda_j|P_j$ is self-adjoint with eigenvalues $|\lambda_j| \geq 0$, so positive, and $f(T)^2 = \left(f^2\right)(T) = \sum \lambda_j^2 P_j = T^2$; by uniqueness it is $\sqrt{T^2}$. (b) $|T|$ is positive; conversely a positive T is a positive square root of T^2. (c) $T_\pm = g_\pm(T)$ with $g_+(\lambda) = \max(\lambda, 0)$, $g_-(\lambda) = \max(-\lambda, 0)$, which are ≥ 0 with $g_+ - g_- = \lambda$ and $g_+ g_- = 0$.

(a) Let $f(\lambda) = |\lambda|$ on the eigenvalues, which are real. By Proposition 8.2.24, $f(T) = \sum_j |\lambda_j| P_j$ is self-adjoint (its values are real) with nonnegative eigenvalues $|\lambda_j|$, so it is positive by Theorem 8.3.5, and $f(T)^2 = \left(f^2\right)(T) = \sum_j \lambda_j^2 P_j = T^2$, because $f^2(\lambda) = \lambda^2$ and $T^2 = \sum_j \lambda_j^2 P_j$ (Proposition 8.2.24(c) with $g(\lambda) = \lambda$). By uniqueness in Theorem 8.3.9, $f(T) = \sqrt{T^2} = |T|$.

(b) $|T|$ is positive, so if $|T| = T$, then T is positive. Conversely, if T is positive, then T is a positive operator whose square is T^2, so $T = \sqrt{T^2} = |T|$ by uniqueness.

(c) By (a), $T_+ = \sum_j g_+(\lambda_j)P_j$ and $T_- = \sum_j g_-(\lambda_j)P_j$, where $g_+(\lambda) = (|\lambda| + \lambda)/2 = \max(\lambda, 0)$ and $g_-(\lambda) = (|\lambda| - \lambda)/2 = \max(-\lambda, 0)$. Both functions are real and nonnegative, so T_+ and T_- are positive (Proposition 8.2.24(d), (e), and Theorem 8.3.5). Clearly $T_+ - T_- = T$. Finally $g_+(\lambda)g_-(\lambda) = 0$ for every real λ, since one of the two factors is 0, so $T_+T_- = (g_+g_-)(T) = 0$ by Proposition 8.2.24(c). This is the operator version of writing a real number as its positive part minus its negative part.

Solution to Exercise 8.3.38

Uses. Theorem 8.3.9 (The positive square root), Proposition 8.2.20 (Spectral projections are polynomials in T), Proposition 8.2.24 (Properties of $f(T)$), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Is $R^2 = \lambda\,\mathrm{id}$ on a space of dimension 2 enough to conclude $R = \sqrt{\lambda}\,\mathrm{id}$? Think of $T = \mathrm{id}_V$.
Hint 2. For (b), every eigenspace is a line, and on a line an operator is a scalar.
Sketch. (a) "$R^2 = \lambda\,\mathrm{id}$ on E_λ, so $R = \sqrt{\lambda}\,\mathrm{id}$" is wrong: R may act as $-\sqrt{\lambda}$, or on a plane as any reflection scaled by $\sqrt{\lambda}$. Counterexamples: $-\sqrt{T}$, and $\mathrm{diag}(1, -1)$ for $T = I_2$. (b) R preserves each eigenline E_{λ_j}, so acts on it as a real μ_j with $\mu_j^2 = \lambda_j$; there are two choices if $\lambda_j > 0$ and one if $\lambda_j = 0$, and every choice works. So 2^k.

(a) Everything up to "$R^2 = \lambda\,\mathrm{id}$ on E_λ" is correct. The error is the conclusion "$R = \sqrt{\lambda}\,\mathrm{id}$ on E_λ". A square root of $\lambda\,\mathrm{id}$ need not be $\sqrt{\lambda}\,\mathrm{id}$: the self-adjoint operator $R_\lambda = R|_{E_\lambda}$ is diagonal in an orthonormal basis of E_λ with entries μ satisfying $\mu^2 = \lambda$, so each entry may be $\sqrt{\lambda}$ or $-\sqrt{\lambda}$. (In the first proof of Theorem 8.3.9 the extra hypothesis $\mu \geq 0$ settles this.) The claim is false: if $T \neq 0$, then $-\sqrt{T}$ is a second self-adjoint square root, and even for $T = I_2$ on $\mathbb{R}^2$, every $\mathrm{diag}(\pm 1, \pm 1)$ and every reflection is a self-adjoint square root.

(b) Let $\lambda_1, \ldots, \lambda_n$ be the distinct eigenvalues, with eigenspaces $E_{\lambda_j} = \mathrm{span}(\mathbf{u}_j)$, which are lines because there are n of them in an n-dimensional space. Let R be self-adjoint with $R^2 = T$. As in the argument, R maps E_{λ_j} into itself, so $R\mathbf{u}_j = \mu_j\mathbf{u}_j$ for a scalar μ_j, which is real by Proposition 8.1.1, and $\mu_j^2\mathbf{u}_j = T\mathbf{u}_j = \lambda_j\mathbf{u}_j$ gives $\mu_j^2 = \lambda_j$. So $\mu_j \in \{\sqrt{\lambda_j}, -\sqrt{\lambda_j}\}$, which is two choices if $\lambda_j > 0$ and one if $\lambda_j = 0$, and R is determined by $\mu_1, \ldots, \mu_n$ because $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis. Conversely, for any such choice, $R = \sum_j \mu_j P_j$ is self-adjoint (its coefficients are real) and $R^2 = \sum_j \mu_j^2 P_j = T$ by Proposition 8.2.24. Different choices give different operators, as they differ on some $\mathbf{u}_j$. Hence T has exactly 2^k self-adjoint square roots, and exactly one of them, $\sqrt{T}$, is positive.

Solution to Exercise 8.3.39

Uses. Proposition 8.3.22 (Properties of the Löwner order), Theorem 8.3.9 (The positive square root), Proposition 8.3.15 (Characterizations of positive definite operators), Theorem 5.4.10 (Multiplicativity), Theorem 7.5.22 (Trace and determinant from eigenvalues)
Techniques. Sandwich between square roots (Technique 8.3.19)
Hint 1. Normalize B to the identity: let $C = B^{-1/2}AB^{-1/2}$. By Proposition 8.3.22(c), $C \preceq I_n$.
Hint 2. By Proposition 8.3.22(d), the eigenvalues of C lie in $(0, 1]$. So the eigenvalues of C^{-1} are at least 1, and $C^{-1} = B^{1/2}A^{-1}B^{1/2}$.
Hint 3. For (b), $\det C = \det A / \det B$. For (c), compute the eigenvalues, or the determinant, of $B - A$ and of $B^2 - A^2$.
Sketch. $C = B^{-1/2}AB^{-1/2} \preceq B^{-1/2}BB^{-1/2} = I_n$, and C is positive definite, so its eigenvalues lie in $(0, 1]$. (a) $C^{-1} = B^{1/2}A^{-1}B^{1/2}$ has eigenvalues ≥ 1, so $I_n \preceq C^{-1}$; congruence by $B^{-1/2}$ gives $B^{-1} \preceq A^{-1}$. (b) $\det A / \det B = \det C \leq 1$. (c) $B - A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is positive, but $B^2 - A^2 = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}$ has determinant -1.

Let $B^{1/2} = \sqrt{B}$ and $B^{-1/2} = \left(\sqrt{B}\right)^{-1}$, both positive definite and Hermitian. Let $C = B^{-1/2}AB^{-1/2}$. By Proposition 8.3.22(c) with $X = B^{-1/2}$,

$$C \preceq B^{-1/2}BB^{-1/2} = I_n.$$

Also $C = X^*AX$ with X invertible, so C is positive definite: $\mathbf{x}^*C\mathbf{x} = (X\mathbf{x})^*A(X\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$. By Proposition 8.3.22(d) and Proposition 8.3.15, every eigenvalue of C lies in $(0, 1]$.

(a) C is invertible with $C^{-1} = B^{1/2}A^{-1}B^{1/2}$, and its eigenvalues are the reciprocals of those of C (if $C\mathbf{v} = \mu\mathbf{v}$, then $C^{-1}\mathbf{v} = \mu^{-1}\mathbf{v}$, and C^{-1} is Hermitian with an orthonormal eigenbasis made of eigenvectors of C). So they are all ≥ 1, and $I_n \preceq C^{-1}$ by Proposition 8.3.22(d). Applying Proposition 8.3.22(c) with $X = B^{-1/2}$,

$$B^{-1} = B^{-1/2}I_nB^{-1/2} \preceq B^{-1/2}B^{1/2}A^{-1}B^{1/2}B^{-1/2} = A^{-1}.$$

(b) $\det C$ is the product of the eigenvalues of C (Theorem 7.5.22), which lie in $(0,1]$, so $\det C \le 1$. By Theorem 5.4.10, $\det C = \det\left(B^{-1/2}\right)^2 \det A$, and $\det\left(B^{-1/2}\right)^2 \det B = \det\left(B^{-1/2}BB^{-1/2}\right) = \det I_n = 1$. So $\det C = \det A / \det B \le 1$, and $\det A \le \det B$ because $\det B > 0$.

(c) A is a projection and B has trace 3 and determinant 1, so both are positive. $B - A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is positive (eigenvalues 2 and 0), so $A \preceq B$. But $A^2 = A$ and $B^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$, so $B^2 - A^2 = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}$, whose determinant is $8 - 9 = -1 < 0$. Its eigenvalues have a negative product, so one of them is negative, and $A^2 \preceq B^2$ fails. The tempting argument "$B^2 - A^2 = (B - A)(B + A)$ is a product of positive operators" is wrong twice: the factorization needs $AB = BA$, and a product of positive operators need not be positive (Proposition 8.3.20).

Solution to Exercise 8.3.40

Uses. Lemma 8.3.8 (Where the form vanishes), Theorem 8.3.5 (Characterizations of positive operators), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators), Exercise 8.3.39
Techniques. A sum of nonnegative terms (Technique 6.1.9), Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. Imitate the third proof of uniqueness in Theorem 8.3.9. Let $X = \sqrt{B} - \sqrt{A}$, which is self-adjoint; it suffices to show that every eigenvalue of X is ≥ 0.
Hint 2. Check that $B - A = \sqrt{B}X + X\sqrt{A}$, apply both sides to an eigenvector $\mathbf{v}$ of X, and take the inner product with $\mathbf{v}$.
Hint 3. You get $\langle (B-A)\mathbf{v}, \mathbf{v}\rangle = \mu\left(\left\langle \sqrt{B}\mathbf{v}, \mathbf{v}\right\rangle + \left\langle \sqrt{A}\mathbf{v}, \mathbf{v}\right\rangle\right)$. The left side is ≥ 0. What if the bracket is 0?
Sketch. With $X = \sqrt{B} - \sqrt{A}$ and $X\mathbf{v} = \mu\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$: $B - A = \sqrt{B}X + X\sqrt{A}$ gives $0 \le \langle (B-A)\mathbf{v}, \mathbf{v}\rangle = \mu\left(\left\langle \sqrt{B}\mathbf{v}, \mathbf{v}\right\rangle + \left\langle \sqrt{A}\mathbf{v}, \mathbf{v}\right\rangle\right)$. If the bracket is positive, $\mu \ge 0$. If it is 0, then $\sqrt{B}\mathbf{v} = \sqrt{A}\mathbf{v} = \mathbf{0}$ by Lemma 8.3.8, so $\mu\mathbf{v} = X\mathbf{v} = \mathbf{0}$ and $\mu = 0$. So X is positive. Converse: square roots $\operatorname{diag}(1,0) \preceq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ of Exercise 8.3.39(c).

Let $X = \sqrt{B} - \sqrt{A}$, which is self-adjoint. Then

$$\sqrt{B}X + X\sqrt{A} = B - \sqrt{B}\sqrt{A} + \sqrt{B}\sqrt{A} - A = B - A.$$

Let μ be an eigenvalue of X, with $X\mathbf{v} = \mu\mathbf{v}$ and $\mathbf{v} \neq \mathbf{0}$; μ is real (Proposition 8.1.1). Using $X^* = X$,

$$\langle (B-A)\mathbf{v}, \mathbf{v}\rangle = \left\langle \sqrt{B}(\mu\mathbf{v}), \mathbf{v}\right\rangle + \left\langle \sqrt{A}\mathbf{v}, X\mathbf{v}\right\rangle = \mu\left(\left\langle \sqrt{B}\mathbf{v}, \mathbf{v}\right\rangle + \left\langle \sqrt{A}\mathbf{v}, \mathbf{v}\right\rangle\right).$$

The left side is ≥ 0 because $A \preceq B$, and the bracket β is a sum of two nonnegative numbers. If $\beta > 0$, then $\mu = \langle (B-A)\mathbf{v}, \mathbf{v}\rangle / \beta \ge 0$. If $\beta = 0$, then both terms are 0, so $\sqrt{B}\mathbf{v} = \sqrt{A}\mathbf{v} = \mathbf{0}$ by Lemma 8.3.8, hence $\mu\mathbf{v} = X\mathbf{v} = \mathbf{0}$ and $\mu = 0$. In either case $\mu \ge 0$, so X is positive by Theorem 8.3.5(II), that is, $\sqrt{A} \preceq \sqrt{B}$.

For the converse, let $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, which are positive with $A_0 \preceq B_0$ (Exercise 8.3.39(c)). Put $A = A_0^2$ and $B = B_0^2$. Then $\sqrt{A} = A_0$ and $\sqrt{B} = B_0$ by uniqueness of positive square roots, so $\sqrt{A} \preceq \sqrt{B}$; but $A \preceq B$ fails, by Exercise 8.3.39(c).

Solution to Exercise 8.3.41

Uses. Algorithm 8.3.24 (Cholesky factorization), Proposition 8.3.25 (Correctness of the Cholesky algorithm), Theorem 8.3.23 (Cholesky factorization)
Hint 1. Work row by row: $r_{kk} = \sqrt{a_{kk} - \sum_{i<k} r_{ik}^2}$, then $r_{kj} = \left(a_{kj} - \sum_{i<k} r_{ik}r_{ij}\right)/r_{kk}$.
Hint 2. In (b) the algorithm stops at $k = 3$. For (c), look for $\mathbf{x} = (x_1, x_2, 1)$ with the first two entries of $B\mathbf{x}$ equal to 0, as in the induction proof of Theorem 8.3.23.
Sketch. (a) $R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{pmatrix}$. (b) $r_{1*} = (1,2,3)$, $r_{22} = 1$, $r_{23} = -2$, then $d = 6 - 9 - 4 = -7 < 0$: not positive definite. (c) Solving $x_1 + 2x_2 = -3$, $2x_1 + 5x_2 = -4$ gives $\mathbf{x} = (-7, 2, 1)$, $B\mathbf{x} = (0, 0, -7)$, and $\mathbf{x}^{\mathsf{T}}B\mathbf{x} = -7$.

(a) $k = 1$: $r_{11} = 1$, $r_{12} = 2$, $r_{13} = 1$. $k = 2$: $d = 13 - 2^2 = 9$, so $r_{22} = 3$, and $r_{23} = (8 - 2 \cdot 1)/3 = 2$. $k = 3$: $d = 14 - 1^2 - 2^2 = 9$, so $r_{33} = 3$. Hence

$$A = R^{\mathsf{T}}R, \quad R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Check the last column of $R^{\mathsf{T}}R$: $(1 \cdot 1, 2 \cdot 1 + 3 \cdot 2, 1 + 4 + 9) = (1, 8, 14)$.

(b) $k=1$: $r_{11}=1$, $r_{12}=2$, $r_{13}=3$. $k=2$: $d=5-4=1$, $r_{22}=1$, $r_{23}=(4-2\cdot 3)/1=-2$. $k=3$: $d=6-3^2-(-2)^2=-7\le 0$, and the algorithm stops. By Proposition 8.3.25, B is not positive definite.

(c) Following the induction proof of Theorem 8.3.23, take $\mathbf{x}=(x_1,x_2,1)$ with x_1,x_2 chosen so that the first two entries of $B\mathbf{x}$ vanish:

$$x_1+2x_2+3=0,\quad 2x_1+5x_2+4=0.$$

This gives $x_2=2$ and $x_1=-7$. Then $B\mathbf{x}=(-7+4+3,-14+10+4,-21+8+6)=(0,0,-7)$, and $\mathbf{x}^{\mathsf{T}}B\mathbf{x}=1\cdot(-7)=-7<0$. The value -7 is exactly the number d at which the algorithm stopped: it is the Schur complement of the leading 2×2 block.

Solution to Exercise 8.3.42

Uses. Exercise 8.3.32, Exercise 8.3.28, Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem 5.4.10 (Multiplicativity)
Techniques. Sandwich between square roots (Technique 8.3.19)
Hint 1. If some $a_{ii}=0$, then $\det A=0$ (Exercise 8.3.32). Otherwise let $D=\operatorname{diag}(a_{11},\dots,a_{nn})$ and $C=D^{-1/2}AD^{-1/2}$. What are the diagonal entries of C?
Hint 2. C is positive with trace n. Apply the AM–GM inequality to its eigenvalues.
Hint 3. For (b), apply (a) to $A=B^*B$.
Sketch. If some $a_{ii}=0$, a row of A vanishes and $\det A=0$. Otherwise $C=D^{-1/2}AD^{-1/2}$ is positive with $c_{ii}=1$, so its eigenvalues $\mu_i\ge 0$ have sum n and $\det C=\prod\mu_i\le(\sum\mu_i/n)^n=1$ by AM–GM. Then $\det A=\det D\det C\le\prod a_{ii}$. (b) $A=B^*B$ has $a_{ii}=\|\mathbf{b}_i\|^2$ and $\det A=|\det B|^2$.

(a) All $a_{ii}\ge 0$. If some $a_{ii}=0$, then row i of A is zero by Exercise 8.3.32(b), so $\det A=0\le\prod_i a_{ii}$. Otherwise let $D=\operatorname{diag}(a_{11},\dots,a_{nn})$ and $D^{-1/2}=\operatorname{diag}\left(a_{11}^{-1/2},\dots,a_{nn}^{-1/2}\right)$, and $C=D^{-1/2}AD^{-1/2}$. Then C is positive by Exercise 8.3.28(b), and $c_{ij}=a_{ij}/\sqrt{a_{ii}a_{jj}}$, so every $c_{ii}=1$ and $\operatorname{tr}C=n$. Let $\mu_1,\dots,\mu_n\ge 0$ be the eigenvalues of C counted with multiplicity. By Theorem 7.5.22 and the AM–GM inequality,

$$\det C=\mu_1\cdots\mu_n\le\left(\frac{\mu_1+\cdots+\mu_n}{n}\right)^n=\left(\operatorname{tr}\frac{C}{n}\right)^n=1.$$

By Theorem 5.4.10, $\det A=\det\left(D^{1/2}\right)\det C\det\left(D^{1/2}\right)=\left(\prod_i a_{ii}\right)\det C\le\prod_i a_{ii}$.

(b) $A=B^*B$ is positive, with $a_{ii}=\mathbf{b}_i^*\mathbf{b}_i=\|\mathbf{b}_i\|^2$ and $\det A=\det(B^*)\det B=\overline{\det B}\det B=|\det B|^2$ (Theorem 5.4.14). So (a) gives $|\det B|^2\le\prod_i\|\mathbf{b}_i\|^2$. For the equality case, see the second solution.

Solution to Exercise 8.3.42

Uses. Theorem 8.3.23 (Cholesky factorization), Proposition 5.3.15 (Triangular matrices), Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose)
Techniques. Sandwich between square roots (Technique 8.3.19)
Hint 1. If A is not invertible, $\det A=0$. Otherwise write $A=R^*R$ (Theorem 8.3.23) and compute $\det A$ and a_{ii} in terms of the entries of R.
Hint 2. a_{ii} is the squared length of column i of R, and $\det A=\prod_i r_{ii}^2$.
Sketch. For singular A, $\det A=0\le\prod a_{ii}$. Otherwise $A=R^*R$, $\det A=\prod r_{ii}^2$, and $a_{ii}=\sum_{k\le i}|r_{ki}|^2\ge r_{ii}^2$. Equality for all i iff R is diagonal iff A is diagonal. In (b), $A=B^*B$ is diagonal iff the columns of B are orthogonal.

(a) If A is not invertible, $\det A=0\le\prod_i a_{ii}$. Otherwise A is positive definite (Proposition 8.3.15), and $A=R^*R$ with R upper triangular with positive diagonal entries (Theorem 8.3.23). By Proposition 5.3.15 and Theorem 5.4.10, $\det A=\overline{\det R}\det R=\prod_i r_{ii}^2$. The (i,i) entry of R^*R is the squared length of column i of R, whose nonzero entries are $r_{1i},\dots,r_{ii}$:

$$a_{ii}=\sum_{k\le i}|r_{ki}|^2\ge r_{ii}^2. \tag{C.1.10}$$

Multiplying these n inequalities between nonnegative numbers gives $\prod_i a_{ii}\ge\prod_i r_{ii}^2=\det A$.

(b) As in the first solution, $A=B^*B$ is positive with $a_{ii}=\|\mathbf{b}_i\|^2$ and $\det A=|\det B|^2$, which gives the inequality. Let B be invertible. Then A is positive definite and every $r_{ii}>0$, so equality $\prod_i a_{ii}=\prod_i r_{ii}^2$ holds

if and only if equality holds in every instance of (C.1.10), that is, if and only if $r_{ki} = 0$ for all $k < i$. This means that R is diagonal, which makes $A = R^*R$ diagonal; conversely, if A is diagonal, then $\det A = \prod_i a_{ii}$. So equality holds if and only if $A = B^*B$ is diagonal, that is, $\mathbf{b}_i^*\mathbf{b}_j = 0$ for $i \neq j$: the columns are orthogonal. Geometrically, the volume $|\det B|$ of the parallelepiped spanned by the columns is at most the product of the edge lengths, with equality for a box.

Solution to Exercise 8.3.43

Uses. Theorem 8.2.12 (Real spectral theorem), Corollary 8.2.7 (Hermitian matrices), Proposition 8.3.15 (Characterizations of positive definite operators), Theorem 8.3.5 (Characterizations of positive operators), Proposition 8.1.1 (Eigenvalues and eigenvectors of self-adjoint operators)
Techniques. Choose a convenient basis (Technique 4.5.33), Test an identity on eigenvectors (Technique 7.4.25)
Hint 1. Diagonalize $A = UDU^*$ and substitute $X = UYU^*$. The equation becomes $DY + YD = U^*CU$; read it entry by entry.
Hint 2. For (b), take the conjugate transpose of the equation and use (a).
Hint 3. For (c), let $X\mathbf{v} = \mu\mathbf{v}$ and compute $\langle C\mathbf{v}, \mathbf{v}\rangle = \mathbf{v}^*(AX + XA)\mathbf{v}$ in terms of μ and $\mathbf{v}^*A\mathbf{v}$.
Hint 4. For (d), take $A = \operatorname{diag}(a_1, \dots, a_n)$ and C the matrix of all ones.
Sketch. (a) With $A = UDU^*$, $D = \operatorname{diag}(d_i)$, $d_i > 0$, the equation is $(d_i + d_j)y_{ij} = (U^*CU)_{ij}$ for $Y = U^*XU$, which has exactly one solution since $d_i + d_j > 0$. (b) X^* solves the same equation. (c) For $X\mathbf{v} = \mu\mathbf{v}$, $\mathbf{v}^*C\mathbf{v} = 2\mu\mathbf{v}^*A\mathbf{v}$, and $\mathbf{v}^*A\mathbf{v} > 0$, so $\mu \geq 0$. (d) With $A = \operatorname{diag}(a_i)$ and $C = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ positive, the solution is $x_{ij} = 1/(a_i + a_j)$.

(a) By Theorem 8.2.12 or Corollary 8.2.7, $A = UDU^*$ with U unitary and $D = \operatorname{diag}(d_1, \dots, d_n)$, and $d_i > 0$ by Proposition 8.3.15. Since $X \mapsto U^*XU$ is a bijection of $\mathbb{F}^{n\times n}$ (with inverse $Y \mapsto UYU^*$), we may write $X = UYU^*$. Then

$$AX + XA = U(DY + YD)U^*,$$

so $AX + XA = C$ if and only if $DY + YD = C'$, where $C' = U^*CU$. The (i,j) entry of $DY + YD$ is $(d_i + d_j)y_{ij}$, and $d_i + d_j > 0$. So the equation holds if and only if $y_{ij} = c'_{ij}/(d_i + d_j)$ for all i, j, and there is exactly one solution.

(b) Taking conjugate transposes of $AX + XA = C$ and using $A^* = A$ gives $X^*A + AX^* = C^* = C$. So X^* is also a solution, and by (a), $X^* = X$.

(c) By (b), X is Hermitian, so its eigenvalues are real (Proposition 8.1.1). Let $X\mathbf{v} = \mu\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{v}^*X = (X\mathbf{v})^* = \overline{\mu}\mathbf{v}^* = \mu\mathbf{v}^*$, and

$$\mathbf{v}^*C\mathbf{v} = \mathbf{v}^*A(X\mathbf{v}) + (\mathbf{v}^*X)A\mathbf{v} = \mu\mathbf{v}^*A\mathbf{v} + \mu\mathbf{v}^*A\mathbf{v} = 2\mu\mathbf{v}^*A\mathbf{v}.$$

The left side is ≥ 0 because C is positive, and $\mathbf{v}^*A\mathbf{v} > 0$ because A is positive definite. So $\mu \geq 0$. Every eigenvalue of the Hermitian matrix X is nonnegative, and X is positive by Theorem 8.3.5(II).

(d) Let $A = \operatorname{diag}(a_1, \dots, a_n)$, which is positive definite, and let $C = \mathbf{1}\mathbf{1}^{\mathsf{T}}$ be the matrix of all ones, where $\mathbf{1} = (1, \dots, 1)$; C is positive, being B^*B for the $1 \times n$ matrix $B = \mathbf{1}^{\mathsf{T}}$. The (i,j) entry of $AX + XA$ is $(a_i + a_j)x_{ij}$, so the solution of $AX + XA = C$ is $x_{ij} = 1/(a_i + a_j)$. By (c) this matrix is positive.

The heart of (c) is the trick from the third proof of Theorem 8.3.9: test the equation on an eigenvector of the unknown, not of the data.

Section 8.4

Solution to Exercise 8.4.8

Uses. Theorem 8.4.4 (Characterizations of isometries), Definition 8.4.1 (Isometry)
Hint 1. Condition (V) of Theorem 8.4.4 asks for the images to be orthonormal, not merely of length 1. Try two basis vectors with the same image.
Hint 2. For (b), every nonzero vector is a positive multiple of a unit vector.
Sketch. (a) False: $S = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ sends $\mathbf{e}_1$ and $\mathbf{e}_2$ to $\mathbf{e}_1$, but $S(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$. (b) True: $\|S\mathbf{v}\| = \|\mathbf{v}\|\|S(\mathbf{v}/\|\mathbf{v}\|)\| = \|\mathbf{v}\|$ for $\mathbf{v} \neq \mathbf{0}$.

(a) False. Let $S \in \mathrm{L}(\mathbb{R}^2)$ have matrix $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Then $S\mathbf{e}_1 = S\mathbf{e}_2 = \mathbf{e}_1$, a unit vector, but $S(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$ while $\|\mathbf{e}_1 - \mathbf{e}_2\| = \sqrt{2}$. The images of the basis have length 1 but are not orthogonal to each other, and (V) of Theorem 8.4.4 needs both.

(b) True. Let $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{v}/\|\mathbf{v}\|$ is a unit vector, so by hypothesis and homogeneity of the norm,

$$\|S\mathbf{v}\| = \|\mathbf{v}\| \cdot \left\| S\left(\frac{\mathbf{v}}{\|\mathbf{v}\|}\right) \right\| = \|\mathbf{v}\|.$$

For $\mathbf{v} = \mathbf{0}$ both sides are 0. So S is an isometry.

Solution to Exercise 8.4.17

Uses. Theorem 8.4.15 (Spectral form of unitary operators), Theorem 2.2.1 (Linear maps are determined by a basis), Proposition 6.3.34 (Characterizations of unitary matrices)
Hint 1. The vectors $(1,1)/\sqrt{2}$ and $(1,-1)/\sqrt{2}$ form an orthonormal basis. Put them in the columns of W.
Hint 2. $U = W\,\mathrm{diag}(i,-1)W^*$. Powers and adjoints act on the diagonal factor.
Sketch. U is determined by its values on a basis. With $W = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = W^*$, $U = W\,\mathrm{diag}(i,-1)W = \frac{1}{2}\begin{pmatrix} i-1 & i+1 \\ i+1 & i-1 \end{pmatrix}$. It is unitary by (III) of Theorem 8.4.15. $U^4 = W\,\mathrm{diag}(1,1)W = I_2$ and $U^* = W\,\mathrm{diag}(-i,-1)W$.

An operator is determined by its values on the basis $(1,1),(1,-1)$ (Theorem 2.2.1), so there is exactly one such U. Let $W = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$; its columns are orthonormal and $W = W^* = W^{-1}$. The columns of W are eigenvectors for i and -1, so $U = W\,\mathrm{diag}(i,-1)W$, and

$$U = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} i-1 & i+1 \\ i+1 & i-1 \end{pmatrix}.$$

It has an orthonormal eigenbasis with eigenvalues i and -1 of absolute value 1, so it is unitary by (III) $\Rightarrow$ (I) of Theorem 8.4.15. Next, $U^4 = W\,\mathrm{diag}\big(i^4,(-1)^4\big)W = WW = I_2$. Finally $U^* = W^*\mathrm{diag}(i,-1)^*W^* = W\,\mathrm{diag}(-i,-1)W$, which is

$$U^* = \frac{1}{2}\begin{pmatrix} -i-1 & -i+1 \\ -i+1 & -i-1 \end{pmatrix},$$

the conjugate transpose of U, as it must be. As a check, $U^* = U^{-1} = U^3$, since $U^4 = I_2$.

Solution to Exercise 8.4.22

Uses. Proposition 8.4.21 (A reflection exchanging two vectors), Example 8.4.20 (Householder reflections)
Hint 1. Both targets have length $3 = \|(2,1,2)\|$, and all entries are real. Use Proposition 8.4.21 with $\mathbf{u}$ parallel to $\mathbf{x} - \mathbf{y}$.
Hint 2. $H_{\mathbf{u}} = I_3 - 2\mathbf{u}\mathbf{u}^{\mathsf{T}} = I_3 - \big(2/\|\mathbf{d}\|^2\big)\mathbf{d}\mathbf{d}^{\mathsf{T}}$ for any nonzero $\mathbf{d}$ parallel to $\mathbf{u}$.
Sketch. For $(3,0,0)$: $\mathbf{d} = (-1,1,2)$, $\|\mathbf{d}\|^2 = 6$, $H = \frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ 2 & -2 & -1 \end{pmatrix}$. For $(-3,0,0)$: $\mathbf{d} = (5,1,2)$, $\|\mathbf{d}\|^2 = 30$, $H = \frac{1}{15}\begin{pmatrix} -10 & -5 & -10 \\ -5 & 14 & -2 \\ -10 & -2 & 11 \end{pmatrix}$.

Let $\mathbf{x} = (2,1,2)$, so $\|\mathbf{x}\| = 3$. Both targets have length 3 and real dot products with $\mathbf{x}$, so Proposition 8.4.21 applies with $\mathbf{u} = \mathbf{d}/\|\mathbf{d}\|$, $\mathbf{d} = \mathbf{x} - \mathbf{y}$, and $H = I_3 - \big(2/\|\mathbf{d}\|^2\big)\mathbf{d}\mathbf{d}^{\mathsf{T}}$.

For $\mathbf{y} = (3,0,0)$: $\mathbf{d} = (-1,1,2)$ and $\|\mathbf{d}\|^2 = 6$, so

$$H = I_3 - \frac{1}{3}\begin{pmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ 2 & -2 & -1 \end{pmatrix}.$$

Indeed

$$H\mathbf{x} = \frac{1}{3}(4+1+4,\, 2+2-4,\, 4-2-2) = (3,0,0).$$

The matrix is symmetric by inspection. Its columns $\frac{1}{3}(2,1,2)$, $\frac{1}{3}(1,2,-2)$, $\frac{1}{3}(2,-2,-1)$ have squared lengths $9/9 = 1$ and pairwise dot products $(2+2-4)/9$, $(4-2-2)/9$, $(2-4+2)/9$, all 0; so H is orthogonal.

For $\mathbf{y} = (-3, 0, 0)$: $\mathbf{d} = (5, 1, 2)$ and $\|\mathbf{d}\|^2 = 30$, so

$$H = I_3 - \frac{1}{15}\begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix} = \frac{1}{15}\begin{pmatrix} -10 & -5 & -10 \\ -5 & 14 & -2 \\ -10 & -2 & 11 \end{pmatrix},$$

and

$$H\mathbf{x} = \frac{1}{15}(-20 - 5 - 20, -10 + 14 - 4, -20 - 2 + 22) = (-3, 0, 0).$$

(In floating-point computation one chooses the sign of the target that makes $\|\mathbf{d}\|$ large, here the second, to avoid subtracting nearly equal numbers.)

Solution to Exercise 8.4.32

Uses. Theorem 8.4.4 (Characterizations of isometries), Theorem 8.4.11 (Characterizations of unitary operators), Definition 8.4.1 (Isometry), Proposition 6.1.16 (Frobenius inner product)
Hint 1. For matrices, check whether the columns are orthonormal. For the polynomial maps, compute $\|Tp\|^2$ from the definition of the inner product.
Hint 2. Finite-dimensional isometries $V \to V$ are unitary; an isometry between spaces of different dimension is not even an operator.
Sketch. (a) Orthonormal columns $\frac{1}{3}(1, 2, 2)$, $\frac{1}{3}(2, 1, -2)$: an isometry, not unitary. (b) Columns $\frac{1}{\sqrt{2}}(1, i)$, $\frac{1}{\sqrt{2}}(i, 1)$ are orthonormal: unitary. (c) $\|Tp\|^2 = p(1)^2 + p(0)^2$: unitary. (d) $p = x$ has $\|Tp\|^2 = 1/4 \neq 1$: no. (e) $\left\|A^{\mathsf{T}}\right\|^2 = \sum |a_{ji}|^2 = \|A\|^2$: unitary.

(a) The matrix of T is $\frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$, with columns of squared length $(1 + 4 + 4)/9 = 1$ and $(4 + 1 + 4)/9 = 1$ and dot product $(2 + 2 - 4)/9 = 0$. So T is an isometry by Theorem 8.4.4 (with the standard bases). It maps $\mathbb{R}^2$ into $\mathbb{R}^3$, so it is not an operator, and not unitary; its image is a plane.

(b) The matrix is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$. Its columns $\frac{1}{\sqrt{2}}(1, i)$ and $\frac{1}{\sqrt{2}}(i, 1)$ have squared length $(1 + 1)/2 = 1$, and their dot product is $\frac{1}{2}\left(1 \cdot \bar{i} + i \cdot \bar{1}\right) = \frac{1}{2}(-i + i) = 0$. So T is an isometry of $\mathbb{C}^2$, hence unitary (Theorem 8.4.11).

(c) $(Tp)(0) = p(1)$ and $(Tp)(1) = p(0)$, so $\|Tp\|^2 = p(1)^2 + p(0)^2 = \|p\|^2$. So T is an isometry of the finite-dimensional space V, hence unitary.

(d) For $p = x$, $\|p\|^2 = 0^2 + 1^2 = 1$, while $Tp = x/2$ has $\|Tp\|^2 = 0 + 1/4$. So T is not an isometry.

(e) T is linear, and by Proposition 6.1.16, $\left\|A^{\mathsf{T}}\right\|^2 = \sum_{i,j} |a_{ji}|^2 = \sum_{i,j} |a_{ij}|^2 = \|A\|^2$. So T is an isometry of a finite-dimensional space, hence unitary. (It also satisfies $T^2 = \mathrm{id}$, so $T^{-1} = T = T^*$: it is self-adjoint as well.)

Solution to Exercise 8.4.33

Uses. Corollary 8.4.6 (Properties of isometries), Theorem 8.4.15 (Spectral form of unitary operators), Theorem 8.2.5 (Complex spectral theorem), Example 8.1.9 (Normal operators)
Hint 1. (a) An isometry $\mathbb{R} \to \mathbb{R}^2$ cannot have an injective adjoint. (b) Look for a non-normal operator with a single eigenvalue 1.
Hint 2. (c) The spectral theorem supplies an orthonormal basis of eigenvectors. (d) A real normal operator may have no eigenvectors at all.
Sketch. (a) False: $Sx = (x, 0)$ has $S^*(y_1, y_2) = y_1$, which kills $\mathbf{e}_2$. (b) False: the shear. (c) True: an orthonormal eigenbasis exists, and $\|T\mathbf{e}_j\| = |\lambda_j| = 1$, so (III) of Theorem 8.4.15 holds. (d) False: $\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$ is normal, has no real eigenvalue, and is not orthogonal.

(a) False. Let $S : \mathbb{R} \to \mathbb{R}^2$, $Sx = (x, 0)$, an isometry. Its adjoint is $S^*(y_1, y_2) = y_1$, since $\langle Sx, (y_1, y_2)\rangle = xy_1$. Then $S^*\mathbf{e}_2 = 0$, so S^* is not injective and not an isometry (Corollary 8.4.6(a)). The statement is true when $\dim V = \dim W$, by Corollary 8.4.6(d).

(b) False. The shear $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ on $\mathbb{C}^2$ has the single eigenvalue 1 but maps $\mathbf{e}_2$ to $(1, 1)$, of length $\sqrt{2}$.

(c) True. By Theorem 8.2.5, V has an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ with $T\mathbf{e}_j = \lambda_j \mathbf{e}_j$. By hypothesis $1 = \|T\mathbf{e}_j\| = |\lambda_j|$, and every eigenvalue of T is one of the λ_j. So T is unitary by (III) $\Rightarrow$ (I) of Theorem 8.4.15.

(d) False. Let $T \in \mathrm{L}(\mathbb{R}^2)$ have matrix $A = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$. It is normal, since $A^{\mathsf{T}}A = 4I_2 = AA^{\mathsf{T}}$ (Example 8.1.9(c)). Its characteristic polynomial $x^2 + 4$ has no real root, so T has no eigenvectors and the hypothesis holds vacuously. But $\|T\mathbf{e}_1\| = 2$, so T is not orthogonal. In (c), the eigenvectors of a complex normal operator span V; over $\mathbb{R}$ they need not exist.

Solution to Exercise 8.4.34

Uses. Theorem 8.4.4 (Characterizations of isometries), Theorem 6.4.20 (Orthonormal bases exist)
Techniques. Plug in clever test vectors (Technique 6.1.40)
Hint 1. Fix an orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$. The images are orthogonal. You want them to have equal lengths.
Hint 2. For $i \neq j$, the vectors $\mathbf{e}_i + \mathbf{e}_j$ and $\mathbf{e}_i - \mathbf{e}_j$ are orthogonal. Apply the hypothesis to them.
Sketch. The images $T\mathbf{e}_i$ are pairwise orthogonal. From $(\mathbf{e}_i + \mathbf{e}_j) \perp (\mathbf{e}_i - \mathbf{e}_j)$ we get $0 = \langle T\mathbf{e}_i + T\mathbf{e}_j, T\mathbf{e}_i - T\mathbf{e}_j \rangle = \|T\mathbf{e}_i\|^2 - \|T\mathbf{e}_j\|^2$. So all $\|T\mathbf{e}_i\|$ equal some c; if $c > 0$, T/c maps an orthonormal basis to an orthonormal list.

If $V = \{\mathbf{0}\}$, then $T = 0$. Otherwise let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be an orthonormal basis of V (Theorem 6.4.20). For $i \neq j$, $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$, so $\langle T\mathbf{e}_i, T\mathbf{e}_j \rangle = 0$ by hypothesis. Also

$$\langle \mathbf{e}_i + \mathbf{e}_j, \mathbf{e}_i - \mathbf{e}_j \rangle = \|\mathbf{e}_i\|^2 - \langle \mathbf{e}_i, \mathbf{e}_j \rangle + \langle \mathbf{e}_j, \mathbf{e}_i \rangle - \|\mathbf{e}_j\|^2 = 1 - 0 + 0 - 1 = 0,$$

so by hypothesis and the orthogonality of $T\mathbf{e}_i$ and $T\mathbf{e}_j$,

$$0 = \langle T\mathbf{e}_i + T\mathbf{e}_j, T\mathbf{e}_i - T\mathbf{e}_j \rangle = \|T\mathbf{e}_i\|^2 - \|T\mathbf{e}_j\|^2.$$

So $\|T\mathbf{e}_i\| = c$ for one $c \geq 0$ and all i. If $c = 0$, then T vanishes on a basis and $T = 0$. If $c > 0$, let $S = c^{-1}T$. Then $S\mathbf{e}_1, \dots, S\mathbf{e}_n$ are pairwise orthogonal unit vectors, so S is an isometry by (V) $\Rightarrow$ (I) of Theorem 8.4.4, and $T = cS$. Conversely, cS preserves orthogonality since $\langle cS\mathbf{u}, cS\mathbf{v} \rangle = c^2 \langle \mathbf{u}, \mathbf{v} \rangle$.

Solution to Exercise 8.4.35

Uses. Definition 6.3.33 (Unitary and orthogonal matrices), Theorem 6.3.29 (Orthonormal columns), Definition 5.2.29 (Permutation matrix), Theorem 5.4.10 (Multiplicativity)
Techniques. A sum of nonnegative terms (Technique 6.1.9)
Hint 1. A column is a unit vector with integer entries. How can integers have squares adding up to 1?
Hint 2. Orthogonal columns with one nonzero entry each must have it in different rows. For the determinant, multiplying the first row by -1 is a bijection between the two kinds.
Sketch. Each column has integer entries with $\sum q_{ij}^2 = 1$, so exactly one entry is ± 1. Two columns with their nonzero entries in the same row would have dot product ± 1, so the rows of these entries are distinct: $Q = DP$ with P a permutation matrix and D diagonal with entries ± 1. Count: $n!$ permutations times 2^n signs. Negating the first row is a bijection between determinant 1 and -1, so $2^{n-1}n!$ have determinant 1.

(a) Let $\mathbf{q}_j$ be the jth column. It is a unit vector (Theorem 6.3.29), so $\sum_i q_{ij}^2 = 1$ with every $q_{ij} \in \mathbb{Z}$. A nonzero integer has square at least 1, so exactly one entry of $\mathbf{q}_j$ is nonzero, and it is ± 1; say it is in row $\sigma(j)$. If $\sigma(j) = \sigma(k)$ with $j \neq k$, then $\mathbf{q}_j \cdot \mathbf{q}_k = \pm 1 \neq 0$, contradicting orthogonality. So σ is injective, hence a permutation of $\{1, \dots, n\}$, and every row also contains exactly one nonzero entry.

(b) Conversely, every signed permutation matrix has orthonormal columns (they are $\pm \mathbf{e}_{\sigma(j)}$ for distinct $\sigma(j)$), so it is orthogonal with integer entries. Such a matrix is determined by the permutation σ ($n!$ choices) and the signs of its n nonzero entries (2^n choices), so there are $2^n n!$ of them. Let $D = \mathrm{diag}(-1, 1, \dots, 1)$. The map $Q \mapsto DQ$ negates the first row, so it maps signed permutation matrices to signed permutation matrices, it is its own inverse, and $\det(DQ) = -\det Q$ (Theorem 5.4.10). So it exchanges those of determinant 1 and those of determinant -1 (the only possible values), and exactly half, $2^{n-1}n!$, have determinant 1.

Solution to Exercise 8.4.36

Uses. Theorem 6.2.1 (Cauchy–Schwarz inequality), Theorem 8.4.11 (Characterizations of unitary operators), Definition 6.3.33 (Unitary and orthogonal matrices), Proposition 6.3.34 (Characterizations of unitary matrices)
Techniques. Apply Cauchy–Schwarz to a well-chosen inner product (Technique 6.2.21)

Hint 1. The sum of all entries is $\mathbf{1}^{\mathsf{T}}Q\mathbf{1} = \langle Q\mathbf{1}, \mathbf{1}\rangle$, where $\mathbf{1} = (1, \ldots, 1)$.
Hint 2. Apply Cauchy–Schwarz, and use $\|Q\mathbf{1}\| = \|\mathbf{1}\| = \sqrt{n}$. For (b), apply Cauchy–Schwarz to each row, a unit vector, and the vector $\mathbf{1}$.
Sketch. (a) $\sum q_{ij} = \langle Q\mathbf{1}, \mathbf{1}\rangle$, and $|\langle Q\mathbf{1}, \mathbf{1}\rangle| \le \|Q\mathbf{1}\|\|\mathbf{1}\| = n$. Equality $= n$ holds iff $Q\mathbf{1}$ is a nonnegative multiple of $\mathbf{1}$ of the same length, i.e. $Q\mathbf{1} = \mathbf{1}$. (b) For row i, $\sum_j |q_{ij}| = (|q_{i1}|, \ldots) \cdot \mathbf{1} \le 1 \cdot \sqrt{n}$, with equality iff all $|q_{ij}|$ are equal. Example: $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Let $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^n$, so $\|\mathbf{1}\| = \sqrt{n}$, and $\|Q\mathbf{1}\| = \sqrt{n}$ since Q is orthogonal.

(a) The ith entry of $Q\mathbf{1}$ is the ith row sum, so $\sum_{i,j} q_{ij} = (Q\mathbf{1}) \cdot \mathbf{1}$. By Cauchy–Schwarz (Theorem 6.2.1),

$$\left|\sum_{i,j} q_{ij}\right| = |(Q\mathbf{1}) \cdot \mathbf{1}| \le \|Q\mathbf{1}\|\|\mathbf{1}\| = n.$$

If every row sums to 1, then $Q\mathbf{1} = \mathbf{1}$ and the sum is $\mathbf{1} \cdot \mathbf{1} = n$. Conversely, if the sum is $n = \|Q\mathbf{1}\|\|\mathbf{1}\|$, then by Theorem 6.2.11(b) (the case of equality $\langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x}\|\|\mathbf{y}\|$), one of $Q\mathbf{1}$ and $\mathbf{1}$ is a nonnegative multiple of the other. They have the same nonzero length, so $Q\mathbf{1} = \mathbf{1}$: every row sums to 1.

(b) Let $\mathbf{r}_i$ be the ith row of Q, a unit vector by Proposition 6.3.34, and let $\mathbf{a}_i = (|q_{i1}|, \ldots, |q_{in}|)$, which has the same length. By Cauchy–Schwarz,

$$\sum_j |q_{ij}| = \mathbf{a}_i \cdot \mathbf{1} \le \|\mathbf{a}_i\|\|\mathbf{1}\| = \sqrt{n},$$

and summing over the n rows gives $\sum_{i,j} |q_{ij}| \le n\sqrt{n}$. Equality holds if and only if it holds for every row, that is (as $\mathbf{a}_i \neq \mathbf{0}$ and both vectors have nonnegative entries), $\mathbf{a}_i$ is a positive multiple of $\mathbf{1}$ for each i. Since $\|\mathbf{a}_i\| = 1$, this says $|q_{ij}| = 1/\sqrt{n}$ for all i, j. For $n = 2$, the matrix $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is orthogonal with all $|q_{ij}| = 1/\sqrt{2}$, and its entries have absolute values adding up to $4/\sqrt{2} = 2\sqrt{2}$.

Solution to Exercise 8.4.37

Uses. Theorem 8.4.15 (Spectral form of unitary operators), Theorem 7.5.22 (Trace and determinant from eigenvalues), Theorem B.1.11 (Triangle inequality in $\mathbb{C}$)
Hint 1. The trace is the sum of the eigenvalues, each on the unit circle.
Hint 2. For equality, the triangle inequality in $\mathbb{C}$ forces all λ_j to point in the same direction. Then use the spectral form.
Sketch. $\operatorname{tr} U = \sum \lambda_j$ with $|\lambda_j| = 1$, so $|\operatorname{tr} U| \le n$. Equality in the triangle inequality for complex numbers of equal modulus forces all λ_j equal to one c; then $U = W(cI_n)W^* = cI_n$. For real Q, $\operatorname{tr} Q = n$ gives $c = 1$.

By Theorem 8.4.15, $U = WDW^*$ with W unitary and $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, $|\lambda_j| = 1$; and $\operatorname{tr} U = \lambda_1 + \cdots + \lambda_n$ (Theorem 7.5.22). By the triangle inequality in $\mathbb{C}$ (Theorem B.1.11),

$$|\operatorname{tr} U| \le |\lambda_1| + \cdots + |\lambda_n| = n.$$

If $U = cI_n$ with $|c| = 1$, then $|\operatorname{tr} U| = n|c| = n$. Conversely, if $|\operatorname{tr} U| = n$, equality holds in the triangle inequality, so by its equality case all nonzero λ_j are positive real multiples of one another. They all have absolute value 1, so they are all equal to one c with $|c| = 1$. Then $D = cI_n$ and $U = W(cI_n)W^* = cWW^* = cI_n$.

If Q is real orthogonal with $\operatorname{tr} Q = n$, it is unitary as a complex matrix, and $|\operatorname{tr} Q| = n$, so $Q = cI_n$ with $nc = \operatorname{tr} Q = n$, that is, $Q = I_n$.

Solution to Exercise 8.4.38

Uses. Theorem 8.4.15 (Spectral form of unitary operators), Theorem 8.2.18 (Spectral decomposition), Definition 8.2.23 (Functions of an operator), Proposition 8.2.24 (Properties of $f(T)$), Theorem 2.2.1 (Linear maps are determined by a basis), Definition 7.2.13 (Eigenspace)
Techniques. Apply a function to the eigenvalues (Technique 8.2.25)
Hint 1. Write $S = \sum_j \lambda_j P_j$ with distinct $\lambda_j = e^{i\theta_j}$, $\theta_j \in [0, 2\pi)$. Take $A = \sum_j \theta_j P_j$.
Hint 2. For (b), take square roots of the eigenvalues: $R = \sum_j e^{i\theta_j/2} P_j$.
Hint 3. For (c), any square root R commutes with $S = R^2$, so it maps each eigenline of S into itself.
Sketch. (a) $A = \sum \theta_j P_j$ is self-adjoint with the same eigenspaces, and $f(A) = \sum e^{i\theta_j} P_j = S$. (b) $R = \sum e^{i\theta_j/2} P_j$ is unitary by Theorem 8.4.15 and $R^2 = S$ by Proposition 8.2.24. (c) $RS = R^3 = SR$, so R preserves each eigenline $\operatorname{span}(\mathbf{e}_j)$: $R\mathbf{e}_j = r_j \mathbf{e}_j$

with $r_j^2 = \lambda_j$, two choices each. Each of the 2^n choices defines a square root with an orthonormal eigenbasis and $|r_j| = 1$, hence unitary.

(a) By Theorem 8.4.15 and Theorem 8.2.18, $S = \sum_{j=1}^{k} \lambda_j P_j$, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues, $P_j = \mathrm{Pr}_{E_{\lambda_j}}$, and $|\lambda_j| = 1$. Write $\lambda_j = e^{i\theta_j}$ with $\theta_j \in [0, 2\pi)$; the θ_j are distinct since the λ_j are. Let $A = \sum_j \theta_j P_j$. By Proposition 8.2.24 (applied to S and the real-valued function $\lambda_j \mapsto \theta_j$), A is self-adjoint, its distinct eigenvalues are $\theta_1, \ldots, \theta_k$, and its eigenspaces are $E_{\lambda_j}(S)$. So by Definition 8.2.23 applied to A, $f(A) = \sum_j e^{i\theta_j} P_j = S$.

(b) Let $R = \sum_j e^{i\theta_j/2} P_j$, the function $\lambda_j \mapsto e^{i\theta_j/2}$ applied to S. By Proposition 8.2.24, $R^2 = \sum_j \left(e^{i\theta_j/2}\right)^2 P_j = S$, and R is normal with eigenvalues $e^{i\theta_j/2}$ of absolute value 1, so R is unitary by Theorem 8.4.15.

(c) Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be an orthonormal eigenbasis of S with $S\mathbf{e}_j = \lambda_j \mathbf{e}_j$ and distinct λ_j. Suppose $R^2 = S$. Then $RS = R^3 = SR$, so $S(R\mathbf{e}_j) = RS\mathbf{e}_j = \lambda_j R\mathbf{e}_j$. Thus $R\mathbf{e}_j \in E_{\lambda_j}(S) = \mathrm{span}(\mathbf{e}_j)$, which is one-dimensional because the eigenvalues are distinct; say $R\mathbf{e}_j = r_j \mathbf{e}_j$. Then $\lambda_j \mathbf{e}_j = R^2 \mathbf{e}_j = r_j^2 \mathbf{e}_j$, so r_j is one of the two square roots of λ_j. Conversely, for each of the 2^n choices of square roots $r_1, \ldots, r_n$, Theorem 2.2.1 gives exactly one R with $R\mathbf{e}_j = r_j \mathbf{e}_j$, and $R^2 \mathbf{e}_j = r_j^2 \mathbf{e}_j = S\mathbf{e}_j$ for all j, so $R^2 = S$. Different choices give different R, so there are exactly 2^n square roots. Each has an orthonormal eigenbasis with eigenvalues r_j, where $|r_j| = |\lambda_j|^{1/2} = 1$, so each is unitary by Theorem 8.4.15.

Solution to Exercise 8.4.39

Uses. Theorem 8.2.12 (Real spectral theorem), Theorem 8.2.5 (Complex spectral theorem), Proposition 8.4.10 (Unitary means $S^* = S^{-1}$), Corollary 8.4.13 (Eigenvalues and determinant), Proposition 8.2.2 (Necessary conditions)
Hint 1. Which spectral theorem applies over $\mathbb{R}$, and what is its hypothesis?
Hint 2. For (b), the real spectral theorem says: orthonormal eigenbasis iff self-adjoint. For orthogonal S, $S^* = S^{-1}$.
Sketch. (a) Over $\mathbb{R}$ the spectral theorem needs self-adjointness, not normality; $R_{\pi/2}$ is orthogonal with no eigenvectors. (b) S has an orthonormal eigenbasis iff $S = S^*$ (Theorem 8.2.12) iff $S = S^{-1}$ iff $S^2 = \mathrm{id}$. Such S fixes E_1 and negates $E_{-1} = E_1^{\perp}$: it is the reflection through E_1.

(a) The normal spectral theorem (Theorem 8.2.5) is a theorem about complex spaces. Over $\mathbb{R}$ the spectral theorem (Theorem 8.2.12) requires S to be self-adjoint, and an orthogonal operator need not be. The rotation $R_{\pi/2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is orthogonal, but its characteristic polynomial $x^2 + 1$ has no real root, so it has no eigenvectors at all.

(b) By Theorem 8.2.12 (with Proposition 8.2.2 for the forward direction), S has an orthonormal eigenbasis if and only if $S^* = S$. Since $S^* = S^{-1}$ (Proposition 8.4.10), this says $S = S^{-1}$, that is, $S^2 = \mathrm{id}$. In that case the eigenvalues are real of absolute value 1 (Corollary 8.4.13(a)), so $V = E_1 \oplus E_{-1}$ with the two eigenspaces orthogonal. So S fixes the subspace E_1 and negates its orthogonal complement E_{-1}; it is the reflection through E_1, that is, $S = 2\,\mathrm{Pr}_{E_1} - \mathrm{id}$ (compare Exercise 8.1.30).

Solution to Exercise 8.4.40

Uses. Example 8.4.12 (The shift), Definition 6.6.8 (Adjoint), Definition 7.2.6 (Eigenvalue, eigenvector), Definition 6.4.11 (Orthogonal projection)
Hint 1. Write $p = \sum a_k x^k$ and $q = \sum b_k x^k$, and compute $\langle xp, q\rangle$ and $\langle p, Lq\rangle$ coefficient by coefficient.
Hint 2. For the eigenvalues, compare degrees: M raises the degree of a nonzero polynomial by 1, and L lowers it by 1 (or kills constants).
Sketch. (a) $\langle Mp, q\rangle = \sum_k a_k \overline{b_{k+1}} = \langle p, Lq\rangle$. (b) $L(xp) = p$; $xL(q) = q - b_0$, and $b_0 = \mathrm{Pr}_{\mathrm{span}(1)}\, q$ since $1, x, x^2, \ldots$ is orthonormal. (c) $\deg(xp) = \deg p + 1$, so $xp = \lambda p$ is impossible for $p \neq \mathbf{0}$. If $Lq = \lambda q$ with $q \neq \mathbf{0}$ of degree d: for $d \geq 1$, $\deg Lq = d - 1 < d$ forces $\lambda = 0$ and $Lq = 0$, a contradiction; so q is a nonzero constant, $Lq = 0$, and $\lambda = 0$.

(a) Let $p = \sum_k a_k x^k$ and $q = \sum_k b_k x^k$ (finite sums, padded with zeros). Then $Mp = \sum_k a_k x^{k+1}$ and $Lq = \sum_k b_{k+1} x^k$, so

$$\langle Mp, q\rangle = \sum_{k \geq 0} a_k \overline{b_{k+1}} = \langle p, Lq\rangle.$$

So L is the adjoint of M (Definition 6.6.8).

(b) $L(Mp) = L\left(\sum_k a_k x^{k+1}\right) = \sum_k a_k x^k = p$, so $LM = \mathrm{id}$. For q as above, $MLq = \sum_{k\geq 1} b_k x^k = q - b_0$. Since $1, x, x^2, \ldots$ is orthonormal, $b_0 = \langle q, 1\rangle$, so $b_0 = \mathrm{Pr}_{\mathrm{span}(1)}\, q$ (Definition 6.4.11), and $ML = \mathrm{id} - \mathrm{Pr}_{\mathrm{span}(1)}$. In particular $ML \neq \mathrm{id}$, as $ML1 = \mathbf{0}$: L is a left inverse of M but not a right inverse.

(c) If $p \neq \mathbf{0}$, then $\deg(xp) = \deg p + 1$, while λp is $\mathbf{0}$ or has degree $\deg p$. So $Mp = \lambda p$ never holds with $p \neq \mathbf{0}$, and M has no eigenvalue. For L: the constant 1 satisfies $L1 = \mathbf{0} = 0 \cdot 1$, so 0 is an eigenvalue. Suppose $Lq = \lambda q$ with $q \neq \mathbf{0}$. If $\deg q = d \geq 1$, then Lq is $\mathbf{0}$ or has degree $d-1$, so λq cannot have degree d; hence $\lambda q = \mathbf{0}$, $\lambda = 0$, and $Lq = \mathbf{0}$, which forces q to be constant, contradicting $d \geq 1$. So q is a nonzero constant, $Lq = \mathbf{0}$, and $\lambda = 0$.

Solution to Exercise 8.4.41

Uses. Definition 8.4.26 (Rotation), Theorem 8.4.27 (Euler's rotation theorem), Theorem 6.3.29 (Orthonormal columns)
Hint 1. Check that the columns are orthonormal and that $\det Q = 1$.
Hint 2. The axis spans $\ker(Q - I_3)$. The angle comes from $\mathrm{tr}\, Q = 1 + 2\cos\theta$.
Sketch. The columns $\frac{1}{9}(1,8,-4)$, $\frac{1}{9}(-4,4,7)$, $\frac{1}{9}(8,1,4)$ are orthonormal and $\det(9Q) = 729$, so $\det Q = 1$. $Q(1,2,2) = (1,2,2)$, so the axis is $\frac{1}{3}(1,2,2)$, and $\mathrm{tr}\, Q = 1 = 1 + 2\cos\theta$ gives $\theta = \pi/2$.

The columns of $9Q$ are $(1,8,-4)$, $(-4,4,7)$, $(8,1,4)$, each of squared length 81, with dot products $-4 + 32 - 28 = 0$, $8 + 8 - 16 = 0$, and $-32 + 4 + 28 = 0$. So the columns of Q are orthonormal and Q is orthogonal. Expanding along the first row,

$$\det(9Q) = 1(16-7) + 4(32+4) + 8(56+16) = 9 + 144 + 576 = 729 = 9^3,$$

so $\det Q = 1$ and Q is a rotation (Definition 8.4.26).

For the axis, $9Q(1,2,2) = (1 - 8 + 16, 8 + 8 + 2, -4 + 14 + 8) = (9, 18, 18)$, so $Q(1,2,2) = (1,2,2)$, and the axis is $\mathbf{a} = \frac{1}{3}(1,2,2)$. (Since $Q \neq I_3$, the fixed space is exactly this line, by the matrix form in Theorem 8.4.27.) Finally $\mathrm{tr}\, Q = (1+4+4)/9 = 1 = 1 + 2\cos\theta$, so $\cos\theta = 0$ and Q is a quarter turn about $\mathbf{a}$.

Solution to Exercise 8.4.42

Uses. Theorem 5.4.10 (Multiplicativity), Theorem 5.4.14 (The transpose), Definition 5.1.1 (Determinant function), Theorem 7.2.9 (Eigenvalues and invertibility), Lemma 8.4.23 (Orthogonal 2×2 matrices), Proposition 7.3.15 (Block-triangular matrices)
Hint 1. Imitate the first proof of Theorem 8.4.27: write $Q + I_n = Q + QQ^{\mathsf{T}}$.
Hint 2. For (c), use rotation blocks. For the second part, recall what every element of $\mathrm{O}(2)$ with determinant -1 is.
Sketch. (a) $\det(Q+I) = \det\left(Q\left(I + Q^{\mathsf{T}}\right)\right) = \det Q \det(I+Q) = -\det(Q+I)$, so $\det(Q+I) = 0$. (b) $\det(Q-I) = \det Q \det\left(I - Q^{\mathsf{T}}\right) = (-1)^n \det(Q-I) = -\det(Q-I)$. (c) $\mathrm{diag}\left(R_{\pi/2}, R_{\pi/2}\right)$ has characteristic polynomial $\left(x^2+1\right)^2$. Every element of $\mathrm{O}(2)$ with determinant -1 is a reflection S_θ, which has eigenvalue 1: none exists.

(a) Using $QQ^{\mathsf{T}} = I_n$, Theorem 5.4.10, and Theorem 5.4.14,

$$\det(Q + I_n) = \det\left(Q + QQ^{\mathsf{T}}\right) = \det Q \cdot \det\left(I_n + Q^{\mathsf{T}}\right) = \det Q \cdot \det(I_n + Q) = -\det(Q + I_n).$$

So $\det(Q - (-1)I_n) = 0$, and -1 is an eigenvalue (Theorem 7.2.9).

(b) Similarly, $\det(Q - I_n) = \det\left(Q - QQ^{\mathsf{T}}\right) = \det Q \cdot \det\left(I_n - Q^{\mathsf{T}}\right) = \det(I_n - Q)$. By multilinearity in the n columns (Definition 5.1.1), $\det(I_n - Q) = (-1)^n \det(Q - I_n) = -\det(Q - I_n)$ for odd n. So $\det(Q - I_n) = 0$, and 1 is an eigenvalue.

(c) Let $Q = \begin{pmatrix} R_{\pi/2} & 0 \\ 0 & R_{\pi/2} \end{pmatrix} \in \mathbb{R}^{4\times 4}$. Its columns are $\mathbf{e}_2, -\mathbf{e}_1, \mathbf{e}_4, -\mathbf{e}_3$, which are orthonormal, and $\det Q = 1 \cdot 1 = 1$. By Proposition 7.3.15, $p_Q = \left(x^2+1\right)^2$, which has no real root, so Q has no real eigenvalue. For the second request, no such matrix exists: by Lemma 8.4.23, an element of $\mathrm{O}(2)$ with determinant -1 is a reflection S_θ, which has the eigenvalue 1.

Solution to Exercise 8.4.43

Uses. Definition 2.4.1 (Matrix product)
Hint 1. Write $S_\alpha = R_\alpha D$ with $D = \mathrm{diag}(1,-1)$, and check that $DR_\beta D = R_{-\beta}$.

Hint 2. For (b), given θ and β, choose α with $\alpha - \beta = \theta$.
Sketch. $S_\alpha = R_\alpha D$, so $S_\alpha S_\beta = R_\alpha D R_\beta D = R_\alpha R_{-\beta} = R_{\alpha-\beta}$. The mirrors are at angles $\alpha/2$ and $\beta/2$, which differ by half the rotation angle. Given θ and any S_β, $R_\theta = S_{\theta+\beta} S_\beta$.

(a) Let $D = \operatorname{diag}(1, -1)$. Multiplying out, $R_\alpha D = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = S_\alpha$, and

$$DR_\beta D = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} = R_{-\beta}.$$

So, since rotations compose by adding angles ($R_\alpha R_\gamma = R_{\alpha+\gamma}$, by the addition formulas),

$$S_\alpha S_\beta = R_\alpha D R_\beta D = R_\alpha R_{-\beta} = R_{\alpha-\beta}.$$

The mirror of S_α is the line at angle $\alpha/2$ (Lemma 8.4.23, or directly), and that of S_β is at angle $\beta/2$. The angle from the second mirror to the first is $(\alpha - \beta)/2$, and the composite rotates by $\alpha - \beta$, twice that.

(b) Given θ and any reflection S_β, let $\alpha = \theta + \beta$. By (a), $S_\alpha S_\beta = R_\theta$. Since every reflection of the plane through $\mathbf{0}$ is some S_β, the reflection applied first can be chosen freely.

Solution to Exercise 8.4.44

Uses. Example 8.4.20 (Householder reflections), Proposition 8.4.21 (A reflection exchanging two vectors), Corollary 6.4.25 (Dimension of the orthogonal complement), Theorem 4.2.11 (Rank–nullity), Theorem 5.4.10 (Multiplicativity), Theorem 8.4.27 (Euler's rotation theorem)
Techniques. Induction on dimension via an invariant subspace (Technique 7.5.12), Parity via an invariant (Technique 5.2.19)
Hint 1. (a) A vector orthogonal to every $\mathbf{u}_j$ is fixed by every factor. How large is the set of such vectors?
Hint 2. (b) Induct on r. If $Q \neq I_n$, pick $\mathbf{v}$ with $Q\mathbf{v} \neq \mathbf{v}$, and let H be the reflection with $HQ\mathbf{v} = \mathbf{v}$ (Proposition 8.4.21). Show that HQ fixes $\mathbf{v}$ and every vector fixed by Q.
Hint 3. For the fixed vectors: if $Q\mathbf{f} = \mathbf{f}$, then $\langle Q\mathbf{v} - \mathbf{v}, \mathbf{f}\rangle = \langle Q\mathbf{v}, Q\mathbf{f}\rangle - \langle \mathbf{v}, \mathbf{f}\rangle = 0$.
Sketch. (a) Q fixes $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}^\perp$, of dimension at least $n - k$, so $n - r = \dim\ker(Q - I) \geq n - k$; and each factor has determinant -1. (b) With $\mathbf{u} \propto Q\mathbf{v} - \mathbf{v}$, HQ fixes $\mathbf{v}$ and all of $F = \ker(Q - I)$ (since $F \perp \mathbf{u}$), so $\operatorname{rank}(HQ - I) \leq r - 1$; by induction HQ is a product of at most $r - 1$ reflections, so $Q = H(HQ)$ is a product of at most r, and by (a) of exactly r. (c) For a rotation $Q \neq I_3$ of $\mathbb{R}^3$, the fixed space is the axis, so $r = 2$, and $(-1)^k = 1$ forces k even.

(a) Let $X = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}^\perp = \operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)^\perp$. Each $H_{\mathbf{u}_j}$ fixes every vector orthogonal to $\mathbf{u}_j$ (Example 8.4.20), so $Q\mathbf{x} = \mathbf{x}$ for $\mathbf{x} \in X$, that is, $X \subseteq \ker(Q - I_n)$. By Corollary 6.4.25, $\dim X = n - \dim\operatorname{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k) \geq n - k$. By rank–nullity (Theorem 4.2.11), $\dim\ker(Q - I_n) = n - r$, so $n - r \geq n - k$, that is, $k \geq r$. Each $H_{\mathbf{u}_j}$ has determinant -1 (Example 8.4.20), so $\det Q = (-1)^k$ by Theorem 5.4.10.

(b) By (a), no product of fewer than r reflections equals Q, so it suffices to show that Q is a product of at most r reflections. We use strong induction on r. If $r = 0$, then $Q = I_n$, the empty product. Let $r \geq 1$, and let $F = \ker(Q - I_n)$, of dimension $n - r < n$. Pick $\mathbf{v} \notin F$, so $Q\mathbf{v} \neq \mathbf{v}$. Since $\|Q\mathbf{v}\| = \|\mathbf{v}\|$ and all entries are real, Proposition 8.4.21 gives a reflection $H = H_{\mathbf{u}}$, $\mathbf{u} = (Q\mathbf{v} - \mathbf{v})/\|Q\mathbf{v} - \mathbf{v}\|$, with $HQ\mathbf{v} = \mathbf{v}$. For $\mathbf{f} \in F$,

$$\langle Q\mathbf{v} - \mathbf{v}, \mathbf{f}\rangle = \langle Q\mathbf{v}, Q\mathbf{f}\rangle - \langle \mathbf{v}, \mathbf{f}\rangle = 0,$$

because $Q\mathbf{f} = \mathbf{f}$ and Q preserves inner products. So $\mathbf{f} \perp \mathbf{u}$, $H\mathbf{f} = \mathbf{f}$, and $HQ\mathbf{f} = \mathbf{f}$. Hence $Q' = HQ \in \mathrm{O}(n)$ fixes every vector of $F + \operatorname{span}(\mathbf{v})$, which has dimension $n - r + 1$ since $\mathbf{v} \notin F$. So $\operatorname{rank}(Q' - I_n) \leq r - 1$, and by the induction hypothesis Q' is a product of $\operatorname{rank}(Q' - I_n) \leq r - 1$ reflections. Since $H^2 = I_n$, $Q = HQ'$ is a product of at most r reflections.

(c) Let $Q \in \mathrm{SO}(3)$, $Q \neq I_3$. By Theorem 8.4.27, in an orthonormal basis starting with an axis Q has matrix $\operatorname{diag}(1, R_\theta)$ with $R_\theta \neq I_2$, that is, $\theta \neq 0$. Then $R_\theta - I_2$ is invertible (its determinant is $(\cos\theta - 1)^2 + \sin^2\theta = 2 - 2\cos\theta \neq 0$), so $\ker(Q - I_3)$ is the axis line and $r = 2$. By (b), Q is a product of two reflections. If $Q = H_{\mathbf{u}_1} \cdots H_{\mathbf{u}_k}$, then $1 = \det Q = (-1)^k$ by (a), so k is even.

Solution to Exercise 8.4.45

Uses. Theorem 8.4.31 (Distance-preserving maps are affine), Lemma 8.4.30 (Points between two points), Theorem 4.7.37 (Affine maps preserve affine combinations), Corollary B.1.12 (Two terms, and the reverse triangle inequality)
Hint 1. $||s| - |t|| \le |s - t|$ (the reverse triangle inequality in $\mathbb{R}$). For non-affineness, look at the midpoint of -1 and 1.
Hint 2. Lemma 8.4.30 uses equality in the triangle inequality. With the max norm, the "midpoint by distances" of $(-1, 1)$ and $(1, 1)$ is not unique.
Sketch. (a) $\|f(s) - f(t)\|_\infty = \max(|s - t|, ||s| - |t||) = |s - t|$. But $f(0) = (0, 0)$ is not the midpoint $(0, 1)$ of $f(-1) = (-1, 1)$ and $f(1) = (1, 1)$. (b) Lemma 8.4.30 fails: the points at max-distance 1 from both $(-1, 1)$ and $(1, 1)$ are $(0, y)$ with $0 \le y \le 2$, a whole segment, and $f(0)$ is one of them.

(a) For $s, t \in \mathbb{R}$, $f(s) - f(t) = (s - t, |s| - |t|)$, and $||s| - |t|| \le |s - t|$ by the reverse triangle inequality for real numbers (Corollary B.1.12). So $\|f(s) - f(t)\|_\infty = |s - t|$. If f were affine, it would preserve affine combinations (Theorem 4.7.37), in particular midpoints: $f(0) = \frac{1}{2}(f(-1) + f(1))$. But $f(0) = (0, 0)$, while $\frac{1}{2}((-1, 1) + (1, 1)) = (0, 1)$.

(b) Step 1 of the proof, which rests on Lemma 8.4.30, fails: in an inner product space the point $\mathbf{z}$ is determined by its distances to $\mathbf{a}$ and $\mathbf{b}$, but not for $\|\cdot\|_\infty$. Let $\mathbf{z} = (x, y)$. The conditions are

$$\max(|x + 1|, |y - 1|) = 1 \quad \text{and} \quad \max(|x - 1|, |y - 1|) = 1.$$

Both force $|x + 1| \le 1$ and $|x - 1| \le 1$, so $x = 0$; then $|x \pm 1| = 1$, and both conditions hold exactly when $|y - 1| \le 1$. So the solutions are the points $(0, y)$ with $0 \le y \le 2$, a whole segment; the midpoint $(0, 1)$ is one of them, and $f(0) = (0, 0)$ is another. The max norm has no strict equality case in the triangle inequality: $\|(1, 1)\|_\infty + \|(1, -1)\|_\infty = 2 = \|(2, 0)\|_\infty$, although neither vector is a multiple of the other.

Solution to Exercise 8.4.46

Uses. Theorem 8.4.31 (Distance-preserving maps are affine), Proposition 4.7.20 (Affine independence through differences), Theorem 2.2.1 (Linear maps are determined by a basis)
Techniques. Choose a base point (Technique 4.7.4)
Hint 1. By Theorem 8.4.31, $f = Q\mathbf{x} + \mathbf{b}$ and $g = Q'\mathbf{x} + \mathbf{b}'$. Measure positions from $\mathbf{p}_0$.
Hint 2. $f(\mathbf{p}_i) - f(\mathbf{p}_0) = Q(\mathbf{p}_i - \mathbf{p}_0)$. The differences $\mathbf{p}_i - \mathbf{p}_0$ form a basis.
Sketch. Write $f(\mathbf{x}) = Q\mathbf{x} + \mathbf{b}$ and $g(\mathbf{x}) = Q'\mathbf{x} + \mathbf{b}'$. Then $Q(\mathbf{p}_i - \mathbf{p}_0) = f(\mathbf{p}_i) - f(\mathbf{p}_0) = g(\mathbf{p}_i) - g(\mathbf{p}_0) = Q'(\mathbf{p}_i - \mathbf{p}_0)$, and these n differences are a basis, so $Q = Q'$; then $\mathbf{b} = f(\mathbf{p}_0) - Q\mathbf{p}_0 = \mathbf{b}'$.

By Theorem 8.4.31, $f(\mathbf{x}) = Q\mathbf{x} + \mathbf{b}$ and $g(\mathbf{x}) = Q'\mathbf{x} + \mathbf{b}'$ with Q, Q' orthogonal. For each i,

$$Q(\mathbf{p}_i - \mathbf{p}_0) = f(\mathbf{p}_i) - f(\mathbf{p}_0) = g(\mathbf{p}_i) - g(\mathbf{p}_0) = Q'(\mathbf{p}_i - \mathbf{p}_0).$$

By Proposition 4.7.20, the n vectors $\mathbf{p}_1 - \mathbf{p}_0, \dots, \mathbf{p}_n - \mathbf{p}_0$ are linearly independent, hence a basis of $\mathbb{R}^n$. The linear maps $\mathbf{x} \mapsto Q\mathbf{x}$ and $\mathbf{x} \mapsto Q'\mathbf{x}$ agree on this basis, so $Q = Q'$ (Theorem 2.2.1). Then $\mathbf{b} = f(\mathbf{p}_0) - Q\mathbf{p}_0 = g(\mathbf{p}_0) - Q'\mathbf{p}_0 = \mathbf{b}'$, so $f = g$.

In $\mathbb{R}^2$, three points are affinely independent exactly when they are not collinear. Applying the first part with $g = \mathrm{id}$ shows that a rigid motion fixing three non-collinear points is the identity. (Two points do not suffice: the reflection across the line through them fixes both.)

Section A.1

Solution to Exercise A.1.15

Uses. Theorem A.1.12 (One-sided inverses of functions), Theorem A.1.14 (Bijections and inverses)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. The construction of a left inverse is forced on $\mathrm{im}\, f$ and free elsewhere. When is the freedom not really a freedom?
Hint 2. Try X with one element.
Sketch. False. If $X = \{1\}$ and $Y = \{1, 2\}$, then every function $Y \to X$ is constant, so the inclusion $f : X \to Y$ has exactly one left inverse, but it is not surjective.

The claim is false. Let $X = \{1\}$, $Y = \{1, 2\}$, and $f(1) = 1$. There is exactly one function $Y \to X$, the constant function g with value 1, and $g(f(1)) = 1$, so f has exactly one left inverse. But $2 \notin \mathrm{im}\, f$, so f is not surjective.

The claim becomes true when X has at least two elements. A left inverse exists only if f is injective (Theorem A.1.12), and it is forced on $\operatorname{im} f$. If some $y \in Y$ is not in $\operatorname{im} f$, changing the value at y to any other element of X gives a second left inverse. So a unique left inverse forces $\operatorname{im} f = Y$, and f is bijective.

Solution to Exercise A.1.23

Uses. Definition A.1.3 (Image and preimage), Proposition A.1.5 (Images and preimages)
Hint 1. Complete the square: $f(x) = (x-1)^2 - 1$.
Hint 2. $f(x) \le 0$ means $(x-1)^2 \le 1$. On $[0,1]$, f decreases from 0 to -1.
Sketch. With $f(x) = (x-1)^2 - 1$: on $[0,3]$ the values fill $[-1,3]$; $f(x) \in [-1,0]$ iff $|x-1| \le 1$, so the preimage is $[0,2]$; $\operatorname{im}\big(f\restriction_{[0,1]}\big) = [-1,0]$, whose preimage is $[0,2] \supsetneq [0,1]$, illustrating (c) for a non-injective f.

Write $f(x) = (x-1)^2 - 1$. For $x \in [0,3]$, $x-1$ ranges over $[-1,2]$, so $(x-1)^2$ ranges over $[0,4]$ and $\operatorname{im}\big(f\restriction_{[0,3]}\big) = [-1,3]$.

Every value of f is at least -1, so $f(x) \in [-1,0]$ if and only if $(x-1)^2 \le 1$, that is, $|x-1| \le 1$. Hence $f^{-1}([-1,0]) = [0,2]$.

For $x \in [0,1]$, $(x-1)^2$ ranges over $[0,1]$, so $\operatorname{im}\big(f\restriction_{[0,1]}\big) = [-1,0]$, and by the previous computation $f^{-1}\big(\operatorname{im}\big(f\restriction_{[0,1]}\big)\big) = [0,2]$. This strictly contains $[0,1]$, as Proposition A.1.5(c) allows: f is not injective, since $f(1-t) = f(1+t)$.

Solution to Exercise A.1.24

Uses. Proposition A.1.5 (Images and preimages), Definition 2.6.1 (Injective)
Techniques. Contrapositive (Technique 1.5.4)
Hint 1. One direction is Proposition A.1.5(b).
Hint 2. For the other, suppose $f(x) = f(x')$ with $x \ne x'$, and choose A and A' as small as possible.
Sketch. If f is injective, this is Proposition A.1.5(b). If $f(x) = f(x')$ with $x \ne x'$, then $A = \{x\}$ and $A' = \{x'\}$ give $\operatorname{im}(f\restriction_{A\cap A'}) = \varnothing$ but $\operatorname{im}(f\restriction_A) \cap \operatorname{im}(f\restriction_{A'}) = \{f(x)\}$.

If f is injective, the equality is Proposition A.1.5(b). We prove the converse by contrapositive. Suppose f is not injective, say $f(x) = f(x')$ with $x \ne x'$. Let $A = \{x\}$ and $A' = \{x'\}$. Then $A \cap A' = \varnothing$, so $\operatorname{im}(f\restriction_{A\cap A'}) = \varnothing$, while

$$\operatorname{im}(f\restriction_A) \cap \operatorname{im}(f\restriction_{A'}) = \{f(x)\} \cap \{f(x')\} = \{f(x)\} \ne \varnothing.$$

So the equality fails for this pair.

Solution to Exercise A.1.25

Uses. Proposition A.1.8 (Composites, injectivity, and surjectivity)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. (b) is Proposition A.1.8. For (a) and (c), make Y bigger than X and Z.
Hint 2. Try $X = Z = \{1\}$ and $Y = \{1,2\}$.
Sketch. (a) False: $X = Z = \{1\}$, $Y = \{1,2\}$, $f(1) = 1$, g constant; $g \circ f = \mathrm{id}$ is injective, g is not. (b) True by Proposition A.1.8(b), (c). (c) False by the same example: f is not surjective and g is not injective.

Let $X = Z = \{1\}$, $Y = \{1,2\}$, $f(1) = 1$, and let $g : Y \to Z$ be the constant function with value 1. Then $g \circ f = \mathrm{id}_X$ is bijective.

(a) False: in the example, $g \circ f$ is injective but $g(1) = g(2)$.

(b) True. If $g \circ f$ is bijective, it is injective, so f is injective by Proposition A.1.8(b), and it is surjective, so g is surjective by Proposition A.1.8(c).

(c) False: in the example, f is not surjective ($2 \notin \operatorname{im} f$) and g is not injective. By (b), what is true is that f is injective and g surjective; if moreover $|X| = |Y| = |Z|$ are finite, then both are bijective by Theorem A.1.20.

Solution to Exercise A.1.26

Uses. Definition 2.6.1 (Injective), Definition 2.6.2 (Surjective), Definition A.1.2 (Function)
Techniques. Fix an arbitrary element and chase it (Technique 1.3.13)
Hint 1. For "injective $\Rightarrow$ cancellable", evaluate $f \circ g = f \circ h$ at each $w \in W$.

Hint 2. For the converse, test cancellability with $W = \{0\}$: a function $\{0\} \to X$ is just a choice of one element of X.
Hint 3. For surjectivity the statement is: f is surjective iff $g \circ f = h \circ f$ implies $g = h$ for all $g, h : Y \to Z$. For the converse, if $y_0 \notin \operatorname{im} f$, compare two functions $Y \to \{0,1\}$ that differ only at y_0.
Sketch. If f is injective and $f(g(w)) = f(h(w))$, then $g(w) = h(w)$. If $f(x) = f(x')$, the functions $g, h : \{0\} \to X$ with values x, x' satisfy $f \circ g = f \circ h$, so $x = x'$. Dually, if f is surjective, every y is $f(x)$, and $g(f(x)) = h(f(x))$; if $y_0 \notin \operatorname{im} f$, the indicator function of $\{y_0\}$ and the zero function $Y \to \{0,1\}$ agree after composing with f.

Suppose f is injective and $f \circ g = f \circ h$. For each $w \in W$, $f(g(w)) = f(h(w))$, so $g(w) = h(w)$. Hence $g = h$. Conversely, suppose f can be cancelled on the left, and let $f(x) = f(x')$. Define $g, h : \{0\} \to X$ by $g(0) = x$ and $h(0) = x'$. Then $f \circ g = f \circ h$, so $g = h$, that is, $x = x'$.

The analogue: f is surjective if and only if, for every set Z and all $g, h : Y \to Z$, $g \circ f = h \circ f$ implies $g = h$. Suppose f is surjective and $g \circ f = h \circ f$. Each $y \in Y$ equals $f(x)$ for some x, so $g(y) = g(f(x)) = h(f(x)) = h(y)$. Conversely, suppose f is not surjective, and let $y_0 \in Y \setminus \operatorname{im} f$. Let $g, h : Y \to \{0,1\}$ with $g(y) = 0$ for all y, and $h(y_0) = 1$, $h(y) = 0$ for $y \neq y_0$. For every $x \in X$, $f(x) \neq y_0$, so $g(f(x)) = 0 = h(f(x))$. Thus $g \circ f = h \circ f$ but $g \neq h$, and f cannot be cancelled on the right.

For matrices, Section 2.5 shows that a left-invertible matrix can be cancelled on the left. For functions, left cancellation is in fact equivalent to having a left inverse when $X \neq \emptyset$, by Theorem A.1.12(a).

Solution to Exercise A.1.27

Uses. Theorem A.1.12 (One-sided inverses of functions), Exercise A.1.28
Hint 1. A left inverse is forced on $\operatorname{im} f$, which has m elements, and arbitrary on the other $n - m$ elements of Y.
Hint 2. A right inverse is exactly a choice of one element $g(y_i) \in f^{-1}(y_i)$ for each i, and the choices are independent.
Sketch. (a) g is determined on $\operatorname{im} f$ (m points) and free on the remaining $n - m$ points, with m options each: m^{n-m}. (b) g is a right inverse iff $g(y_i) \in f^{-1}(y_i)$ for all i, and the choices are independent. (c) $2^2 = 4$ left inverses; $2 \cdot 2 = 4$ right inverses.

(a) A function $g : Y \to X$ is a left inverse of f if and only if $g(f(x)) = x$ for all x. Since f is injective, this fixes g at each of the m elements of $\operatorname{im} f$ and says nothing about the $n - m$ elements of $Y \setminus \operatorname{im} f$, where g may take any of the m values in X. By Exercise A.1.28 there are m^{n-m} such g.

(b) A function $g : Y \to X$ is a right inverse of f if and only if $f(g(y_i)) = y_i$ for each i, that is, $g(y_i) \in f^{-1}(y_i)$. So a right inverse is a choice of an element in each of the k fibers, independently, and there are $\left|f^{-1}(y_1)\right| \cdots \left|f^{-1}(y_k)\right|$ ways to choose.

(c) For f, $m = 2$ and $n = 4$, so there are $2^{4-2} = 4$ left inverses: $g(1) = 1$, $g(2) = 2$, and $g(3), g(4) \in \{1, 2\}$ arbitrary. For p, the fibers are $p^{-1}(1) = \{1, 3\}$ and $p^{-1}(2) = \{2, 4\}$, so there are $2 \cdot 2 = 4$ right inverses.

Solution to Exercise A.1.28

Uses. Corollary A.1.18 (Cardinality is well defined), Theorem A.1.14 (Bijections and inverses), Remark 1.2.11 (Tuples and matrices are functions)
Hint 1. For the step in (a), with $X = X' \cup \{x_0\}$ and $x_0 \notin X'$, a function $X \to Y$ is a function $X' \to Y$ together with a value at x_0.
Hint 2. The map $(h, y) \mapsto$ "h extended by $x_0 \mapsto y$" is a bijection $Y^{X'} \times Y \to Y^X$, and $|A \times B| = |A||B|$ (list the pairs row by row).
Hint 3. In (b), send A to its indicator function. In (c), an element of F^n is a function $[n] \to F$.
Sketch. (a) Induction: if $X = \emptyset$, Y^X has one element. Otherwise remove x_0; restricting to X' and recording the value at x_0 is a bijection $Y^X \to Y^{X'} \times Y$, giving $|Y|^{|X|-1} \cdot |Y|$. (b) $A \mapsto \mathbb{1}_A$ with inverse $h \mapsto h^{-1}(1)$. (c) $F^n = F^{[n]}$ and $F^{m \times n} = F^{[m] \times [n]}$.

(a) If $|X| = 0$, then $X = \emptyset$, and there is exactly one function $\emptyset \to Y$, so $\left|Y^X\right| = 1 = |Y|^0$. Let $|X| = k \geq 1$, pick $x_0 \in X$, and let $X' = X \setminus \{x_0\}$, so $|X'| = k - 1$ (remove x_0 from a listing of X). The map

$$Y^X \to Y^{X'} \times Y, \quad h \mapsto (h|_{X'}, h(x_0)),$$

is a bijection, with inverse sending (h', y) to the function that agrees with h' on X' and takes the value y at x_0. If $|A| = r$ and $|B| = s$, then $|A \times B| = rs$: listing $A = \{\alpha_1, \ldots, \alpha_r\}$ and $B = \{\beta_1, \ldots, \beta_s\}$, the pair (α_i, β_j)

corresponds to $(i-1)s+j \in [rs]$. Hence, by the induction hypothesis and Corollary A.1.18, $\left|Y^X\right| = |Y|^{k-1} \cdot |Y| = |Y|^k$.

(b) For $A \subseteq X$, let $\mathbb{1}_A : X \to \{0,1\}$ be its **indicator function**: $\mathbb{1}_A(x) = 1$ if $x \in A$ and 0 otherwise. The map $A \mapsto \mathbb{1}_A$ has inverse $h \mapsto h^{-1}(1)$: indeed $\mathbb{1}_A^{-1}(1) = A$, and h and $\mathbb{1}_{h^{-1}(1)}$ agree at every x. So it is a bijection (Theorem A.1.14), and $|P(X)| = \left|\{0,1\}^X\right| = 2^{|X|}$ by (a).

(c) An element of F^n is a function $[n] \to F$, and an element of $F^{m\times n}$ is a function $[m]\times[n] \to F$ (Remark 1.2.11). So $|F^n| = q^n$ and $|F^{m\times n}| = q^{mn}$ by (a), using $|[m]\times[n]| = mn$.

Solution to Exercise A.1.29

> *Uses.* Theorem A.1.17 (Pigeonhole principle), Definition 1.1.19 (Integers modulo n)
> *Techniques.* Pigeonhole (Technique A.1.21)
> *Hint 1.* Blocks of consecutive terms are differences of partial sums $s_j = a_1 + \cdots + a_j$, with $s_0 = 0$.
> *Hint 2.* There are $n+1$ partial sums $s_0, \ldots, s_n$ and only n remainders mod n.
> *Sketch.* The $n+1$ partial sums $s_0, \ldots, s_n$ have only n possible remainders mod n, so $s_i \bmod n = s_j \bmod n$ for some $i < j$, and then n divides $s_j - s_i = a_{i+1} + \cdots + a_j$.

Let $s_0 = 0$ and $s_j = a_1 + \cdots + a_j$ for $1 \le j \le n$. The function $\{0,1,\ldots,n\} \to \{0,1,\ldots,n-1\}$, $j \mapsto s_j \bmod n$, goes from a set with $n+1$ elements to one with n elements, so it is not injective (Theorem A.1.17). Hence $s_i \bmod n = s_j \bmod n$ for some $i < j$, which means n divides $s_j - s_i = a_{i+1} + \cdots + a_j$.

The bound is sharp: for $a_1 = \cdots = a_{n-1} = 1$, which has only $n-1$ terms, every block sum lies between 1 and $n-1$.

Solution to Exercise A.1.30

> *Uses.* Definition 2.6.2 (Surjective)
> *Techniques.* Extremal choice and minimal counterexamples (Technique 1.5.9)
> *Hint 1.* Given $f : X \to P(X)$, build a subset of X that differs from every $f(x)$, by making it disagree with $f(x)$ at the element x.
> *Hint 2.* Let $D = \{x \in X \mid x \notin f(x)\}$. Could $D = f(d)$? Ask whether $d \in D$.
> *Sketch.* Diagonal argument: $D = \{x \mid x \notin f(x)\}$ is not a value of f, since $D = f(d)$ would give $d \in D \Leftrightarrow d \notin D$.

Let $f : X \to P(X)$ and $D = \{x \in X \mid x \notin f(x)\} \in P(X)$. Suppose $D = f(d)$ for some $d \in X$. If $d \in D$, then $d \notin f(d) = D$; if $d \notin D$, then $d \notin f(d)$, so $d \in D$ by definition of D. Both cases are contradictions, so D is not a value of f, and f is not surjective. In particular there is no bijection $\mathbb{N} \to P(\mathbb{N})$, although $\mathbb{N}$ and $P(\mathbb{N})$ are both infinite.

The set D is built to disagree with $f(x)$ about the single element x. The same diagonal idea shows that the real numbers cannot be listed as a sequence.

Section A.2

Solution to Exercise A.2.7

> *Uses.* Theorem A.2.1 (Induction, strong induction, and well-ordering), Definition 2.4.1 (Matrix product)
> *Hint 1.* Induct on n. Write $J^{n+1} = J^n J$.
> *Hint 2.* Multiply out $\begin{pmatrix}\lambda^n & n\lambda^{n-1}\\ 0 & \lambda^n\end{pmatrix}\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$; the top right entry is $\lambda^n + n\lambda^n$.
> *Sketch.* For $n = 1$ the formula reads J. If it holds for n, then $J^{n+1} = J^n J$ has entries λ^{n+1}, $\lambda^n + n\lambda^{n-1}\lambda = (n+1)\lambda^n$, 0, λ^{n+1}.

For $n = 1$ the right side is $\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} = J$, since $1\lambda^0 = 1$. Suppose the formula holds for some $n \ge 1$. Then

$$J^{n+1} = J^n J = \begin{pmatrix}\lambda^n & n\lambda^{n-1}\\ 0 & \lambda^n\end{pmatrix}\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} = \begin{pmatrix}\lambda^{n+1} & \lambda^n + n\lambda^n\\ 0 & \lambda^{n+1}\end{pmatrix} = \begin{pmatrix}\lambda^{n+1} & (n+1)\lambda^n\\ 0 & \lambda^{n+1}\end{pmatrix},$$

which is the formula for $n+1$. By induction it holds for all $n \ge 1$.

Over a field of characteristic p, the top right entry of J^p is $p\lambda^{p-1} = 0$, so $J^p = \lambda^p I$. For instance over $\mathbb{F}_2$, $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^2 = I$.

Solution to Exercise A.2.8

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering), Definition 2.4.1 (Matrix product)
Hint 1. (a) is an induction on k: $A^{k+1}B = A\big(A^kB\big)$.
Hint 2. For (b), multiply the formula for m by $A+B$ on the left, use (a) to write $BA^kB^{m-k} = A^kB^{m-k+1}$, and combine terms with $\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}$.
Hint 3. For (c), compute $(A+B)^2$ and $A^2 + 2AB + B^2$.
Sketch. (a) $A^{k+1}B = ABA^k = BA^{k+1}$. (b) $(A+B)^{m+1} = \sum_k \binom{m}{k}\big(A^{k+1}B^{m-k} + A^kB^{m+1-k}\big)$ by (a), and Pascal's rule collects the terms. (c) $(A+B)^2 = I$ but $A^2 + 2AB + B^2 = 2AB = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

(a) For $k = 0$ both sides are B. If $A^kB = BA^k$, then $A^{k+1}B = A\big(A^kB\big) = A\big(BA^k\big) = (AB)A^k = (BA)A^k = BA^{k+1}$.

(b) For $m = 0$ both sides are I. Suppose the formula holds for m. Then

$$(A+B)^{m+1} = (A+B)\sum_{k=0}^{m}\binom{m}{k}A^kB^{m-k} = \sum_{k=0}^{m}\binom{m}{k}A^{k+1}B^{m-k} + \sum_{k=0}^{m}\binom{m}{k}BA^kB^{m-k}.$$

By (a), $BA^kB^{m-k} = A^kB^{m+1-k}$. Shifting the index in the first sum ($k+1 \to k$),

$$(A+B)^{m+1} = \sum_{k=1}^{m+1}\binom{m}{k-1}A^kB^{m+1-k} + \sum_{k=0}^{m}\binom{m}{k}A^kB^{m+1-k} = \sum_{k=0}^{m+1}\binom{m+1}{k}A^kB^{m+1-k},$$

using $\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}$ for $1 \le k \le m$ and $\binom{m}{m} = \binom{m+1}{m+1} = \binom{m}{0} = \binom{m+1}{0} = 1$ for the end terms.

(c) Here $A + B = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, so $(A+B)^2 = I$. On the other hand $A^2 = B^2 = 0$ and $AB = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, so $A^2 + 2AB + B^2 = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right) \neq I$ (also in characteristic 2, where it is 0). The induction in (b) used $AB = BA$ in (a), and indeed $BA = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) \neq AB$.

Solution to Exercise A.2.9

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering)
Techniques. Uniqueness via two candidates (Technique 1.1.8), Strong induction and well-ordering (Technique 5.2.12)
Hint 1. Let 2^k be the largest power of 2 with $2^k \le n$. Apply strong induction to $n - 2^k$.
Hint 2. Why can 2^k not appear in the representation of $n - 2^k$? Use $n < 2^{k+1}$.
Hint 3. For uniqueness, show that 2^k must appear in every representation, because $1 + 2 + \cdots + 2^{k-1} = 2^k - 1 < n$.
Sketch. Strong induction. With $2^k \le n < 2^{k+1}$, the number $m = n - 2^k < 2^k$ is 0 or a sum of distinct powers of 2, each less than 2^k, so adding 2^k gives n. Every representation of n uses 2^k (smaller powers sum to at most $2^k - 1$) and no larger power, so removing 2^k leaves a representation of m, which is unique by hypothesis.

We use strong induction on $n \ge 1$. Let $k \ge 0$ be the largest integer with $2^k \le n$ (it exists because $2^j > n$ for $j \ge n$), so that $2^k \le n < 2^{k+1}$, and let $m = n - 2^k$, so $0 \le m < 2^k$.

Existence. If $m = 0$, then $n = 2^k$. Otherwise $1 \le m < n$, so by the induction hypothesis m is a sum of distinct powers of 2, each at most $m < 2^k$. Adding the term 2^k gives a representation of n with distinct terms.

Uniqueness. Take any representation of n as a sum of distinct powers of 2. No term exceeds $n < 2^{k+1}$, so all terms are among $1, 2, \ldots, 2^k$. If 2^k were not a term, the sum would be at most $1 + 2 + \cdots + 2^{k-1} = 2^k - 1 < n$, by Example A.2.2 for 1×1 matrices over $\mathbb{Q}$ with $A = 2$. So 2^k is a term, and the remaining terms form a representation of m (empty if $m = 0$). If $m \ge 1$, this representation is unique by the induction hypothesis, since $m < n$. So the representation of n is unique.

Solution to Exercise A.2.10

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering), Proposition A.2.4 (Division with remainder)
Techniques. Double containment (Technique 1.3.14), Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. First show that $0 \in S$, that S is closed under negation and addition, and that S contains a positive element.
Hint 2. Take d least positive in S by well-ordering. Show $d\mathbb{Z} \subseteq S$ by induction on $|k|$, or by closure under addition and negation.
Hint 3. For $S \subseteq d\mathbb{Z}$, divide $a \in S$ by d: $a = qd + r$ with $0 \le r < d$. Then $r = a - qd \in S$.

Sketch. $0 = a - a$, $-b = 0 - b$ and $a + b = a - (-b)$ lie in S, and some nonzero $a \in S$ gives a positive element $|a|$. Let d be the least positive element. Then $d\mathbb{Z} \subseteq S$. For $a \in S$, $a = qd + r$ with $0 \le r < d$ and $r = a - qd \in S$, so $r = 0$ by minimality of d.

Pick $a \in S$. Then $0 = a - a \in S$, so for every $b \in S$ also $-b = 0 - b \in S$, and for $a, b \in S$, $a + b = a - (-b) \in S$. Since $S \neq \{0\}$, it contains some $a \neq 0$, and one of a and $-a$ is a positive element of S. By well-ordering (Theorem A.2.1(III)), S has a least positive element d.

By closure under addition, $d, 2d, 3d, \ldots \in S$ (induction on k: $(k+1)d = kd + d$), and by closure under negation, so are $-d, -2d, \ldots$; also $0 \in S$. So $d\mathbb{Z} \subseteq S$.

Conversely, let $a \in S$. By Proposition A.2.4, $a = qd + r$ with $0 \le r < d$. Since $qd \in S$, also $r = a - qd \in S$. As d is the least positive element of S and $0 \le r < d$, we get $r = 0$, so $a = qd \in d\mathbb{Z}$.

A similar argument, with the degree in place of the absolute value, Theorem 7.1.14 in place of Proposition A.2.4, and the fact that $p(T) = 0$ implies $(qp)(T) = 0$ for every polynomial q, shows that for an operator T on a finite-dimensional space, the polynomials p with $p(T) = 0$ are exactly the multiples of one of least degree (Section 7.6).

Solution to Exercise A.2.11

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering)
Techniques. Strengthen the induction hypothesis (Technique A.2.5)
Hint 1. Let p_n be the left side, so $p_{n+1} = p_n(2n+1)/(2n+2)$. Check that $p_n \le 1/\sqrt{3n}$ does not imply $p_{n+1} \le 1/\sqrt{3n+3}$: square and compare.
Hint 2. Prove instead the stronger bound $p_n \le 1/\sqrt{3n+1}$.
Hint 3. The step reduces to $(2n+1)^2(3n+4) \le (2n+2)^2(3n+1)$. Expand both sides.
Sketch. Strengthen to $p_n \le 1/\sqrt{3n+1}$. Base: $p_1 = 1/2$. Step: $p_{n+1} \le (2n+1)/\left((2n+2)\sqrt{3n+1}\right)$, and this is at most $1/\sqrt{3n+4}$ iff $(2n+1)^2(3n+4) \le (2n+2)^2(3n+1)$, which expands to $12n^3 + 28n^2 + 19n + 4 \le 12n^3 + 28n^2 + 20n + 4$.

Let p_n be the left side, so $p_1 = 1/2$ and $p_{n+1} = p_n(2n+1)/(2n+2)$.

The direct induction fails: from $p_n \le 1/\sqrt{3n}$ we get $p_{n+1} \le (2n+1)/\left((2n+2)\sqrt{3n}\right)$, and this is at most $1/\sqrt{3n+3}$ only if $(2n+1)^2(3n+3) \le (2n+2)^2(3n)$, that is, $12n^3 + 24n^2 + 15n + 3 \le 12n^3 + 24n^2 + 12n$, which is false. (Technique A.2.5: the hypothesis leaves no room.)

We prove instead that $p_n \le 1/\sqrt{3n+1}$ for all $n \ge 1$, which implies the claim since $3n + 1 > 3n$. For $n = 1$, $p_1 = 1/2 = 1/\sqrt{4}$. Suppose $p_n \le 1/\sqrt{3n+1}$. Then

$$p_{n+1} \le \frac{2n+1}{(2n+2)\sqrt{3n+1}},$$

and this is at most $1/\sqrt{3n+4}$ if and only if $(2n+1)^2(3n+4) \le (2n+2)^2(3n+1)$. Expanding,

$$(2n+1)^2(3n+4) = 12n^3 + 28n^2 + 19n + 4, \quad (2n+2)^2(3n+1) = 12n^3 + 28n^2 + 20n + 4,$$

and $19n \le 20n$. So $p_{n+1} \le 1/\sqrt{3(n+1)+1}$, completing the induction.

Solution to Exercise A.2.12

Uses. Theorem A.2.1 (Induction, strong induction, and well-ordering)
Hint 1. Is the step correct? Is anything missing from the induction?
Hint 2. Try $n = 0$.
Sketch. The step is valid, but there is no base case, and it fails: $0^2 + 0 + 1 = 1$ is odd. In fact $n^2 + n = n(n+1)$ is always even, so $n^2 + n + 1$ is always odd.

The induction step is correct: it shows that $P(n)$ implies $P(n+1)$. But (I) of Theorem A.2.1 also requires $P(0)$, which was never checked, and it is false: $0^2 + 0 + 1 = 1$ is odd. In fact the claim is false for every n: $n^2 + n = n(n+1)$ is a product of consecutive integers, one of which is even, so $n^2 + n + 1$ is odd. The step proves "if $P(n)$ holds for some n, then it holds for all larger n", which is vacuous here.

Solution to Exercise A.2.13

Uses. Definition 7.2.1 (Invariant subspace), Definition 7.2.6 (Eigenvalue, eigenvector)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. The induction hypothesis is about operators on n-dimensional spaces. Is the restriction of T to W an operator on W?
Hint 2. Test the argument on the rotation of $\mathbb{R}^2$ by a right angle.
Sketch. The restriction of T to W is a map $W \to V$, and it is an operator on W only if W is invariant under T. The rotation $T(x, y) = (-y, x)$ of $\mathbb{R}^2$ has no invariant line, and no eigenvector, so the claim itself is false.

The induction hypothesis applies to operators on n-dimensional spaces, that is, to linear maps $W \to W$. The restriction of T to W is a linear map $W \to V$, and it maps W into W only if W is invariant under T (Definition 7.2.1). The argument never chooses W invariant, and in general no invariant subspace of dimension n exists. Indeed the claim is false: the rotation $T(x, y) = (-y, x)$ of $\mathbb{R}^2$ has no eigenvector, since $T(x, y) = \lambda(x, y)$ gives $-y = \lambda x$ and $x = \lambda y$, so $x = -\lambda^2 x$ and $y = -\lambda^2 y$, which forces $x = y = 0$ because $-\lambda^2 \neq 1$ for real λ.

This is the general warning of Section A.2.2. Inductions on dimension pass to a smaller space, and what is passed down must be an object of the kind the hypothesis is about. In Section 7.5 the induction works over $\mathbb{C}$ because an eigenvector exists, and the induction hypothesis is applied to an operator on a smaller space built from it.

Section A.3

Solution to Exercise A.3.14

Uses. Definition A.3.1 (Equivalence relation), Example A.3.2 (Equivalence relations)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Relations of the form "$g(x) = g(y)$" are always equivalence relations (Example A.3.2(c)). Which of these have that form?
Hint 2. Check reflexivity at the "degenerate" elements: 0 in (b) and $\emptyset$ in (e).
Sketch. (a) Yes, classes $x + \mathbb{Z}$, one for each $x \in [0, 1)$. (b) No: $0 \sim 0$ fails. (c) Yes, $\underset{\text{rank}}{\sim}$; classes are the matrices of rank r, $0 \leq r \leq \min(m, n)$. (d) Yes, $\underset{g}{\sim}$ for $g(\mathbf{v}) = \operatorname{span}(\mathbf{v})$; classes are the lines through $\mathbf{0}$ with $\mathbf{0}$ removed. (e) No: $\emptyset \sim \emptyset$ fails, and so does transitivity.

(a) Yes. $x - x = 0 \in \mathbb{Z}$; if $x - y \in \mathbb{Z}$, then $y - x \in \mathbb{Z}$; and $x - z = (x - y) + (y - z)$. The class of x is $x + \mathbb{Z}$. Each class contains exactly one element of $[0, 1)$, namely $x - \lfloor x \rfloor$, so the classes are indexed by $[0, 1)$.

(b) No: $0 \cdot 0 = 0$, so $0 \sim 0$ fails. (On $\mathbb{Z} \setminus \{0\}$ it is an equivalence relation, "same sign", with two classes.)

(c) Yes: it is $\underset{g}{\sim}$ for $g = \text{rank}$. The classes are the sets of matrices of rank r for $r = 0, 1, \ldots, \min(m, n)$, each nonempty (the matrix $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ has rank r).

(d) Yes: it is $\underset{g}{\sim}$ for $g(\mathbf{v}) = \operatorname{span}(\mathbf{v})$. The class of $\mathbf{v}$ is $\operatorname{span}(\mathbf{v}) \setminus \{\mathbf{0}\} = \{c\mathbf{v} \mid c \neq 0\}$, a line through $\mathbf{0}$ with $\mathbf{0}$ removed. Compare Non-example A.3.3(b).

(e) No: $\emptyset \cap \emptyset = \emptyset$, so $\emptyset \sim \emptyset$ fails. Transitivity fails too: for distinct $a, b \in X$, $\{a\} \sim \{a, b\}$ and $\{a, b\} \sim \{b\}$, but $\{a\} \cap \{b\} = \emptyset$.

Solution to Exercise A.3.15

Uses. Theorem A.3.7 (Equivalence relations are partitions)
Hint 1. By Theorem A.3.7, count partitions instead.
Hint 2. Sort the partitions by their block sizes. For $[4]$ the possible block sizes are 4; $3 + 1$; $2 + 2$; $2 + 1 + 1$; $1 + 1 + 1 + 1$.
Sketch. Count partitions. $[3]$: one with one block, three of type $2 + 1$, one into singletons: 5. $[4]$: $1 + 4 + 3 + 6 + 1 = 15$ by block sizes 4, $3 + 1$, $2 + 2$, $2 + 1 + 1$, $1 + 1 + 1 + 1$.

By Theorem A.3.7, equivalence relations on X correspond to partitions of X, so we count partitions by their block sizes.

For $[3]$: one partition with a single block; three of type $2 + 1$ (choose the singleton); one into three singletons. Total 5.

For [4]: one of type 4; four of type $3+1$ (choose the singleton); three of type $2+2$ (the block containing 1 is $\{1,j\}$ for $j=2,3,4$); six of type $2+1+1$ (choose the pair, $\binom{4}{2}=6$); one into singletons. Total $1+4+3+6+1=15$.

Solution to Exercise A.3.16

Uses. Definition A.3.1 (Equivalence relation), Theorem A.3.7 (Equivalence relations are partitions), Exercise A.1.28, Proposition 1.2.35 (Zero products)
Hint 1. For symmetry use c^{-1}; for transitivity multiply the scalars.
Hint 2. Each class has the same number of elements. How many nonzero multiples does a nonzero vector have?
Hint 3. The classes partition a set of $q^{n+1}-1$ elements into blocks of size $q-1$.
Sketch. Reflexive ($c=1$), symmetric ($\mathbf{v}=c^{-1}\mathbf{w}$), transitive ($\mathbf{u}=cd\mathbf{v}$). The class of $\mathbf{v}$ is $\operatorname{span}(\mathbf{v})\setminus\{\mathbf{0}\}$, and the map $c\mapsto c\mathbf{v}$ from $F\setminus\{0\}$ onto it is injective, so each class has $q-1$ elements. So there are $\big(q^{n+1}-1\big)/(q-1)$ classes.

(a) Reflexivity: $\mathbf{v}=1\mathbf{v}$. Symmetry: if $\mathbf{w}=c\mathbf{v}$ with $c\neq 0$, then $\mathbf{v}=c^{-1}\mathbf{w}$. Transitivity: if $\mathbf{w}=c\mathbf{v}$ and $\mathbf{u}=d\mathbf{w}$, then $\mathbf{u}=(dc)\mathbf{v}$ with $dc\neq 0$. The class of $\mathbf{v}$ is $\{c\mathbf{v}\mid c\neq 0\}=\operatorname{span}(\mathbf{v})\setminus\{\mathbf{0}\}$, the line through $\mathbf{v}$ with $\mathbf{0}$ removed (by Proposition 1.2.35, $c\mathbf{v}=\mathbf{0}$ only for $c=0$). So the classes correspond to the one-dimensional subspaces of F^{n+1}. The quotient set is the **projective space** of F^{n+1}, and the coordinates of a representative are the homogeneous coordinates of its class (Definition 4.7.41).

(b) For $\mathbf{v}\neq\mathbf{0}$, the map $F\setminus\{0\}\to[\mathbf{v}]$, $c\mapsto c\mathbf{v}$, is surjective by definition and injective since $c\mathbf{v}=d\mathbf{v}$ gives $(c-d)\mathbf{v}=\mathbf{0}$, so $c=d$. So every class has $q-1$ elements. The set $F^{n+1}\setminus\{\mathbf{0}\}$ has $q^{n+1}-1$ elements (Exercise A.1.28), and by Theorem A.3.7 it is the disjoint union of the classes. If there are N classes, then $N(q-1)=q^{n+1}-1$, so $N=\big(q^{n+1}-1\big)/(q-1)=1+q+\cdots+q^n$. For example, $\mathbb{F}_2^3$ has 7 one-dimensional subspaces.

Solution to Exercise A.3.17

Uses. Definition A.3.1 (Equivalence relation)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Each axiom is a closure condition. Which of them survive intersections?
Hint 2. For unions, take congruence mod 2 and congruence mod 3 on $\mathbb{Z}$, and look at 0, 2, 5.
Sketch. (a) True: each axiom for $R\cap S$ follows from the same axiom for R and for S. (b) False: with R congruence mod 2 and S mod 3, $(0,2)\in R$ and $(2,5)\in S$, but $(0,5)\notin R\cup S$.

(a) True. (x,x) lies in R and in S. If $(x,y)\in R\cap S$, then $(y,x)\in R$ and $(y,x)\in S$. If $(x,y),(y,z)\in R\cap S$, then $(x,z)\in R$ by transitivity of R, and $(x,z)\in S$ likewise. In terms of classes, the class of x under $R\cap S$ is the intersection of its classes under R and S.

(b) False. On $\mathbb{Z}$, let R be congruence modulo 2 and S congruence modulo 3. Then $(0,2)\in R$ and $(2,5)\in S$, so both lie in $R\cup S$, but $5-0$ is divisible by neither 2 nor 3, so $(0,5)\notin R\cup S$. Transitivity fails. ($R\cup S$ is still reflexive and symmetric.)

Solution to Exercise A.3.18

Uses. Definition A.3.1 (Equivalence relation)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Is there always a y with $x\sim y$?
Hint 2. Try the empty relation on a nonempty set.
Sketch. The step "pick y with $x\sim y$" assumes such a y exists. The empty relation on $\{1\}$ is symmetric and transitive (vacuously) but not reflexive. The argument does show $x\sim x$ for every x related to something.

The argument assumes that some y with $x\sim y$ exists, and nothing guarantees this. For instance, the empty relation $R=\emptyset$ on $X=\{1\}$ is symmetric and transitive, since their hypotheses are never satisfied, but $1\sim 1$ fails. What the argument does prove is that a symmetric and transitive relation is reflexive on the set of elements that are related to at least one element. So in checking an equivalence relation, reflexivity cannot be skipped.

Solution to Exercise A.3.19

Uses. Theorem A.3.10 (Functions on a quotient set), Example A.3.9 ($\mathbb{Z}/n\mathbb{Z}$ as a quotient set)

Techniques. Check that a definition does not depend on the representative (Technique 4.6.18)
Hint 1. By Theorem A.3.10, the formula defines a function iff $a \equiv a' (\mathrm{mod}\, n)$ implies $a \equiv a' (\mathrm{mod}\, m)$.
Hint 2. For "only if", use the pair $a = 0$, $a' = n$.
Hint 3. Surjectivity: $[a]_m = f([a]_n)$. Count the elements for bijectivity.
Sketch. Well defined iff $n \mid a - a'$ implies $m \mid a - a'$. If $m \mid n$, this holds; conversely, $a = n$, $a' = 0$ gives $m \mid n$. f is surjective since $q_m = f \circ q_n$ is. A bijection needs $n = |\mathbb{Z}/n\mathbb{Z}| = |\mathbb{Z}/m\mathbb{Z}| = m$.

Let $g : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$, $g(a) = [a]_m$. By Theorem A.3.10(a), the formula defines a function f with $f([a]_n) = g(a)$ if and only if g is constant on classes mod n, that is, if and only if $n \mid a - a'$ implies $m \mid a - a'$ for all $a, a' \in \mathbb{Z}$.

If $m \mid n$ and $n \mid a - a'$, then $m \mid a - a'$. Conversely, if the implication holds, apply it to $a = n$ and $a' = 0$: since $n \mid n$, we get $m \mid n$.

Suppose $m \mid n$. Since g is surjective, so is f (Theorem A.3.10(c)). By Example A.3.9, $\mathbb{Z}/n\mathbb{Z}$ has n elements and $\mathbb{Z}/m\mathbb{Z}$ has m, so if f is bijective, then $n = m$ (Corollary A.1.18). Conversely, for $m = n$, f is the identity. For example, reduction mod 2 is a well-defined function $\mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, but "reduction mod 3" is not: $[0]_4 = [4]_4$ while $[0]_3 \neq [4]_3$.

Solution to Exercise A.3.20

Uses. Definition A.3.1 (Equivalence relation), Definition 3.2.21 (Row equivalence), Proposition 3.2.10 (Inverses of elementary matrices)
Techniques. Double containment (Technique 1.3.14)
Hint 1. For (a): the empty chain ($k = 0$) gives reflexivity; reversing a chain gives symmetry; concatenating gives transitivity.
Hint 2. For minimality, if $\sim$ is an equivalence relation containing R, show by induction on k that each $x_0 \sim x_k$, using symmetry for the backward steps.
Hint 3. For (b), elementary row operations can be undone by elementary row operations (Proposition 3.2.10), so a backward step is also a forward step.
Sketch. (a) Chains of length 0, reversed chains, and concatenated chains give the three axioms; a one-step chain shows $R \subseteq \approx$. If $\sim \supseteq R$ is an equivalence relation, each step of a chain is in $\sim$ (by symmetry for backward steps), and transitivity along the chain gives $x \sim y$. (b) By (a) the generated relation is the smallest equivalence relation containing R; row equivalence contains R and is an equivalence relation, and conversely a row equivalence $B = E_k \cdots E_1 A$ is a chain of forward R-steps.

(a) Reflexivity: the chain $x_0 = x$ with $k = 0$. Symmetry: a chain $x_0, \ldots, x_k$ from x to y, read backwards, is a chain from y to x, since the condition on consecutive pairs is symmetric. Transitivity: a chain from x to y followed by one from y to z is a chain from x to z. If $(x, y) \in R$, then x, y is a chain with $k = 1$, so $R \subseteq \approx$.

Now let $\sim$ be an equivalence relation with $R \subseteq \sim$, and let $x = x_0, \ldots, x_k = y$ be a chain. We show $x_0 \sim x_j$ by induction on j. For $j = 0$ this is reflexivity. If $x_0 \sim x_{j-1}$, then $(x_{j-1}, x_j) \in R$ gives $x_{j-1} \sim x_j$, and $(x_j, x_{j-1}) \in R$ gives $x_j \sim x_{j-1}$, hence $x_{j-1} \sim x_j$ by symmetry. Either way $x_0 \sim x_j$ by transitivity. So $x \approx y$ implies $x \sim y$.

(b) Row equivalence is an equivalence relation (Proposition 3.2.22) and contains R, so by (a) it contains the generated relation $\approx$. Conversely, if $B = E_k \cdots E_1 A$ with E_i elementary, then $A, E_1 A, E_2 E_1 A, \ldots, B$ is a chain in which each step is one elementary row operation, so $A \approx B$. Hence the two relations are equal. The backward steps allowed in $\approx$ add nothing here because each elementary operation is undone by another one (Proposition 3.2.10). This is exactly why row equivalence is symmetric.

Section A.4

Solution to Exercise A.4.8

Uses. Definition A.4.1 (Partial order), Definition A.4.2 (Upper bound, maximal element), Corollary A.1.18 (Cardinality is well defined)
Techniques. Extremal choice and minimal counterexamples (Technique 1.5.9)
Hint 1. Look at strictly increasing sequences $p_1 < p_2 < \cdots < p_k$. Why are their entries distinct, and why is $k \leq |P|$?
Hint 2. Take such a sequence with k as large as possible. Its last entry is maximal.

Sketch. The entries of a strictly increasing sequence are distinct (by transitivity and antisymmetry), so its length is at most $|P|$. A longest one exists, since one-term sequences exist, and its last entry p_k is maximal: $p_k < p$ would give a longer sequence.

In a strictly increasing sequence $p_1 < p_2 < \cdots < p_k$, the entries are distinct: if $i < j$, then $p_i < p_j$ by transitivity, and if $p_i = p_j$ this would read $p_i \leq p_{i+1}$ and $p_{i+1} \leq p_j = p_i$, so $p_i = p_{i+1}$ by antisymmetry, contradicting $p_i < p_{i+1}$. So $i \mapsto p_i$ is an injection $[k] \to P$, and $k \leq |P|$ (Corollary A.1.18). Since $P \neq \emptyset$, there are sequences of length 1, so there is a strictly increasing sequence of maximal length k. Its last entry p_k is maximal: if $p_k \leq p$ with $p \neq p_k$, then $p_1 < \cdots < p_k < p$ would be longer.

Solution to Exercise A.4.9

Uses. Theorem A.4.4 (Zorn's lemma), Example A.4.3 (Chains and maximal elements)
Hint 1. An upper bound has to be an element of P. Is the union of a chain of proper subspaces proper?
Hint 2. Look at Example A.4.3(c).
Sketch. The union is a subspace, but it need not be proper, so it need not lie in P: in $F[x]$ the union of the chain $F[x]_{\leq n}$ is $F[x]$. The claim is still true (Exercise A.4.10), but the proof must keep the union away from V, for instance by fixing a vector it must avoid.

The union of the chain is a subspace containing every member, but an upper bound must be an element of P, that is, a *proper* subspace, and the argument never checks this. It can fail: in $V = F[x]$, the subspaces $F[x]_{\leq n}$, $n \geq 0$, form a chain of proper subspaces whose union is $F[x]$ (Example A.4.3(c)). So the hypothesis of Zorn's lemma has not been verified.

The claim itself is true. One way to repair the proof is to fix $\mathbf{v} \neq \mathbf{0}$ and let P be the set of subspaces that do not contain $\mathbf{v}$. It contains $\{\mathbf{0}\}$, and the union of a nonempty chain in P is a subspace that does not contain $\mathbf{v}$, so it lies in P. Let M be a maximal element. If some $\mathbf{x}$ were not in $M + \operatorname{span}(\mathbf{v})$, then $M + \operatorname{span}(\mathbf{x})$ would not contain $\mathbf{v}$ (from $\mathbf{v} = \mathbf{m} + c\mathbf{x}$ we would get $c \neq 0$ and $\mathbf{x} = c^{-1}(\mathbf{v} - \mathbf{m}) \in M + \operatorname{span}(\mathbf{v})$), contradicting maximality. So $V = M + \operatorname{span}(\mathbf{v})$. Every subspace strictly containing M contains $\mathbf{v}$, by maximality, and hence equals V; so M is a maximal proper subspace. Exercise A.4.10 gives another route.

Solution to Exercise A.4.10

Uses. Theorem A.4.7 (Complements exist), Theorem 4.4.7 (Criteria for direct sums), Definition 4.4.4 (Direct sum), Definition 2.1.1 (Linear map)
Techniques. Check that a definition does not depend on the representative (Technique 4.6.18)
Hint 1. The sum $U + \operatorname{span}(\mathbf{v})$ is direct. Take a complement W of it.
Hint 2. Every vector is uniquely $\mathbf{u} + c\mathbf{v} + \mathbf{w}$. Let φ read off c.
Hint 3. For the last part, $\ker\varphi$ is a proper subspace. Why is it maximal?
Sketch. $U \cap \operatorname{span}(\mathbf{v}) = \{\mathbf{0}\}$, so $U' = U \oplus \operatorname{span}(\mathbf{v})$; by Theorem A.4.7, $V = U \oplus \operatorname{span}(\mathbf{v}) \oplus W$. Define $\varphi(\mathbf{u} + c\mathbf{v} + \mathbf{w}) = c$, well defined by uniqueness of decompositions and linear by comparing decompositions. With $U = \{\mathbf{0}\}$, $\ker\varphi$ is a proper subspace with $V = \ker\varphi \oplus \operatorname{span}(\mathbf{v})$, and any subspace strictly containing it contains $\mathbf{v}$.

If $c\mathbf{v} \in U$ with $c \neq 0$, then $\mathbf{v} \in U$; so $U \cap \operatorname{span}(\mathbf{v}) = \{\mathbf{0}\}$, and the sum $U' = U + \operatorname{span}(\mathbf{v})$ is direct (Theorem 4.4.7). By Theorem A.4.7 there is W with $V = U' \oplus W$. So every $\mathbf{x} \in V$ has exactly one decomposition $\mathbf{x} = \mathbf{u} + c\mathbf{v} + \mathbf{w}$ with $\mathbf{u} \in U$, $c \in F$, $\mathbf{w} \in W$ (the coefficient c is unique because $\mathbf{v} \neq \mathbf{0}$). Define $\varphi(\mathbf{x}) = c$. This is a function because the decomposition is unique. It is linear: if $\mathbf{x} = \mathbf{u} + c\mathbf{v} + \mathbf{w}$ and $\mathbf{x}' = \mathbf{u}' + c'\mathbf{v} + \mathbf{w}'$, then $a\mathbf{x} + \mathbf{x}' = (a\mathbf{u} + \mathbf{u}') + (ac + c')\mathbf{v} + (a\mathbf{w} + \mathbf{w}')$ is the decomposition of $a\mathbf{x} + \mathbf{x}'$, so $\varphi(a\mathbf{x} + \mathbf{x}') = ac + c'$. Clearly $\varphi(\mathbf{u}) = 0$ for $\mathbf{u} \in U$ and $\varphi(\mathbf{v}) = 1$.

Now let $V \neq \{\mathbf{0}\}$, pick $\mathbf{v} \neq \mathbf{0}$, and apply this with $U = \{\mathbf{0}\}$. Then $M = \ker\varphi$ is a subspace not containing $\mathbf{v}$, so it is proper. Every $\mathbf{x} \in V$ equals $(\mathbf{x} - \varphi(\mathbf{x})\mathbf{v}) + \varphi(\mathbf{x})\mathbf{v}$ with the first term in M, so $V = M + \operatorname{span}(\mathbf{v})$. If M' is a subspace with $M \subsetneq M'$, pick $\mathbf{x} \in M' \setminus M$; then $\mathbf{x} - \varphi(\mathbf{x})\mathbf{v} \in M \subseteq M'$ and $\varphi(\mathbf{x}) \neq 0$, so $\mathbf{v} \in M'$ and $M' \supseteq M + \operatorname{span}(\mathbf{v}) = V$. So M is a maximal proper subspace.

Solution to Exercise A.4.11

Uses. Theorem A.4.4 (Zorn's lemma), Lemma A.4.6 (Finitely many elements of a union of a chain), Theorem A.1.12 (One-sided inverses of functions)

Techniques. Apply Zorn's lemma (Technique A.4.5)
Hint 1. Let P be the set of partial right inverses: pairs (A, g) with $A \subseteq Y$ and $g : A \to X$ such that $f(g(a)) = a$ for all $a \in A$. Order them by extension: $(A, g) \leq (A', g')$ if $A \subseteq A'$ and g' agrees with g on A.
Hint 2. For a nonempty chain, define g^* on the union of the domains by "use any member of the chain whose domain contains the point". Why is this well defined?
Hint 3. If a maximal (A, g) has $A \neq Y$, extend g to one more point y by choosing one preimage of y. A single choice needs no axiom.
Sketch. Partial right inverses, ordered by extension, form a partially ordered set containing $(\emptyset, \text{empty function})$. The union of a chain is a partial right inverse, well defined because any two members agree where both are defined. A maximal (A, g) has $A = Y$: otherwise send one $y \notin A$ to one of its preimages, which extends g strictly.

Let P be the set of pairs (A, g) with $A \subseteq Y$ and $g : A \to X$ such that $f(g(a)) = a$ for every $a \in A$. Order P by $(A, g) \leq (A', g')$ if $A \subseteq A'$ and $g'(a) = g(a)$ for all $a \in A$. This is a partial order: reflexivity and transitivity are clear, and if $(A, g) \leq (A', g') \leq (A, g)$, then $A = A'$ and $g = g'$. P contains $(\emptyset, g_\emptyset)$ with $g_\emptyset$ the empty function, so $P \neq \emptyset$.

Let $\Gamma \subseteq P$ be a nonempty chain, and let A^* be the union of the domains A of its members. For $y \in A^*$, define $g^*(y) = g(y)$ for any $(A, g) \in \Gamma$ with $y \in A$. This does not depend on the member chosen: if $y \in A$ and $y \in A'$ for $(A, g), (A', g') \in \Gamma$, these are comparable, say $(A, g) \leq (A', g')$, and then $g'(y) = g(y)$. So $g^* : A^* \to X$ is a function with $f(g^*(y)) = f(g(y)) = y$, and $(A^*, g^*) \in P$ is an upper bound of Γ.

By Zorn's lemma, P has a maximal element (A, g). Suppose $A \neq Y$ and pick $y_0 \in Y \setminus A$. Since f is surjective, there is some x_0 with $f(x_0) = y_0$; this is one choice, which needs no axiom. Extending g by $g(y_0) = x_0$ gives an element of P strictly larger than (A, g), a contradiction. So $A = Y$, and $f \circ g = \mathrm{id}_Y$.

Section B.1

Solution to Exercise B.1.6

Uses. Proposition B.1.2 (Conjugation respects the field operations), Proposition 1.1.17 (No zero divisors)
Hint 1. Move everything to one side and factor the difference of squares.
Hint 2. $(z - \overline{z})(z + \overline{z}) = 0$. Now use Proposition B.1.2(c).
Sketch. $z^2 - \overline{z}^2 = (z - \overline{z})(z + \overline{z})$, so the equation holds if and only if $\overline{z} = z$ or $\overline{z} = -z$. The solutions are the real numbers and the purely imaginary numbers.

Since multiplication in $\mathbb{C}$ is commutative, $(z - \overline{z})(z + \overline{z}) = z^2 - \overline{z}^2$. So $z^2 = \overline{z}^2$ if and only if $(z - \overline{z})(z + \overline{z}) = 0$, which by Proposition 1.1.17 happens if and only if $\overline{z} = z$ or $\overline{z} = -z$. By Proposition B.1.2(c), the solutions are exactly the real numbers and the purely imaginary numbers, that is, the points of the two coordinate axes.

Solution to Exercise B.1.10

Uses. Proposition B.1.8 (Properties of the modulus)
Hint 1. Use Proposition B.1.8(d) for the quotient.
Hint 2. To compute z, multiply the numerator and the denominator by $\overline{2 - i} = 2 + i$; this is Proposition B.1.8(c).
Sketch. $|z| = |1 + 2i|/|2 - i| = \sqrt{5}/\sqrt{5} = 1$. And $z = (1 + 2i)(2 + i)/5 = 5i/5 = i$, which has modulus 1.

By Proposition B.1.8(d), $|z| = |1 + 2i|/|2 - i| = \sqrt{1 + 4}/\sqrt{4 + 1} = 1$. By Proposition B.1.8(c), $(2 - i)^{-1} = (2 + i)/5$, so

$$z = \frac{(1 + 2i)(2 + i)}{5} = \frac{2 + i + 4i + 2i^2}{5} = \frac{5i}{5} = i,$$

and indeed $|i| = 1$.

Solution to Exercise B.1.23

Uses. Proposition B.1.15 (Existence and uniqueness of polar form), Theorem B.1.20 (De Moivre's formula), Proposition B.1.17 (Multiplication in polar form)
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. The moduli are $2\sqrt{2}$ and 2. Find angles with the right cosine and sine.

Hint 2. $-2+2i = 2\sqrt{2}e^{3\pi i/4}$ and $-1-\sqrt{3}i = 2e^{-2\pi i/3}$. Now apply Theorem B.1.20 and divide.
Sketch. $(-2+2i)^5 = \left(2\sqrt{2}\right)^5 e^{15\pi i/4} = 128\sqrt{2}e^{-i\pi/4} = 128-128i$ and $\left(-1-\sqrt{3}i\right)^3 = 8e^{-2\pi i} = 8$, so the quotient is $16-16i$.

We have $|-2+2i| = \sqrt{8} = 2\sqrt{2}$ and $(-2+2i)/\left(2\sqrt{2}\right) = -1/\sqrt{2}+i/\sqrt{2} = e^{3\pi i/4}$, so $-2+2i = 2\sqrt{2}e^{3\pi i/4}$. Similarly $\left|-1-\sqrt{3}i\right| = 2$ and $-1-\sqrt{3}i = 2\left(-1/2-\left(\sqrt{3}/2\right)i\right) = 2e^{-2\pi i/3}$. By Theorem B.1.20,

$$(-2+2i)^5 = \left(2\sqrt{2}\right)^5 e^{15\pi i/4} = 128\sqrt{2}e^{-i\pi/4} = 128\sqrt{2}\left(\frac{1}{\sqrt{2}}-\frac{i}{\sqrt{2}}\right) = 128-128i,$$

using $\left(2\sqrt{2}\right)^5 = 32\cdot 4\sqrt{2}$ and $15\pi/4 = 4\pi-\pi/4$. Also $\left(-1-\sqrt{3}i\right)^3 = 8e^{-2\pi i} = 8$. So the quotient is $(128-128i)/8 = 16-16i$.

Solution to Exercise B.1.33

Uses. Proposition B.1.15 (Existence and uniqueness of polar form), Theorem B.1.20 (De Moivre's formula), Exercise 1.1.36
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. Write each base in polar form and apply Theorem B.1.20. Then reduce the argument modulo 2π.
Hint 2. $1-i = \sqrt{2}e^{-i\pi/4}$ and $1+\sqrt{3}i = 2e^{i\pi/3}$. For (d), first simplify the base; it is a fourth root of unity.
Sketch. (a) $32e^{-5\pi i/2} = 32e^{-i\pi/2} = -32i$. (b) $2^{-4}e^{-4\pi i/3} = (1/16)e^{2\pi i/3} = -1/32+\left(\sqrt{3}/32\right)i$. (c) $e^{4\pi i/3} = -1/2-\left(\sqrt{3}/2\right)i$. (d) The base is i, and $i^{2027} = i^3 = -i$.

(a) $1-i = \sqrt{2}e^{-i\pi/4}$, so $(1-i)^{10} = 2^5 e^{-10\pi i/4} = 32e^{-5\pi i/2}$. Since $-5\pi/2 = -\pi/2-2\pi$, this is $32e^{-i\pi/2} = -32i$.

(b) $1+\sqrt{3}i = 2e^{i\pi/3}$, so $\left(1+\sqrt{3}i\right)^{-4} = 2^{-4}e^{-4\pi i/3} = (1/16)e^{2\pi i/3} = (1/16)\left(-1/2+\left(\sqrt{3}/2\right)i\right) = -1/32+\left(\sqrt{3}/32\right)i$.

(c) By Theorem B.1.20, the power is $e^{12\pi i/9} = e^{4\pi i/3} = -1/2-\left(\sqrt{3}/2\right)i$.

(d) By Exercise 1.1.36(c), $(1+i)/(1-i) = i$. Since $i^4 = 1$ and $2027 = 4\cdot 506+3$, the power is $i^3 = -i$.

Solution to Exercise B.1.34

Uses. Theorem B.1.26 (nth roots), Proposition B.1.4 (Roots of real polynomials come in conjugate pairs)
Techniques. Use polar form for products, powers, and roots (Technique B.1.21)
Hint 1. $-64 = 64e^{i\pi}$. Apply Theorem B.1.26 with $n=6$.
Hint 2. The arguments are $\pi/6+k\pi/3$ for $k = 0,\dots,5$, and the modulus is 2.
Sketch. The solutions are $2e^{i(\pi+2\pi k)/6}$: $\pm\sqrt{3}\pm i$ and $\pm 2i$, the vertices of a regular hexagon of radius 2. Of these, $\pm 2i$ satisfy $z^2 = -4$.

Since $-64 = 64e^{i\pi}$ and $64^{1/6} = 2$, Theorem B.1.26 gives the six solutions $2e^{i(\pi+2\pi k)/6}$, $k = 0,\dots,5$, with arguments $\pi/6, \pi/2, 5\pi/6, 7\pi/6, 3\pi/2, 11\pi/6$:

$$\sqrt{3}+i, \quad 2i, \quad -\sqrt{3}+i, \quad -\sqrt{3}-i, \quad -2i, \quad \sqrt{3}-i.$$

They are the vertices of a regular hexagon inscribed in the circle of radius 2, with vertices on the imaginary axis. As z^6+64 has real coefficients, they come in conjugate pairs (Proposition B.1.4). The solutions of $z^2 = -4$ are $\pm 2i$, and both are on the list, as they must be: if $z^2 = -4$, then $z^6 = \left(z^2\right)^3 = -64$.

Solution to Exercise B.1.35

Uses. Proposition B.1.8 (Properties of the modulus), Proposition B.1.2 (Conjugation respects the field operations)
Techniques. Square the modulus (Technique B.1.9)
Hint 1. Apply Proposition B.1.8(f) to $z+w$ and to $z+(-w)$.
Hint 2. $\operatorname{Re}(z\overline{w}) = ac+bd$ for $z = a+bi$ and $w = c+di$.
Sketch. $|z\pm w|^2 = |z|^2 \pm 2\operatorname{Re}(z\overline{w})+|w|^2$. Adding gives (a): the squares of the diagonals add up to the squares of the four sides. Subtracting gives (b), and $\operatorname{Re}(z\overline{w}) = ac+bd$ is the dot product of (a,b) and (c,d), so the diagonals are equal exactly when the parallelogram is a rectangle.

By Proposition B.1.8(f), $|z+w|^2 = |z|^2+2\operatorname{Re}(z\overline{w})+|w|^2$. Applied to z and $-w$, with $\overline{-w} = -\overline{w}$ and $|-w| = |w|$, it gives $|z-w|^2 = |z|^2-2\operatorname{Re}(z\overline{w})+|w|^2$.

(a) Adding the two identities gives $|z+w|^2+|z-w|^2 = 2|z|^2+2|w|^2$. The diagonals of the parallelogram are the segments from 0 to $z+w$ and from w to z, of lengths $|z+w|$ and $|z-w|$, and its four sides have

lengths $|z|, |w|, |z|, |w|$. So the sum of the squares of the diagonals equals the sum of the squares of the sides. This is the parallelogram law, the case $V = \mathbb{C}$ of Proposition 6.1.30.

(b) Subtracting gives $|z+w|^2 - |z-w|^2 = 4\,\mathrm{Re}(z\overline{w})$. Both lengths are nonnegative, so $|z+w| = |z-w|$ if and only if $\mathrm{Re}(z\overline{w}) = 0$. For $z = a + bi$ and $w = c + di$, $z\overline{w} = (ac + bd) + (bc - ad)i$, so $\mathrm{Re}(z\overline{w}) = ac + bd$ is the dot product of (a, b) and (c, d). So the diagonals are equally long exactly when the sides z and w are perpendicular (or one of them is 0), that is, when the parallelogram is a rectangle.

Solution to Exercise B.1.36

Uses. Proposition B.1.8 (Properties of the modulus), Proposition B.1.2 (Conjugation respects the field operations)
Techniques. Square the modulus (Technique B.1.9)
Hint 1. Expand both squared moduli with Proposition B.1.8(f), or as $u\overline{u}$.
Hint 2. The cross terms are $-2\,\mathrm{Re}(\overline{z}w)$ and $-2\,\mathrm{Re}(z\overline{w})$. These are equal, because $\overline{z}w$ is the conjugate of $z\overline{w}$.
Hint 3. For (b), the right side of (a) is positive. For (c), it is zero.
Sketch. $|1 - \overline{z}w|^2 = 1 - 2\,\mathrm{Re}(\overline{z}w) + |z|^2|w|^2$ and $|z-w|^2 = |z|^2 - 2\,\mathrm{Re}(z\overline{w}) + |w|^2$; the real parts agree, and the difference is $1 - |z|^2 - |w|^2 + |z|^2|w|^2$. In (b), $|\overline{z}w| < 1$, so the denominator is nonzero, and the identity gives $|z-w|^2 < |1-\overline{z}w|^2$. In (c) the right side vanishes.

(a) By Proposition B.1.8(f) (applied to 1 and $-\overline{z}w$, and to z and $-w$) and (d),

$$|1 - \overline{z}w|^2 = 1 - 2\,\mathrm{Re}(\overline{z}w) + |z|^2|w|^2,$$
$$|z-w|^2 = |z|^2 - 2\,\mathrm{Re}(z\overline{w}) + |w|^2.$$

Since $\overline{z}w = \overline{z\overline{w}}$ (Proposition B.1.2), the two numbers have the same real part. Subtracting,

$$|1-\overline{z}w|^2 - |z-w|^2 = 1 - |z|^2 - |w|^2 + |z|^2|w|^2 = \left(1-|z|^2\right)\left(1-|w|^2\right).$$

(b) $|\overline{z}w| = |z||w| < 1$, so $\overline{z}w \neq 1$ and $1 - \overline{z}w \neq 0$. Both factors on the right of (a) are positive, so $|z-w|^2 < |1-\overline{z}w|^2$. Taking square roots and dividing by $|1-\overline{z}w| > 0$ gives the claim (Proposition B.1.8(d)).

(c) Now $1 - |w|^2 = 0$, so (a) gives $|1-\overline{z}w| = |z-w|$, and the denominator is nonzero by assumption. Hence the quotient has modulus 1.

So for fixed z with $|z| < 1$, the map $w \mapsto (z-w)/(1-\overline{z}w)$ sends the open unit disk into itself, by (b), and the unit circle into itself, by (c); the condition $\overline{z}w \neq 1$ holds automatically, since $|\overline{z}w| < 1$.

Solution to Exercise B.1.37

Uses. Proposition B.1.8 (Properties of the modulus), Proposition B.1.2 (Conjugation respects the field operations)
Techniques. Square the modulus (Technique B.1.9)
Hint 1. For (a) and (b), square both sides and write $z = x + yi$.
Hint 2. For (c), multiply the numerator and the denominator by $\overline{z+1}$. The denominator becomes a positive real number, and the numerator is $(z-1)(\overline{z}+1)$.
Sketch. (a) Squaring, $-2x + 1 = 2y + 1$: the line $y = -x$, the perpendicular bisector of 1 and $-i$. (b) Squaring gives $3x^2 + 10x + 3 + 3y^2 = 0$, the circle with centre $-5/3$ and radius $4/3$. (c) $(z-1)(\overline{z}+1) = |z|^2 - 1 + (z - \overline{z})$ has real part $|z|^2 - 1$, so the set is the unit circle without -1.

Write $z = x + yi$. Moduli are nonnegative, so each equation between moduli is equivalent to the equation between their squares.

(a) $|z-1|^2 = (x-1)^2 + y^2$ and $|z+i|^2 = x^2 + (y+1)^2$. They are equal if and only if $-2x + 1 = 2y + 1$, that is, $y = -x$. The set is the line through 0 with slope -1: the perpendicular bisector of the segment from 1 to $-i$, as it should be, since it consists of the points equidistant from 1 and $-i$.

(b) The condition is $(x-1)^2 + y^2 = 4\left((x+1)^2 + y^2\right)$, that is, $3x^2 + 10x + 3 + 3y^2 = 0$. Dividing by 3 and completing the square, $(x + 5/3)^2 + y^2 = 25/9 - 1 = 16/9$. The set is the circle with centre $-5/3$ and radius $4/3$.

(c) For $z \neq -1$,

$$\frac{z-1}{z+1} = \frac{(z-1)\overline{(z+1)}}{|z+1|^2} = \frac{(z-1)(\overline{z}+1)}{|z+1|^2}$$

by Proposition B.1.8(c), and $|z+1|^2 > 0$ is real. The numerator is $z\overline{z} + z - \overline{z} - 1 = \left(|z|^2 - 1\right) + (z - \overline{z})$, where $z - \overline{z} = 2yi$ is purely imaginary. So the real part of the quotient is $\left(|z|^2 - 1\right)/|z+1|^2$, which is 0 if and only if $|z| = 1$. The set is the unit circle with -1 removed. Geometrically, $(z-1)/(z+1)$ is purely imaginary exactly when the segments from z to 1 and to -1 are perpendicular (or $z = 1$), and the set of such z is the circle with diameter from -1 to 1 (Thales).

Solution to Exercise B.1.38

Uses. Proposition B.1.8 (Properties of the modulus), Proposition B.1.2 (Conjugation respects the field operations)
Techniques. Finding counterexamples and minimal examples (Technique 1.4.18)
Hint 1. Test each claim on $z, w \in \{1, i, 1+i\}$ before trying to prove it.
Hint 2. (b), (e), and (f) are true; for (e), expand $|z \pm w|^2$ with Proposition B.1.8(f).
Sketch. (a) False: $z = w = i$. (b) True: $|zz| = |z||z|$. (c) False: $z = i$. (d) False: $z = 1$, $w = i$. (e) True: $|z+w|^2 - |z-w|^2 = 4\,\mathrm{Re}(z\overline{w})$. (f) True: Proposition B.1.2(e).

(a) False. For $z = w = i$, $\mathrm{Re}(zw) = \mathrm{Re}(-1) = -1$, but $\mathrm{Re}\, z\, \mathrm{Re}\, w = 0$. The correct formula is $\mathrm{Re}(zw) = \mathrm{Re}\, z\, \mathrm{Re}\, w - \mathrm{Im}\, z\, \mathrm{Im}\, w$.

(b) True, by Proposition B.1.8(d) with $w = z$.

(c) False: $i^2 = -1$ is real, but i is not.

(d) False: $|1| = |i| = 1$, but $1 \neq \pm i$. Numbers of equal modulus differ by a factor of modulus 1, of which ± 1 are only two.

(e) True. By Proposition B.1.8(f), $|z \pm w|^2 = |z|^2 \pm 2\,\mathrm{Re}(z\overline{w}) + |w|^2$, so the hypothesis gives $4\,\mathrm{Re}(z\overline{w}) = 0$.

(f) True, by Proposition B.1.2(e).

Solution to Exercise B.1.39

Uses. Theorem B.1.20 (De Moivre's formula), Proposition B.1.29 (Square roots in $\mathbb{C}$), Proposition B.1.15 (Existence and uniqueness of polar form)
Hint 1. For which exponents n did we prove Theorem B.1.20? Was $z^{1/2}$ ever defined?
Hint 2. What does de Moivre's formula with $n = 2$ say about $\cos(\theta/2) + i\sin(\theta/2)$?
Sketch. Theorem B.1.20 holds for integer n only, and $1^{1/2}$ has no meaning: 1 has two square roots, ± 1. The formula $\cos(\theta/2) + i\sin(\theta/2)$ depends on the argument θ chosen, not only on the number $e^{i\theta}$, and the arguments 0 and 2π of 1 produce its two square roots.

The first sentence is the error. Theorem B.1.20 was proved for integers n, by induction and by inverting, and it says nothing for $n = 1/2$. Worse, the symbol $1^{1/2}$ was never defined: 1 has two square roots, 1 and -1 (Proposition B.1.29), and no rule has been given for choosing one.

What de Moivre's formula does give, with $n = 2$, is $(\cos(\theta/2) + i\sin(\theta/2))^2 = \cos\theta + i\sin\theta$. So $\cos(\theta/2) + i\sin(\theta/2)$ is *a* square root of $e^{i\theta}$. But it depends on θ, not only on the number $e^{i\theta}$, and a number has many arguments (Proposition B.1.15). The arguments 0 and 2π of 1 give its two different square roots, 1 and -1, exactly as in Theorem B.1.26 with $n = 2$. The rest of the computation is correct; it shows only that 1 and -1 are both square roots of 1.

Solution to Exercise B.1.40

Uses. Proposition 2.3.11 (Entries of $A\mathbf{x}$), Proposition B.1.2 (Conjugation respects the field operations)
Techniques. Conjugate the equation (Technique B.1.3)
Hint 1. Compute the ith entry of $A\overline{\mathbf{z}}$ with Proposition 2.3.11 and use that each a_{ij} is real.
Hint 2. For (b) and (c), conjugate the given equation and apply (a). For the real solution in (c), average a solution with its conjugate.
Sketch. (a) $\sum_j a_{ij}\overline{z_j} = \overline{\sum_j a_{ij} z_j}$ because $\overline{a_{ij}} = a_{ij}$. (b) $A\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \overline{\lambda}\,\overline{\mathbf{v}}$. (c) $A\overline{\mathbf{z}} = \overline{\mathbf{b}} = \mathbf{b}$, and then $(\mathbf{z} + \overline{\mathbf{z}})/2$, which has real entries $\mathrm{Re}\, z_j$, is a solution.

(a) By Proposition 2.3.11 and Proposition B.1.2, the ith entry of $A\overline{\mathbf{z}}$ is

$$\sum_{j=1}^{n} a_{ij}\overline{z_j} = \sum_{j=1}^{n} \overline{a_{ij}z_j} = \overline{\sum_{j=1}^{n} a_{ij}z_j},$$

since $\overline{a_{ij}} = a_{ij}$ for real a_{ij}. This is the ith entry of $\overline{A\mathbf{z}}$.

(b) By (a), $A\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \overline{\lambda}\,\overline{\mathbf{v}}$, where the last step conjugates each entry λv_j.

(c) If $A\mathbf{z} = \mathbf{b}$, then by (a), $A\overline{\mathbf{z}} = \overline{\mathbf{b}} = \mathbf{b}$, since $\mathbf{b}$ is real. So $\overline{\mathbf{z}}$ is also a solution. Then $\mathbf{x} = \frac{1}{2}(\mathbf{z} + \overline{\mathbf{z}})$ satisfies $A\mathbf{x} = \frac{1}{2}(\mathbf{b} + \mathbf{b}) = \mathbf{b}$, and its entries $\left(z_j + \overline{z_j}\right)/2 = \operatorname{Re} z_j$ are real (Proposition B.1.2(d)).

In the language of Section 7.2, (b) says that the nonreal eigenvalues of a real matrix come in conjugate pairs $\lambda, \overline{\lambda}$, with conjugate eigenvectors.

Solution to Exercise B.1.41

Uses. Proposition B.1.18 (Multiplication by w as a map of the plane), Theorem 2.4.2 (The matrix of a composite), Theorem B.1.25 (The nth roots of unity)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. By Proposition B.1.18, complex numbers act on $\mathbb{R}^2$ through the matrices $J_w = [M_w]_{\mathcal{BB}}$. How does J_wJ_z compare with J_{wz}?
Hint 2. $J_w^k = J_{w^k}$. So look for $w \in \mathbb{C}$ with $w^3 = -1$ but $w \neq -1$, and in (b), for an nth root of unity whose smaller powers are not 1.
Sketch. $M_wM_z = M_{wz}$, so $J_w^k = J_{w^k}$, and $J_w = J_{w'}$ only if $w = w'$. (a) $w = e^{i\pi/3}$ has $w^3 = -1$, giving $A = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$. (b) $w = e^{2\pi i/n}$, the rotation by $2\pi/n$.

For $w = c + di$ let $J_w = [M_w]_{\mathcal{BB}} = \begin{pmatrix} c & -d \\ d & c \end{pmatrix}$, as in Proposition B.1.18. Since $M_w \circ M_z = M_{wz}$, Theorem 2.4.2 gives $J_wJ_z = J_{wz}$, and by induction $J_w^k = J_{w^k}$ for $k \geq 1$. Also J_w determines w (its first column is (c, d)), and $J_1 = I_2$, $J_{-1} = -I_2$.

(a) Let $w = e^{i\pi/3} = 1/2 + \left(\sqrt{3}/2\right)i$. Then $w^3 = e^{i\pi} = -1$ by Theorem B.1.20, so

$$A = J_w = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$$

satisfies $A^3 = J_{-1} = -I_2$, and $A \neq -I_2$. Geometrically, A is the rotation by $\pi/3$, and three of them make a half turn.

(b) Let $\omega = e^{2\pi i/n}$ and $A = J_\omega = \begin{pmatrix} \cos(2\pi/n) & -\sin(2\pi/n) \\ \sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix}$, the rotation by $2\pi/n$. Then $A^k = J_{\omega^k}$, which equals $I_2 = J_1$ if and only if $\omega^k = 1$, that is, $n \mid k$ (Theorem B.1.25(a)). So $A^n = I_2$ and $A^k \neq I_2$ for $1 \leq k < n$.

Solution to Exercise B.1.42

Uses. Definition 2.4.1 (Matrix product), Proposition 2.4.14 (Algebra of matrices), Definition 2.4.21 (Transpose), Definition 5.1.14 (The 2×2 determinant), Proposition B.1.18 (Multiplication by w as a map of the plane)
Techniques. Translate between maps and matrices (Technique 2.3.27)
Hint 1. Compute $\varphi(z)\varphi(w)$ directly, or note that $\varphi(z) = [M_z]_{\mathcal{BB}}$ and $M_z \circ M_w = M_{zw}$.
Hint 2. For (c), most field axioms hold because they hold for all matrices (Proposition 2.4.14). Commutativity of multiplication and inverses can be transported from $\mathbb{C}$ with (a).
Sketch. (a) Entrywise for sums; $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{pmatrix} = \varphi(zw)$. (b) Swap b and $-b$; $\det = a^2 + b^2$. (c) K contains 0 and I_2 and is closed under sums, negatives, and products; $\varphi(z)\varphi(w) = \varphi(zw) = \varphi(w)\varphi(z)$; and $\varphi(z)\varphi\left(z^{-1}\right) = \varphi(1) = I_2$.

Let $z = a + bi$ and $w = c + di$.

(a) $\varphi(z + w) = \begin{pmatrix} a+c & -(b+d) \\ b+d & a+c \end{pmatrix} = \varphi(z) + \varphi(w)$. By Definition 2.4.1,

$$\varphi(z)\varphi(w) = \begin{pmatrix} ac - bd & -ad - bc \\ bc + ad & -bd + ac \end{pmatrix} = \varphi((ac - bd) + (ad + bc)i) = \varphi(zw).$$

(Equivalently, $\varphi(z) = [M_z]_{\mathcal{BB}}$ by Proposition B.1.18, and $M_z \circ M_w = M_{zw}$.) Also $\varphi(1) = I_2$, and $\varphi(z)$ determines z through its first column (a, b), so φ is injective.

(b) $\varphi(\overline{z}) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \varphi(z)^{\mathsf{T}}$, and $\det \varphi(z) = a \cdot a - (-b) \cdot b = a^2 + b^2 = |z|^2$ (Definition 5.1.14).

(c) By (a), K is closed under addition and multiplication, and it contains $\varphi(0) = 0$, $\varphi(1) = I_2$, and the negative $-\varphi(z) = \varphi(-z)$ of each of its elements. Addition of matrices is commutative and associative, and multiplication is associative and distributes over addition (Proposition 2.4.14), so these laws hold in K; 0 and I_2 are identities, and $I_2 \neq 0$. Multiplication in K is commutative, because $\varphi(z)\varphi(w) = \varphi(zw) = \varphi(wz) = \varphi(w)\varphi(z)$. Finally, a nonzero element of K is $\varphi(z)$ with $z \neq 0$, because φ is injective and $\varphi(0) = 0$; then $\varphi(z)\varphi\left(z^{-1}\right) = \varphi(1) = I_2 = \varphi\left(z^{-1}\right)\varphi(z)$, so $\varphi\left(z^{-1}\right) \in K$ is its inverse. So K is a field, and $\varphi : \mathbb{C} \to K$ is a bijection that respects sums and products: K is a copy of $\mathbb{C}$ made of real matrices, in which i is the quarter turn $\varphi(i) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and in which conjugation is transposition.

Solution to Exercise B.1.43

Uses. Exercise 1.1.44, Proposition B.1.2 (Conjugation respects the field operations)
Hint 1. σ is determined by $\sigma(i)$, since $\sigma(a + bi) = \sigma(a) + \sigma(b)\sigma(i)$.
Hint 2. Apply σ to $i^2 = -1$. Which complex numbers square to -1?
Sketch. $\sigma(i)^2 = \sigma\left(i^2\right) = \sigma(-1) = -1$, so $\sigma(i) = \pm i$. Then $\sigma(a + bi) = a + b\sigma(i)$ is $a + bi$ or $a - bi$.

Since σ respects products and fixes -1, $\sigma(i)^2 = \sigma\left(i^2\right) = \sigma(-1) = -1$. By Exercise 1.1.44(a), $\sigma(i) = i$ or $\sigma(i) = -i$. For $a, b \in \mathbb{R}$, since σ respects sums and products and fixes a and b,

$$\sigma(a + bi) = \sigma(a) + \sigma(b)\sigma(i) = a + b\sigma(i).$$

If $\sigma(i) = i$, this is $a + bi$, so σ is the identity; if $\sigma(i) = -i$, it is $a - bi$, so σ is conjugation. Both are indeed automorphisms fixing $\mathbb{R}$ (Proposition B.1.2).

The hypothesis that σ fixes $\mathbb{R}$ matters. Every automorphism fixes $\mathbb{Q}$ (it fixes 1, hence every integer and every quotient of integers), but $\mathbb{C}$ has other, very discontinuous, automorphisms that move real numbers; proving that they exist needs the axiom of choice, for instance in the form of Zorn's lemma (Section A.4).

Solution to Exercise B.1.44

Uses. Proposition B.1.8 (Properties of the modulus), Proposition B.1.28 (Sum and product of the roots of unity), Theorem B.1.25 (The nth roots of unity)
Techniques. Rotate the phase (Technique 6.2.8), Square the modulus (Technique B.1.9)
Hint 1. Rotate so that $z_1 = 1$: divide everything by z_1. The hypotheses are unchanged.
Hint 2. With $w_k = z_k/z_1$, you have $w_3 = -1 - w_2$. Take the squared modulus of both sides to find $\operatorname{Re} w_2$.
Sketch. Let $w_k = z_k/z_1$: then $w_1 = 1$, $|w_2| = |w_3| = 1$, and $1 + w_2 + w_3 = 0$. Now $1 = |w_3|^2 = |1 + w_2|^2 = 2 + 2\operatorname{Re} w_2$, so $\operatorname{Re} w_2 = -1/2$ and $w_2 \in \{\omega, \omega^2\}$; then $w_3 = -1 - w_2$ is the other one. So $z_k \in \{z_1, z_1\omega, z_1\omega^2\}$.

Since $|z_1| = 1$, $z_1 \neq 0$. Let $w_k = z_k/z_1$. Then $w_1 = 1$, $|w_2| = |w_3| = 1$ by Proposition B.1.8(d), and dividing $z_1 + z_2 + z_3 = 0$ by z_1 gives $1 + w_2 + w_3 = 0$. So $w_3 = -(1 + w_2)$, and by Proposition B.1.8(f),

$$1 = |w_3|^2 = |1 + w_2|^2 = 1 + 2\operatorname{Re} w_2 + |w_2|^2 = 2 + 2\operatorname{Re} w_2.$$

Hence $\operatorname{Re} w_2 = -1/2$, and $(\operatorname{Im} w_2)^2 = 1 - 1/4 = 3/4$, so $w_2 = -1/2 \pm \left(\sqrt{3}/2\right)i$. These two numbers are ω and $\omega^2 = \overline{\omega}$. By Proposition B.1.28, $1 + \omega + \omega^2 = 0$, so if $w_2 = \omega$ then $w_3 = -1 - \omega = \omega^2$, and if $w_2 = \omega^2$ then $w_3 = \omega$. Either way $\{w_1, w_2, w_3\} = \{1, \omega, \omega^2\}$, and multiplying by $u = z_1$ gives $\{z_1, z_2, z_3\} = \{u, u\omega, u\omega^2\}$.

For $j \neq k$ in $\{0, 1, 2\}$, $\left|u\omega^j - u\omega^k\right| = \left|\omega^j\right|\left|1 - \omega^{k-j}\right| = \left|1 - \omega^{k-j}\right|$, and $\omega^{k-j} \in \{\omega, \omega^2 = \overline{\omega}\}$, so this is $|1 - \omega| = \left|3/2 - \left(\sqrt{3}/2\right)i\right| = \sqrt{3}$ in every case (Proposition B.1.8(b)). So all three sides have length $\sqrt{3}$.

Solution to Exercise B.1.45

Uses. Proposition B.1.28 (Sum and product of the roots of unity), Proposition 7.1.22 (Counting roots), Proposition B.1.8 (Properties of the modulus), Remark B.1.13 (Facts about cos and sin)
Techniques. Polynomial identity arguments (Technique 5.5.24)

Hint 1. Approach 2 of Proposition B.1.28 shows $x^n - 1 = (x-1)(x-\omega)\cdots(x-\omega^{n-1})$. Also $x^n - 1 = (x-1)(1 + x + \cdots + x^{n-1})$. You cannot simply cancel $x - 1$ yet; compare values instead.
Hint 2. Both sides of (a) take the same value at every $c \neq 1$. Two polynomials of degree at most $n-1$ that agree at infinitely many points are equal (Proposition 7.1.22).
Hint 3. For (c), $\left|1 - e^{i\theta}\right|^2 = 2 - 2\cos\theta = 4\sin^2(\theta/2)$.
Sketch. Let $P = 1 + x + \cdots + x^{n-1}$ and $Q = \prod_{k=1}^{n-1}(x - \omega^k)$. For every $c \neq 1$, $(c-1)P(c) = c^n - 1 = (c-1)Q(c)$, so $P(c) = Q(c)$; as $P - Q$ has degree at most $n-1$ and infinitely many roots, $P = Q$. Evaluating at 1 gives (b). Taking moduli, $\prod\left|1 - \omega^k\right| = n$, and $\left|1 - \omega^k\right| = 2\sin(k\pi/n)$.

(a) Let $P = 1 + x + \cdots + x^{n-1}$ and $Q = (x-\omega)(x-\omega^2)\cdots(x-\omega^{n-1})$. By Approach 2 of Proposition B.1.28, $x^n - 1 = (x-1)Q$, and multiplying out, $x^n - 1 = (x-1)P$. Evaluating at $c \in \mathbb{C}$ (evaluation respects products, Section 7.1),

$$(c-1)P(c) = c^n - 1 = (c-1)Q(c).$$

For $c \neq 1$ we may cancel $c - 1$, so $P(c) = Q(c)$ for every $c \neq 1$. Both P and Q have degree $n-1$, so $P - Q$ has degree at most $n-1$ and infinitely many roots, and $P - Q = \mathbf{0}$ by Proposition 7.1.22.

(b) Evaluating (a) at 1 gives $n = P(1) = Q(1) = (1-\omega)(1-\omega^2)\cdots(1-\omega^{n-1})$.

(c) The vertices of a regular n-gon inscribed in the unit circle are equally spaced on it, so they are $u, u\omega, \ldots, u\omega^{n-1}$, where u is the chosen vertex and $|u| = 1$. Multiplication by $\overline{u}$ preserves distances, since $|\overline{u}a - \overline{u}b| = |\overline{u}||a-b| = |a-b|$, and it carries the vertices to $1, \omega, \ldots, \omega^{n-1}$, with the chosen vertex going to 1. Taking moduli in (b), with Proposition B.1.8(d), the product of the distances is $\prod_{k=1}^{n-1}\left|1-\omega^k\right| = n$.

For $\theta \in \mathbb{R}$, by (T1),

$$\left|1 - e^{i\theta}\right|^2 = (1-\cos\theta)^2 + \sin^2\theta = 2 - 2\cos\theta = 4\sin^2\left(\frac{\theta}{2}\right),$$

where the last step is $\cos\theta = \cos^2(\theta/2) - \sin^2(\theta/2) = 1 - 2\sin^2(\theta/2)$. For $\theta = 2\pi k/n$ with $1 \leq k \leq n-1$, we have $0 < \theta/2 < \pi$, so $\sin(\theta/2) > 0$ and $\left|1-\omega^k\right| = 2\sin(k\pi/n)$. Hence $2^{n-1}\prod_{k=1}^{n-1}\sin(k\pi/n) = n$.

Solution to Exercise B.1.46

Uses. Theorem B.1.25 (The nth roots of unity), Euclid's lemma: if $n \mid km$ and $\gcd(k,n) = 1$, then $n \mid m$
Hint 1. By Theorem B.1.25(a), $\left(\omega^k\right)^m = 1$ if and only if $n \mid km$.
Hint 2. If $g = \gcd(k,n) > 1$, try $m = n/g$. If $g = 1$, use Euclid's lemma.
Hint 3. For (b), show that the n powers are distinct, and count.
Sketch. (a) $\left(\omega^k\right)^m = 1$ if and only if $n \mid km$. If $g = \gcd(k,n) > 1$, then $m = n/g < n$ works, so ω^k is not primitive. If $g = 1$, then $n \mid km$ forces $n \mid m$, so $m \geq n$. (b) If $\zeta^j = \zeta^l$ with $0 \leq j < l < n$, then $\zeta^{l-j} = 1$, which is impossible; so the powers are n distinct nth roots of unity. (c) $k = 1, 5, 7, 11$: $\pm\sqrt{3}/2 \pm i/2$.

(a) By Theorem B.1.25(a), $\left(\omega^k\right)^m = \omega^{km} = 1$ if and only if $n \mid km$. Let $g = \gcd(k,n)$. If $g > 1$, then $m = n/g$ satisfies $1 \leq m < n$ and $km = (k/g)n$, so $\left(\omega^k\right)^m = 1$ and ω^k is not primitive. If $g = 1$ and $\left(\omega^k\right)^m = 1$ with $m \geq 1$, then $n \mid km$, so $n \mid m$ by Euclid's lemma, and $m \geq n$. So ω^k is primitive.

(b) Each ζ^j is an nth root of unity, by Theorem B.1.25(c). If $\zeta^j = \zeta^l$ with $0 \leq j < l \leq n-1$, then $\zeta^{l-j} = 1$ with $1 \leq l - j < n$, contradicting primitivity. So $1, \zeta, \ldots, \zeta^{n-1}$ are n distinct nth roots of unity, and by Theorem B.1.25(b) there are only n, so these are all of them.

(c) For $n = 12$, the k in $\{0, \ldots, 11\}$ with $\gcd(k,12) = 1$ are $1, 5, 7, 11$. So the primitive 12th roots of unity are

$$e^{i\pi/6} = \frac{\sqrt{3}}{2} + \frac{i}{2}, \quad e^{5\pi i/6} = -\frac{\sqrt{3}}{2} + \frac{i}{2}, \quad e^{7\pi i/6} = -\frac{\sqrt{3}}{2} - \frac{i}{2}, \quad e^{11\pi i/6} = \frac{\sqrt{3}}{2} - \frac{i}{2}.$$

Solution to Exercise B.1.47

Uses. Theorem 7.1.14 (Division with remainder), Proposition B.1.28 (Sum and product of the roots of unity), Theorem B.1.25 (The nth roots of unity)

Hint 1. Let $p = x^{2n} + x^n + 1$ and $\omega = e^{2\pi i/3}$, a root of $x^2 + x + 1$. Divide p by $x^2 + x + 1$ with remainder in $\mathbb{R}[x]$, and evaluate at ω.
Hint 2. The remainder is $cx + d$ with $c, d \in \mathbb{R}$, and $p(\omega) = c\omega + d$. Since ω is not real, $c\omega + d = 0$ forces $c = d = 0$.
Hint 3. Compute $p(\omega) = 1 + \zeta + \zeta^2$ with $\zeta = \omega^n$. When is $\zeta = 1$?
Sketch. Write $p = q\big(x^2 + x + 1\big) + (cx + d)$ with $q \in \mathbb{R}[x]$ and $c, d \in \mathbb{R}$. Since $\omega^2 + \omega + 1 = 0$, $p(\omega) = c\omega + d$, whose imaginary part is $c\sqrt{3}/2$; so the remainder vanishes if and only if $p(\omega) = 0$. With $\zeta = \omega^n$, $p(\omega) = 1 + \zeta + \zeta^2$, which is 3 if $3 \mid n$ and 0 otherwise, since then ζ is a cube root of unity other than 1.

Let $p = x^{2n} + x^n + 1$ and $\omega = e^{2\pi i/3} = -1/2 + \big(\sqrt{3}/2\big)i$. By Proposition B.1.28, $\omega^2 + \omega + 1 = 0$. By the division algorithm in $\mathbb{R}[x]$ (Theorem 7.1.14),

$$p = q\big(x^2 + x + 1\big) + (cx + d)$$

for some $q \in \mathbb{R}[x]$ and $c, d \in \mathbb{R}$, and $x^2 + x + 1$ divides p if and only if $c = d = 0$. Evaluating at ω (evaluation respects sums and products, Section 7.1) gives $p(\omega) = q(\omega) \cdot 0 + c\omega + d = c\omega + d$. Now $c\omega + d = (d - c/2) + \big(c\sqrt{3}/2\big)i$ is 0 if and only if $c = 0$ and $d = 0$. So $x^2 + x + 1$ divides p if and only if $p(\omega) = 0$.

Let $\zeta = \omega^n$. Then $p(\omega) = 1 + \zeta + \zeta^2$, and $\zeta^3 = \big(\omega^3\big)^n = 1$. If $3 \mid n$, then $\zeta = 1$ by Theorem B.1.25(a), and $p(\omega) = 3 \neq 0$. If $3 \nmid n$, then $\zeta \neq 1$ by Theorem B.1.25(a), so $(\zeta - 1)\big(1 + \zeta + \zeta^2\big) = \zeta^3 - 1 = 0$ gives $1 + \zeta + \zeta^2 = 0$. So $p(\omega) = 0$ if and only if $3 \nmid n$.

For instance, $x^4 + x^2 + 1 = \big(x^2 + x + 1\big)\big(x^2 - x + 1\big)$, while $x^6 + x^3 + 1$ is not divisible by $x^2 + x + 1$.

Solution to Exercise B.1.48

Uses. Proposition B.1.8 (Properties of the modulus), Definition 1.1.29 (Complex numbers)
Hint 1. For (a), use the product formula of Definition 1.1.29 and Proposition B.1.8(d).
Hint 2. For (b), take norms in $zw = 1$. Which integers a, b have $a^2 + b^2 = 1$?
Hint 3. For (c), compute $N((2 + i)(3 + 2i))$ and $N((2 + i)(3 - 2i))$.
Sketch. (a) $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$ has integer coordinates, $N(a + bi) = a^2 + b^2$, and $N(zw) = |zw|^2 = |z|^2|w|^2$. (b) $N(z)N(w) = 1$ forces $N(z) = 1$, so $z \in \{\pm 1, \pm i\}$; each of these has an inverse in $\mathbb{Z}[i]$. (c) $(2 + i)(3 + 2i) = 4 + 7i$ and $(2 + i)(3 - 2i) = 8 - i$, so $65 = 4^2 + 7^2 = 8^2 + 1^2$.

(a) For $z = a + bi$ and $w = c + di$ with integer coordinates, $z + w$, $-z$, and $zw = (ac - bd) + (ad + bc)i$ have integer coordinates. $N(z) = a^2 + b^2$ is a nonnegative integer, and $N(zw) = |zw|^2 = |z|^2|w|^2 = N(z)N(w)$ by Proposition B.1.8(d).

(b) If $zw = 1$ with $z, w \in \mathbb{Z}[i]$, then $N(z)N(w) = N(1) = 1$ by (a), and $N(z), N(w)$ are nonnegative integers, so $N(z) = 1$. If $z = a + bi$, then $a^2 + b^2 = 1$ with $a, b \in \mathbb{Z}$, so one of a, b is ± 1 and the other is 0: $z \in \{1, -1, i, -i\}$. Conversely, $1 \cdot 1 = (-1)(-1) = i(-i) = 1$, so all four are units.

(c) By (a), $N((2 + i)(3 + 2i)) = 5 \cdot 13 = 65$, and $(2 + i)(3 + 2i) = 6 + 4i + 3i + 2i^2 = 4 + 7i$, so $65 = 4^2 + 7^2$. Also $N(3 - 2i) = 13$, and $(2 + i)(3 - 2i) = 6 - 4i + 3i - 2i^2 = 8 - i$, so $65 = 8^2 + 1^2$. In coordinates, (a) is the identity $\big(a^2 + b^2\big)\big(c^2 + d^2\big) = (ac - bd)^2 + (ad + bc)^2$ noted after Proposition B.1.8.

Solution to Exercise B.1.49

Uses. Proposition B.1.8 (Properties of the modulus), Notation B.1.19 ($e^{i\theta}$), integrals of continuous functions
Techniques. Rotate the phase (Technique 6.2.8)
Hint 1. For a direction $e^{i\theta}$, keep the terms that make an acute angle with it: $S_\theta = \big\{k \;\big|\; \mathrm{Re}\big(e^{-i\theta} z_k\big) > 0\big\}$. Bound $\big|\sum_{k \in S_\theta} z_k\big|$ from below by a real part.
Hint 2. $\big|\sum_{k \in S_\theta} z_k\big| \geq f(\theta) = \sum_k \max\big(0, \mathrm{Re}\big(e^{-i\theta} z_k\big)\big)$. You cannot choose the best θ directly; average f over $\theta \in [0, 2\pi]$ instead.
Hint 3. With $z_k = |z_k| e^{i\varphi_k}$, $\mathrm{Re}\big(e^{-i\theta} z_k\big) = |z_k| \cos(\theta - \varphi_k)$, and the integral of $\max(0, \cos t)$ over a period is 2.
Sketch. For each θ, $\big|\sum_{S_\theta} z_k\big| \geq \mathrm{Re}\big(e^{-i\theta} \sum_{S_\theta} z_k\big) = f(\theta)$. Since $\int_0^{2\pi} f = \sum_k |z_k| \int_{-\pi/2}^{\pi/2} \cos t \, dt = 2 \sum_k |z_k|$, the average of f over $[0, 2\pi]$ is $(1/\pi) \sum_k |z_k|$, and the continuous function f reaches its average somewhere.

For $\theta \in \mathbb{R}$ let $S_\theta = \big\{k \;\big|\; \mathrm{Re}\big(e^{-i\theta} z_k\big) > 0\big\}$ and

$$f(\theta) = \sum_{k=1}^{n} \max\big(0, \mathrm{Re}\big(e^{-i\theta} z_k\big)\big) = \sum_{k \in S_\theta} \mathrm{Re}\big(e^{-i\theta} z_k\big).$$

By Proposition B.1.8(d) and (e),

$$\left|\sum_{k\in S_\theta} z_k\right| = \left|e^{-i\theta}\sum_{k\in S_\theta} z_k\right| \geq \operatorname{Re}\left(\sum_{k\in S_\theta} e^{-i\theta} z_k\right) = f(\theta).$$

So it suffices to find a θ with $f(\theta) \geq (1/\pi)\sum_k |z_k|$.

Write $z_k = |z_k| e^{i\varphi_k}$, with any φ_k if $z_k = 0$. Then $\operatorname{Re}\left(e^{-i\theta} z_k\right) = |z_k| \operatorname{Re}\left(e^{i(\varphi_k - \theta)}\right) = |z_k| \cos(\theta - \varphi_k)$. So f is continuous and 2π-periodic, and integrating each term over a period centred at φ_k,

$$\int_0^{2\pi} f(\theta)\,\mathrm{d}\theta = \sum_{k=1}^{n} |z_k| \int_{\varphi_k-\pi}^{\varphi_k+\pi} \max(0, \cos(\theta - \varphi_k))\,\mathrm{d}\theta = \sum_{k=1}^{n} |z_k| \int_{-\pi/2}^{\pi/2} \cos t\,\mathrm{d}t = 2\sum_{k=1}^{n} |z_k|.$$

If we had $f(\theta) < (1/\pi)\sum_k |z_k|$ for every $\theta \in [0, 2\pi]$, then, as f is continuous, the integral would be less than $2\pi \cdot (1/\pi)\sum_k |z_k| = 2\sum_k |z_k|$, a contradiction. So some θ has $f(\theta) \geq (1/\pi)\sum_k |z_k|$, and $S = S_\theta$ works.

This is an averaging argument. We cannot say in advance which direction θ is best, but the average of f over all directions is computable, and some direction does at least as well as the average.