Finite-Dimensional Linear Algebra
Solutions to selected odd-numbered exercises

Mark S. Gockenbach

September 19, 2018

Errata for the first printing

The following corrections will be made in the second printing of the text, expected in 2011. These solutions are written as if they have already been made.

Page 65: Exercise 14: belongs in Section 2.7.
Page 65: Exercise 16: should read "(cf. Exercise 2.3.21)", not "(cf. Exercise 2.2.21)".
Page 71: Exercise 9 (b): $\mathbf{Z}_4^5$ should be $\mathbf{Z}_5^4$.
Page 72: Exercise 11: "over V" should be "over F".
Page 72: Exercise 15: "$i = 1, 2, \ldots, k$" should be "$j = 1, 2, \ldots, k$" (twice).
Page 79: Exercise 1: "$x_3 = 2$" should be "$x_3 = 3$".
Page 82: Exercise 14(a): "Each A_i and B_i has degree $2n + 1$" should read "$A_i, B_i \in \mathcal{P}_{2n+1}$ for all $i = 0, 1, \ldots, n$".
Page 100, Exercise 11: "$K : C[a, b] \to C[a, b]$" should be "$K : C[c, d] \to C[a, b]$"
Page 114, Line 9: "$L : F^n \to \mathbf{R}^m$" should be "$L : F^n \to F^m$".
Page 115: Exercise 8:

$$\mathcal{S} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \ \mathcal{X} = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}.$$

should be

$$\mathcal{S} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \ \mathcal{X} = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}.$$

Page 116, Exercise 17(b): "$\mathcal{F}^{mn}$" should be "F^{mn}".
Page 121, Exercise 3: "$T : \mathbf{R}^4 \to \mathbf{R}^3$" should be "$T : \mathbf{R}^4 \to \mathbf{R}^4$".
Page 124, Exercise 15: "$T : X/\mathrm{ker}(L) \to \mathcal{R}(U)$" should be "$T : X/\mathrm{ker}(L) \to \mathcal{R}(L)$".
Page 124, Exercise 15:

$$T([x]) = T(x) \text{ for all } [x] \in X/\mathrm{ker}(L)$$

should be

$$T([x]) = L(x) \text{ for all } [x] \in X/\mathrm{ker}(L).$$

Page 129, Exercise 4(b): Period is missing at the end of the sentence.
Page 130, Exercise 8: $L : \mathbf{Z}_3^3 \to \mathbf{Z}_3^3$ should read $L : \mathbf{Z}_5^3 \to \mathbf{Z}_5^3$
Page 130, Exercise 13(b): "T defines ..." should be "S defines ...".
Page 131, Exercise 15: "$K : C[a, b] \times C[c, d] \to C[a, b]$" should be "$K : C[c, d] \to C[a, b]$".
Page 138, Exercise 7(b): "define" should be "defined".
Page 139, Exercise 12: In the last line, "$\mathrm{sp}\{x_1, x_2, \ldots, x_n\}$" should be "$\mathrm{sp}\{x_1, x_2, \ldots, x_k\}$".
Page 139, Exercise 12: The proposed plan for the proof is not valid. Instead, the instructions should read: Choose vectors $x_1, \ldots, x_k \in X$ such that $\{T(x_1), \ldots, T(x_k)\}$ is a basis for $\mathcal{R}(T)$, and choose a basis $\{y_1, \ldots, y_\ell\}$ for $\mathrm{ker}(T)$. Prove that $\{x_1, \ldots, x_k, y_1, \ldots, y_\ell\}$ is a basis for X. (Hint: First show that $\mathrm{ker}(T) \cap \mathrm{sp}\{x_1, \ldots, x_k\}$ is trivial.)
Page 140, Exercise 15: In the displayed equation, $|A_{ii}$ should be $|A_{ii}|$.
Page 168: Definition 132 defines the adjacency matrix of a graph, not the incidence matrix (which is something different). The correct term (adjacency matrix) is used throughout the rest of the section. (Change "incidence" to "adjacency" in three places: the title of Section 3.10.1, Page 168 line -2, Page 169 line 1.)
Page 199, Equation (3.41d): "$x_1, x_2 \leq 0$" should be "$x_1, x_2 \geq 0$".

Page 204, Exercise 10: "$\alpha_1,\ldots,\alpha_k \in \mathbf{R}$" should be "$\alpha_1,\ldots,\alpha_k \geq 0$". Also, C should not be boldface in the displayed formula.
Page 221, Exercise 9: "$m > n$" should be "$m < n$".
Page 242, Corollary 194: "for each $i = 1, 2, \ldots, t$" should be "for each $i = 1, 2, \ldots, m$".
Page 251, Exercise 18(e):

$$w = \left[\frac{0}{v}\right],$$

should be

$$w = \left[\frac{0}{v}\right].$$

(That is, the comma should be a period.)
Page 256, Exercise 13: First line should read "Let X be a finite-dimensional vector space over $\mathbf{C}$ with basis…". References in part (b) to $F^{n\times n}$, $F^{k\times k}$, $F^{k\times \ell}$, $F^{\ell\times\ell}$ should be replaced with $\mathbf{C}^{n\times n}$, etc. Also, in part (b), "Prove that $[T]_{\mathcal{X}}$" should be replaced with "Prove that $[T]_{\mathcal{X},\mathcal{X}}$".
Page 264, Exercise 3: Add "Assume $\{p, q\}$ is linearly independent."
Page 271, Exercise 3: "…we introduced the incidence matrix …" should be "…we introduced the adjacency matrix …".
Page 282, Exercise 6: $S = \mathrm{sp}\{(1, 3, -3, 2), (3, 7, -11, -4)\}$ should be $S = \mathrm{sp}\{(1, 4, -1, 3), (4, 7, -19, 3)\}$.
Page 282, Exercise 7(b): "$\mathcal{N}(A) \cap \mathrm{col}(A)$" should be "$\mathcal{N}(A) \cap \mathrm{col}(A) = \{0\}$".
Page 283, Exercise 12: "Lemma 5.1.2" should be "Lemma 229".
Page 306, Example 252: "$\mathcal{B} = \{p_0, D(p_0), D^2(p_0)\} = \{x^2, 2x, 2\}$" should be "$\mathcal{B} = \{D^2(p_0), D(p_0), p_0\} = \{2, 2x, x^2\}$". Also, "$[T]_{\mathcal{B},\mathcal{B}}$" should be "$[D]_{\mathcal{B},\mathcal{B}}$" (twice). Similarly, $\mathcal{A}$ should be defined as $\{2, -1+2x, 1-x+x^2\}$ and "$[T]_{\mathcal{A},\mathcal{A}}$" should be "$[D]_{\mathcal{A},\mathcal{A}}$".
Page 308, Exercise 3: "Suppose X is a vector space…" should be "Suppose X is a finite-dimensional vector space…".
Page 311, Line 7: "corresponding to λ" should be "corresponding to λ_i".
Page 316, Exercise 6(f): Should end with a ";" instead of a ".".
Page 317, Exercise 15: "$\ker((T - \lambda I)^2) = \ker(A - \lambda I)$" should be "$\ker((T - \lambda I)^2) = \ker(T - \lambda I)$".
Page 322, displayed equation (5.21): The last line should read $v_r' = \lambda v_r$.
Page 325, Exercise 9: "If $U(t_0)$ is singular, say $U(t)c = 0$ for some $c \in \mathbf{C}^n$, $c \neq 0$" should be "If $U(t_0)$ is singular, say $U(t_0)c = 0$ for some $c \in \mathbf{C}^n$, $c \neq 0$".
Page 331, Line 16: "…is at least $t + 1$" should be "…is at least $s + 1$".
Page 356, Exercise 9: "…such that $\{x_1, x_2, x_3, x_4\}$." should be "…such that $\{x_1, x_2, x_3, x_4\}$ is an orthogonal basis for $\mathbf{R}^4$."
Page 356, Exercise 13: "…be a linearly independent subset of V" should be "…be an orthogonal subset of V"
Page 356, Exercise 14: "…be a linearly independent subset of V" should be "…be an orthogonal subset of V"
Pages 365–368: Miscellaneous exercises 1–21 should be numbered 2–22.
Page 365, Exercise 6 (should be 7): "…under the $L^2(0, 1)$ norm" should be "…under the $L^2(0, 1)$ inner product".
Page 383, Line 1: "$\mathrm{col}(T)$" should be "$\mathrm{col}(A)$" and "$\mathrm{col}(T)^\perp$" should be "$\mathrm{col}(A)^\perp$".
Page 383, Exercise 3: "…a basis for $\mathbf{R}^4$" should be "…a basis for $\mathbf{R}^3$".
Page 384, Exercise 6: "basis" should be "bases".
Page 385, Exercise 14: "Exercise 6.4.13" should be "Exercise 6.4.1". "That exercise also" should be "Exercise 6.4.13".
Page 385, Exercise 15: "See Exercise 6.4" should be "See Exercise 6.4.14".
Page 400, Exercise 4:

$$\tilde{f}(x) = f\left(a + \frac{b-a}{2}(t+1)\right).$$

should be

$$\tilde{f}(x) = f\left(a + \frac{b-a}{2}(x+1)\right).$$

Page 410, Exercise 1: The problem should specify $\ell = 1$, $k(x) = x + 1$, $f(x) = -4x - 1$.
Page 411, Exercise 6: "$u(\ell) = 0$." should be "$u(\ell) = 0$" (i.e. there should not be a period after 0).
Page 424, Exercise 1: "…prove (1)" should be "…prove (6.50)".
Page 432, Exercise 9: "$G^{-1/2}$ is the inverse of $G^{-1/2}$" should be "$G^{-1/2}$ is the inverse of $G^{1/2}$".
Page 433, Exercise 16: "…so we will try to estimate the values $u(x_1), u(x_2), \ldots, u(x_n)$" should be "…so we will try to estimate the values $u(x_1), u(x_2), \ldots, u(x_{n-1})$".
Page 438, Exercise 3: "…define $T : \mathbf{R}^n \to F^n$" should be "…define $T : F^n \to F^n$".
Page 448, Exercise 8: In the formula for f, $-200x_1^2x2$ should be $-200x_1^2x_2$. Also, $(-1.2, 1)$ should be $(1, 1)$.
Page 453, Exercise 6: Add: "Assume $\nabla g(x(0))$ has full rank."
Page 475, Exercise 10: "$A = GH$" should be "$A = GQ$".
Page 476, Exercise 15(a):

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{ij}^2} \text{ for all } A \in \mathbf{C}^{m\times n}.$$

should be

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{ij}|^2} \text{ for all } A \in \mathbf{C}^{m\times n}.$$

Page 476, Exercise 15: No need to define $\|C\|_F$ again.
Page 501, last paragraph: The text fails to define $k \equiv \ell \pmod p$ for general $k, \ell \in \mathbf{Z}$. The following text should be added: "In general, for $k, \ell \in \mathbf{Z}$, we say that $k \equiv \ell \pmod p$ if p divides $k - \ell$, that is, if there exists $m \in \mathbf{Z}$ with $k = \ell + mp$. It is easy to show that, if r is the congruence class of $k \in \mathbf{Z}$, then p divides $k - r$, and hence this is consistent with the earlier definition. Moreover, it is a straightforward exercise to show that $k \equiv \ell \pmod p$ if and only if k and ℓ have the same congruence class modulo p."
Page 511, Theorem 381: "$A_{ij}^{(k)} = A_{ij}$" should be "$M_{ij}^{(k)} = A_{ij}$".
Page 516, Exercise 8: "$A^{(1)}, A^{(2)}, \ldots, A^{(n-1)}$" should be "$M^{(1)}, M^{(2)}, \ldots, M^{(n-1)}$".
Page 516, Exercise 10: $n^2/2 - n/2$ should be $n^2 - n$.
Page 523, Exercise 6(b): "…the columns of AP are…" should be "…the columns of AP^T are…".
Page 535, Theorem 401: $\|A\|_1$ should be $\|A\|_\infty$ (twice).
Page 536, Exercise 1: "…be any matrix norm…" should be "…be any induced matrix norm…".
Page 554, Exercise 4: "…b is consider the data …" should be "…b is considered the data …".
Page 554, Exercise 5: "…b is consider the data …" should be "…b is considered the data …".
Page 563, Exercise 7: "Let $v \in \mathbf{R}^m$ be given and define $\alpha = \pm\|x\|_2$, let $\{u_1, u_2, \ldots, u_m\}$ be an orthonormal basis for $\mathbf{R}^m$, where $u_1 = x/\|x\|_2$ …" should be "Let $v \in \mathbf{R}^m$ be given, define $\alpha = \pm\|v\|_2$, $x = \alpha e_1 - v$, $u_1 = x/\|x\|_2$, and let $\{u_1, u_2, \ldots, u_m\}$ be an orthonormal basis for $\mathbf{R}^m$,…".
Page 571, Exercise 3: "Prove that the angle between A^kv_0 and x_1 converges to zero as $k \to \infty$" should be "Prove that the angle between A^kv_0 and $\mathrm{sp}\{x_1\} = E_A(\lambda_1)$ converges to zero as $k \to \infty$.
Page 575, line 15: $3n^2 - n$ should be $3n^2 + 2n - 5$.
Page 575, line 16: "n square roots" should be "$n - 1$ square roots".
Page 580, Exercise 3: "…requires $3n^2 - n$ arithmetic operations, plus the calculation of n square roots,…" should be "…requires $3n^2 + 2n - 5$ arithmetic operations, plus the calculation of $n - 1$ square roots,…".
Page 585, line 19: "original subsequence" should be "original sequence".
Page 585, line 20: "original subsequence" should be "original sequence".
Page 604, Exercise 4: "Theorem 4" should be "Theorem 451".
Page 608, line 18: "…exists a real number…" should be "…exists as a real number…".

Chapter 2

Fields and vector spaces

2.1 Fields

3. Let F be a field and let $\alpha \in F$ be nonzero. We wish to show that the multiplicative inverse of α is unique. Suppose $\beta \in F$ satisfies $\alpha\beta = 1$. Then, multiplying both sides of the equation by α^{-1}, we obtain $\alpha^{-1}(\alpha\beta) = \alpha^{-1} \cdot 1$, or $(\alpha^{-1}\alpha)\beta = \alpha^{-1}$, or $1 \cdot \beta = \alpha^{-1}$. It follows that $\beta = \alpha^{-1}$, and thus α has a unique multiplicative inverse.

7. Let F be a field and let α, β be elements of F. We wish to show that the equation $\alpha + x = \beta$ has a unique solution. The proof has two parts. First, if x satisfies $\alpha + x = \beta$, then adding $-\alpha$ to both sides shows that x must equal $-\alpha + \beta = \beta - \alpha$. This shows that the equation has at most one solution. On the other hand, $x = -\alpha + \beta$ is a solution since $\alpha + (-\alpha + \beta) = (\alpha - \alpha) + \beta = 0 + \beta = \beta$. Therefore, $\alpha + x = \beta$ has a unique solution, namely, $x = -\alpha + \beta$.

13. Let $F = \{(\alpha, \beta) : \alpha, \beta \in \mathbf{R}\}$, and define addition and multiplication on F by $(\alpha,\beta)+(\gamma,\delta) = (\alpha+\gamma, \beta+\delta)$, $(\alpha, \beta) \cdot (\gamma, \delta) = (\alpha\gamma, \beta\delta)$. With these definitions, F is not a field because multiplicative inverses do not exists. It is straightforward to verify that $(0, 0)$ is an additive identity and $(1, 1)$ is a multiplicative identity. Then $(1, 0) \neq (0, 0)$, yet $(1, 0) \cdot (\alpha, \beta) = (\alpha, 0) \neq (1, 1)$ for all $(\alpha, \beta) \in F$. Since F contains a nonzero element with no multiplicative inverse, F is not a field.

15. Suppose F is a set on which are defined two operations, addition and multiplication, such that all the properties of a field are satisfied except that addition is not assumed to be commutative. We wish to show that, in fact, addition must be commutative, and therefore F must be a field. We first note that it is possible to prove that $0 \cdot \gamma = 0$, $-1 \cdot \gamma = -\gamma$, and$-(-\gamma) = \gamma$ for all $\gamma \in F$ without invoking commutativity of addition. Moreover, for all $\alpha, \beta \in F$, $-\beta + (-\alpha) = -(\alpha + \beta)$ since $(\alpha + \beta) + (-\beta + (-\alpha)) = ((\alpha + \beta) + (-\beta)) + (-\alpha) = (\alpha + (\beta + (-\beta))) + (-\alpha) = (\alpha + 0) + (-\alpha) = \alpha + (-\alpha) = 0$. We therefore conclude that $-1 \cdot (\alpha + \beta) = -\beta + (-\alpha)$ for all $\alpha, \beta \in F$. But, by the distributive property, $-1 \cdot (\alpha + \beta) = -1 \cdot \alpha + (-1) \cdot \beta = -\alpha + (-\beta)$, and therefore $-\alpha + (-\beta) = -\beta + (-\alpha)$ for all $\alpha, \beta \in F$. Applying this property to $-\alpha$, $-\beta$ in place of α, β, respectively, yields $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in F$, which is what we wanted to prove.

19. Let F be a finite field.

    (a) Consider the elements $1, 1+1, 1+1+1, \ldots$ in F. Since F contains only finitely many elements, there must exist two terms in this sequence that are equal, say $1 + 1 + \cdots + 1$ (ℓ terms) and $1 + 1 + \cdots + 1$ (k terms), where $k > \ell$. We can then add -1 to both sides ℓ times to show that $1 + 1 + \cdots + 1$ ($k - \ell$ terms) equals 0 in F. Since at least one of the sequence $1, 1+1, 1+1+1, \ldots$ equals 0, we can define n to be the smallest integer greater than 1 such that $1 + 1 + \cdots + 1 = 0$ (n terms). We call n the characteristic of the field.

    (b) Given that the characteristic of F is n, for any $\alpha \in F$, we have $\alpha + \alpha + \cdots + \alpha = \alpha(1 + 1 + \cdots + 1) = \alpha \cdot 0 = 0$ if the sum has n terms.

(c) We now wish to show that the characteristic n is prime. Suppose, by way of contradiction, that $n = k\ell$, where $1 < k, \ell < n$. Define $\alpha = 1 + 1 + \cdots + 1$ (k terms) and $\beta = 1 + 1 + \cdots + 1$ (ℓ terms). Then $\alpha\beta = 1 + 1 + \cdots + 1$ (n terms), so that $\alpha\beta = 0$. But this implies that $\alpha = 0$ or $\beta = 0$, which contradicts the definition of the characteristic n. This contradiction shows that n must be prime.

2.2 Vector spaces

7. (a) The elements of $\mathcal{P}_1(\mathbf{Z}_2)$ are the polynomials 0, 1, x, $1 + x$, which define distinct functions on $\mathbf{Z}_2$. We have $0+0 = 0$, $0+1 = 1$, $0+x = x$, $0+(1+x) = 1+x$, $1+1 = 0$, $1+x = 1+x$, $1+(1+x) = x$, $x+x = (1+1)x = 0x = 0$, $x+(1+x) = 1+(x+x) = 1$, $(1+x)+(1+x) = (1+1)+(x+x) = 0+0 = 0$.

   (b) Nominally, the elements of $\mathcal{P}_2(\mathbf{Z}_2)$ are $0, 1, x, 1 + x, x^2, 1 + x^2, x + x^2, 1 + x + x^2$. However, since these elements are interpreted as functions mapping $\mathbf{Z}_2$ into $\mathbf{Z}_2$, it turns out that the last four functions equal the first four. In particular, $x^2 = x$ (as functions), since $0^2 = 0$ and $1^2 = 1$. Then $1 + x^2 = 1 + x$, $x + x^2 = x + x = 0$, and $1 + x + x^2 = 1 + 0 = 1$. Thus we see that the function spaces $\mathcal{P}_2(\mathbf{Z}_2)$ and $\mathcal{P}_1(\mathbf{Z}_2)$ are the same.

   (c) Let V be the vector space consisting of all functions from $\mathbf{Z}_2$ into $\mathbf{Z}_2$. To specify $f \in V$ means to specify the two values $f(0)$ and $f(1)$. There are exactly four ways to do this: $f(0) = 0$, $f(1) = 0$ (so $f(x) = 0$); $f(0) = 1$, $f(1) = 1$ (so $f(x) = 1$); $f(0) = 0$, $f(1) = 1$ (so $f(x) = x$); and $f(0) = 1$, $f(1) = 0$ (so $f(x) = 1 + x$). Thus we see that $V = \mathcal{P}_1(\mathbf{Z}_2)$.

9. Let $V = \mathbf{R}^2$ with the usual scalar multiplication and the following nonstandard vector addition: $u \oplus v = (u_1 + v_1, u_2 + v_2 + 1)$ for all $u, v \in \mathbf{R}^2$. It is easy to check that commutativity and associativity of $\oplus$ hold, that $(0, -1)$ is an additive identity, and that each $u = (u_1, u_2)$ has an additive inverse, namely, $(-u_1, -u_2 - 2)$. Also, $\alpha(\beta u) = (\alpha\beta)u$ for all $u \in V$, $\alpha, \beta \in \mathbf{R}$ (since scalar multiplication is defined in the standard way). However, if $\alpha \in \mathbf{R}$, then $\alpha(u + v) = \alpha(u_1 + v_1, u_2 + v_2 + 1) = (\alpha u_1 + \alpha v_1, \alpha u_2 + \alpha v_2 + \alpha)$, while $\alpha u + \alpha v = (\alpha u_1, \alpha u_2) + (\alpha v_1, \alpha v_2) = (\alpha u_1 + \alpha v_1, \alpha u_2 + \alpha v_2 + 1)$, and these are unequal if $\alpha \neq 1$. Thus the first distributive property fails to hold, and V is not a vector space over $\mathbf{R}$. (In fact, the second distributive property also fails.)

15. Suppose U and V are vector spaces over a field F, and define addition and scalar multiplication on $U \times V$ by $(u, v) + (w, z) = (u + w, v + z)$, $\alpha(u, v) = (\alpha u, \alpha v)$. We wish to prove that $U \times V$ is a vector space over F. In fact, the verifications of all the defining properties of a vector space are straightforward. For instance, $(u, v) + (w, z) = (u + w, v + z) = (w + u, z + v) = (w, z) + (u, v)$ (using the commutativity of addition in U and V), and therefore addition in $U \times V$ is commutative. Note that the additive identity in $U \times V$ is $(0, 0)$, where the first 0 is the zero vector in U and the second is the zero vector in V. We will not verify the remaining properties here.

2.3 Subspaces

3. Let V be a vector space over $\mathbf{R}$, and let $v \in V$ be nonzero. We wish to prove that $S = \{0, v\}$ is not a subspace of V. If S were a subspace, then $2v$ would lie in S. But $2v \neq 0$ by Theorem 5, and $2v \neq v$ (since otherwise adding $-v$ to both sides would imply that $v = 0$). Hence $2v \notin S$, and therefore S is not a subspace of V.

7. Define $S = \{x \in \mathbf{R}^2 : ax_1 + bx_2 = 0\}$, where $a, b \in \mathbf{R}$ are constants. We will show that S is subspace of $\mathbf{R}^2$. First, $(0, 0) \in S$, since $a \cdot 0 + b \cdot 0 = 0$. Next, suppose $x \in S$ and $\alpha \in \mathbf{R}$. Then $ax_1 + bx_2 = 0$, and therefore $a(\alpha x_1) + b(\alpha x_2) = \alpha(ax_1 + bx_2) = \alpha \cdot 0 = 0$. This shows that $\alpha x \in S$, and therefore S is closed under scalar multiplication. Finally, suppose $x, y \in S$, so that $ax_1 + bx_2 = 0$ and $ay_1 + by_2 = 0$. Then $a(x_1 + y_1) + b(x_2 + y_2) = (ax_1 + bx_2) + (ay_1 + by_2) = 0 + 0 = 0$, which shows that $x + y \in S$, and therefore that S is closed under addition. This completes the proof.

11. Let $\mathbf{R}$ be regarded as a vector space over $\mathbf{R}$. We wish to prove that $\mathbf{R}$ has no proper subspaces. It suffices to prove that if S is a nontrivial subspace of $\mathbf{R}$, then $S = \mathbf{R}$. So suppose S is a nontrivial subspace,

which means that there exists $x \neq 0$ belonging to S. But then, given any $y \in \mathbf{R}$, $y = (yx^{-1})x$ belongs to S because S is closed under scalar multiplication. Thus $\mathbf{R} \subset S$, and hence $S = \mathbf{R}$.

17. Let $S = \left\{ u \in C[a,b] \,:\, \int_a^b u(x)\,dx = 0 \right\}$. We will show that S is a subspace of $C[a,b]$. First, since the integral of the zero function is zero, we see that the zero function belongs to S. Next, suppose $u \in S$ and $\alpha \in \mathbf{R}$. Then $\int_a^b (\alpha u)(x)\,dx = \int_a^b \alpha u(x)\,dx = \alpha \int_a^b u(x)\,dx = \alpha \cdot 0 = 0$, and therefore $\alpha u \in S$. Finally, suppose $u, v \in S$. Then $\int_a^b (u+v)(x)\,dx = \int_a^b (u(x)+v(x))\,dx = \int_a^b u(x)\,dx + \int_a^b v(x)\,dx = 0+0 = 0$. This shows that $u + v \in S$, and we have proved that S is a subspace of $C[a,b]$.

19. Let V be a vector space over a field F, and let X and Y be subspaces of V.

    (a) We will show that $X \cap Y$ is also a subspace of V. First of all, since $0 \in X$ and $0 \in Y$, it follows that $0 \in X \cap Y$. Next, suppose $x \in X \cap Y$ and $\alpha \in F$. Then, by definition of intersection, $x \in X$ and $x \in Y$. Since X and Y are subspaces, both are closed under scalar multiplication and therefore $\alpha x \in X$ and $\alpha x \in Y$, from which it follows that $\alpha \in X \cap Y$. Thus $X \cap Y$ is closed under scalar multiplication. Finally, suppose $x, y \in X \cap Y$. Then $x, y \in X$ and $x, y \in Y$. Since X and Y are closed under addition, we have $x + y \in X$ and $x + y \in Y$, from which we see that $x + y \in X \cap Y$. Therefore, $X \cap Y$ is closed under addition, and we have proved that $X \cap Y$ is a subspace of V.

    (b) It is not necessarily that case that $X \cup Y$ is a subspace of V. For instance, let $V = \mathbf{R}^2$, and define $X = \{x \in \mathbf{R}^2 \,:\, x_2 = 0\}$, $Y = \{x \in \mathbf{R}^2 \,:\, x_1 = 0\}$. Thus $X \cup Y$ is not closed under addition, and hence is not a subspace of $\mathbf{R}^2$. For instance, $(1,0) \in X \subset X \cup Y$ and $(0,1) \in Y \subset X \cup Y$; however, $(1,0) + (0,1) = (1,1) \notin X \cup Y$.

2.4 Linear combinations and spanning sets

3. Let $S = \mathrm{sp}\{1 + 2x + 3x^2, x - x^2\} \subset \mathcal{P}_2$.

    (a) There is a (unique) solution $\alpha_1 = 2$, $\alpha_2 = 1$ to $\alpha_1(1 + 2x + 3x^2) + \alpha_2(x - x^2) = 2 + 5x + 5x^2$. Therefore, $2 + 5x + 5x^2 \in S$.

    (b) There is no solution α_1, α_2 to $\alpha_1(1+2x+3x^2)+\alpha_2(x-x^2) = 1-x+x^2$. Therefore, $1-x+x^2 \notin S$.

7. Let $u = (1,1,-1)$, $v = (1,0,2)$ be vectors in $\mathbf{R}^3$. We wish to show that $S = \mathrm{sp}\{u,v\}$ is a plane in $\mathbf{R}^3$. First note that if $S = \{x \in \mathbf{R}^3 \,:\, ax_1 + bx_2 + cx_3 = 0\}$, then (taking $x = u$, $x = v$) we see that a, b, c must satisfy $a + b - c = 0$, $a + 2c = 0$. One solution is $a = 2$, $b = -3$, $c = -1$. We will now prove that $S = \{x \in \mathbf{R}^3 \,:\, 2x_1 - 3x_2 - x_3 = 0\}$. First, suppose $x \in S$. Then there exist $\alpha, \beta \in \mathbf{R}$ such that $x = \alpha u + \beta v = \alpha(1,1,-1) + \beta(1,0,2) = (\alpha+\beta, \alpha, -\alpha+2\beta)$, and $2x_1 - 3x_2 - x_3 = 2(\alpha+\beta) - 3\alpha - (-\alpha+2\beta) = 2\alpha + 2\beta - 3\alpha + \alpha - 2\beta = 0$. Therefore, $x \in \{x \in \mathbf{R}^3 \,:\, 2x_1 - 3x_2 - x_3 = 0\}$. Conversely, suppose $x \in \{x \in \mathbf{R}^3 \,:\, 2x_1 - 3x_2 - x_3 = 0\}$. If we solve the equation $\alpha u + \beta v = x$, we see that it has the solution $\alpha = x_2$, $\beta = x_1 - x_2$, and therefore $x \in S$. (Notice that $x_2(1,1,-1) + (x_1 - x_2)(1,0,2) = (x_1, x_2, 2x_1 - 3x_2)$, and the assumption $2x_1 - 3x_2 - x_3 = 0$ implies that $2x_1 - 3x_2 = x_3$.) This completes the proof.

11. Let $S = \mathrm{sp}\{(-1,-3,3), (-1,-4,3), (-1,-1,4)\} \subset \mathbf{R}^3$. We wish to determine if $S = \mathbf{R}^3$ or if S is a proper subspace of $\mathbf{R}^3$. Given an arbitrary $x \in \mathbf{R}^3$, we solve $\alpha_1(-1,-3,3) + \alpha_2(-1,-4,3) + \alpha_3(-1,-1,4) = (x_1, x_2, x_3)$ and find that there is a unique solution, namely, $\alpha_1 = -13x_1 + x_2 - 3x_3$, $\alpha_2 = 9x_1 - x_2 + 2x_3$, $\alpha_3 = 3x_1 + x_3$. This shows that every $x \in \mathbf{R}^3$ lies in S, and therefore $S = \mathbf{R}^3$.

15. Let V be a vector space over a field F, and let $u \in V$, $u \neq 0$, $\alpha \in F$. We wish to prove that $\mathrm{sp}\{u\} = \mathrm{sp}\{u, \alpha u\}$. First, if $x \in \mathrm{sp}\{u\}$, then $x = \beta u$ for some $\beta \in F$, in which case we can write $x = \beta u + 0(\alpha u)$, which shows that x also belongs to $\mathrm{sp}\{u, \alpha u\}$. Conversely, if $x \in \mathrm{sp}\{u, \alpha u\}$, then there exist scalars $\beta, \gamma \in F$ such that $x = \beta u + \gamma(\alpha u)$. But then $x = (\beta + \gamma\alpha)u$, and therefore $x \in \mathrm{sp}\{u\}$. Thus $\mathrm{sp}\{u\} = \mathrm{sp}\{u, \alpha u\}$.

2.5 Linear independence

3. Let V be a vector space over a field F, and let $u_1, \ldots, u_n \in V$. Suppose $u_i = 0$ for some i, $1 \leq i \leq n$, and define scalars $\alpha_1, \ldots, \alpha_n \in F$ by $\alpha_k = 0$ if $k \neq i$, $\alpha_i = 1$. Then $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \cdot u_1 + \cdots + 0 \cdot u_{i-1} + 1 \cdot 0 + 0 \cdot u_{i+1} + \cdots + 0 \cdot u_n = 0$, and hence there is a nontrivial solution to $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$. This shows that $\{u_1, \ldots, u_n\}$ is linearly dependent.

9. We wish to show that $\{1, x, x^2\}$ is linearly dependent in $\mathcal{P}_2(\mathbf{Z}_2)$. The equation $\alpha_1 \cdot 1 + \alpha_2 x + \alpha_3 x^2 = 0$ has the nontrivial solution $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = 1$. To verify this, we must simply verify that $x + x^2$ is the zero function in $\mathcal{P}_2(\mathbf{Z}_2)$. Substituting $x = 0$, we obtain $0 + 0^2 = 0 + 0 = 0$, and with $x = 1$, we obtain $1 + 1^2 = 1 + 1 = 0$.

13. We wish to show that $\{p_1, p_2, p_3\}$, where $p_1(x) = 1 - x^2$, $p_2(x) = 1 + x - 6x^2$, $p_3(x) = 3 - 2x^2$, is linearly independent and spans $\mathcal{P}_2$. We first verify that the set is linearly independent by solving $\alpha_1(1-x^2) + \alpha_2(1+x-6x^2) + \alpha_3(3-2x^2) = 0$. This equation is equivalent to the system $\alpha_1 + \alpha_2 + 3\alpha_2 = 0$, $\alpha_2 = 0$, $-\alpha_1 - 6\alpha_2 - 2\alpha_3 = 0$, and a direct calculation shows that the only solution is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. To show that the set spans $\mathcal{P}_2$, we take an arbitrary $p \in \mathcal{P}_2$, say $p(x) = c_0 + c_1 x + c_2 x^2$, and solve $\alpha_1(1 - x^2) + \alpha_2(1 + x - 6x^2) + \alpha_3(3 - 2x^2) = c_0 + c_1 x + c_2 x^2$. This is equivalent to the system $\alpha_1 + \alpha_2 + 3\alpha_2 = c_0$, $\alpha_2 = c_1$, $-\alpha_1 - 6\alpha_2 - 2\alpha_3 = c_2$. There is a unique solution: $\alpha_1 = -2c_0 - 16c_1 - 3c_2$, $\alpha_2 = c_1$, $\alpha_3 = c_0 + 5c_1 + c_2$. This shows that $p \in \mathrm{sp}\{p_1, p_2, p_3\}$, and, since p was arbitrary, that $\{p_1, p_2, p_3\}$ spans all of $\mathcal{P}_2$.

17. (a) Let V be a vector space over $\mathbf{R}$, and suppose $\{x, y, z\}$ is a linearly independent subset of V. We wish to show that $\{x+y, y+z, x+z\}$ is also linearly independent. Let $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{R}$ satisfy $\alpha_1(x+y) + \alpha_2(y+z) + \alpha_3(x+z) = 0$. This equation is equivalent to $(\alpha_1 + \alpha_3)x + (\alpha_1 + \alpha_2)y + (\alpha_2 + \alpha_3)z = 0$. Since $\{x, y, z\}$ is linearly independent, it follows that $\alpha_1 + \alpha_3 = \alpha_1 + \alpha_2 = \alpha_2 + \alpha_3 = 0$. This system can be solved directly to show that $\alpha_1 = \alpha_2 = \alpha_3 = 0$, which proves that $\{x + y, y + z, x + z\}$ is linearly independent.

    (b) We now show, by example, that the previous result is not necessarily true if V is a vector space over some field $F \neq \mathbf{R}$. Let $V = \mathbf{Z}_2^3$, and define $x = (1, 0, 0)$, $y = (0, 1, 0)$, and $z = (0, 0, 1)$. Obviously $\{x, y, z\}$ is linearly independent. On the other hand, we have $(x+y) + (y+z) + (x+z) = (1,1,0) + (0,1,1) + (1,0,1) = (1+0+1, 1+1+0, 0+1+1) = (0,0,0)$, which shows that $\{x+y, y+z, x+z\}$ is linearly dependent.

21. Let V be a vector space over a field F, and suppose $\{u_1, u_2, \ldots, u_n\}$ is linearly dependent. We wish to prove that, given any i, $1 \leq i \leq n$, either u_i is a linear combination of $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n$ or these vectors form a linearly dependent set. By assumption, there exist scalars $\alpha_1, \ldots, \alpha_n \in F$, not all zero, such that $\alpha_1 u_1 + \cdots + \alpha_i u_i + \cdots + \alpha_n u_n = 0$. We now consider two cases. If $\alpha_i \neq 0$, the we can solve the latter equation for u_i to obtain $u_i = -\alpha_i^{-1}\alpha_1 u_1 - \cdots - \alpha_i^{-1}\alpha_{i-1}u_{i-1} - \alpha_i^{-1}\alpha_{i+1}u_{i+1} - \cdots - \alpha_i^{-1}\alpha_n u_n$. In this case, u_i is a linear combination of the remaining vectors. The second case is that $\alpha_i = 0$, in which case at least one of $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$ is nonzero, and we have $\alpha_1 u_1 + \cdots + \alpha_{i-1}u_{i-1} + \alpha_{i+1}u_{i+1} + \cdots + \alpha_n u_n = 0$. This shows that $\{u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n\}$ is linearly dependent.

2.6 Basis and dimension

3. We now repeat the previous exercise for the vectors $v_1 = (-1, 3, -1)$, $v_2 = (1, -2, -2)$, $v_3 = (-1, 7, -13)$. If we try to solve $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = x$ for an arbitrary $x \in \mathbf{R}^3$, we find that this equation is equivalent to the following system:

$$\begin{aligned} -\alpha_1 + \alpha_2 - \alpha_3 &= x_1 \\ \alpha_2 + 4\alpha_3 &= 3x_1 + x_2 \\ 0 &= 8x_1 + 3x_2 + x_3. \end{aligned}$$

Since this system is inconsistent for most $x \in \mathbf{R}^3$ (the system is consistent only if x happens to satisfy $8x_1 + 3x_2 + x_3 = 0$), $\{v_1, v_2, v_3\}$ does not span $\mathbf{R}^3$ and therefore is not a basis.

7. Consider the subspace $S = \mathrm{sp}\{p_1, p_2, p_3, p_4, p_5\}$ of $\mathcal{P}_3$, where

$$\begin{aligned}
&p_1(x) = -1 + 4x - x^2 + 3x^3,\ p_2(x) = 2 - 8x + 2x^2 - 5x^3,\\
&p_3(x) = 3 - 11x + 3x^2 - 8x^3,\ p_4(x) = -2 + 8x - 2x^2 - 3x^3,\\
&p_5(x) = 2 - 8x + 2x^2 + 3x^3.
\end{aligned}$$

(a) The set $\{p_1, p_2, p_3, p_4, p_5\}$ is linearly dependent (by Theorem 34) because it contains five elements and the dimension of $\mathcal{P}_3$ is only four.

(b) As illustrated in Example 39, we begin by solving

$$\alpha_1 p_1(x) + \alpha_2 p_2(x) + \alpha_3 p_3(x) + \alpha_4 p_4(x) + \alpha_5 p_5(x) = 0;$$

this is equivalent to the system

$$\begin{aligned}
-\alpha_1 + 2\alpha_2 + 3\alpha_3 - 2\alpha_4 + 2\alpha_5 &= 0,\\
4\alpha_1 - 8\alpha_2 - 11\alpha_3 + 8\alpha_4 - 8\alpha_5 &= 0,\\
-\alpha_1 + 2\alpha_2 + 3\alpha_3 - 2\alpha_4 + 2\alpha_5 &= 0,\\
3\alpha_1 - 5\alpha_2 - 8\alpha_3 - 3\alpha_4 + 3\alpha_5 &= 0,
\end{aligned}$$

which reduces to

$$\begin{aligned}
\alpha_1 &= 16\alpha_4 - 16\alpha_5,\\
\alpha_2 &= 9\alpha_4 - 9\alpha_5,\\
\alpha_3 &= 0.
\end{aligned}$$

Since there are nontrivial solutions, $\{p_1, p_2, p_3, p_4, p_5\}$ is linearly dependent (which we already knew), but we can deduce more than that. By taking $\alpha_4 = 1$, $\alpha_5 = 0$, we see that $\alpha_1 = 16$, $\alpha_2 = 9$, $\alpha_3 = 0$, $\alpha_4 = 1$, $\alpha_5 = 0$ is one solution, which means that

$$16p_1(x) + 9p_2(x) + p_4(x) = 0 \ \Rightarrow\ p_4(x) = -16p_1(x) - 9p_2(x).$$

This shows that $p_4 \in \mathrm{sp}\{p_1, p_2\} \subset \mathrm{sp}\{p_1, p_2, p_3\}$. Similarly, taking $\alpha_4 = 0$, $\alpha_5 = 0$, we find that

$$-16p_1(x) - 9p_2(x) + p_5(x) = 0 \ \Rightarrow\ p_5(x) = 16p_1(x) + 9p_2(x),$$

and hence $p_5 \in \mathrm{sp}\{p_1, p_2\} \subset \mathrm{sp}\{p_1, p_2, p_3\}$. It follows from Lemma 19 that $\mathrm{sp}\{p_1, p_2, p_3, p_4, p_5\} = \mathrm{sp}\{p_1, p_2, p_3\}$. Our calculations above show that $\{p_1, p_2, p_3\}$ is linearly independent (if $\alpha_4 = \alpha_5 = 0$, then also $\alpha_1 = \alpha_2 = \alpha_3 = 0$). Therefore, $\{p_1, p_2, p_3\}$ is a linearly independent spanning set of S and hence a basis for S.

13. Suppose V is a vector space over a field F, and S, T are two n-dimensional subspaces of V. We wish to prove that if $S \subset T$, then in fact $S = T$. Let $\{s_1, s_2, \ldots, s_n\}$ be a basis for S. Since $S \subset T$, this implies that $\{s_1, s_2, \ldots, s_n\}$ is a linearly independent subset of T. We will now show that $\{s_1, s_2, \ldots, s_n\}$ also spans T. Let $t \in T$ be arbitrary. Since T has dimension n, the set $\{s_1, s_2, \ldots, s_n, t\}$ is linearly dependent by Theorem 34. But then, by Lemma 33, t must be a linear combination of $s_1, s_2, \ldots, s_n$ (since no s_k is a linear combination of $s_1, s_2, \ldots, s_{k-1}$). This shows that $t \in \mathrm{sp}\{s_1, s_2, \ldots, s_n\}$, and hence we have shown that $\{s_1, s_2, \ldots, s_n\}$ is a basis for T. But then

$$T = \mathrm{sp}\{s_1, s_2, \ldots, s_n\} = S,$$

as desired.

2.7 Properties of bases

1. Consider the following vectors in $\mathbf{R}^3$: $v_1 = (1,5,4)$, $v_2 = (1,5,3)$, $v_3 = (17,85,56)$, $v_4 = (1,5,2)$, $v_5 = (3,16,13)$.

   (a) We wish to show that $\{v_1, v_2, v_3, v_4, v_5\}$ spans $\mathbf{R}^3$. Given an arbitrary $x \in \mathbf{R}^3$, the equation

   $$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5 = x$$

   is equivalent to the system

   $$\begin{aligned} \alpha_1 + \alpha_2 + 17\alpha_3 + \alpha_4 + 3\alpha_5 &= x_1, \\ 5\alpha_1 + 5\alpha_2 + 85\alpha_3 + 5\alpha_4 + 16\alpha_5 &= x_2, \\ 4\alpha_1 + 3\alpha_2 + 56\alpha_3 + 2\alpha_4 + 13\alpha_5 &= x_3. \end{aligned}$$

   Applying Gaussian elimination, this system reduces to

   $$\begin{aligned} \alpha_1 &= 17x_1 - 4x_2 + x_3 - 5\alpha_3 + \alpha_4, \\ \alpha_2 &= x_2 - x_1 - x_3 - 12\alpha_3 - 2\alpha_5, \\ \alpha_5 &= x_2 - 5x_1. \end{aligned}$$

   This shows that there are solutions regardless of the value of x; that is, each $x \in \mathbf{R}^3$ can be written as a linear combination of v_1, v_2, v_3, v_4, v_5. Therefore, $\{v_1, v_2, v_3, v_4, v_5\}$ spans $\mathbf{R}^3$.

   (b) Now we wish to find a subset of $\{v_1, v_2, v_3, v_4, v_5\}$ that is a basis for $\mathbf{R}^3$. According to the calculations given above, each $x \in \mathbf{R}^3$ can be written as a linear combination of $\{v_1, v_2, v_5\}$ (just take $\alpha_3 = \alpha_4 = 0$ in the system solved above). Since $\dim(\mathbf{R}^3) = 3$, any three vectors spanning $\mathbf{R}^3$ form a basis for $\mathbf{R}^3$ (by Theorem 45). Hence $\{v_1, v_2, v_5\}$ is a basis for $\mathbf{R}^3$.

5. Let $u_1 = (1,4,0,-5,1)$, $u_2 = (1,3,0,-4,0)$, $u_3 = (0,4,1,1,4)$ be vectors in $\mathbf{R}^5$.

   (a) To show that $\{u_1, u_2, u_3\}$ is linearly independent, we solve the equation $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 = 0$, which is equivalent to the system

   $$\begin{aligned} \alpha_1 + \alpha_2 &= 0, \\ 4\alpha_1 + 3\alpha_2 + 4\alpha_3 &= 0, \\ \alpha_3 &= 0, \\ -5\alpha_1 - 4\alpha_2 + \alpha_3 &= 0, \\ \alpha_1 + 4\alpha_3 &= 0. \end{aligned}$$

   A direct calculation shows that this system has only the trivial solution.

   (b) To extend $\{u_1, u_2, u_3\}$ to a basis for $\mathbf{R}^5$, we need two more vectors. We will try $u_4 = (0,0,0,1,0)$ and $u_5 = (0,0,0,0,1)$. We solve $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 + \alpha_5 u_5 = 0$ and find that the only solution is the trivial one. This implies that $\{u_1, u_2, u_3, u_4, u_5\}$ is linearly independent and hence, by Theorem 45, a basis for $\mathbf{R}^5$.

9. Consider the vectors $u_1 = (3,1,0,4)$ and $u_2 = (1,1,1,4)$ in $\mathbf{Z}_5^4$.

   (a) It is obvious that $\{u_1, u_2\}$ is linearly independent, since neither vector is a multiple of the other.

   (b) To extend $\{u_1, u_2\}$ to a basis for $\mathbf{Z}_5^4$, we must find vectors u_3, u_4 such that $\{u_1, u_2, u_3, u_4\}$ is linearly independent. We try $u_3 = (0,0,1,0)$ and $u_4 = (0,0,0,1)$. A direct calculation then shows that $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 = 0$ has only the trivial solution. Therefore $\{u_1, u_2, u_3, u_4\}$ is linearly independent and hence, since $\dim(\mathbf{Z}_5^4) = 4$, it is a basis for $\mathbf{Z}_5^4$.

15. Let V be a vector space over a field F, and let $\{u_1, \ldots, u_n\}$ be a basis for V. Let $v_1, \ldots, v_k$ be vectors in V, and suppose

$$v_j = \alpha_{1,j}u_1 + \ldots + \alpha_{n,j}u_n, \ j = 1, 2, \ldots, k.$$

Define the vectors $x_1, \ldots, x_k$ in F^n by

$$x_j = (\alpha_{1,j}, \ldots, \alpha_{n,j}), \ j = 1, 2, \ldots, k.$$

(a) We first prove that $\{v_1, \ldots, v_k\}$ is linearly independent if and only if $\{x_1, \ldots, x_k\}$ is linearly independent. We will do this by showing that $c_1v_1 + \cdots + c_kv_k = 0$ in V is equivalent to $c_1x_1 + \cdots + c_kx_k = 0$ in F^n. Then the first equation has only the trivial solution if and only if the second equation does, and the result follows. The proof is a direct manipulation, for which summation notation is convenient:

$$\begin{aligned}
\sum_{j=1}^{k} c_j v_j = 0 &\Leftrightarrow \sum_{j=1}^{k} c_j \left(\sum_{i=1}^{n} \alpha_{ij} u_i \right) \\
&\Leftrightarrow \sum_{j=1}^{k} \sum_{i=1}^{n} c_j \alpha_{ij} u_i = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \sum_{j=1}^{k} c_j \alpha_{ij} u_i = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left(\sum_{j=1}^{k} c_j \alpha_{ij} \right) u_i = 0.
\end{aligned}$$

Since $\{u_1, \ldots, u_n\}$ is linearly independent, the last equation is equivalent to

$$\sum_{j=1}^{k} c_j \alpha_{ij} = 0, \ i = 1, 2, \ldots, n,$$

which, by definition of x_j and of addition in F^n, is equivalent to

$$\sum_{j=1}^{k} c_j x_j = 0.$$

This completes the proof.

(b) Now we show that $\{v_1, \ldots, v_k\}$ spans V if and only if $\{x_1, \ldots, x_k\}$ spans F^n. Since each vector in V can be represented uniquely as a linear combination of $u_1, \ldots, u_n$, there is a one-to-one correspondence between V and F^n:

$$w = c_1u_1 + \cdots + c_nu_n \in V \ \longleftrightarrow \ x = (c_1, \ldots, c_n) \in F^n.$$

Mimicking the manipulations in the first part of the exercise, we see that

$$\sum_{j=1}^{k} c_j v_j = w \ \Leftrightarrow \ \sum_{j=1}^{k} c_j x_j = x.$$

Thus the first equation has a solution for every $v \in V$ if and only if the second equation has a solution for every $x \in F^n$. The result follows.

2.8 Polynomial interpolation and the Lagrange basis

1. (a) The Lagrange polynomials for the interpolation nodes $x_0 = 1$, $x_1 = 2$, $x_3 = 3$ are

$$\begin{aligned}
L_0(x) &= \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{1}{2}(x-2)(x-3),\\
L_1(x) &= \frac{(x-1)(x-3)}{(2-1)(2-3)} = -(x-1)(x-3),\\
L_2(x) &= \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{1}{2}(x-1)(x-2).
\end{aligned}$$

(b) The quadratic polynomial interpolating $(1, 0)$, $(2, 2)$, $(3, 1)$ is

$$\begin{aligned}
p(x) &= 0L_0(x) + 2L_1(x) + L_2(x)\\
&= -2(x-1)(x-3) + \frac{1}{2}(x-1)(x-2)\\
&= -\frac{3}{2}x^2 + \frac{13}{2}x - 5.
\end{aligned}$$

5. We wish to write $p_2(x) = 2 + x - x^2$ as a linear combination of the Lagrange polynomials constructed on the nodes $x_0 = -1$, $x_1 = 1$, $x_2 = 3$. The graph of p passes through the points $(-1, p(-1))$, $(1, p(1))$, $(3, p(3))$, that is, $(-1, 0)$, $(1, 2)$, $(3, -4)$. The Lagrange polynomials are

$$\begin{aligned}
L_0(x) &= \frac{(x-1)(x-3)}{(-1-1)(-1-3)} = \frac{1}{8}(x-1)(x-3),\\
L_1(x) &= \frac{(x+1)(x-3)}{(1+1)(1-3)} = -\frac{1}{4}(x+1)(x-3),\\
L_2(x) &= \frac{(x+1)(x-1)}{(3+1)(3-1)} = \frac{1}{8}(x+1)(x-1),
\end{aligned}$$

and therefore,

$$\begin{aligned}
p(x) &= 0L_0(x) + 2L_1(x) - 4L_2(x)\\
&= -\frac{1}{2}(x+1)(x-3) - \frac{1}{2}(x+1)(x-1).
\end{aligned}$$

11. Consider a secret sharing scheme in which five individuals will receive information about the secret, and any two of them, working together, will have access to the secret. Assume that the secret is a two-digit integer, and that p is chosen to be 101. The degree of the polynomial will be one, since then the polynomial will be uniquely determined by two data points. Let us suppose that the secret is $N = 42$ and we choose the polynomial to be $p(x) = N + c_1x$, where $c_1 = 71$ (recall that c_1 is chosen at random). We also choose the five interpolation nodes at random to obtain $x_1 = 9$, $x_2 = 14$, $x_3 = 39$, $x_4 = 66$, and $x_5 = 81$. We then compute

$$\begin{aligned}
y_1 &= p(x_1) = 42 + 71 \cdot 9 = 75,\\
y_2 &= p(x_2) = 42 + 71 \cdot 14 = 26,\\
y_3 &= p(x_3) = 42 + 71 \cdot 39 = 84,\\
y_4 &= p(x_4) = 42 + 71 \cdot 66 = 82,\\
y_5 &= p(x_5) = 42 + 71 \cdot 81 = 36
\end{aligned}$$

(notice that all arithmetic is done modulo 101). The data points, to be distributed to the five individuals, are $(9, 75)$, $(14, 26)$, $(39, 84)$, $(66, 82)$, $(81, 36)$.

2.9 Continuous piecewise polynomial functions

1. The following table shows the maximum errors obtained in approximating $f(x) = e^x$ on the interval $[0, 1]$ by polynomial interpolation and by piecewise linear interpolation, each on a uniform grid with n nodes.

n	Poly. interp. err.	PW linear interp. err.
1	$2.1187 \cdot 10^{-1}$	$2.1187 \cdot 10^{-1}$
2	$1.4420 \cdot 10^{-2}$	$6.6617 \cdot 10^{-2}$
3	$9.2390 \cdot 10^{-4}$	$3.2055 \cdot 10^{-2}$
4	$5.2657 \cdot 10^{-5}$	$1.8774 \cdot 10^{-2}$
5	$2.6548 \cdot 10^{-6}$	$1.2312 \cdot 10^{-2}$
6	$1.1921 \cdot 10^{-7}$	$8.6902 \cdot 10^{-3}$
7	$4.8075 \cdot 10^{-9}$	$6.4596 \cdot 10^{-3}$
8	$1.7565 \cdot 10^{-10}$	$4.9892 \cdot 10^{-3}$
9	$5.8575 \cdot 10^{-12}$	$3.9692 \cdot 10^{-3}$
10	$1.8119 \cdot 10^{-13}$	$3.2328 \cdot 10^{-3}$

For this example, polynomial interpolation is very effective.

Chapter 3

Linear operators

3.1 Linear operators

1. Let m, b be real numbers, and define $f : \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = mx + b$. If b is nonzero, then this function is not linear. For example,

$$f(2 \cdot 1) = f(2) = 2m + b,$$
$$2f(1) = 2(m + b) = 2m + 2b.$$

Since $b \neq 0$, $2m + b \neq 2m + 2b$, which shows that f is not linear. On the other hand, if $b = 0$, then f is linear:

$$f(x + y) = m(x + y) = mx + my = f(x) + f(y) \text{ for all } x, y \in \mathbf{R},$$
$$f(ax) = m(ax) = a(mx) = af(x) \text{ for all } a, x \in \mathbf{R}.$$

Thus we see that $f(x) = mx + b$ is linear if and only if $b = 0$.

5. We wish to determine which of the following real-valued functions defined on $\mathbf{R}^n$ is linear.

   (a) $f : \mathbf{R}^n \to \mathbf{R}$, $f(x) = \sum_{i=1}^n x_i$.
   (b) $g : \mathbf{R}^n \to \mathbf{R}$, $g(x) = \sum_{i=1}^n |x_i|$.
   (c) $h : \mathbf{R}^n \to \mathbf{R}$, $h(x) = \prod_{i=1}^n x_i$.

The function f is linear:

$$f(x + y) = \sum_{i=1}^n (x + y)_i = \sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i = f(x) + f(y),$$
$$f(\alpha x) = \sum_{i=1}^n (\alpha x)_i = \sum_{i=1}^n \alpha x_i = \alpha \sum_{i=1}^n x_i = \alpha f(x).$$

However, g and h are both nonlinear. For instance, if $x \neq 0$, then $g(-x) \neq -g(x)$ (in fact, $g(-x) = g(x)$ for all $x \in \mathbf{R}^n$). Also, if no component of x is zero, then $h(2x) = 2^n h(x) \neq 2h(x)$ (of course, we are assuming $n > 1$).

9. (a) If $A \in \mathbf{C}^{2\times 3}$, $x \in \mathbf{C}^3$ are defined by

$$A = \begin{bmatrix} 1+i & 1-i & 2i \\ 2-i & 1+2i & 3 \end{bmatrix}, \quad x = \begin{bmatrix} 3 \\ 2+i \\ 1-3i \end{bmatrix},$$

then

$$Ax = \begin{bmatrix} 12+4i \\ 9-7i \end{bmatrix}.$$

(b) If $A \in \mathbf{Z}_2^{3\times 3}$, $x \in \mathbf{Z}_2^3$ are defined by

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \ x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

then

$$Ax = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

13. We now wish to give a formula for the ith row of AB, assuming $A \in F^{m\times n}$, $B \in F^{n\times p}$. As pointed out in the text, we have a standard notation for the columns of a matrix, but no standard notation for the rows. Let us suppose that the rows of B are $r_1, r_2, \dots, r_n$. Then

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}(r_k)_j = \left(\sum_{k=1}^{n} A_{ik} r_k\right)_j$$

(once again using the componentwise definition of the operations). This shows that the ith row of AB is

$$\sum_{k=1}^{n} A_{ik} r_k,$$

that is, the linear combination of the rows of B, with the weights taken from the ith row of A.

3.2 More properties of linear operators

3. Consider the linear operator mapping $\mathbf{R}^2$ into itself that sends each vector (x, y) to its projection onto the x-axis, namely, $(x, 0)$. We see that e_1 is mapped to itself and $e_2 = (0, 1)$ is mapped to $(0, 0)$. Therefore, the matrix representing the linear operator is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

9. Let $x \in \mathbf{R}^N$ be denoted as $x = (x_0, x_1, \dots, x_{N-1})$. Given $x, y \in \mathbf{R}^N$, the *convolution* of x and y is the vector $x * y \in \mathbf{R}^N$ defined by

$$(x * y)_n = \sum_{m=0}^{N-1} x_m y_{n-m}, \ n = 0, 1, \dots, N-1.$$

In this formula, y is regarded as defining a periodic vector of period N; therefore, if $n - m < 0$, we take

$y_{n-m} = y_{N+n-m}$. The linearity of the convolution operator is obvious:

$$\begin{aligned}
((x+z)*y)_n = \sum_{m=0}^{N-1}(x+z)_m y_{n-m} &= \sum_{m=0}^{N-1}(x_m+z_m)y_{n-m} \\
&= \sum_{m=0}^{N-1}(x_m y_{n-m} + z_m y_{n-m}) \\
&= \sum_{m=0}^{N-1} x_m y_{n-m} + \sum_{m=0}^{N-1} z_m y_{n-m} \\
&= (x*y)_m + (z*y)_m, \\
((\alpha x)*y)_n = \sum_{m=0}^{N-1}(\alpha x)_m y_{n-m} = \sum_{m=0}^{N-1}(\alpha x_m)y_{n-m} &= \sum_{m=0}^{N-1}\alpha(x_m y_{n-m}) \\
&= \alpha\sum_{m=0}^{N-1} x_m y_{n-m} \\
&= \alpha(x*y)_n.
\end{aligned}$$

Therefore, if y is fixed and $F : \mathbf{R}^N \to \mathbf{R}^N$, then $F(x+z) = F(x) + F(z)$ for all $x, z \in \mathbf{R}^N$ and $F(\alpha x) = \alpha F(x)$ for all $x \in \mathbf{R}^N$, $\alpha \in \mathbf{R}$. This proves that F is linear.

Next, notice that, if e_k is the kth standard basis vector, then

$$F(e_k)_n = y_{n-k},\ n = 0, 1, \ldots, N-1.$$

It follows that

$$\begin{aligned}
F(e_0) &= (y_0, y_1, \ldots, Y_{N-1}), \\
F(e_1) &= (y_{N-1}, y_0, \ldots, Y_{N-2}), \\
F(e_2) &= (y_{N-2}, y_{N-1}, \ldots, Y_{N-3}), \\
&\vdots \qquad\qquad \vdots \\
F(e_{N-1}) &= (y_1, y_2, \ldots, Y_0).
\end{aligned}$$

Therefore, $F(x) = Ax$ for all $x \in \mathbf{R}^N$, where

$$A = \begin{bmatrix} y_0 & y_{N-1} & y_{N-2} & \cdots & y_1 \\ y_1 & y_0 & y_{N-1} & \cdots & y_2 \\ y_2 & y_1 & y_0 & \cdots & y_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{N-1} & y_{N-2} & y_{N-3} & \cdots & y_0 \end{bmatrix}.$$

3.3 Isomorphic vector spaces

1. (a) The function $f : \mathbf{R} \to \mathbf{R}$, $f(x) = 2x + 1$ is invertible since $f(x) = y$ has a unique solution for each $y \in \mathbf{R}$:

$$f(x) = y \ \Leftrightarrow\ 2x + 1 = y \ \Leftrightarrow\ x = \frac{y-1}{2}.$$

   We see that $f^{-1}(y) = (y-1)/2$.

   (b) The function$f : \mathbf{R} \to (0, \infty)$, $f(x) = e^x$ is also invertible, and the inverse is $f^{-1}(y) = \ln(y)$:

$$e^{\ln(y)} = y \text{ for all } y \in \mathbf{R},\ \ln(e^x) = x \text{ for all } x \in \mathbf{R}.$$

(c) The function $f : \mathbf{R}^2 \to \mathbf{R}^2$, $f(x) = (x_1 + x_2, x_1 - x_2)$ is invertible since $f(x) = y$ has a unique solution for each $y \in \mathbf{R}^2$:

$$f(x) = y \Leftrightarrow \begin{array}{c} x_1 + x_2 = y_1, \\ x_1 - x_2 = y_2 \end{array} \Leftrightarrow \begin{array}{l} x_1 = \frac{y_1+y_2}{2}, \\ x_2 = \frac{y_1-y_2}{2}. \end{array}$$

We see that

$$f^{-1}(y) = \left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2} \right).$$

(d) The function $f : \mathbf{R}^2 \to \mathbf{R}^2$, $f(x) = (x_1 - 2x_2, -2x_1 + 4x_2)$ fails to be invertible, and in fact is neither injective nor surjective. For example, $f(0) = 0$ and also $f((2,1)) = (0,0) = 0$, which shows that f is not injective. The equation $f(x) = (1,1)$ has no solution:

$$f(x) = (1,1) \Leftrightarrow \begin{array}{r} x_1 - 2x_2 = 1, \\ -2x_1 + 4x_2 = 1 \end{array} \Leftrightarrow \begin{array}{r} x_1 - 2x_2 = 1, \\ 0 = 3. \end{array}$$

This shows that f is not surjective.

5. Let X, Y, and Z be sets, and suppose $f : X \to Y$, $g : Y \to Z$ are given functions.

(a) If f and $g \circ f$ are invertible, then g is invertible. We can prove this using the previous exercise and the fact that f^{-1} is invertible. We have $g = (g \circ f) \circ f^{-1}$:

$$((g \circ f) \circ f^{-1})(y) = g(f(f^{-1}(y))) = g(y) \text{ for all } y \in Y.$$

Therefore, since $g \circ f$ and f^{-1} are invertible, the previous exercise implies that g is invertible.

(b) Similarly, if g and $g \circ f$ are invertible, then f is invertible since $f = g^{-1} \circ (g \circ f)$.

(c) If we merely know that $g \circ f$ is invertible, we cannot conclude that either f or g is invertible. For example, let $f : \mathbf{R} \to \mathbf{R}^2$, $g : \mathbf{R}^2 \to \mathbf{R}$ be defined by $f(x) = (x, 0)$ for all $x \in \mathbf{R}$ and $g(y) = y_1 + y_2$ for all $y \in \mathbf{R}^2$, respectively. Then

$$(g \circ f)(x) = g(f(x)) = g((x,0)) = x + 0 = x,$$

and it is obvious that $g \circ f$ is invertible (in fact, $(g \circ f)^{-1} = g \circ f)$). However, f is not surjective and g is not injective, so neither is invertible.

13. Let $\mathcal{X}$ be the basis for $\mathbf{Z}_2^3$ from the previous exercise, let $A \in \mathbf{Z}_2^{3\times 3}$ be defined by

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

and define $L : \mathbf{Z}_2^3 \to \mathbf{Z}_2^3$ by $L(x) = Ax$. We wish to find $[L]_{\mathcal{X},\mathcal{X}}$. The columns of this matrix are $[L(x_1)]_{\mathcal{X}}$, $[L(x_2)]_{\mathcal{X}}$, $[L(x_3)]_{\mathcal{X}}$. We have

$$\begin{aligned} [L(x_1)]_{\mathcal{X}} &= [(1,1,0)]_{\mathcal{X}} = (0,1,0), \\ [L(x_2)]_{\mathcal{X}} &= [(0,1,1)]_{\mathcal{X}} = (1,0,1), \\ [L(x_3)]_{\mathcal{X}} &= [(0,0,0)]_{\mathcal{X}} = (0,0,0), \end{aligned}$$

and hence

$$[L]_{\mathcal{X},\mathcal{X}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

19. We wish to determine if the operator D defined in Example 79 is an isomorphism. In fact, D is not an isomorphism since it is not injective. For example, if $p(x) = x$ and $q(x) = x + 1$, then $p \neq q$ but $D(p) = D(q)$.

3.4 Linear operator equations

1. Suppose $L : \mathbf{R}^3 \to \mathbf{R}^3$ is linear, $b \in \mathbf{R}^3$ is given, and $y = (1,0,1)$, $z = (1,1,-1)$ are two solutions to $L(x) = b$. We are asked to find two more solutions to $L(x) = b$. We know that $y - z = (0,-1,2)$ satisfies $L(y-z) = 0$. Therefore, $z + \alpha(y-z)$ satisfies $L(z+\alpha(y-z)) = L(z) + \alpha L(y-z) = b + \alpha \cdot 0 = b$ for all $\alpha \in \mathbf{R}$. Thus two more solutions of $L(x) = b$ are

$$\begin{aligned} z + 2(y-z) &= (1,1,-1) + (0,-2,4) = (1,-1,3), \\ z + 3(y-z) &= (1,1,-1) + (0,-3,6) = (1,-2,5). \end{aligned}$$

5. Let $L : \mathbf{R}^3 \to \mathbf{R}^3$ satisfy $\ker(L) = \mathrm{sp}\{(1,1,1)\}$ and $L(u) = v$, where $u = (1,1,0)$ and $v = (2,-1,2)$. A vector x satisfies $L(x) = v$ if and only if there exists α such that $x = u + \alpha z$ (where $z = (1,1,1)$), that is, if and only if $x - u = \alpha z$ for some $\alpha \in \mathbf{R}$.

   (a) For $x = (1,2,1)$, $x - u = (0,1,1)$, which is not a multiple of z. Thus $L(x) \neq v$.

   (b) For $x = (3,3,2)$, $x - u = (2,2,2) = 2z$. Thus $L(x) = v$.

   (c) For $x = (-3,-3,-2)$, $x - u = (-4,-4,-2)$, which is not a multiple of z. Thus $L(x) \neq v$.

7. Let $A \in \mathbf{Z}_2^{3\times 3}$, $b \in \mathbf{Z}_2^3$ be defined by

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

   If we solve $Ax = 0$ directly, we obtain two solutions: $x = (0,0,0)$ or $x = (1,1,1)$. Thus the solution space of $Ax = 0$ is $\{(0,0,0),(1,1,1)\}$. If we solve $Ax = b$, we also obtain two solutions: $x = (0,1,0)$ or $x = (1,0,1)$. If $L : \mathbf{Z}_2^3 \to \mathbf{Z}_2^3$ is the linear operator defined by $L(x) = Ax$ for all $x \in \mathbf{Z}_2^3$, then $\ker(L) = \{(0,0,0),(1,1,1)\}$ and the solution set of $L(x) = b$ is $\{(0,1,0),(1,0,1)\} = (0,1,0) + \ker(L)$, as predicted by Theorem 90.

3.5 Existence and uniqueness of solutions

3. Each part of this exercise describes an operator with certain properties. We wish to determine if such an operator can exist.

   (a) We wish to find a linear operator $T : \mathbf{R}^3 \to \mathbf{R}^2$ such that $T(x) = b$ has a solution for all $b \in \mathbf{R}^2$. Any surjective operator will do, such as $T : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $T(x) = (x_1, x_2)$.

   (b) No linear operator $T : \mathbf{R}^2 \to \mathbf{R}^3$ has the property that $T(x) = b$ has a solution for all $b \in \mathbf{R}^3$. This would imply that T is surjective, but then Theorem 99 would imply that $\dim(\mathbf{R}^3) \leq \dim(\mathbf{R}^2)$, which is obviously not true.

   (c) Every linear operator $T : \mathbf{R}^3 \to \mathbf{R}^2$ has the property that, for some $b \in \mathbf{R}^2$, the equation $T(x) = b$ has infinitely many solutions. In fact, since $\dim(\mathbf{R}^2) < \dim(\mathbf{R}^3)$, any such $T : \mathbf{R}^3 \to \mathbf{R}^2$ fails to be injective and hence has a nontrivial kernel. Therefore, $T(x) = 0$ necessarily has infinitely many solutions.

   (d) We wish to construct a linear operator $T : \mathbf{R}^2 \to \mathbf{R}^3$ such that, for some $b \in \mathbf{R}^3$, the equation $T(x) = b$ has infinitely many solutions. We will take T defined by $T(x) = (x_1 - x_2, x_1 - x_2, 0)$. Then, for every $\alpha \in \mathbf{R}$, $x = (\alpha, \alpha)$ satisfies $T(x) = 0$.

   (e) We wish to find a linear operator $T : \mathbf{R}^2 \to \mathbf{R}^3$ with the property that $T(x) = b$ does not have a solution for all $b \in \mathbf{R}^3$, but when there is a solution, it is unique. Any nonsingular T will do, since $\mathcal{R}(T)$ is necessarily a proper subspace of $\mathbf{R}^3$ (that is, $T : \mathbf{R}^2 \to \mathbf{R}^3$ cannot be surjective), and nonsingularity implies that $T(x) = b$ has a unique solution for each $b \in \mathcal{R}(T)$ (by Definition 96 and Theorem 92). For example, $T : \mathbf{R}^2 \to \mathbf{R}^3$ defined by $T(x) = (x_1, x_2, 0)$ has the desired properties.

7. Define $S : \mathcal{P}_n \to \mathcal{P}_n$ by $S(p)(x) = p(2x+1)$. We wish to find the rank and nullity of S. We first find the kernel of S. We have that $S(p) = 0$ if and only if $p(2x+1) = 0$ for all $x \in \mathbf{R}$. Consider an arbitrary $y \in \mathbf{R}$ and define $x = (y-1)/2$. Then

$$p(2x+1) = 0 \;\Rightarrow\; p\left(2 \cdot \frac{y-1}{2} + 1\right) = 0 \;\Rightarrow\; p(y) = 0.$$

Since $p(y) = 0$ for all $y \in \mathbf{R}$, it follows that p must be the zero polynomial, and hence the kernel of S is trivial. Thus nullity$(S) = 0$.

Now consider any $q \in \mathcal{P}_n$, and define $p \in \mathcal{P}_n$ by $p(x) = q((x-1)/2)$. Then

$$(S(p))(x) = p(2x+1) = q(\frac{2x+1-1}{2}) = q(x).$$

This shows that S is surjective, and hence rank$(S) = \dim(\mathcal{P}_n) = n+1$.

13. (a) Suppose X and U are finite-dimensional vector spaces over a field F and $T : X \to U$ is an injective linear operator. Then $\mathcal{R}(T)$ is a subspace of U, and we can define $T_1 : X \to \mathcal{R}(T)$ by $T_1(x) = T(x)$ for all $x \in X$. Obviously T_1 is surjective, and it is injective since T is injective. Thus T_1 is an isomorphism between X and $\mathcal{R}(T)$.

(b) Consider the operator $S : \mathcal{P}_{n-1} \to \mathcal{P}_n$ of Example 100 (and the previous exercise), which we have seen to be injective. By the previous part of this exercise, S defines an isomorphism between $\mathcal{P}_{n-1}$ and $\mathcal{R}(S) = \text{sp}\{x, x^2, \ldots, x^n\} \subset \mathcal{P}_n$.

3.6 The fundamental theorem; inverse operators

3. Let F be a field, let $A \in F^{n\times n}$, and let $T : F^n \to F^n$ be defined by $T(x) = Ax$. We first wish to show that A is invertible if and only if T is invertible. We begin by noting that if $M \in F^{n\times n}$ and $Mx = x$ for all $x \in F^n$, then $M = I$. This follows because each linear operator mapping F^n into F^n is represented by a unique matrix (with respect to the standard basis on F^n). The condition $Mx = x$ for all $x \in F^n$ shows that M represents the identity operator, as does the identity matrix I; hence $M = I$.

Now suppose that T is invertible, and let B be the matrix of T^{-1} under the standard basis. We then have

$$(AB)x = A(Bx) = T(T^{-1}(x)) = x \text{ for all } x \in F^n \;\Rightarrow\; AB = I$$

and

$$(BA)x = B(Ax) = T^{-1}(T(x)) = x \text{ for all } x \in F^n \;\Rightarrow\; BA = I.$$

This shows that A is invertible and that $B = A^{-1}$.

Conversely, suppose that A is invertible, and define $S : F^n \to F^n$ by $S(x) = A^{-1}x$. Then

$$S(T(x)) = A^{-1}(Ax) = (A^{-1}A)x = Ix = x \text{ for all } x \in F^n$$

and

$$T(S(x)) = A(A^{-1}x) = (AA^{-1})x = Ix = x \text{ for all } x \in F^n.$$

This shows that T is invertible and that $S = T^{-1}$.

Notice that the above also shows that if A is invertible, then A^{-1} is the matrix defining T^{-1}.

7. We repeat the previous exercise for the operators defined below.

(a) $M : \mathbf{R}^2 \to \mathbf{R}^3$ defined by $M(x) = Ax$, where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since the dimension of the domain is less than the dimension of the co-domain, Theorem 99 implies that M is not surjective. The range of M is spanned by the columns of A (which are linearly independent), so

$$\mathcal{R}(M) = \mathrm{sp}\{(1,1,0),(1,0,1)\}$$

and $\mathrm{rank}(M) = 2$. By the fundamental theorem, we see that $\mathrm{nullity}(M) = 0$; thus $\ker(M)$ is trivial and M is injective.

(b) $M : \mathbf{R}^3 \to \mathbf{R}^2$ define by $M(x) = Ax$, where

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix}.$$

Since the dimension of the domain is greater than the dimension of the co-domain, Theorem 93 implies that M cannot be injective. A direct calculation shows that $\ker(M) = \mathrm{sp}\{(1,-1,1)\}$, and hence $\mathrm{nullity}(M) = 1$. By the fundamental theorem, we see that $\mathrm{rank}(M) = 2 = \dim(\mathbf{R}^2)$. Hence M must be surjective.

13. Let F be a field and suppose $A \in F^{m\times n}$, $B \in F^{n\times p}$. We wish to show that that $\mathrm{rank}(AB) \le \mathrm{rank}(A)$. For all $x \in F^p$, we have $(AB)x = A(Bx)$, where $Bx \in F^n$. It follows that $(AB)x \in \mathrm{col}(A)$ for all $x \in F^p$, which shows that $\mathrm{col}(AB) \subset \mathrm{col}(A)$. Hence, by Exercise 2.6.13, $\dim(\mathrm{col}(AB)) \le \dim(\mathrm{col}(A))$, that is, $\mathrm{rank}(AB) \le \mathrm{rank}(A)$.

15. Suppose $A \in \mathbf{C}^{n\times n}$ is called strictly diagonally dominant:

$$|A_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{ij}|,\ i = 1,2,\ldots,n.$$

We wish to prove that a A is nonsingular. Suppose $x \in \mathbf{R}^n$ satisfies $x \neq 0$, $Ax = 0$. Let j be the index with the property that $|x_j| \ge |x_k|$ for all $k = 1,2,\ldots,n$, $k \neq j$. Then

$$(Ax)_i = \sum_{j=1}^{n} A_{ij}x_j = A_{ii}x_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} A_{ij}x_j.$$

Since $(Ax)_i = 0$, we obtain

$$A_{ii}x_i = -\sum_{\substack{j=1 \\ j\neq i}}^{n} A_{ij}x_j.$$

But

$$\begin{aligned}\left| -\sum_{\substack{j=1 \\ j\neq i}}^{n} A_{ij}x_j \right| \le \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{ij}||x_j| &\le \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{ij}||x_i| \\ &= |x_i| \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{ij}| < |x_i||A_{ii}|.\end{aligned}$$

This is a contradiction, which shows that x cannot be nonzero. Thus the only solution of $Ax = 0$ is $x = 0$, which shows that A must be nonzero.

3.7 Gaussian elimination

3. The matrix A has no inverse; its rank is only 2.

7. The matrix A is not invertible; its rank is only 2.

3.8 Newton's method

1. The two solutions of (3.13) are

$$\left(\sqrt{\frac{-1+\sqrt{5}}{2}}, \frac{-1+\sqrt{5}}{2}\right),\ \left(-\sqrt{\frac{-1+\sqrt{5}}{2}}, \frac{-1+\sqrt{5}}{2}\right).$$

3. We apply Newton's method to solve $F(x)=0$, where $F:\mathbf{R}^3\to\mathbf{R}^3$ is defined by

$$F(x)=\left[\begin{array}{c} x_1^2+x_2^2+x_3^2-1 \\ x_1^2+x_2^2-x_3 \\ 3x_1+x_2+3x_3 \end{array}\right].$$

There are two solutions:

$$(-0.721840, 0.311418, 0.6180340),\ (-0.390621, -0.682238, 0.6180340)$$

(to six digits).

3.9 Linear ordinary differential equations

1. If $x_1(t)=e^{rt}$, $x_2(t)=te^{rt}$, where $r\in\mathbf{R}$, then the Wronskian matrix is

$$W=\left[\begin{array}{cc} e^{rt_0} & t_0e^{rt_0} \\ re^{rt_0} & e^{rt_0}+rt_0e^{rt_0} \end{array}\right].$$

Choosing $t_0=0$, we obtain

$$W=\left[\begin{array}{cc} 1 & 0 \\ r & 1 \end{array}\right],$$

which is obvious nonsingular. Thus $\{e^{rt}, te^{rt}\}$ is a linearly independent subset of $C(\mathbf{R})$.

7. Consider the set $\{x_1,x_2,x_3\}\subset C(\mathbf{R})$, where $x_1(t)=t$, $x_2(t)=t^2$, $x_3(t)=t^3$.

   (a) The Wronskian matrix of x_1,x_2,x_3 at $t_0=1$ is

   $$W=\left[\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 2 & 6 \end{array}\right].$$

   A direct calculation shows that $\ker(W)=\{0\}$, and hence W is nonsingular. By Theorem 129, this implies that $\{x_1,x_2,x_3\}$ is linearly independent.

   (b) The Wronskian matrix of x_1,x_2,x_3 at $t_0=0$ is

   $$W=\left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{array}\right],$$

   which is obviously singular.

   Theorem 130 states that, if x_1,x_2,x_3 are all solutions of a third-order linear ODE, then W is nonsingular for all t_0 if and only if $\{x_1,x_2,x_3\}$ is linearly independent. In this example, $\{x_1,x_2,x_3\}$ is linearly independent, but W is singular for $t_0=0$. This does not violate Theorem 130, because x_1,x_2,x_3 are not solutions of a common third-order ODE.

3.10 Graph theory

1. Let G be a graph, and let v_i, v_j be two nodes in V_G. Since $(A_G^\ell)_{ij}$ is the number of walks of length ℓ joining v_i and v_j, the distance between v_i and v_j is the smallest ℓ ($\ell = 1, 2, 3, \ldots$) such that $(A_G^\ell)_{ij} \neq 0$.

3. Let G be a graph, and let A_G be the adjacency matrix of G. We wish to prove that $\left(A_G^2\right)_{ii}$ is the degree of v_i for each $i = 1, 2, \ldots, n$. Since A_G is symmetric ($(A_G)_{ij} = (A_G)_{ji}$ for all $i, j = 1, 2, \ldots, n$), we have

$$\left(A_G^2\right)_{ii} = \sum_{j=1}^{n} (A_G)_{ij}^2.$$

Now, $(A_G)_{ij}^2$ is 1 if an edge joins v_i and v_j, and it equals 0 otherwise. Thus $\left(A_G^2\right)_{ii}$ counts the number of edges having v_i as an endpoint.

3.11 Coding theory

3. The following message is received.

010110 000101 010110 010110 010001 100100 010110 010001

It is known that the code of Example 141 is used. Let B be the 8×6 binary matrix whose rows are the above codewords. If we try to solve $MG = B$, we obtain

$$M = \begin{bmatrix} 0 & 1 & 0 & 1 \\ x & x & x & x \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

where the second "codeword" cannot be decoded because it is not a true codeword (that is $mG = b_2$ is inconsistent, where $b_2 = 000101$). In fact, both 000111 ($= mG$ for $m = 0001$) and 001101 ($= mG$ for $m = 0011$) are distance 1 from b_2. This means that the first ASCII character, with code $0101xxxx$ could be 01010001 (Q) or 01010011 (S). The remaining characters are 01010101 (U), 01001001 (I), 01010100 (T), so the message is either "QUIT" or "SUIT".

3.12 Linear programming

5. The LP is unbounded; every point of the form $(x_1, x_2) = (3 + t, t/3)$, $t \geq 0$, is feasible, and for such points z increases without bound as $t \to \infty$.

11. If we apply the simplex method to the LP of Example 158, using the smallest subscript rule to choose both the entering and leaving variables, the method terminates after 7 iterations with an optimal solution of $x = (1, 0, 1, 0)$ and $z = 1$. The basic variables change as follows:

$$\begin{aligned} &\{x_5, x_6, x_7\} \to \{x_1, x_6, x_7\} \to \{x_1, x_2, x_7\} \to \{x_3, x_2, x_7\} \to \{x_3, x_4, x_7\} \\ &\to \{x_5, x_4, x_7\} \to \{x_5, x_1, x_7\} \to \{x_5, x_1, x_3\}. \end{aligned}$$

Chapter 4

Determinants and eigenvalues

4.1 The determinant function

3. Define

$$A = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}, \; B = \begin{bmatrix} a & b & c & d \\ 0 & e & f & g \\ 0 & 0 & h & i \\ 0 & 0 & 0 & j \end{bmatrix}.$$

Then

$$\begin{aligned}
\det(A) = \det(ae_1, be_2, ce_3, de_4) &= a\det(e_1, be_2, ce_3, de_4) \\
&= ab\det(e_1, e_2, ce_3, de_4) \\
&= abc\det(e_1, e_2, e_3, de_4) \\
&= abcd\det(e_1, e_2, e_3, e_4) \\
&= abcd.
\end{aligned}$$

Here we have repeatedly used the second part of the definition of the determinant. To compute $\det(B)$, we note that we can add any multiple of one column to another without changing the determinant (the third part of the definition of the determinant). We have

$$\begin{aligned}
\det(B) &= \det(ae_1, be_1 + ee_2, ce_1 + fe_2 + he_3, de_1 + ge_2 + ie_3 + je_4) \\
&= a\det(e_1, be_1 + ee_2, ce_1 + fe_2 + he_3, de_1 + ge_2 + ie_3 + je_4) \\
&= a\det(e_1, ee_2, fe_2 + he_3, ge_2 + ie_3 + je_4) \\
&= ae\det(e_1, e_2, fe_2 + he_3, ge_2 + ie_3 + je_4) \\
&= ae\det(e_1, e_2, he_3, ie_3 + je_4) \\
&= aeh\det(e_1, e_2, e_3, ie_3 + je_4) \\
&= aeh\det(e_1, e_2, e_3, je_4) \\
&= aehj\det(e_1, e_2, e_3, e_4) = aehj.
\end{aligned}$$

5. Consider the permutation $\tau = (4, 3, 2, 1) \in S_4$. We can write

$$\begin{aligned}
\tau &= [2, 3][1, 4], \\
\tau &= [3, 4][2, 4][2, 3][1, 4][1, 3][1, 2].
\end{aligned}$$

Since τ is the product of an even number of permutations, we see that $\sigma(\tau) = 1$.

11. Let n be a positive integer, and let i and j be integers satisfying

$$1 \leq i, j \leq n, \ i \neq j.$$

For any $\tau \in S_n$, define τ' by $\tau' = \tau[i,j]$ (that is, τ' is the composition of τ and the transposition $[i,j]$. Finally, define $f : S_n \to S_n$ by $f(\tau) = \tau'$. We wish to prove that f is a bijection. First, let $\gamma \in S_n$, and define $\theta \in S_n$ by $\theta = \gamma[i,j]$. Then $f(\theta) = (\gamma[i,j])[i,j] = \gamma([i,j][i,j])$. It is obvious that $[i,j][i,j]$ is the identity permutation, and hence $f(\theta) = \gamma$. This shows that f is surjective. Similarly, if $\theta_1, \theta_2 \in S_n$ and $f(\theta_1) = f(\theta_2)$, then $\theta_1[i,j] = \theta_2[i,j]$. But then

$$\begin{aligned} \theta_1[i,j] = \theta_2[i,j] \ &\Rightarrow \ (\theta_1[i,j])[i,j] = (\theta_2[i,j])[i,j] \\ \Rightarrow \theta_1([i,j][i,j]) = \theta_2([i,j][i,j]) \ &\Rightarrow \ \theta_1 = \theta_2. \end{aligned}$$

This shows that f is injective, and hence bijective.

4.2 Further properties of the determinant function

3. Let F be a field and let $A \in F^{n\times n}$. We wish to show that A^TA is singular if and only if A is singular. We have

$$\det(A^TA) = \det(A^T)\det(A) = \det(A)^2,$$

since $\det(A^T) = \det(A)$. It follows that $\det(A^TA) = 0$ if and only if $\det(A) = 0$, that is, A^TA is singular if and only if A is singular.

9. Let $A \in F^{m\times n}$, $B \in F^{n\times m}$, where $m < n$. We will show by example that both $\det(AB) = 0$ and $\det(AB) \neq 0$ are possible. First, note that

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow AB = \begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix}$$
$$\Rightarrow \det(AB) = 4 \cdot 2 - 2 \cdot 4 = 0.$$

On the other hand,

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \Rightarrow AB = \begin{bmatrix} 6 & 2 \\ 2 & 2 \end{bmatrix}$$
$$\Rightarrow \det(AB) = 6 \cdot 2 - 2 \cdot 2 = 8.$$

4.3 Practical computation of $\det(A)$

7. Suppose $A \in \mathbf{R}^{n\times n}$ is invertible and has integer entries, and assume $\det(A) = \pm 1$. Notice that the determinant of a matrix with integer entries is obviously an integer. We can compute the jth column of A^{-1} by solving $Ax = e_j$. By Cramer's rule,

$$(A^{-1})_{ij} = \frac{\det(A_i(e_j))}{\det(A)} = \pm\det(A_i(e_j)), \ i = 1, \ldots, n.$$

Since $\det(A_i(e_j))$ is an integer, so is $(A^{-1})_{ij}$. This holds for all i, j, and hence A^{-1} has integer entries.

8. The solution is $x = (2, 1, 1)$.

4.5 Eigenvalues and the characteristic polynomial

1. (a) The eigenvalues are $\lambda_1 = 1$ (algebraic multiplicity 2) and $\lambda_2 = -1$ (algebraic multiplicity 1). Bases for the eigenspaces are $\{(0,1,0),(1,0,2)\}$ and $\{(4,0,7)\}$, respectively.
   (b) The eigenvalues are $\lambda_1 = 2$ (algebraic multiplicity 2) and $\lambda_2 = 1$ (algebraic multiplicity 1). Bases for the eigenspaces are $\{(1,1,-2)\}$ and $\{(5,5,-9)\}$, respectively.

5. Suppose $A \in \mathbf{R}^{n\times n}$ has a real eigenvalue λ and a corresponding eigenvector $z \in \mathbf{C}^n$. We wish to show that either the real or imaginary part of z is an eigenvector of A. We are given that $Az = \lambda z$, $z \neq 0$. Write $z = x + iy$, $x, y \in \mathbf{R}^n$. Then

$$\begin{aligned} Az = \lambda z \ \Rightarrow \ A(x+iy) = \lambda(x+iy) &\Rightarrow Ax + iAy = \lambda x + i\lambda y \\ &\Rightarrow Ax = \lambda x \text{ and } Ay = \lambda y. \end{aligned}$$

   Since $z \neq 0$, it follows that $x \neq 0$ or $y \neq 0$. If $x \neq 0$, then the real part of z is an eigenvector for A corresponding to λ; otherwise, $y \neq 0$ and the imaginary part of z is an eigenvector.

9. Let $q(r) = r^n + c_{n-1}r^{n-1} + \cdots + c_0$ be an arbitrary polynomial with coefficients in a field F, and let

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & -c_0 \\ 1 & 0 & 0 & \cdots & -c_1 \\ 0 & 1 & 0 & \cdots & -c_2 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix}.$$

   We wish to prove that $p_A(r) = q(r)$. We argue by induction on n. The case $n = 1$ is trivial, since A is 1×1 in that case: $A = [-c_0]$, $|rI - A| = r + c_0 = q(r)$. Suppose the result holds for polynomials of degree $n-1$, let $q(r) = r^n + c_{n-1}r^{n-1} + \ldots + c_0$, and let A be defined as above. Then

$$|rI - A| = \begin{vmatrix} r & 0 & 0 & \cdots & c_0 \\ -1 & r & 0 & \cdots & c_1 \\ 0 & -1 & r & \cdots & c_2 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & r + c_{n-1} \end{vmatrix},$$

   and cofactor expansion along the first row yields

$$|rI - A| = r \begin{vmatrix} r & 0 & \cdots & c_1 \\ -1 & r & \cdots & c_2 \\ \ddots & \ddots & \vdots & \\ 0 & \cdots & -1 & r + c_{n-1} \end{vmatrix} +$$

$$(-1)^{n+1}c_0 \begin{vmatrix} -1 & r & & \cdots & \\ & -1 & r & \cdots & \\ & & \ddots & \ddots & \\ & & & -1 & r \\ & & & & -1 \end{vmatrix}.$$

   By the induction hypothesis, the first determinant is

$$c_1 + c_2 r + \cdots + c_{n-1}r^{n-2} + r^{n-1},$$

   and the second is simply $(-1)^{n-1}$. Thus

$$\begin{aligned} p_A(r) &= |rI - A| \\ &= r\left(c_1 + c_2 r + \cdots + c_{n-1}r^{n-2} + r^{n-1}\right) + (-1)^{n+1}c_0(-1)^{n-1} \\ &= c_1 r + c_2 r^2 + \cdots + c_{n-1}r^{n-1} + r^n + c_0 = q(r). \end{aligned}$$

This completes the proof by induction.

4.6 Diagonalization

3. A is diagonalizable: $A = XDX^{-1}$, where

$$D = \begin{bmatrix} 1 - i\sqrt{2} & 0 \\ 0 & 1 + i\sqrt{2} \end{bmatrix}, \; X = \begin{bmatrix} i\sqrt{2} & -i\sqrt{2} \\ 1 & 1 \end{bmatrix}.$$

7. A is diagonalizable: $A = XDX^{-1}$, where

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \; X = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

11. Let F be a finite field. We will show that F is not algebraically closed by constructing a polynomial $p(x)$ with coefficients in F such that $p(x) \neq 0$ for all $x \in F$. Let the elements of F be $\alpha_1, \alpha_2, \ldots, \alpha_q$. Define

$$p(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_q) + 1.$$

Then p is a polynomial of degree q, and $p(x) = 1$ for all $x \in F$. Thus $p(x)$ has no roots, which shows that F is not algebraically closed.

4.7 Eigenvalues of linear operators

1. Let $T : \mathbf{R}^3 \to \mathbf{R}^3$ be defined by

$$T(x) = (ax_1 + bx_2, bx_1 + ax_2 + bx_3, bx_2 + ax_3),$$

where $a, b \in \mathbf{R}$ are constants. Notice that $T(x) = Ax$ for all $x \in \mathbf{R}^3$, where

$$A = \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix}.$$

The eigenvalues of A are $a, a+\sqrt{2}b, a-\sqrt{2}b$. If $b = 0$, then A is already diagonal, which means that $[T]_{\mathcal{X},\mathcal{X}}$ is diagonal if $\mathcal{X}$ is the standard basis. If $b \neq 0$, then A (and hence T) has three distinct eigenvalues, and hence three linearly independent eigenvectors. It follows that there exists a basis $\mathcal{X}$ such that $[T]_{\mathcal{X},\mathcal{X}}$ is diagonal.

9. Let $L : \mathbf{C}^3 \to \mathbf{C}^3$ be defined by

$$L(z) = (z_1, 2z_2, z_1 + z_3).$$

Then $L(z) = Az$ for all $z \in \mathbf{C}^3$, where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The eigenvalues of A are the diagonal entries, $\lambda_1 = 1$ (with algebraic multiplicity 2) and $\lambda_2 = 2$. A straightforward calculation shows that $E_A(1) = \mathrm{sp}\{(0,0,1)\}$, and hence A is defective and not diagonalizable. Notice that $A = [L]_{\mathcal{S},\mathcal{S}}$, where $\mathcal{S}$ is the standard basis for $\mathbf{C}^3$. Since $[L]_{\mathcal{X},\mathcal{X}}$ is either diagonalizable for every basis $\mathcal{X}$ or diagonalizable for no basis $\mathcal{X}$, we see that $[L]_{\mathcal{X},\mathcal{X}}$ is not diagonalizable for every basis $\mathcal{X}$ of $\mathbf{C}^3$.

11. Let $T:\mathbf{Z}_2^2\to\mathbf{Z}_2^2$ be defined by $T(x)=(0,x_1+x_2)$. Then $T(x)=Ax$ for all $x\in\mathbf{Z}_2$, where

$$A=\left[\begin{array}{cc}0&0\\1&1\end{array}\right].$$

The eigenvalues of A are the diagonal entries, $\lambda_1=0$ and $\lambda_2=1$. Corresponding eigenvectors are $x_1=(1,1)$ and $x_2=(0,1)$, respectively. It follows that if $\mathcal{X}=\{x_1,x_2\}$, then $[T]_{\mathcal{X},\mathcal{X}}$ is diagonal:

$$[T]_{\mathcal{X},\mathcal{X}}=\left[\begin{array}{cc}0&0\\0&1\end{array}\right].$$

4.8 Systems of linear ODEs

5. Let $A\in\mathbf{R}^{2\times 2}$ be defined by

$$A=\left[\begin{array}{cc}1&4\\9&1\end{array}\right].$$

We wish to solve the IVP $u'=Au$, $u(0)=v$, where $v=(1,2)$. The eigenvalues of A are $\lambda_1=7$ and $\lambda_2=-5$. Corresponding eigenvectors are $(2,3)$ and $(2,-3)$, respectively. Therefore, the general solution of the ODE $u'=Au$ is

$$x(t)=c_1e^{7t}\left[\begin{array}{c}2\\3\end{array}\right]+c_2e^{-5t}\left[\begin{array}{c}2\\-3\end{array}\right].$$

We solve $x(0)=(1,2)$ to obtain $c_1=7/12$, $c_2=-1/12$, and thus the solution of the IVP is

$$x(t)=\frac{7}{12}e^{7t}\left[\begin{array}{c}2\\3\end{array}\right]-\frac{1}{12}e^{-5t}\left[\begin{array}{c}2\\-3\end{array}\right].$$

7.
$$e^{tA}=\left[\begin{array}{cc}e^t\cos(2t)&e^t\sin(2t)\\-e^t\sin(2t)&e^t\cos(2t)\end{array}\right].$$

4.9 Integer programming

1. Each of the matrices is totally unimodular by Theorem 219. The sets S_1 and S_2 are given below.

   (a) $S_1=\{1,2\}$, $S_2=\{3,4\}$.

   (b) $S_1=\{1,2,3,4\}$, $S_2=\emptyset$.

   (c) $S_1=\{2,3\}$, $S_2=\{1,4\}$.

Chapter 5

The Jordan canonical form

5.1 Invariant subspaces

1. (a) S is not invariant under A (in fact, A does not map either basis vector into S).
   (b) T is invariant under A.

5. Let $A \in \mathbf{R}^{3\times 3}$ be defined by

$$A = \begin{bmatrix} 3 & 0 & -1 \\ -6 & 1 & 3 \\ 2 & 0 & 0 \end{bmatrix},$$

and let $S = \mathrm{sp}\{s_1, s_2\}$, where $s_1 = (0,1,0)$, $s_2 = (1,0,1)\}$. A direct calculation shows that

$$As_1 = s_1,\ As_2 = -3s_1 + 2s_2.$$

It follows that S is invariant under A. We extend $\{s_1, s_2\}$ to a basis $\{s_1, s_2, s_3\}$ of $\mathbf{R}^3$ by defining $s_3 = (0,0,1)$, and define $X = [s_1|s_2|s_3]$. Then

$$X^{-1}AX = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

is block upper triangular (in fact, simply upper triangular).

9. Let U be a finite-dimensional vector space over a field F, and let $T : U \to U$ be a linear operator. Let $\mathcal{U} = \{u_1, u_2, \ldots, u_n\}$ be a basis for U and define $A = [T]_{\mathcal{U},\mathcal{U}}$. Suppose $X \in F^{n\times n}$ is an invertible matrix, and define $J = X^{-1}AX$. Finally, define

$$v_j = \sum_{i=1}^{n} X_{ij} u_i,\ j = 1, 2, \ldots, n.$$

   (a) We wish to show that $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ is a basis for U. Since $n = \dim(V)$, it suffices to prove that $\mathcal{V}$ is linearly independent. First notice that

$$\begin{aligned} \sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{n} X_{ij} u_i \right) = \sum_{j=1}^{n}\sum_{i=1}^{n} X_{ij} c_j u_i &= \sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij} c_j u_i \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} X_{ij} c_j \right) u_i \\ &= \sum_{j=1}^{n} (Xc)_i u_i. \end{aligned}$$

Also, $\sum_{j=1}^{n}(Xc)_i u_i = 0$ implies that $Xc = 0$ since $\mathcal{U}$ is linearly independent. Therefore,

$$\sum_{j=1}^{n} c_j v_j = 0 \;\Rightarrow\; \sum_{j=1}^{n}(Xc)_i u_i = 0 \;\Rightarrow\; Xc = 0 \;\Rightarrow\; c = 0,$$

where the last step follows from the fact that X is invertible. Thus $\mathcal{V}$ is linearly independent.

(b) Now we wish to prove that $[T]_{\mathcal{V},\mathcal{V}} = J$. The calculation above shows that if $[v]_{\mathcal{V}} = c$, then $[v]_{\mathcal{U}} = Xc$, that is,

$$[v]_{\mathcal{U}} = X[v]_{\mathcal{V}}.$$

Therefore, for all $v \in V$,

$$\begin{aligned}[T]_{\mathcal{V},\mathcal{V}}[v]_{\mathcal{V}} = [T(v)]_{\mathcal{V}} &\Leftrightarrow [T]_{\mathcal{V},\mathcal{V}}X^{-1}[v]_{\mathcal{U}} = X^{-1}[T(v)]_{\mathcal{U}} \\ &\Leftrightarrow \left(X[T]_{\mathcal{V},\mathcal{V}}X^{-1}\right)[v]_{\mathcal{U}} = [T(v)]_{\mathcal{U}},\end{aligned}$$

which shows that $[T]_{\mathcal{U},\mathcal{U}} = X[T]_{\mathcal{V},\mathcal{V}}X^{-1}$, or $[T]_{\mathcal{V},\mathcal{V}} = X^{-1}[T]_{\mathcal{U},\mathcal{U}}X = X^{-1}AX = J$, as desired.

5.2 Generalized eigenspaces

5. For the given $A \in \mathbf{R}^{5\times 5}$, we have $p_A(r) = (r-1)^3(r+1)^2$, so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$. Direct calculation shows that

$$\dim(\mathcal{N}(A-I)) = 1,\ \dim(\mathcal{N}((A-I)^2)) = 2,\ \dim(\mathcal{N}((A-I)^3)) = 3,\ \dim(\mathcal{N}((A-I)^4)) = 3$$

and

$$\dim(\mathcal{N}(A+I)) = 1,\ \dim(\mathcal{N}((A+I)^2)) = 2,\ \dim(\mathcal{N}((A+I)^3)) = 2.$$

These results show that the generalized eigenspaces are $\mathcal{N}((A-I)^3)$ and $\mathcal{N}((A+I)^2)$. Bases for these subspaces are

$$\begin{aligned}\mathcal{N}((A-I)^3) &= \mathrm{sp}\{(0,1,0,0,0),(0,0,1,0,0),(0,0,0,1,0)\},\\ \mathcal{N}((A+I)^2) &= \mathrm{sp}\{(-1,0,12,4,0),(1,4,0,0,2)\}.\end{aligned}$$

A direct calculation shows that the union of the two bases is linearly independent and hence a basis for $\mathbf{R}^5$, and it then follows from Theorem 226 that $\mathbf{R}^5 = \mathcal{N}((A-I)^3) + \mathcal{N}((A+I)^2)$.

9. Let F be a field, let $\lambda \in F$ be an eigenvalue of $A \in F^{n\times n}$, and suppose that the algebraic and geometric multiplicities of λ are equal, say to m. We wish to show that $\mathcal{N}((A-\lambda I)^2) = \mathcal{N}(A-\lambda I)$. By Theorem 235, there exists a positive integer k such that

$$\dim(\mathcal{N}((A-\lambda I)^{k+1})) = \dim(\mathcal{N}((A-\lambda I)^k))$$

and $\dim(\mathcal{N}((A-\lambda I)^k)) = m$. We know that

$$\mathcal{N}(A-\lambda I) \subset \mathcal{N}((A-\lambda I)^2) \subset \cdots \subset \mathcal{N}((A-\lambda I)^k),$$

and, by hypothesis, $\dim(\mathcal{N}(A-\lambda I)) = m = \dim(\mathcal{N}((A-\lambda I)^k))$. This implies that $\mathcal{N}(A-\lambda I) = \mathcal{N}((A-\lambda I)^k)$, and hence that $\mathcal{N}((A-\lambda I)^2) = \mathcal{N}(A-\lambda I)$ (since $\mathcal{N}(A-\lambda I) \subset \mathcal{N}((A-\lambda I)^2) \subset \mathcal{N}((A-\lambda I)^k)$).

13. Let F be a field and suppose $A \in F^{n\times n}$ has distinct eigenvalues $\lambda_1,\ldots,\lambda_t$. We wish to show that A is diagonalizable if and only if

$$m_A(r) = (r-\lambda_1)\cdots(r-\lambda_t).$$

First, suppose A is diagonalizable and write $p(r) = (r-\lambda_1)\cdots(r-\lambda_t)$. Every vector $x \in \mathbf{R}^n$ can be written as a linear combination of eigenvectors of A, from which it is easy to prove that $p(A)x = 0$ for all $x \in F^n$ (recall that the factors $(A-\lambda_i I)$ commute with one another). Hence $p(A)$ is the zero matrix.

It follows from Theorem 239 that $m_A(r)$ divides $p(r)$. But every eigenvalue of A is a root of $m_A(r)$, and hence $p(r)$ divides $m_A(r)$. Since both $p(r)$ and $m_A(r)$ are monic, it follows that $m_A(r) = p(r)$, as desired.

Conversely, suppose the minimal polynomial of A is $m_A(r) = (r - \lambda_1) \cdots (r - \lambda_t)$. But then Corollary 244 implies that F^n is the direct sum of the subspaces

$$\mathcal{N}(A - \lambda_1 I), \ldots, \mathcal{N}(A - \lambda_t I),$$

which are the eigenspaces

$$E_A(\lambda_1), \ldots, E_A(\lambda_t).$$

Since F^n is the direct sum of the eigenspaces, it follows that there is a basis of F^n consisting of eigenvectors of A (see the proof of Theorem 226 given in Exercise 5.1.10), and hence A is diagonalizable.

5.3 Nilpotent operators

5. Let $A \in \mathbf{C}^{n \times n}$ be nilpotent. We wish to prove that the index of nilpotency of A is at most n. By Exercise 2, the only eigenvalue of A is $\lambda = 0$, which means that the characteristic polynomial of A must be $p_A(r) = r^n$. By the Cayley-Hamilton theorem, $p_A(A) = 0$, and hence $A^n = 0$. It follows that the index of nilpotency of A is at most n.

9. Suppose $A \in \mathbf{R}^{n \times n}$ has 0 as its only eigenvalue, so that it must be nilpotent of index k for some k satisfying $1 \leq k \leq 5$. The possibilities are

   - If $k = 1$, then $A = 0$ and $\dim(\mathcal{N}(A)) = 5$.
   - If $k = 2$, there is at least one chain x_1, Ax_1 of nonzero vectors. There could be a second such chain, x_2, Ax_2, in which case the fifth basis vector must come from $\mathcal{N}(A)$, we have $\dim(\mathcal{N}(A)) = 3$, $\dim(\mathcal{N}(A^2)) = 5$, and A is similar to

     $$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

     It is also possible that there is only one such chain, in which case $\dim(\mathcal{N}(A)) = 4$, $\dim(\mathcal{N}(A^2)) = 5$, and A is similar to

     $$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

   - If $k = 3$, there is a chain x_1, Ax_1, A^2x_1 of nonzero vectors. There are two possibilities for the other vectors needed for the basis. There could be a chain x_2, Ax_2 (with A^2x_20), in which case $\dim(\mathcal{N}(A)) = 2$, $\dim(\mathcal{N}(A^2)) = 4$, and $\dim(\mathcal{N}(A^3)) = 5$, and A is similar to

     $$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Altenatively, there could be two more independent vectors in $\mathcal{N}(A)$, in which case $\dim(\mathcal{N}(A)) = 3$, $\dim(\mathcal{N}(A^2)) = 4$, $\dim(\mathcal{N}(A^3)) = 5$, and A is similar to

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- If $k = 4$, then there is a chain $x_1, Ax_1, A^2x_1, A^3x_1$ of nonzero vectors, and there must be a second independent vector in $\mathcal{N}(A)$. Then $\dim(\mathcal{N}(A)) = 2$, $\dim(\mathcal{N}(A^2)) = 3$, $\dim(\mathcal{N}(A^3)) = 4$, $\dim(\mathcal{N}(A^4)) = 5$, and A is similar to

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- If $k = 5$, then there is a chain $x_1, Ax_1, A^2x_1, A^3x_1, A^4x_1$ of nonzero vectors, $\dim(\mathcal{N}(A)) = 1$, $\dim(\mathcal{N}(A^2)) = 2$, $\dim(\mathcal{N}(A^3)) = 3$, $\dim(\mathcal{N}(A^4)) = 4$, $\dim(\mathcal{N}(A^5)) = 5$, and A is similar to

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

15. Let F be a field and suppose $A \in F^{n\times n}$ is nilpotent. We wish to prove that $\det(I + A) = 1$. By Theorem 251 and the following discussion, we know there exists an invertible matrix $X \in F^n$ such that $X^{-1}AX$ is upper triangular with zeros on the diagonal. But then

$$X^{-1}(I + A)X = I + X^{-1}AX$$

is upper triangular with ones on the diagonal. From this, we conclude that $\det(X^{-1}(I+A)X) = 1$. Since $I + A$ is similar to $X^{-1}(I + A)X$, it has the same determinant, and hence $\det(I + A) = 1$.

5.4 The Jordan canonical form of a matrix

1. Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbf{R}^{3\times 3}.$$

The only eigenvalue of A is $\lambda = 1$, and $\dim(\mathcal{N}(A - I)) = 1$. Therefore, there must be a single chain of the form $(A - I)^2x_1, (A - I)x_1, x_1$, where $x_1 \in \mathcal{N}((A - I)^3) \setminus \mathcal{N}((A - I)^2)$. We have

$$(A - I)^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and thus it is easy to see that $x_1 = (0, 0, 1) \notin \mathcal{N}((A - I)^2)$. We define

$$X = [(A - I)^2x_1|(A - I)x_1|x_1] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and then

$$X^{-1}AX = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

5. Let $A \in \mathbf{R}^{4\times 4}$ be defined by

$$A = \begin{bmatrix} -3 & 1 & -4 & -4 \\ -17 & 1 & -17 & -38 \\ -4 & -1 & -3 & -14 \\ 4 & 0 & 4 & 10 \end{bmatrix}.$$

Then $p_A(r) = (r-1)^3(r-2)$. A direct calculation shows that $\dim(\mathcal{N}(A-I)) = 1$, $\dim(\mathcal{N}((A-I)^2)) = 2$, $\dim(\mathcal{N}((A-I)^3)) = 3$ (and, of course, $\dim(\mathcal{N}(A-2I)) = 1$. Therefore, the Jordan canonical form of A is

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

7. Let $A \in \mathbf{R}^{5\times 5}$ be defined by

$$A = \begin{bmatrix} -7 & 1 & 24 & 4 & 7 \\ -9 & 4 & 21 & 3 & 6 \\ -2 & -1 & 11 & 2 & 3 \\ -7 & 13 & -18 & -6 & -8 \\ 3 & -5 & 6 & 3 & 5 \end{bmatrix}.$$

Then $p_A(r) = (r-1)^3(r-2)^2$. A direct calculation shows that $\dim(\mathcal{N}(A-I)) = 2$, $\dim(\mathcal{N}((A-I)^2)) = 3$, $\dim(\mathcal{N}(A-2I)) = 1$, and $\dim(\mathcal{N}((A-2I)^2)) = 2$. Therefore, the Jordan canonical form of A is

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

5.5 The matrix exponential

3. We wish to find the matrix exponential e^{tA} for the matrix given in Exercise 5.4.5. The similarity transformation defined by

$$X = \begin{bmatrix} 1 & 20 & -9 & 0 \\ 0 & 1 & 0 & -4 \\ -1 & -20 & 0 & -2 \\ 0 & 0 & 4 & 1 \end{bmatrix}$$

puts A in Jordan canonical form:

$$X^{-1}AX = J = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

We have

$$e^{tJ} = \begin{bmatrix} e^t & te^t & \frac{t^2}{2}e^t & 0 \\ 0 & e^t & te^t & 0 \\ 0 & 0 & e^t & 0 \\ 0 & 0 & 0 & e^{2t} \end{bmatrix},$$

and $e^{tA} = Xe^{tJ}X^{-1}$ is

$$\begin{bmatrix} e^t\left(1-4t-\frac{t^2}{2}\right) & te^t & e^t\left(-4t-\frac{t^2}{2}\right) & e^t\left(-4t-t^2\right) \\ e^t(16-t)-16e^{2t} & e^t & e^t(16-t)-16e^{2t} & e^t(36-2t)-36e^{2t} \\ e^t\left(8-4t+\frac{t^2}{2}\right)-8e^{2t} & -te^t & e^t\left(9+4t+\frac{t^2}{2}\right)-8e^{2t} & e^t\left(18+4t+t^2\right)-18e^{2t} \\ -4e^t+4e^{2t} & 0 & -4e^t+4e^{2t} & -8e^t+4e^{2t} \end{bmatrix}.$$

5. Let $A, B \in \mathbf{C}^{n\times n}$.

   (a) We first show that if A and B commute, then so do e^{tA} and B. We define $U(t) = e^{tA}B - Be^{tA}$. Then

   $$\begin{aligned} U'(t) = Ae^{tA}B - BAe^{tA} &= Ae^{tA}B - ABe^{tA} \\ &= A\left(e^{tA}B - Be^{tA}\right) = AU(t). \end{aligned}$$

   Also, $U(0) = IB - BI = B - B = 0$. Thus U satisfies $U' = AU$, $U(0) = 0$. But the unique solution of this IVP is $U(t) = 0$; hence $e^{tA}B = Be^{tA}$.

   (b) Use the preceding result to show that if A and B commute, the $e^{t(A+B)} = e^{tA}e^{tB}$ holds. We define $U(t) = e^{tA}e^{tB}$. Then

   $$\begin{aligned} U'(t) = Ae^{tA}e^{tB} + e^{tA}Be^{tB} &= Ae^{tA}e^{tB} + Be^{tA}e^{tB} \\ &= (A+B)e^{tA}e^{tB} \end{aligned}$$

   (notice how we used the first part of the exercise), and $U(0) = I$. But $e^{t(A+B)}$ also satisfies the IVP $U' = (A+B)U$, $U(0) = I$. Hence $e^{t(A+B)}$ and $e^{tA}e^{tB}$ must be equal.

   (c) Let

   $$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

   Then

   $$e^{tA} = \begin{bmatrix} \frac{1+e^2}{2} & \frac{-1+e^2}{2e} \\ \frac{-1+e^2}{2e} & \frac{1+e^2}{2e} \end{bmatrix},\ e^{tB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

   and

   $$e^{A+B} = \begin{bmatrix} \frac{1}{2}\left(e^{\sqrt{2}} + e^{-\sqrt{2}}\right) & \frac{1}{\sqrt{2}}\left(e^{\sqrt{2}} - e^{-\sqrt{2}}\right) \\ \frac{1}{2\sqrt{2}}\left(e^{\sqrt{2}} - e^{-\sqrt{2}}\right) & \frac{1}{2}\left(e^{\sqrt{2}} + e^{-\sqrt{2}}\right) \end{bmatrix}.$$

   It is easy to verify that $e^Ae^B \neq e^{A+B}$.

5.6 Graphs and eigenvalues

3. The adjacency matrix is

$$A_G = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

and its eigenvalues are $0, \pm 1, \pm 2$. Notice that the adjacency matrix shows that every vertex has degree 2, and hence G is 2-regular. Theorem 263 states that the largest eigenvalue of G should be 2, as indeed it is. Also, since the multiplicity of $\lambda = 2$ is 2, G must have two connected components. It does; the vertices v_1, v_2, v_3 form one connected component, and the vertices v_4, v_5, v_6, v_7 form the other.

Chapter 6

Orthogonality and best approximation

6.1 Norms and inner products

1. Let V be a normed vector space over $\mathbf{R}$, and suppose $u, v \in V$ with $v = \alpha u$, $\alpha \geq 0$. Then

$$\begin{aligned} \|u+v\| = \|u+au\| = \|(1+a)u\| &= (1+a)\|u\| \\ &= \|u\| + a\|u\| \\ &= \|u\| + \|au\| = \|u\| + \|v\| \end{aligned}$$

(note the repeated use of the second property of a norm from Definition 265). Thus, if $v = au$, $a \geq 0$, then equality holds in the triangle inequality ($\|u+v\| = \|u\| + \|v\|$).

5. We wish to derive relationships among the $L^1(a,b)$, $L^2(a,b)$, and $L^\infty(a,b)$ norms. Three such relationships exist:

$$\begin{aligned} \|f\|_1 &\leq (b-a)\|f\|_\infty \text{ for all } f \in L^\infty(a,b), \\ \|f\|_2 &\leq \sqrt{b-a}\|f\|_\infty \text{ for all } f \in L^\infty(a,b), \\ \|f\|_1 &\leq \sqrt{b-a}\|f\|_2 \text{ for all } f \in L^2(a,b). \end{aligned}$$

The first two are simple; we have $|f(x)| \leq \|f\|_\infty$ for (almost) all $x \in [a,b]$, and hence

$$\begin{aligned} \|f\|_1 &= \int_a^b |f(x)|\,dx \leq \int_a^b \|f\|_\infty\,dx = (b-a)\|f\|_\infty, \\ \|f\|_2^2 &= \int_a^b |f(x)|^2\,dx \leq \int_a^b \|f\|_\infty^2\,dx = (b-a)\|f\|_\infty^2. \end{aligned}$$

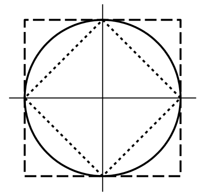
To prove the third relationship, define $v(x) = 1$ for all $x \in [a,b]$. Then

$$\int_a^b |f(x)|\,dx = \int_a^b |v(x)||f(x)|\,dx = \langle |v|, |f| \rangle_2 \leq \|v\|_2 \|f\|_2,$$

where the last step follows from the Cauchy-Schwarz inequality. Since $\|v\|_2 = \sqrt{b-1}$, the desired result follows.

Note that it is not possible to bound $\|f\|_\infty$ by a multiple of either $\|f\|_1$ or $\|f\|_2$, nor to bound $\|f\|_2$ by a multiple of $\|f\|_1$.

9. The following graph shows the (boundaries of the) unit balls in the ℓ^2 norm (solid curve), the ℓ^1 norm (dotted curve), and the ℓ^∞ norm (dashed curve).



Notice that ℓ^1 unit ball is contained in the other two, which is consistent with $\|x\|_1 \geq \|x\|_2, \|x\|_\infty$, while the ℓ^2 unit ball is contained in the ℓ^∞ unit ball, consistent with $\|x\|_\infty \leq \|x\|_2$.

11. Suppose V is an inner product space and $\|\cdot\|$ is the norm defined by the inner product $\langle\cdot,\cdot\rangle$ on V. Then, for all $u, v \in V$,

$$\begin{aligned}
\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle \\
&= \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle + \langle u,u\rangle - 2\langle u,v\rangle + \langle v,v\rangle \\
&= 2\langle u,u\rangle + 2\langle v,v\rangle \\
&= 2\|u\|^2 + 2\|v\|^2.
\end{aligned}$$

Thus the parallelogram law holds in V.

If $u = (1,1)$ and $v = (1,-1)$ in $\mathbf{R}^2$, then a direct calculation shows that

$$\|u+v\|_1^2 + \|u-v\|_1^2 = 8,\ 2\|u\|_1^2 + 2\|v\|_1^2 = 16,$$

and hence the parallelogram law does not hold for $\|\cdot\|_1$. Therefore, $\|\cdot\|_1$ cannot be defined by an inner product. For the same u and v, we have

$$\|u+v\|_\infty^2 + \|u-v\|_\infty^2 = 8,\ 2\|u\|_\infty^2 + 2\|v\|_\infty^2 = 4,$$

and hence the parallelogram law does not hold for $\|\cdot\|_\infty$. Therefore, $\|\cdot\|_\infty$ cannot be defined by an inner product.

6.2 The adjoint of a linear operator

9. Let $M : \mathcal{P}_2 \to \mathcal{P}_3$ be defined by $M(p) = q$, where $q(x) = xp(x)$. We wish to find M^*, assuming that the $L^2(0,1)$ inner product is imposed on both $\mathcal{P}_2$ and $\mathcal{P}_3$. Following the technique of Example 277 (and the previous exercise), we compute the matrix N, defined by $N_{ij} = \langle M(p_i), q_j\rangle_{\mathcal{P}_3}$, where $\mathcal{S}_1 = \{p_1, p_2, p_3\} = \{1, x, x^2\}$ and $\mathcal{S}_2 = \{q_1, q_2, q_3, q_4\} = \{1, x, x^2, x^3\}$:

$$N = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}.$$

The Gram matrix is the same as in the previous exercise, and we obtain

$$[M]_{\mathcal{S}_3,\mathcal{S}_2} = G^{-1}N = \begin{bmatrix} 0 & 0 & \frac{1}{20} & \frac{3}{35} \\ 1 & 0 & -\frac{3}{5} & -\frac{32}{35} \\ 0 & 1 & \frac{3}{2} & \frac{12}{7} \end{bmatrix}.$$

This shows that M^* maps $a_0 + a_1x + a_2x^2 + a_3x^3$ to

$$\left(\frac{1}{20}a_2 + \frac{3}{35}a_3\right) + \left(a_0 - \frac{3}{5}a_2 - \frac{32}{35}a_3\right)x + \left(a_1 + \frac{3}{2}a_2 + \frac{12}{7}a_3\right)x^2.$$

11. Let X and U be finite-dimensional inner product spaces over $\mathbf{R}$, and suppose $T : X \to U$ is linear. Define $S : \mathcal{R}(T^*) \to \mathcal{R}(T)$ by $S(x) = T(x)$ for all $x \in \mathcal{R}(T^*)$. We will prove that S is an isomorphism between $\mathcal{R}(T^*)$ and $\mathcal{R}(T)$.

(a) First, suppose $x_1, x_2 \in \mathcal{R}(T^*)$ satisfy $S(x_1) = S(x_2)$. Since $x_1, x_2 \in \mathcal{R}(T^*)$, there exist $u_1, u_2 \in U$ such that $x_1 = T^*(u_1)$, $x_2 = T^*(u_2)$. Then we have

$$\begin{aligned} S(x_1) = S(x_2) &\Rightarrow T(x_1) = T(x_2) \\ &\Rightarrow T(T^*(u_1)) = T(T^*(u_2)) \\ &\Rightarrow T(T^*(u_1 - u_2)) = 0 \\ &\Rightarrow \langle u_1 - u_2, T(T^*(u_1 - u_2))\rangle_U = 0 \\ &\Rightarrow \langle T^*(u_1 - u_2), T^*(u_1 - u_2)\rangle_X = 0 \\ &\Rightarrow T^*(u_1 - u_2) = 0 \\ &\Rightarrow T^*(u_1) = T^*(u_2) \\ &\Rightarrow x_1 = x_2. \end{aligned}$$

Therefore, S is injective.

(b) Since S injective, Theorem 93 implies that $\dim(\mathcal{R}(T)) \geq \dim(\mathcal{R}(T^*))$. This results holds for any linear operator mapping one finite-dimensional inner product space to another, and hence it applies to the operator T^*. Hence $\dim(\mathcal{R}(T^*)) \geq \dim(\mathcal{R}((T^*)^*))$. Since $(T^*)^* = T$, it follows that $\dim(\mathcal{R}(T^*)) \geq \dim(\mathcal{R}(T))$, and therefore that $\dim(\mathcal{R}(T^*)) \geq \dim(\mathcal{R}(T))$ (that is, $\operatorname{rank}(T^*) = \operatorname{rank}(T)$).

(c) Now we see that S is an injective linear operator mapping one finite-dimensional vector space to another of the same dimension. It follows from Corollary 105 that S is also surjective and hence is an isomorphism.

6.3 Orthogonal vectors and bases

3. The equation $\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = q$ is equivalent to

$$\left(\alpha_1 - \frac{1}{2}\alpha_2 + \frac{1}{6}\alpha_3\right) + (\alpha_2 - \alpha_3)x + \alpha_3 x^2 = 3 + 2x + x^2,$$

which in turn is equivalent to the system

$$\begin{aligned} \alpha_1 - \frac{1}{2}\alpha_2 + \frac{1}{6}\alpha_3 &= 3, \\ \alpha_2 - \alpha_3 &= 2, \\ \alpha_3 &= 1. \end{aligned}$$

The solution is $\alpha = (13/3, 3/1)$, and thus

$$q(x) = \frac{13}{3}p_1(x) + 3p_2(x) + p_3(x).$$

7. Consider the functions e^x and e^{-x} to be elements of $C[0,1]$, and regard $C[0,1]$ as an inner product space under the $L^2(0,1)$ inner product. Define $S = \operatorname{sp}\{e^x, e^{-x}\}$. We wish to find an orthogonal basis $\{f_1, f_2\}$ for S. We will take $f_1(x) = e^x$. The function f_2 must be of the form $f(x) = c_1 e^x + c_2 e^{-x}$ and satisfy

$$\int_0^1 f(x)e^x \, dx = 0.$$

This last condition leads to the equation

$$c_1 \int_0^1 e^{2x}\, dx + c_2 \int_0^1 1\, dx = 0 \Leftrightarrow \frac{1}{2}(e^2-1)c_1 + c_2 = 0.$$

One solution is $c_1 = 2$, $c_2 = 1 - e^2$. Thus if $f_2(x) = 2e^x + (1-e^2)e^{-x}$, then $\{f_1, f_2\}$ is an orthonal basis for S.

15. Let V be an inner product space over $\mathbf{R}$, and let u, v be vectors in V.

   (a) Assume u and v are nonzero. We wish to prove that $v \in \mathrm{sp}\{u\}$ if and only if $|\langle u, v\rangle| = \|u\|\|v\|$. First suppose $v \in \mathrm{sp}\{u\}$. Then, by Exercise 13,

   $$v = \frac{\langle v, u\rangle}{\langle u, u\rangle} u \Rightarrow \|v\| = \left\| \frac{\langle v, u\rangle}{\langle u, u\rangle} u \right\| = \frac{|\langle v, u\rangle|}{\|u\|^2}\|u\| = \frac{|\langle v, u\rangle|}{\|u\|}.$$

   This yields $\|u\|\|v\| = |\langle v, u\rangle|$, as desired.

   On the other hand, if $v \notin \mathrm{sp}\{u\}$, then, by the previous exercise,

   $$\|v\| > \left\| \frac{\langle v, u\rangle}{\langle u, u\rangle} u \right\| = \frac{|\langle v, u\rangle|}{\|u\|},$$

   and hence $\|u\|\|v\| > |\langle v, u\rangle|$. Thus $v \in \mathrm{sp}\{u\}$ if and only if $\|u\|\|v\| = |\langle v, u\rangle|$.

   (b) If u and v are nonzero, then the first part of the exercise shows that equality holds in the Cauchy-Schwarz inequality if and only if $v \in \mathrm{sp}\{u\}$, that is, if and only if v is a multiple of u. If $v = 0$ or $u = 0$, then equality trivially holds in the Cauchy-Schwarz inequality (both sides are zero). Thus we can say that equality holds in the Cauchy-Schwarz inequality if and only if $u = 0$ or v is a multiple of u.

6.4 The projection theorem

1. Let $A \in \mathbf{R}^{m\times n}$.

   (a) We wish to prove that $\mathcal{N}(A^TA) = \mathcal{N}(A)$. First, if $x \in \mathcal{N}(A)$, then $Ax = 0$, which implies that $A^TAx = 0$, and hence that $x \in \mathcal{N}(A^TA)$. Thus $\mathcal{N}(A) \subset \mathcal{N}(A^TA)$.

   Conversely, suppose $x \in \mathcal{N}(A^TA)$. Then

   $$A^TAx = 0 \;\Rightarrow\; x \cdot A^TAx = 0 \;\Rightarrow\; (Ax)\cdot(Ax) = 0 \;\Rightarrow\; Ax = 0.$$

   Therefore, $x \in \mathcal{N}(A)$, and we have shown that $\mathcal{N}(A^TA) \subset \mathcal{N}(A)$. This completes the proof.

   (b) If A has full rank, then the null space of A is trivial (by the fundamental theorem of linear algebra) and hence so is the null space of $\mathcal{N}(A^TA)$. Since A^TA is square, this shows that A^TA is invertible.

   (c) Thus, if A has full rank, then A^TA is invertible and, for any $y \in \mathbf{R}^m$ there is a unique solution $x = (A^TA)^{-1}A^Ty$ of the normal equations $A^TAx = A^Ty$. Thus, by Theorem 291, there is a unique least-squares solution to $Ax = y$.

9. Consider the following data points: $(0, 3.1)$, $(1, 1.4)$, $(2, 1.0)$, $(3, 2.2)$, $(4, 5.2)$, $(5, 15.0)$. We wish to find the function of the form $f(x) = a_1e^x + a_2e^{-x}$ that fits the data as nearly as possible in the least-squares sense. We wish to solve the equations

$$a_1e^{x_i} + a_2e^{-x_i} = y_i,\; i = 1, 2, 3, 4, 5, 6$$

in the least-squares sense. These equations are equivalent to the system $Ma = y$, where

$$M = \begin{bmatrix} e^{x_1} & e^{-x_1} \\ e^{x_2} & e^{-x_2} \\ e^{x_3} & e^{-x_3} \\ e^{x_4} & e^{-x_4} \\ e^{x_5} & e^{-x_5} \\ e^{x_6} & e^{-x_6} \end{bmatrix}, \quad y = \begin{bmatrix} 3.1 \\ 1.4 \\ 1.0 \\ 2.2 \\ 5.2 \\ 15.0 \end{bmatrix}.$$

The solution is $a \doteq (0.10013, 2.9878)$, and the approximating function is $0.10013e^x + 2.9878e^{-x}$.

15. Let $A \in \mathbf{R}^{m\times n}$, where $m < n$ and $\operatorname{rank}(A) = m$. Let $y \in \mathbf{R}^m$.

(a) Since $\operatorname{rank}(A) = \dim(\operatorname{col}(A))$, the fact that $\operatorname{rank}(A) = m$ proves that $\operatorname{col}(A) = \mathbf{R}^m$. Thus $Ax = y$ has a solution for all $y \in \mathbf{R}^m$. Moreover, by Theorem 93, $\mathcal{N}(A)$ is nontrivial (a linear operator cannot be injective unless the dimension of the co-domain is at least as large as the dimension of the domain), and hence $Ax = y$ has infinitely many solutions.

(b) Consider the matrix $AA^T \in \mathbf{R}^{m\times m}$. We know from Exercise 1(a) (applied to A^T) that $\mathcal{N}(AA^T) = \mathcal{N}(A^T)$. By Exercise 6.2.11, $\operatorname{rank}(A^T) = \operatorname{rank}(A) = m$, and hence $\operatorname{nullity}(A^T) = 0$ by the fundamental theorem of linear algebra. Since $\mathcal{N}(AA^T) = \mathcal{N}(A^T)$ is trivial, this proves that the square matrix AA^T is invertible.

(c) If $\overline{x} = A^T\left(AA^T\right)^{-1} y$, then

$$A\overline{x} = A\left(A^T\left(AA^T\right)^{-1}y\right) = \left(AA^T\right)\left(AA^T\right)^{-1}y = y,$$

and hence $\overline{x}$ is a solution of $Ax = y$.

6.5 The Gram-Schmidt process

5. (a) The best cubic approximation, in the $L^2(-1,1)$ norm, to the function $f(x) = e^x$, is the polynomial $q(x) = \alpha_1+\alpha_2x+\alpha_3x^2+\alpha_4x^3$, where $G\alpha = b$. G is the Gram matrix, and b is defined by $b_i = \langle f, p_1\rangle_2$, where $p_i(x) = x^{i-1}$, $i = 1, 2, 3, 4$:

$$G = \begin{bmatrix} 2 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} \end{bmatrix}, \quad b = \begin{bmatrix} e - e^{-1} \\ 2e^{-1} \\ e - 5e^{-1} \\ 16e^{-1} - 2e \end{bmatrix}.$$

We obtain

$$\alpha = \begin{bmatrix} \frac{33}{4e} - \frac{3e}{4} \\ \frac{105e}{4} - \frac{765}{4e} \\ \frac{15e}{4} - \frac{105}{4e} \\ \frac{1295}{4e} - \frac{175e}{4} \end{bmatrix}.$$

(b) Applying the Gram-Schmidt process to the standard basis $\{1, x, x^2, x^3\}$, we obtain the orthogonal basis

$$\left\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\right\}.$$

(c) Using the orthogonal basis for $\mathcal{P}_3$, we compute the best approximation to $f(x) = e^x$ from $\mathcal{P}_3$ (relative to the $L^2(-1,1)$ norm) to be

$$\frac{e - e^{-1}}{2} + \frac{3}{e}x + \left(\frac{15e}{4} - \frac{105}{4e}\right)\left(x^2 - \frac{1}{3}\right) + \left(\frac{1295}{4e} - \frac{175e}{4}\right)\left(x^3 - \frac{3}{5}x\right).$$

9. Let P be the plane in $\mathbf{R}^3$ defined by the equation $3x - y - z = 0$.
   (a) A basis for P is $\{(1,3,0),(1,0,3)\}$; applying the Gram-Schmidt process to this basis yields the orthogonal basis $\{(1,3,0),(9/10,-3/10,3)\}$.
   (b) The projection of $u = (1,1,1)$ onto P is $(8/11, 12/11, 12/11)$.

13. Define an inner product on $C[0,1]$ by
$$\langle f, g\rangle = \int_0^1 (1+x)f(x)g(x)\,dx.$$
   (a) We will first verify that $\langle\cdot,\cdot\rangle$ really does define an inner product on $C[0,1]$. For all $f, g \in C[0,1]$, we have
$$\langle f, g\rangle = \int_0^1 (1+x)f(x)g(x)\,dx = \int_0^1 (1+x)g(x)f(x)\,dx = \langle g, f\rangle\,,$$
   and thus the first property of an inner product is satisfied. If $f, g, h \in C[0,1]$ and $\alpha, \beta \in \mathbf{R}$, then
$$\begin{aligned}\langle \alpha f + \beta g, h\rangle &= \int_0^1 (1+x)(\alpha f(x) + \beta g(x))h(x)\,dx \\ &= \int_0^1 \{\alpha(1+x)f(x)h(x) + \beta(1+x)g(x)h(x)\}\,dx \\ &= \alpha\int_0^1 (1+x)f(x)h(x)\,dx + \beta\int_0^1 (1+x)g(x)h(x)\,dx \\ &= \alpha\langle f, h\rangle + \beta\langle g, h\rangle\,.\end{aligned}$$
   This verifies the second property of an inner product. Finally, for any $f \in C[0,1]$,
$$\langle f, f\rangle = \int_0^1 (1+x)f(x)^2\,dx \geq 0$$
   (since $(1+x)f(x)^2 \geq 0$ for all $x \in [0,1]$). Also, if $\langle f, f\rangle = 0$, then $(1+x)f(x)^2 = 0$ for all $x \in [0,1]$. Since $1 + x > 0$ for all $x \in [0,1]$, this implies that $f(x)^2 \equiv 0$, or $f(x) \equiv 0$. Therefore, $\langle f, f\rangle = 0$ if and only if $f = 0$, and we have verified that $\langle\cdot,\cdot\rangle$ defines an inner product on $C[0,1]$.
   (b) Applying the Gram-Schmidt process to the standard basis yields the orthogonal basis
$$\left\{1, x - \frac{5}{9}, x^2 - \frac{68}{65}x + \frac{5}{26}\right\}.$$

6.6 Orthogonal complements

1. Let $S = \mathrm{sp}\{(1,2,1,-1),(1,1,2,0)\}$. We wish to find a basis for $S^\perp$. Let $A \in \mathbf{R}^{2\times 4}$ be the matrix whose rows are the given vectors (the basis vectors for S). Then $x \in S^\perp$ if and only if $Ax = 0$; that is, $S^\perp = \mathcal{N}(A)$. A direct calculation shows that $S^\perp = \mathcal{N}(A) = \{(-3,1,1,0),(-1,1,0,1)\}$.

5. (a) Since $\mathcal{N}(A)$ is orthogonal to $\mathrm{col}(A^T)$, it suffices to orthogonalize the basis of $\mathrm{col}(A^T)$ by (a single step of) the Gram-Schmidt process, yielding $\{(1,4,-4),(-16/33,-97/33,-101/33)\}$. Then
$$\{(24,-5,1),(1,4,-4),(-16/33,-97/33,-101/33)\}$$
   is an orthogonal basis for $\mathbf{R}^3$.
   (b) A basis for $\mathcal{N}(A^T)$ is $\{(1,-1,1,0),(-2,1,0,1)\}$ and a basis for $\mathrm{col}(A)$ is $\{(1,1,0,1),(4,3,-1,5)\}$. Applying the Gram-Schmidt process to each of the bases individually yields $\{(1,-1,1,0),(-1,0,1,1)\}$ and $\{(1,1,0,1),(0,-1,-1,1)\}$, respectively. The union of these two bases,
$$\{(1,1,0,1),(-1,0,1,1),(1,1,0,1),(0,-1,-1,1)\},$$
   is an orthogonal basis for $\mathbf{R}^4$.

9. (a) Let $A \in \mathbf{R}^{n\times n}$ be symmetric. Then $\mathcal{N}(A)^\perp = \mathrm{col}(A)$ and $\mathrm{col}(A)^\perp = \mathcal{N}(A)$.
   (b) It follows that $y \in \mathrm{col}(A)$ if and only if $y \in \mathcal{N}(A)^\perp$; in other words, $Ax = y$ has a solution if and only if
$$Az = 0 \ \Rightarrow\ y\cdot z = 0.$$

6.7 Complex inner product spaces

1. The projection of v onto S is
$$w = \left(\frac{2}{3} - \frac{4}{9}i, \frac{4}{9} + \frac{2}{3}i, \frac{1}{3} + \frac{4}{9}i\right).$$

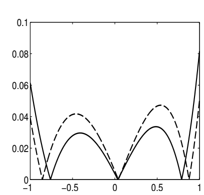
5. The best approximation to f from $\mathcal{P}_2$ is
$$\frac{2i}{\pi} - \frac{24}{\pi^2}\left(x - \frac{1}{2}\right) + \frac{60i(\pi^2-12)}{\pi^3}\left(x^2 - x + \frac{1}{6}\right).$$

9. (a) Let $u = (1,1), v = (1, -1+i) \in \mathbf{C}^2$. Then a direct calculation shows that $\|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2$ and $\langle u, v\rangle_2 = -i \neq 0$.
   (b) Suppose V is a complex inner product space. If $u, v \in V$ and $\|u+v\|^2 = \|u\|^2 + \|v\|^2$, then
$$\begin{aligned} & \langle u,u\rangle + \langle v,v\rangle + \langle u,v\rangle + \langle v,u\rangle = \langle u,u\rangle + \langle v,v\rangle \\ \Leftrightarrow\ & \langle u,v\rangle + \langle v,u\rangle = 0 \\ \Leftrightarrow\ & \langle u,v\rangle + \overline{\langle u,v\rangle} = 0. \end{aligned}$$
   For any $z = x + iy \in \mathbf{C}$, $z + \overline{z} = 0$ is equivalent to $2x = 0$ or, equivalently, $x = 0$. Thus $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ holds if and only if the real part of $\langle u, v\rangle$ is zero.

6.8 More on polynomial approximation

1. (a) The best quadratic approximation to f in the (unweighted) $L^2(-1,1)$ norm is
$$\frac{e - e^{-1}}{2} + \frac{3}{e}x + \frac{15e - 105e^{-1}}{4}\left(x^2 - \frac{1}{3}\right).$$
   (b) The best quadratic approximation to f in the weighted $L^2(-1,1)$ norm is (approximately)
$$1.2660659 + 1.1303182x + 0.27149534(2x^2 - 1)$$
   (note that the integrals were computed numerically).

   The following graph shows the error in the L^2 approximation (solid curve) and the error in the weighted L^2 approximation (dashed) curve.



   We see that the weighted L^2 approximation has the smaller maximum error.

3. (a) The orthogonal basis for $\mathcal{P}_3$ on $[-1,1]$ is

$$\left\{1, x, x^2-\frac{1}{3}, x^3-\frac{3}{5}x\right\}.$$

Transforming this to the interval $[0,\pi]$, we obtain the following orthogonal basis for $\mathcal{P}_3$ as a subspace of $L^2(0,\pi)$:

$$\left\{1, \frac{2}{\pi}t-1, \frac{4}{\pi^2}t^2-\frac{4}{\pi}t+\frac{2}{3}, \frac{8}{\pi^3}t^3-\frac{12}{\pi^2}t^2+\frac{24}{5\pi}t-\frac{2}{5}\right\}.$$

The best cubic approximation to $f(t)=\sin(t)$ on $[0,\pi]$ is

$$p(t)=\frac{2}{\pi}+\frac{15(\pi^2-12)}{\pi^3}\left(\frac{4}{\pi^2}t^2-\frac{4}{\pi}t+\frac{2}{3}\right).$$

(b) The orthogonal basis for $\mathcal{P}_3$ on $[-1,1]$, in the weighted L^2 inner product, is

$$\left\{1, x, 2x^2-1, 4x^3-3x\right\}.$$

Transforming this to the interval $[0,\pi]$, we obtain the following orthogonal basis for $\mathcal{P}_3$ as a subspace of $L^2(0,\pi)$:

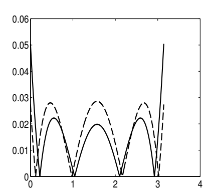
$$\left\{1, \frac{2}{\pi}t-1, \frac{8}{\pi^2}t^2-\frac{8}{\pi}t+1, \frac{32}{\pi^3}t^3-\frac{48}{\pi^2}t^2+\frac{18}{\pi}t-1\right\}.$$

The best cubic approximation to $f(t)=\sin(t)$ on $[0,\pi]$, in the weighted L^2 norm, is (approximately)

$$q(t)=0.47200122-0.49940326\left(\frac{8}{\pi^2}t^2-\frac{8}{\pi}t+1\right)$$

(the integrals were computed numerically).

(c) The following graph shows the error in the ordinary L^2 approximation (solid curve) and the error in the weighted L^2 approximation (dashed curve):



The second approximation has a smaller maximum error.

6.9 The energy inner product and Galerkin's method

5. The variational form of the BVP is to find $u\in V$ such that

$$\int_0^\ell \left\{k(x)\frac{du}{dx}(x)\frac{dv}{dx}(x)+p(x)u(x)v(x)\right\}dx=\int_0^\ell f(x)v(x)\,dx \text{ for all } v\in V.$$

If the basis for V_n is $\{\phi_1,\ldots,\phi_n\}$, then Galerkin's method results in the system $(K+M)U=F$, where K and F are the same stiffness matrix and load vector as before, and M is the *mass matrix:*

$$M_{ij}=\int_0^\ell p(x)\phi_j(x)\phi_i(x)\,dx,\ i,j=1,\ldots,n.$$

7. The projection of f onto V can be computed just as described by the projection theorem. The projection is $\sum_{i=1}^{n-1} V_i\phi_i$, where $V \in \mathbf{R}^{n-1}$ satisfies the system $MV = B$. The Gram matrix is called the mass matrix in this context (see the solution of Exercise 5):

$$M_{ij} = \int_0^\ell \phi_j(x)\phi_i(x)\,dx,\ i,j = 1,\ldots,n-1.$$

The vector B is defined by $B_i = \int_0^\ell f(x)\phi_i(x)\,dx$, $i = 1,\ldots,n-1$.

6.10 Gaussian quadrature

1. We wish to find the Gaussian quadrature rule with $n = 3$ quadrature nodes (on the reference interval $[-1,1]$). We know that the nodes are the roots of the third orthogonal polynomial, $p_3(x)$. Therefore, the nodes are $x_1 = -\sqrt{3/5}$, $x_2 = 0$, $x_3 = \sqrt{3/5}$. To find the weights, we need the Lagrange polynomials defined by these nodes, which are

$$L_1(x) = \frac{5}{6}x(x-\sqrt{3/5}),\ L_2(x) = -\frac{5}{3}x^2+1,\ L_3(x) = \frac{5}{6}x(x+\sqrt{3/5}).$$

Then

$$\begin{aligned} w_1 &= \int_{-1}^1 L_1(x)\,dx = \frac{5}{9}, \\ w_2 &= \int_{-1}^1 L_2(x)\,dx = \frac{5}{9}, \\ w_3 &= \int_{-1}^1 L_3(x)\,dx = \frac{8}{9}. \end{aligned}$$

3. Let $w(x) = 1/\sqrt{1-x^2}$. We wish to find the weighted Gaussian quadrature rule

$$\int_{-1}^1 w(x)f(x)\,dx \doteq \sum_{i=1}^n w_i f(x_i),$$

where $n = 3$. The nodes are the roots of the third orthogonal polynomial under this weight function, which is $T_3(x) = 4x^3 - 3x$. The nodes are thus $x_1 = -\sqrt{3/4}$, $x_2 = 0$, $x_3 = \sqrt{3/4}$. The corresponding Lagrange polynomials are

$$L_1(x) = \frac{2}{3}x(x-\sqrt{3/4}),\ L_2(x) = -\frac{4}{3}x^2+1,\ L_3(x) = \frac{2}{3}x(x+\sqrt{3/4}).$$

The weights are then Then

$$\begin{aligned} w_1 &= \int_{-1}^1 w(x)L_1(x)\,dx = \frac{\pi}{3}, \\ w_2 &= \int_{-1}^1 w(x)L_2(x)\,dx = \frac{\pi}{3}, \\ w_3 &= \int_{-1}^1 w(x)L_3(x)\,dx = \frac{\pi}{3}. \end{aligned}$$

6.11 The Helmholtz decomposition

1. Let Ω be a domain in $\mathbf{R}^3$, and let ϕ, u be a scalar field and a vector field, respectively, defined on Ω. We have

$$\nabla\cdot(\phi u) = \sum_{i=1}^3 \frac{\partial}{\partial x_i}(\phi u_i) = \sum_{i=1}^3\left(\frac{\partial\phi}{\partial x_i}u_i + \phi\frac{\partial u_i}{\partial x_i}\right) = \sum_{i=1}^3 \frac{\partial\phi}{\partial x_i}u_i + \phi\sum_{i=1}^3\frac{\partial u_i}{\partial x_i} = \nabla\phi\cdot u + \phi\nabla\cdot u.$$

3. Let $\phi : \Omega \to \mathbf{R}$ be a smooth scalar field. Then

$$\nabla \cdot \nabla \phi = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(\frac{\partial \phi}{\partial x_i} \right) = \sum_{i=1}^{3} \frac{\partial^2 \phi}{\partial x_i^2} = \frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2}.$$

Chapter 7

The spectral theory of symmetric matrices

7.1 The spectral theorem for symmetric matrices

1. Let $A \in \mathbf{R}^{m\times n}$. Then $(A^TA)^T = A^T(A^T)^T = A^TA$ and, for every $x \in \mathbf{R}^n$,

$$x \cdot A^TAx = (Ax)\cdot(Ax) = \|Ax\|_2^2 \geq 0.$$

Therefore, A^TA is symmetric and positive semidefinite.

5. Suppose $A \in \mathbf{R}^{n\times n}$ satisfies

$$(Ax)\cdot(Ay) = x \cdot y \text{ for all } x, y \in \mathbf{R}^n.$$

Then

$$\begin{aligned}
& x \cdot (A^TAy) = x \cdot y \text{ for all } x, y \in \mathbf{R}^n \\
\Rightarrow\ & A^TAy = y \text{ for all } y \in \mathbf{R}^n \text{ (by Corollary 275)} \\
\Rightarrow\ & A^TA = I.
\end{aligned}$$

The last step follows from the uniqueness of the matrix representing a linear operator on $\mathbf{R}^n$. Since $A^TA = I$, we see that A is orthogonal.

9. Let X be a finite-dimensional inner product space over $\mathbf{R}$ with basis $\mathcal{X} = \{x_1, \ldots, x_n\}$, and assume that $T : X \to X$ is a self-adjoint linear operator ($T^* = T$). Define $A = [T]_{\mathcal{X},\mathcal{X}}$. Let G be the Gram matrix for the basis $\mathcal{X}$, and define $B \in \mathbf{R}^{n\times n}$ by

$$B = G^{1/2}AG^{-1/2},$$

where $G^{1/2}$ is the square root of G (see Exercise 7) and $G^{-1/2}$ is the inverse of $G^{1/2}$.

(a) We wish to prove that B is symmetric. Let $x, y \in X$ be given, and define $\alpha = [x]_{\mathcal{X}}$, $\beta = [y]_{\mathcal{X}}$. Then, since $[T(x)]_{\mathcal{X}} = A[x]_{\mathcal{X}} = A\alpha$, we have $T(x) = \sum_{i=1}^n (A\alpha)_i x_i$. Therefore,

$$\begin{aligned}
\langle T(x), y\rangle = \left\langle \sum_{i=1}^n (A\alpha)_i x_i, \sum_{j=1}^n \beta_j x_j \right\rangle &= \sum_{i=1}^n \sum_{j=1}^n (A\alpha)_i \beta_j \langle x_i, x_j\rangle \\
&= \sum_{i=1}^n (A\alpha)_i \left(\sum_{j=1}^n G_{ij}\beta_j\right) \\
&= (A\alpha)\cdot G\beta = \alpha \cdot (A^TG)\beta.
\end{aligned}$$

Similarly, $[T(y)]_{\mathcal{X}} = A\beta$, and an analogous calculation shows that $\langle x, T(y)\rangle = \alpha \cdot GA\beta$. Since $\langle T(x), y\rangle = \langle x, T(y)\rangle$, it follows that $\alpha \cdot (A^T G)\beta = \alpha \cdot GA\beta$ for all $\alpha, \beta \in \mathbf{R}^n$, and hence $GA = A^T G$. This implies that $(GA)^T = A^T G^T = A^T G = GA$ (using the fact that G is symmetric), and hence GA is symmetric. (Thus the fact that T is self-adjoint does not imply that A is symmetric, but rather than GA is symmetric.)

Now, it is easy to see that $G^{-1/2}$ is symmetric, and if C is symmetric, then so is XCX^T for any square matrix X. Therefore,

$$G^{-1/2}GA(G^{-1/2})^T = G^{-1/2}GAG^{-1/2} = G^{1/2}AG^{-1/2}$$

is symmetric, which is what we wanted to prove.

(b) If λ, u is an eigenvalue/eigenvector pair of B, then

$$Bu = \lambda u \;\Rightarrow\; G^{1/2}AG^{-1/2}u = \lambda u \;\Rightarrow\; AG^{-1/2}u = \lambda G^{-1/2}u.$$

Since $u \neq 0$ and $G^{-1/2}$ is nonsingular, $G^{-1/2}u$ is also nonzero, and hence $\lambda, G^{-1/2}u$ is an eigenpair of A.

(c) Since B is symmetric, there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathbf{R}^n$ consisting of eigenvectors of B. Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. From above, we know that

$$\{G^{-1/2}u_1, \ldots, G^{-1/2}u_n\}$$

is a basis of $\mathbf{R}^n$ consisting of eigenvectors of A. Define $\{y_1, \ldots, y_n\} \subset X$ by $[y_i]_{\mathcal{X}} = G^{-1/2}u_i$, that is,

$$y_i = \sum_{k=1}^n \left(G^{-1/2}u_i\right)_k x_k,\; i = 1, \ldots, n.$$

Notice that

$$\begin{aligned}
\langle y_i, y_j\rangle &= \left\langle \sum_{k=1}^n \left(G^{-1/2}u_i\right)_k x_k, \sum_{\ell=1}^n \left(G^{-1/2}u_j\right)_\ell x_\ell \right\rangle \\
&= \sum_{k=1}^n \sum_{\ell=1}^n \left(G^{-1/2}u_i\right)_k \left(G^{-1/2}u_j\right)_\ell \langle x_k, x_\ell\rangle \\
&= \left(G^{-1/2}u_i\right) \cdot G\left(G^{-1/2}u_j\right) \\
&= u_i \cdot \left(G^{-1/2}GG^{-1/2}\right) u_j \\
&= u_i \cdot u_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}
\end{aligned}$$

This shows that $\{y_1, \ldots, y_n\}$ is an orthonormal basis for X. Also,

$$[T(y_i)]_{\mathcal{X}} = A[y_i]_{\mathcal{X}} = AG^{-1/2}u_i = \lambda_i G^{-1/2}u_i = \lambda_i [y_i]_{\mathcal{X}} = [\lambda_i y_i]_{\mathcal{X}},$$

which proves that $T(y_i) = \lambda_i y_i$. Thus each y_i is an eigenvector of T, and we have proved that there exists an orthonormal basis of X consisting of eigenvectors of T.

7.2 The spectral theorem for normal matrices

1. The matrix A is normal since $A^T A = AA^T = 2I$. We have $A = UDU^*$, where

$$U = \begin{bmatrix} \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix},\; D = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}.$$

7. Let $A \in \mathbf{R}^{n\times n}$ be skew-symmetric.
   (a) Since $A^TA = AA^T = -A^2$, it follows that A is normal
   (b) Hence there exist a unitary matrix $X \in \mathbf{C}^{n\times n}$ and a diagonal matrix $D \in \mathbf{C}^{n\times n}$ such that $A = XDX^*$, or, equivalently, $D = X^*AX$. We then have
   $$D^* = X^*A^*X = X^*(-A)X = -X^*AX = -D.$$
   Therefore, each diagonal entry λ of D, that is, each eigenvalue of A, satisfies $\overline{\lambda} = -\lambda$. This implies that λ is of the form $i\theta$, $\theta \in \mathbf{R}$. Thus a skew-symmetric matrix has only purely imaginary eigenvalues.

11. Suppose $A, B \in \mathbf{C}^{n\times n}$ are normal and commute ($AB = BA$). Then, by Exercise 5.4.18, A and B are simultaneously diagonalizable; that is, there exists a unitary matrix $X \in \mathbf{C}^{n\times n}$ such that $X^*AX = D$ and $X^*BX = C$ are both diagonal matrices in $\mathbf{C}^{n\times n}$. It follows that
$$A + B = XDX^* + XCX^* = X(D + C)X^*,\ (A + B)^* = X(D + C)^*X^*.$$
Since $A + B$ and $(A + B)^*$ are simultaneously diagonalizable, it follows that they commute and hence $A + B$ is normal.

15. Let $A \in F^{m\times n}$ and $B \in F^{n\times p}$, where F represents $\mathbf{R}$ or $\mathbf{C}$. We wish to find a formula for the product AB in terms of outer products of the columns of A and the rows of B. Let $A = [c_1|\cdots|c_n]$,
$$B = \begin{bmatrix} r_1 \\ \hline \vdots \\ \hline r_n \end{bmatrix}.$$
Then
$$AB = \sum_{k=1}^{n} c_k r_k^T.$$
To verify this, notice that $(c_k r_k^T)_{ij} = A_{ik}B_{kj}$, and hence
$$\left(\sum_{k=1}^{n} c_k r_k^T\right)_{ij} = \sum_{k=1}^{n} \left(c_k r_k^T\right)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = (AB)_{ij}.$$
This holds for all i, j, and thus verifies the formula above. It follows that the linear operator defined by AB can be written as
$$\sum_{k=1}^{n} c_k \otimes r_k.$$

7.3 Optimization and the Hessian matrix

1. Suppose $A \in \mathbf{R}^{n\times n}$ and define $A_{sym} = (1/2)\left(A + A^T\right)$. We have
$$(A_{sym})^T = \frac{1}{2}\left(A + A^T\right)^T = \frac{1}{2}\left(A^T + A\right) = \frac{1}{2}\left(A + A^T\right) = A_{sym},$$
and hence A_{sym} is symmetric. Also, for any $x \in \mathbf{R}^n$,
$$\begin{aligned} x \cdot A_{sym}x = x \cdot \left(\frac{1}{2}\left(A + A^T\right)\right)x &= \frac{1}{2}x \cdot \left(Ax + A^Tx\right) \\ &= \frac{1}{2}x \cdot Ax + \frac{1}{2}x \cdot A^Tx \\ &= \frac{1}{2}x \cdot Ax + \frac{1}{2}Ax \cdot x \\ &= \frac{1}{2}x \cdot Ax + \frac{1}{2}x \cdot Ax = x \cdot Ax. \end{aligned}$$

5. The eigenvalues of A are $\lambda = -1$, $\lambda = 3$. Since A is indefinite, q has no global minimizer (or global maximizer).

7. The eigenvalues of A are $\lambda = 0$, $\lambda = 5$, and therefore A is positive semidefinite and singular. The vector b belongs to $\text{col}(A)$, and therefore, by Exercise 2, every vector in $x^* + \mathcal{N}(A)$, where x^* is any solution of $Ax = -b$, is a global minimizer of q. (In other words, every solution of $Ax = -b$ is a global minimizer.) We can take $x^* = (-1, 0)$, and $\mathcal{N}(A) = \text{sp}\{(2, -1)\}$.

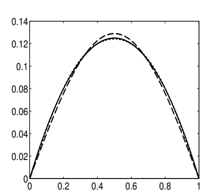
7.4 Lagrange multipliers

3. The maximizer is $x \doteq (-0.058183, 0.73440, -0.67622)$, with Lagrange multiplier $\lambda \doteq (3.1883, -5.7454)$ and $f(x) \doteq 17.923$, while the minimizer is $x \doteq (0.67622, -0.73440, 0.058183)$, with Lagrange multiplier $\lambda \doteq (0.81171, -5.7454)$ and $f(x) \doteq 14.077$.

7. The minimizer, and associated Lagrange multiplier, of $f(x)$ subject to $g(x) = u$ is
$$x(u) = \left(-\frac{\sqrt{1+u}}{\sqrt{3}}, -\frac{\sqrt{1+u}}{\sqrt{3}}, -\frac{\sqrt{1+u}}{\sqrt{3}}\right),\ \lambda(u) = -\frac{\sqrt{3}}{2\sqrt{1+u}}.$$
We have $p(u) = f(x(u)) = -\sqrt{3}\sqrt{1+u}$. Therefore, by direct calculation,
$$\nabla p(u) = -\frac{\sqrt{3}}{2\sqrt{1+u}} \ \Rightarrow\ \nabla p(0) = -\frac{\sqrt{3}}{2}.$$
On the other hand, the Lagrange multiplier associated with $x^* = x(0)$ is $\lambda^* = \lambda(0) = -\sqrt{3}/2$. Thus $\nabla p(0) = \lambda^*$, as implied by the previous exercise.

7.5 Spectral methods for differential equations

1. The exact solution is $u(x) = (x - x^2)/2$, and the solution obtained by the method of Fourier series is
$$u(x) = \sum_{n=1}^{\infty} \frac{2(1-(-1)^n)}{n^3\pi^3} \sin{(n\pi x)}.$$
The following graph shows the exact solution u (the solid curve), together with the partial Fourier series with 1 (the dashed curve) and 4 terms (the dotted curve).



Notice that the partial Fourier series with 4 terms is already indistinguishable from the exact solution on this scale.

3. Consider the operator $M : C_D^2[0, \ell] \to C[0, \ell]$ defined by $M(u) = -u'' + u$.

(a) For any $u,v\in C_D^2[0,\ell]$, we have

$$\begin{aligned}
\langle M(u),v\rangle_2 &= \int_0^1 (-u''(x)+u(x))v(x)\,dx\\
&= -\int_0^1 u''(x)v(x)\,dx + \int_0^1 u(x)v(x)\,dx\\
&= -\left. u'(x)v(x)\right|_0^1 + \int_0^1 u'(x)v'(x)\,dx + \int_0^1 u(x)v(x)\,dx\\
&= \int_0^1 u'(x)v'(x)\,dx + \int_0^1 u(x)v(x)\,dx \text{ (since } v(0)=v(1)=0)\\
&= \left. u(x)v'(x)\right|_0^1 - \int_0^1 u(x)v''(x)\,dx + \int_0^1 u(x)v(x)\,dx\\
&= -\int_0^1 u(x)v''(x)\,dx + \int_0^1 u(x)v(x)\,dx \text{ (since } u(0)=u(1)=0)\\
&= \int_0^1 u(x)(-v''(x)+v(x))\,dx\\
&= \langle u, M(v)\rangle_2\,.
\end{aligned}$$

This shows that M is a symmetric operator. Also, notice from the above calculation that

$$\langle M(u),u\rangle_2 = \int_0^1 (u'(x))^2\,dx + \int_0^1 (u(x))^2\,dx,$$

which shows that $\langle M(u),u\rangle_2 > 0$ for all $u\in C_D^2[0,\ell]$, $u\neq 0$. Then, if λ is an eigenvalue of M and u is a corresponding eigenfunction with $\langle u,u\rangle_2 = 1$, then

$$\lambda = \lambda\langle u,u\rangle_2 = \langle \lambda u,u\rangle_2 = \langle M(u),u\rangle_2 > 0,$$

which shows that all the eigenvalues of M are positive.

(b) It is easy to show that the eigenvalues of M are $n^2\pi^2+1$, $n=1,2,\ldots$, with corresponding eigenfunctions $\sin(n\pi x)$ (the calculation is essentially the same as in Section 7.5.1).

Chapter 8

The singular value decomposition

8.1 Introduction to the SVD

3. The SVD of A is $U\Sigma V^T$, where

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & -1 & 0 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 4\sqrt{3} & 0 & 0 \\ 0 & \sqrt{11} & 0 \\ 0 & 0 & 0 \end{bmatrix}, V = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{11}} & \frac{7}{\sqrt{66}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{11}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{6}} & \frac{3}{\sqrt{11}} & \frac{1}{\sqrt{66}} \end{bmatrix}.$$

The outer product form simplifies to

$$A = 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}\begin{bmatrix} -1 & -1 & 3 \end{bmatrix}.$$

7. If the columns of A are $A_1, \dots, A_n$, $U = \left[\|A_1\|^{-1}A_1|\cdots|\|A_n\|^{-1}A_n\right]$, $V = I$, and Σ is the diagonal matrix with diagonal entries $\|A_1\|, \dots, \|A_n\|$, then $A = U\Sigma V^T$ is the SVD of A.

11. Suppose $A \in \mathbf{C}^{n\times n}$ is invertible and $A = U\Sigma V^*$ is the SVD of A.

    (a) We have $(U\Sigma V^*)^* = V\Sigma^* U^* = V\Sigma U^*$ (since Σ is a real, diagonal matrix), and hence $A^* = V\Sigma U^*$ is the SVD of A.

    (b) Since A is invertible, all the diagonal entries of Σ are positive, and hence Σ^{-1} exists. We have $(U\Sigma V^*)^{-1} = V\Sigma^{-1}U^*$; however, the diagonal entries of Σ^{-1}, $\sigma_1^{-1}, \dots, \sigma_n^{-1}$, are ordered from smallest to largest. We obtain the SVD of A^{-1} by re-ordering. Define $W = [v_n|\cdots|v_1]$, $Z = [u_n|\cdots|u_1]$, and let T be the diagonal matrix with diagonal entries $\sigma_n^{-1}, \dots, \sigma_1^{-1}$. Then $A^{-1} = WTZ^*$ is the SVD of A^{-1}.

    (c) The SVD of A^{-*} is ZTW^*, where W, T, and Z are defined above.

    In outer product form,

$$A^* = \sum_{i=1}^{n} \sigma_i (v_i \otimes u_i), \ A^{-1} = \sum_{i=1}^{n} \sigma_i^{-1}(v_i \otimes u_i), \ A^{-*} = \sum_{i=1}^{n} \sigma_i^{-1}(u_i \otimes v_i).$$

8.2 The SVD for general matrices

3. We have

$$\text{proj}_{\text{col}(A)} b = (1, 5/2, 5/2, 4)\,, \ \text{proj}_{\mathcal{N}(A^T)} b = (0, -1/2, 1/2, 0)\,.$$

5. Referring to the solution of Exercise 8.1.4, if $U = [u_1|u_2|u_3|u_4]$, $V = [v_1|v_2|v_3|v_4]$, then $\{u_1, u_2, u_3\}$ is a basis for $\mathrm{col}(A)$, $\{u_4\}$ is a basis for $\mathcal{N}(A^T)$, $\{v_1, v_2, v_3\}$ is a basis for $\mathrm{col}(A^T)$, and $\{v_4\}$ is a basis for $\mathcal{N}(A)$.

9. Let $A \in \mathbf{R}^{m\times n}$ be nonsingular. We wish to compute $\min\{\|Ax\|_2 : x \in \mathbf{R}^n, \|x\|_2 = 1\}$. Note that $\|Ax\|_2 = \sqrt{x \cdot A^T Ax}$. By Exercise 7.4.5, the minimum value of $x \cdot A^T Ax$, where $\|x\|_2 = 1$, is the smallest eigenvalue of $A^T A$, which is σ_n^2. Therefore, $\min\{\|Ax\|_2 : x \in \mathbf{R}^n, \|x\|_2 = 1\} = \sigma_n$, and the value of x yielding the minimum is v_n, the right singular vector corresponding to σ_n, the smallest singular value.

15. (a) Let $U \in \mathbf{C}^{m\times m}$, $V \in \mathbf{C}^{n\times n}$ be unitary. We wish to prove that $\|UA\|_F = \|A\|_F$ and $\|AV\|_F = \|A\|_F$ for all $A \in \mathbf{C}^{m\times n}$. We begin with two preliminary observations. By definition of the Frobenius norm, for any $A \in \mathbf{C}^{m\times n}$,

$$\|A\|_F^2 = \sum_{j=1}^{n}\left(\sum_{i=1}^{m}|A_{ij}|^2\right) = \sum_{j=1}^{n}\|A_j\|_2^2,$$

where A_j is the jth column of A. Also, it is obvious that $\|A^T\|_F = \|A\|_F$ for all $A \in \mathbf{C}^{m\times n}$. We thus have

$$\|UA\|_F^2 = \sum_{j=1}^{n}\|(UA)_j\|_F^2 = \sum_{j=1}^{n}\|UA_j\|_F^2 = \sum_{j=1}^{n}\|A_j\|_F^2 = \|A\|_F^2$$

and

$$\|AV\|_F = \|V^T A^T\|_F = \|A^T\|_F = \|A\|_F,$$

as desired.

(b) Let $A \in \mathbf{C}^{m\times n}$ be given, and let r be a positive integer with $r < \mathrm{rank}(A)$. We wish to find the matrix $B \in \mathbf{C}^{m\times n}$ of rank r such that $\|A - B\|_F$ is as small as possible. If we define $B = U\Sigma_r V^T$, where $\Sigma_r \in \mathbf{R}^{m\times n}$ is the diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_r, 0, \ldots, 0$, then

$$\begin{aligned}\|A - B\|_F = \|U\Sigma V^* - U\Sigma_r V^*\|_F = \|U(\Sigma - \Sigma_r)V^*\|_F = \|(\Sigma - \Sigma_r)V^*\|_F = \|\Sigma - \Sigma_r\|_F \\ = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_t^w},\end{aligned}$$

where $t \le \min\{m, n\}$ is the rank of A. Notice that the rank of a matrix is the number of positive singular values, so $\mathrm{rank}(B) = r$, as desired. Thus we can make $\|A-B\|_F$ as small as $\sqrt{\sigma_{r+1}^2 + \cdots + \sigma_t^w}$. Moreover, for any $B \in \mathbf{C}^{m\times n}$, we have

$$\begin{aligned}\|A - B\|_F^2 &= \|\Sigma - U^*BV\|_F^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\Sigma_{ij} - (U^*BV)_{ij}\right)^2 \\ &= \sum_{i=1}^{t}(\sigma_i - (U^*BV)_{ii})^2 + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{t}(U^*BV)_{ij}^2 + \sum_{i=1}^{m}\sum_{j=t+1}^{n}(U^*BV)_{ij}^2.\end{aligned}$$

Now, we are free to choose all the entries of U^*BV (since U^* and V are invertible, given any $C \in \mathbf{C}^{m\times n}$, there exists a unique $B \in \mathbf{C}^{m\times n}$ with $U^*BV = C$) to make the above sum as small as possible. Since all three summations are nonnegative, we should choose $(U^*BV)_{ij} = 0$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$, $j \neq i$ for $i = 1, \ldots, t$. This causes the second two summations to vanish, and yields

$$\|A - B\|_F^2 = \sum_{i=1}^{t}(\sigma_i - (U^*BV)_{ii})^2.$$

The rank of U^*BV (and hence the rank of B) is the number of nonzero diagonal entries

$$(U^*BV)_{11}, \ldots, (U^*BV)_{tt}.$$

Since the rank of B must be r, it is clear that $U^*BV = \Sigma_r$, where Σ_r is defined above, will make $\|A - B\|_F$ as small as possible. This shows that $B = U\Sigma_r V^*$ is the desired matrix.

8.3 Solving least-squares problems using the SVD

3. The matrix A has two positive singular values, $\sigma_1 = 6$ and $\sigma_2 = 2$. The corresponding right singular vectors are

$$v_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad v_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

and the left singular vectors are

$$u_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad u_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

The minimum-norm least-squares solution to $Ax = b$ is then given by

$$x = \frac{u_1 \cdot b}{\sigma_1} v_1 + \frac{u_2 \cdot b}{\sigma_2} v_2.$$

(a) $x = (2/3, -1/3, 1/12)$;

(b) $x = (13/18, -5/18, 1/9)$;

(c) $x = (0, 0, 0)$ (b is orthogonal to $\mathrm{col}(A)$).

7. Let $A \in \mathbf{R}^{m\times n}$ have rank r, and let $\sigma_1, \ldots, \sigma_r$ be the positive singular values of A, with corresponding right singular vectors $v_1, \ldots, v_r \in \mathbf{R}^n$ and left singular vectors $u_1, \ldots, u_r \in \mathbf{R}^m$. Then, for all $x \in \mathbf{R}^n$,

$$Ax = \sum_{i=1}^{r} \sigma_i (v_i \cdot x) u_i$$

and, for all $b \in \mathbf{R}^m$,

$$A^\dagger b = \sum_{i=1}^{r} \frac{u_i \cdot b}{\sigma_i} v_i.$$

8.4 The SVD and linear inverse problems

1. (a) Let $A \in \mathbf{R}^{m\times n}$ be given, let $I \in \mathbf{R}^{n\times n}$ be the identity matrix, and let ϵ be a positive number. For any $\hat{b} \in \mathbf{R}^m$, we can solve the equation

$$\left[\begin{array}{c} A \\ \hline \epsilon I \end{array}\right] x = \left[\begin{array}{c} \hat{b} \\ \hline 0 \end{array}\right] \tag{8.1}$$

in the least-square sense, which is equivalent to minimizing

$$\left\| \left[\begin{array}{c} A \\ \hline \epsilon I \end{array}\right] x - \left[\begin{array}{c} \hat{b} \\ \hline 0 \end{array}\right] \right\|_2^2 = \left\| \left[\begin{array}{c} Ax - \hat{b} \\ \hline \epsilon x \end{array}\right] \right\|_2^2.$$

Now, for any Euclidean vector $w \in \mathbf{R}^k$, partitioned as $w = (u, v)$, $u \in \mathbf{R}^p$, $v \in \mathbf{R}^q$, $p + q = k$, we have $\|w\|_2^2 = \|u\|_2^2 + \|v\|_2^2$, and hence

$$\left\| \left[\frac{Ax - \hat{b}}{\epsilon x} \right] \right\|_2^2 = \|Ax - \hat{b}\|_2^2 + \|\epsilon x\|_2^2 = \|Ax - \hat{b}\|_2^2 + \epsilon^2 \|x\|_2^2.$$

Thus solving (8.1) in the least-squares sense is equivalent to choosing $x \in \mathbf{R}^n$ to minimize $\|Ax - \hat{b}\|_2^2 + \epsilon^2 \|x\|_2^2$.

(b) We have

$$\left[\frac{A}{\epsilon I} \right]^T = \left[\begin{array}{c|c} A^T & \epsilon I \end{array} \right],$$

which implies that

$$\left[\frac{A}{\epsilon I} \right]^T \left[\frac{A}{\epsilon I} \right] = A^T A + \epsilon^2 I, \ \left[\frac{A}{\epsilon I} \right]^T \left[\frac{\hat{b}}{0} \right] = A^T \hat{b}.$$

Hence the normal equations for (8.1) take the form $(A^T A + \epsilon^2 I)x = A^T \hat{b}$.

(c) We have

$$\begin{bmatrix} A \\ \epsilon I \end{bmatrix} x = 0 \ \Rightarrow \ \begin{bmatrix} Ax \\ \epsilon x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which yields $\epsilon x = 0$, that is, $x = 0$. Thus the matrix is nonsingular and hence, by the fundamental theorem, it has full rank. It follows from Exercise 6.4.1 that $A^T A + \epsilon^2 I$ is invertible, and hence there is a unique solution x_ϵ to $(A^T A + \epsilon^2 I)x = A^T \hat{b}$.

3. With the errors drawn from a normal distribution with mean zero and standard deviation 10^{-4}:

- truncated SVD works best with $k = 3$ singular values/vectors;
- Tikhonov regularization works best with ϵ around 10^{-3}.

8.5 The Smith normal form of a matrix

1. We have $A = USV$, where

$$U = \begin{bmatrix} 4 & 2 & -1 \\ 6 & 1 & 0 \\ 3 & 2 & -1 \end{bmatrix}, \ S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{bmatrix}, \ V = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

3. We have $A = USV$, where

$$U = \begin{bmatrix} 4 & 0 & 1 \\ 5 & 0 & 1 \\ 7 & -1 & 0 \end{bmatrix}, \ S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ V = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 7 \\ 0 & 0 & 1 \end{bmatrix}.$$

Chapter 9

Matrix factorizations and numerical linear algebra

9.1 The LU factorization

1.
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -1 & -4 & 1 & 0 \\ -2 & 0 & 5 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

7. Let
$$L = \begin{bmatrix} 1 & 0 \\ \ell & 1 \end{bmatrix}, \; U = \begin{bmatrix} u & v \\ 0 & w \end{bmatrix},$$
and notice that
$$LU = \begin{bmatrix} u & v \\ \ell u & \ell v + w \end{bmatrix}.$$

(a) We wish to show that there do not exist matrices L, U of the above forms such that $LU = A$, where
$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$
This is straightforward, since $LU = A$ implies $u = 0$ (comparing the $1,1$ entries) and also $\ell u = 1$ (comparing the $2,1$ entries). No choice of ℓ, u, v, w can make both of these true, and hence there do not exist such L and U.

(b) If
$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$
then $LU = A$ is equivalent to $u = 0$, $v = 1$, and $\ell + w = 1$. There are infinitely many choices of ℓ and w that will work, and hence there exist infinitely many L, U satisfying $LU = A$.

11. Computing A^{-1} is equivalent to solving $Ax = e_j$ for $j = 1, 2, \dots, n$. If we compute the LU factorization and then solve the n systems $LU = e_j$, $j = 1, \dots, n$, the operation count is
$$\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{1}{6}n + n\left(2n^2 - n\right) = \frac{8}{3}n^3 + \frac{1}{2}n^2 - \frac{7}{6}n.$$

We can reduce the above operation count by taking advantage of the zeros in the vectors e_j. Solving $Lc = e_j$ takes the following form:

$$\begin{aligned}
c_i &= 0,\ i = 1, \ldots, j-1, \\
c_j &= 1, \\
c_i &= -\sum_{k=j}^{i-1} L_{ik} c_k,\ i = j+1, \ldots, n.
\end{aligned}$$

Thus solving $Lc - e_j$ requires $\sum_{i=j+1}^{n} 2(i-j) = (n-j)^2 + (n-j)$ operations. The total for solving all n of the lower triangular systems $Lc = e_j$, $j = 1, \ldots, n$, is

$$\sum_{j=1}^{n} \left\{ (n-j)^2 + (n-j) \right\} = \frac{1}{3}n^3 - \frac{1}{3}n$$

(instead of $n(n^2 - n) = n^3 - n$ if we ignore the structure of the right-hand side). We still need n^3 operations to solve the n upper triangular systems $UA_j = L^{-1}e_j$ (since we perform back substition, there is no simplification from the fact that the first $j-1$ entries in $L^{-1}e_j$ are zero). Hence the total is

$$\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{1}{6}n + \frac{1}{3}n^3 - \frac{1}{3}n + n^3 = 2n^3 + \frac{3}{2}n^2 - \frac{3}{2}n.$$

Notice the reduction in the leading term from $(8/3)n^3$ to $2n^3$.

9.2 Partial pivoting

1. The solution is $x = (2, 1, -1)$; partial pivoting requires interchanging rows 1 and 2 on the first step, and interchanging rows 2 and 3 on the second step.

5. Suppose $A \in \mathbf{R}^{n \times n}$ has an LU decomposition, $A = LU$. We know that the determinant of a square matrix is the product of the eigenvalues, and also that the determinant of a triangular matrix is the product of the diagonal entries. We therefore have $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A (listed according to multiplicity), and also

$$\det(A) = \det(LU) = \det(L)\det(U) = (1\,1 \cdots 1)(U_{11}U_{22}\cdots U_{nn}) = U_{11}U_{22}\cdots U_{nn}.$$

This shows that $\lambda_1 \lambda_2 \cdots \lambda_n = U_{11}U_{22}\cdots U_{nn}$.

9. On the first step, partial pivoting requires that rows 1 and 3 be interchanged. No interchange is necessary on step 2, and rows 3 and 4 must be interchanged on step 3. Thus

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The LU factorization of PA is $PA = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.25 & -0.5 & 1 & 0 \\ -0.25 & -0.1 & 0.25 & 1 \end{bmatrix},\quad U = \begin{bmatrix} 6 & 2 & -1 & 4 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

9.3 The Cholesky factorization

1. (a) The Cholesky factorization is $A = R^T R$, where

$$R = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 2 \end{bmatrix}.$$

(b) We also have $A = LDL^T$, where $L = (D^{-1/2}R)^T$ and the diagonal entries of D are the diagonal entries of R:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \; L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & \sqrt{2} & 0 \\ 2 & -\sqrt{2} & \sqrt{2} \end{bmatrix}.$$

Alternatively, we can write $A = LU$, where $U = DR$ and $L = (D^{-1}R)^T$:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}, \; U = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 4 & -4 \\ 0 & 0 & 4 \end{bmatrix}.$$

5. Let $A \in \mathbf{R}^{n\times n}$ be SPD. The algorithm described on pages 527–528 of the text (that is, the equations derived on those pages) shows that there is a unique upper triangular matrix R with positive diagonal entries such that $R^T R = A$. (The only freedom in solving those equations for the entries of R lies in choosing the positive or negative square root when computing R_{ii}. If R_{ii} is constrained to be positive, then the entries of R are uniquely determined.)

9.4 Matrix norms

1. Let $\|\cdot\|$ be any induced matrix norm on $\mathbf{R}^{n\times n}$. If $\lambda \in \mathbf{C}^{n\times n}$, $x \in \mathbf{C}^n$, $x \neq 0$, is an eigenvalue/eigenvector pair of A, then

$$\|Ax\| \le \|A\|\|x\| \;\Rightarrow\; \|\lambda x\| \le \|A\|\|x\| \;\Rightarrow\; |\lambda|\|x\| \le \|A\|\|x\| \;\Rightarrow\; |\lambda| \le \|A\|$$

(the last step follows from the fact that $x \neq 0$). Since this holds for every eigenvalue of A, it follows that

$$\rho(A) = \max\{|\lambda| \,:\, \lambda \text{ is an eigenvalue of } A\} \le \|A\|.$$

7. Let $A \in \mathbf{R}^{m\times n}$. We wish to prove that $\|A^T\|_2 = \|A\|_2$. This follows immediately from Theorem 403 and Exercise 4.5.14. Theorem 403 implies that $\|A\|_2 = \sqrt{\lambda_{max}(A^TA)}$, $\|A^T\|_2 = \sqrt{\lambda_{max}(AA^T)}$. By Exercise 4.5.14, A^TA and AA^T have the same nonzero eigenvalues, and hence $\lambda_{max}(A^TA) = \lambda_{max}(AA^T)$, from which it follows that $\|A\|_2 = \|A^T\|_2$.

9.5 The sensitivity of linear systems to errors

1. Let $A \in \mathbf{R}^{n\times n}$ be nonsingular.

(a) Suppose $b \in \mathbf{R}^n$ is given. Choose $c \in \mathbf{R}^n$ such that

$$\frac{\|A^{-1}c\|}{\|c\|} = \sup\left\{\frac{\|A^{-1}x\|}{\|x\|} \,:\, x \in \mathbf{R}^n, x \neq 0\right\} = \|A^{-1}\|,$$

and notice that $\|A^{-1}c\| = \|A^{-1}\|\|c\|$. (Such a c exists. For the common norms—the ℓ^1, ℓ^∞, and Euclidean norms—we have seen in the text how to compute such a c; for an arbitrary norm, it can be shown that such a c exists, although some results from analysis concerning continuity and compactness are needed.) Define $\hat{b} = b + c$ and $x = A^{-1}b$, $\hat{x} = A^{-1}(b+c)$. Then

$$\|\hat{x} - x\| = \|A^{-1}(b+c) - A^{-1}b\| = \|A^{-1}c\| = \|A^{-1}\|\|c\| = \|A^{-1}\|\|\hat{b} - b\|.$$

Notice that, for the Euclidean norm $\|\cdot\|_2$ and induced matrix norm, c should be chosen to be a right singular vector corresponding to the smallest singular value of A.

(b) Let $A \in \mathbf{R}^{n\times n}$ be given, and let $x, c \in \mathbf{R}^n$ be chosen so that

$$\begin{aligned}\frac{\|Ax\|}{\|x\|} &= \sup\left\{\frac{\|Ay\|}{\|y\|} \,:\, y \in \mathbf{R}^n, y \neq 0\right\} = \|A\|,\\ \frac{\|A^{-1}c\|}{\|c\|} &= \sup\left\{\frac{\|A^{-1}y\|}{\|y\|} \,:\, y \in \mathbf{R}^n, y \neq 0\right\} = \|A^{-1}\|.\end{aligned}$$

Define $b = Ax$, $\hat{b} = b + c$, and $\hat{x} = A^{-1}(b+c) = x + A^{-1}c$. We then have

$$\|b\| = \|Ax\| = \|A\|\|x\| \;\Rightarrow\; \|x\| = \frac{\|b\|}{\|A\|}$$

and

$$\|\hat{x} - x\| = \|A^{-1}c\| = \|A^{-1}\|\|c\| = \|A^{-1}\|\|\hat{b} - b\|.$$

It follows that

$$\frac{\|\hat{x} - x\|}{\|x\|} = \frac{\|A^{-1}\|\|\hat{b} - b\|}{\frac{\|b\|}{\|A\|}} = \|A\|\|A^{-1}\|\frac{\|\hat{b} - b\|}{\|b\|}.$$

5. Let $A \in \mathbf{R}^{n\times n}$ be invertible, and let $\|\cdot\|$ denote any norm on $\mathbf{R}^n$ and the corresponding induced matrix norm.

(a) Let $B \in \mathbf{R}^{n\times n}$ be any singular matrix. We have seen (Exercise 9.4.9) that $\|Ax\| \geq \|x\|/\|A^{-1}\|$ for all $x \in \mathbf{R}^n$. Let $x \in \mathcal{N}(B)$ with $\|x\| = 1$. Then

$$\|A - B\| \geq \|(A - B)x\| = \|Ax\| \geq \frac{\|x\|}{\|A^{-1}\|} = \frac{1}{\|A^{-1}\|}.$$

(b) It follows that

$$\begin{aligned}&\inf\left\{\frac{\|A - B\|}{\|A\|} \,:\, B \in \mathbf{R}^{n\times n}, \det(B) = 0\right\}\\ &\geq \inf\left\{\frac{1/\|A^{-1}\|}{\|A\|} \,:\, B \in \mathbf{R}^{n\times n}, \det(B) = 0\right\}\\ &= \frac{1}{\mathrm{cond}(A)}.\end{aligned}$$

(c) Consider the special case of the Euclidean norm on $\mathbf{R}^n$ and induced norm $\|\cdot\|_2$ on $\mathbf{R}^{n\times n}$. Let $A = U\Sigma V^T$ be the SVD of A, and define $A' = U\Sigma' V^T$, where Σ' is the diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_{n-1}, 0$ ($\sigma_1 \geq \cdots \sigma_{n-1} \geq \sigma_n > 0$ are the singular values of A). Then

$$\begin{aligned}\|A - A'\|_2 = \|U\Sigma V^T - U\Sigma' V^T\|_2 &= \|U(\Sigma - \Sigma')V^T\|_2\\ &= \|\Sigma - \Sigma'\|_2\\ &= \sigma_n = \frac{1}{\|A^{-1}\|}\end{aligned}$$

($\|\Sigma - \Sigma'\|_2$ is the largest singular value of $\Sigma - \Sigma'$, which is a diagonal matrix with a single nonzero entry, σ_n). It follows that

$$\frac{\|A - A'\|_2}{\|A\|_2} = \frac{1}{\|A\|_2\|A^{-1}\|_2} = \frac{1}{\mathrm{cond}_2(A)}.$$

Hence the inequality derived in part (b) is an equality in this case.

9.6 Numerical stability

1. (a) Suppose x and y are two real numbers and $\hat{x}$ and $\hat{y}$ are perturbations of x and y, respectively. We have

$$\hat{x}\hat{y} - xy = \hat{x}\hat{y} - \hat{x}y + \hat{x}y - xy = \hat{x}(\hat{y} - y) + (\hat{x} - x)y,$$

which implies that

$$|\hat{x}\hat{y} - xy| \le |\hat{x}||\hat{y} - y| + |\hat{x} - x||y|.$$

This gives a bound on the absolute error in approximating xy by $\hat{x}\hat{y}$. Dividing by xy yields

$$\frac{|\hat{x}\hat{y} - xy|}{|xy|} \le \frac{|\hat{x}|}{|x|}\frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|}.$$

This yields a bound on the relative error in $\hat{x}\hat{y}$ in terms of the relative errors in $\hat{x}$ and $\hat{y}$, although it would be preferable if the bound did not contain $\hat{x}$ (except in the expression for the relative error in $\hat{x}$). We can manipulate the bound as follows:

$$\begin{aligned}
\frac{|\hat{x}\hat{y} - xy|}{|xy|} \le \frac{|\hat{x}|}{|x|}\frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|} &= \frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|} + \left(\frac{|\hat{x}|}{|x|} - 1\right)\frac{|\hat{y} - y|}{|y|} \\
&\le \frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|} + \left|\frac{|\hat{x}|}{|x|} - 1\right|\frac{|\hat{y} - y|}{|y|} \\
&= \frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|} + \frac{||\hat{x}| - |x||}{|x|}\frac{|\hat{y} - y|}{|y|} \\
&\le \frac{|\hat{y} - y|}{|y|} + \frac{|\hat{x} - x|}{|x|} + \frac{|\hat{x} - x|}{|x|}\frac{|\hat{y} - y|}{|y|}.
\end{aligned}$$

When the relative errors in $\hat{x}$ and $\hat{y}$ are small (that is, much less than 1), then their product is much smaller, and we see that the relative error in $\hat{x}\hat{y}$ as an approximation to xy is approximately bounded by the sum of the errors in $\hat{x}$ and $\hat{y}$.

(b) If x and y are floating point numbers, then $\mathrm{fl}(xy) = xy(1+\epsilon)$, where $|\epsilon| \le \mathbf{u}$. Therefore, $\mathrm{fl}(xy)$ is the exact product of $\tilde{x}$ and $\tilde{y}$, where $\tilde{x} = x$ and $\tilde{y} = y(1+\epsilon)$. This shows that the computed product is the exact product of nearby numbers, and therefore that floating point multiplication is backward stable.

9.7 The sensitivity of the least-squares problem

3. Let $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$ be given, and let x be a least-squares solution of $Ax = b$. We have

$$b = Ax + (b - Ax) \ \Rightarrow\ b\cdot Ax = (Ax)\cdot(Ax) + (b - Ax)\cdot Ax = \|Ax\|_2^2,$$

since $(b - Ax)\cdot Ax = 0$ ($b - Ax$ is orthogonal to $\mathrm{col}(A)$). It follows that

$$\|b\|_2\|Ax\|_2\cos(\theta) = \|Ax\|_2^2 \ \Rightarrow\ \|Ax\|_2 = \|b\|_2\cos(\theta),$$

where θ is the angle between Ax and b. Also, since Ax, $b - Ax$ are orthogonal, the Pythagorean theorem implies

$$\|b\|_2^2 = \|Ax\|_2^2 + \|b - Ax\|_2^2.$$

Dividing both sides by $\|b\|_2^2$ yields

$$1 = \frac{\|Ax\|_2^2}{\|b\|_2^2} + \frac{\|b - Ax\|_2^2}{\|b\|_2^2} \ \Rightarrow\ 1 = \cos^2(\theta) + \frac{\|b - Ax\|_2^2}{\|b\|_2^2}.$$

Therefore,

$$\frac{\|b - Ax\|_2^2}{\|b\|_2^2} = \sin^2(\theta) \ \Rightarrow\ \|b - Ax\|_2 = \|b\|_2\sin(\theta).$$

9.8 The QR factorization

1. Let $x, y \in \mathbf{R}^3$ be defined by $x = (1, 2, 1)$ and $y = (2, 1, 1)$. We define

$$u = \frac{x - y}{\|x - y\|_2} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),\ Q = I - 2uu^T = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then $Qx = y$.

3. We have $A = QR$, where

$$Q = \begin{bmatrix} -0.42640 & 0.65970 & 0.61885 \\ -0.63960 & -0.70368 & 0.30943 \\ 0.63960 & -0.26388 & 0.72199 \end{bmatrix},$$
$$R = \begin{bmatrix} -4.6904 & -3.8376 & 2.7716 \\ 0 & 2.0671 & -2.1110 \\ 0 & 0 & 9.2828 \end{bmatrix}$$

(correct to the digits shown). The Householder vectors are

$$u_1 = (0.84451, 0.37868, -0.37868),\ u_2 = (-0.99987, 0.015968).$$

9.9 Eigenvalues and simultaneous iteration

1. We apply n iterations of the power method, normalizing the approximate eigenvectors in the Euclidean norm at each iteration, and estimating the eigenvalue by $\lambda = (x \cdot Ax)/(x \cdot x) = x \cdot Ax$. We use a random starting vector. With $n = 10$, we have

$$\begin{aligned} x_0 &= (0.37138, -0.22558, 1.1174), \\ x_{10} &= (0.40779, -0.8165, 0.40871), \\ \lambda &\doteq 3.9999991538580488. \end{aligned}$$

With $n = 20$, we obtain

$$\begin{aligned} x_0 &= (-1.0891, 0.032557, 0.55253), \\ x_{20} &= (-0.40825, 0.8165, -0.40825), \\ \lambda &\doteq 3.9999999999593752. \end{aligned}$$

It seems clear that the dominant eigenvalue is $\lambda = 4$.

5. Let $A \in \mathbf{R}^{n\times n}$. We wish to prove that there exists an orthogonal matrix Q such that Q^TAQ is block upper triangular, with each diagonal block of size 1×1 or 2×2. If all the eigenvalues of A are real, then the proof of Theorem 413 can be given using only real numbers and vectors, and the result is immediate. Therefore, let $\lambda, \overline{\lambda}$ be a complex conjugate pair of eigenvalues of A. By an induction argument similar to that in the proof of Theorem 413, it suffices to prove that there exists an orthogonal matrix $\hat{Q} \in \mathbf{R}^{n\times n}$ such that

$$\hat{Q}^T A\hat{Q} = \begin{bmatrix} T & B \\ 0 & C \end{bmatrix},$$

where $T \in \mathbf{R}^{2\times 2}$ has eigenvalues $\lambda, \overline{\lambda}$. Suppose $z = x + iy$, $z \neq 0$, $x, y \in \mathbf{R}^n$, satisfies $Az = \lambda z$. In the solution of the previous exercise, we saw that $\{x, y\}$ is linearly independent, so define $S = \mathrm{sp}\{x, y\} \subset \mathbf{R}^n$ and $\hat{S} = \mathrm{sp}\{x, y\} \subset \mathbf{C}^n$ ($\hat{S} = \mathrm{sp}\{z, \overline{z}\}$). Let $\{q_1, q_2\}$ be an orthonormal basis for S, and extend it to an orthonormal basis $\{q_1, q_2, \ldots, q_n\}$ for $\mathbf{R}^n$. Define $\hat{Q}_1 = [q_1|q_2]$, $\hat{Q}_2 = [q_3|\cdots|q_n]$, and $\hat{Q} = [\hat{Q}_1|\hat{Q}_2]$. We then have

$$\hat{Q}^T A\hat{Q} = \left[\begin{array}{c} \hat{Q}_1^T \\ \hline \hat{Q}_2^T \end{array}\right] [A\hat{Q}_1|A\hat{Q}_2] = \left[\begin{array}{c|c} \hat{Q}_1^T A\hat{Q}_1 & \hat{Q}_1^T A\hat{Q}_2 \\ \hline \hat{Q}_2^T A\hat{Q}_1 & \hat{Q}_2^T A\hat{Q}_2 \end{array}\right].$$

Since both columns of $A\hat{Q}_1$ belong to S and each column of $\hat{Q}_2$ belongs to $S^\perp$, it follows that $\hat{Q}_2^T A\hat{Q}_1 = 0$. Thus it remains only to prove that the eigenvalues of $T = \hat{Q}_1^T A\hat{Q}_1$ are $\lambda, \overline{\lambda}$. We see that $\eta \in \mathbf{C}$, $u \in \mathbf{C}^2$ form an eigenpair of T if and only if

$$Tu = \eta u \;\Leftrightarrow\; \hat{Q}_1^T A\hat{Q}_1 u = \eta u \;\Leftrightarrow\; A(\hat{Q}_1 u) = \eta(\hat{Q}_1 u).$$

Since both z and $\overline{z}$ can be written as $\hat{Q}_1 u$ for $u \in \mathbf{C}^2$, it follows that both λ and $\overline{\lambda}$ are eigenvalues of T; moreover, since T is 2×2, these are the only eigenvalues of T. This completes the proof.

9.10 The QR algorithm

1. Two steps are required to reduce A to upper Hessenberg form. The result is

$$H = \begin{bmatrix} -2.0000 & 4.8507 \cdot 10^{-1} & -1.1021 \cdot 10^{-1} & 8.6750 \cdot 10^{-1} \\ -4.1231 & 3.1765 & 1.3851 & -7.6145 \cdot 10^{-1} \\ 0 & -1.4240 & 1.9481 & 1.6597 \cdot 10^{-1} \\ 0 & 0 & 8.9357 \cdot 10^{-1} & -3.1246 \end{bmatrix}$$

and the vectors defining the two Householder transformations are

$$u_1 = (8.6171 \cdot 10^{-1}, -2.8146 \cdot 10^{-1}, 4.2219 \cdot 10^{-1}),$$
$$u_2 = (9.3689 \cdot 10^{-1}, 3.4964 \cdot 10^{-1}).$$

5. The inequality

$$\frac{|\lambda_{k+1} - \mu|}{|\lambda_k - \mu|} < \frac{|\lambda_{k+1}|}{|\lambda_k|} \tag{9.1}$$

is equivalent to

$$\frac{|\lambda_{k+1} - \mu|}{|\lambda_{k+1}|} < \frac{|\lambda_k - \mu|}{|\lambda_k|},$$

and hence (9.1) holds if and only if the relative error in μ as an estimate of λ_{k+1} is less than the relative error in μ as an estimate of λ_k.

Chapter 10

Analysis in vector spaces

10.1 Analysis in $\mathbf{R}^n$

3. Let $\|\cdot\|$ and $\|\cdot\|_*$ be two norms on $\mathbf{R}^n$. Since $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent, there exist positive constants c_1, c_2 such that $c_1\|x\| \le \|x\|_* \le c_2\|x\|$ for all $x \in \mathbf{R}^n$. Suppose that $x_k \to x$ under $\|\cdot\|$, and let $\epsilon > 0$. Then there exists a positive integer N such that $\|x_k - x\| < \epsilon/c_2$ for all $k \ge N$. It follows that

$$\|x_k - x\|_* \le c_2\|x_k - x\| < c_2\frac{\epsilon}{c_2} = \epsilon$$

for all $k \ge N$. Therefore, $x_k \to x$ under $\|\cdot\|_*$.

Conversely, if $x_k \to x$ under $\|\cdot\|_*$ and $\epsilon > 0$ is given, there exists a positive integer N such that $\|x_k - x\|_* < c_1\epsilon$ for all $k \ge N$. It follows that $\|x_k - x\| \le c_1^{-1}\|x_k - x\|_* < c_1^{-1}c_1\epsilon = \epsilon$ for all $k \ge N$. Therefore, $x_k \to x$ under $\|\cdot\|$.

7. Let $\|\cdot\|$ and $\|\cdot\|_*$ be two norms on $\mathbf{R}^n$, let S be a nonempty subset of $\mathbf{R}^n$, let $f : S \to \mathbf{R}^n$ be a function, and let y be an accumulation point of S. Since $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent, there exist positive constants c_1, c_2 such that $c_1\|x\| \le \|x\|_* \le c_2\|x\|$ for all $x \in \mathbf{R}^n$. Suppose first that $\lim_{x\to y} f(x) = L$ under $\|\cdot\|$, and let $\epsilon > 0$ be given. Then there exists $\delta > 0$ such that if $x \in S$ and $\|x - y\| < \delta$, then $|f(x) - L| < \epsilon$. But then $\|x - y\|_* < c_1\delta \Rightarrow \|y - x\| \le c_1^{-1}\|x - y\| < c^{-1}c_1\delta = \delta$. Therefore, if $x \in S$ and $\|x - y\|_* < c_1\delta$, it follows that $\|x - y\| < \delta$, and hence that $|f(x) - L| < \epsilon$. This shows that $\lim_{x\to y} f(x) = L$ under $\|\cdot\|_*$.

Conversely, suppose that $\lim_{x\to y} f(x) = L$ under $\|\cdot\|_*$, and let $\epsilon > 0$ be given. Then there exists $\delta > 0$ such that if $x \in S$ and $\|x - y\|_* < \delta$, then $|f(x) - L| < \epsilon$. But then

$$\|x - y\| < c_2^{-1}\delta \ \Rightarrow\ \|y - x\|_* \le c_2\|x - y\| < c_2c_2^{-1}\delta = \delta.$$

Therefore, if $x \in S$ and $\|x - y\| < c_2^{-1}\delta$, it follows that $\|x - y\|_* < \delta$, and hence that $|f(x) - L| < \epsilon$. This shows that $\lim_{x\to y} f(x) = L$ under $\|\cdot\|$.

11. Let $\|\cdot\|$ and $\|\cdot\|_*$ be two norms on $\mathbf{R}^n$, and let $\{x_k\}$ be a sequence in $\mathbf{R}^n$. Since $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent, there exist positive constants c_1, c_2 such that $c_1\|x\| \le \|x\|_* \le c_2\|x\|$ for all $x \in \mathbf{R}^n$. Suppose first that $\{x_k\}$ is Cauchy under $\|\cdot\|$, and let $\epsilon > 0$ be given. Then there exists a positive integer N such that $m, n \ge N$ implies that $\|x_m - x_n\| < c_2^{-1}\epsilon$. But then $m, n \ge N$ implies that

$$\|x_m - x_n\|_* \le c_2\|x_m - x_n\| < c_2c_2^{-1}\epsilon = \epsilon,$$

and hence $\{x_k\}$ is Cauchy under $\|\cdot\|_*$.

Conversely, suppose $\{x_k\}$ is Cauchy under $\|\cdot\|_*$, and let $\epsilon > 0$ be given. Then there exists a positive integer N such that $m, n \ge N$ implies that $\|x_m - x_n\|_* < c_1\epsilon$. But then $m, n \ge N$ implies that

$$\|x_m - x_n\| \le c_1^{-1}\|x_m - x_n\|_* < c_1^{-1}c_1\epsilon = \epsilon,$$

and hence $\{x_k\}$ is Cauchy under $\|\cdot\|$.

10.2 Infinite-dimensional vector spaces

3. Suppose $\{f_k\}$ is a Cauchy sequence in $C[a,b]$ (under the L^∞ norm) that converges pointwise to $f : [a,b] \to \mathbf{R}$. We wish to prove that $f_k \to f$ in the L^∞ norm. By Theorem 442, $C[a,b]$ is complete under $\|\cdot\|_\infty$, and hence there exists a function $g \in C[a,b]$ such that $\|f_k - g\|_\infty \to 0$ as $k \to \infty$. By the previous exercise, $\{f_k\}$ converges uniformly to g and hence, in particular, $g(x) = \lim_{k\to\infty} f_k(x)$ for all $x \in [a,b]$ (cf. the discussion on page 593 in the text). However, by assumption, $f(x) = \lim_{k\to\infty} f_k(x)$ for all $x \in [a,b]$. This proves that $g(x) = f(x)$ for all $x \in [a,b]$, that is, that $g = f$. Thus $\|f_k - f\|_\infty \to 0$ as $k \to \infty$.

10.3 Functional analysis

3. Let V be a normed vector space. We wish to prove that V^* is complete. Let $\{f_k\}$ be a Cauchy sequence in V^*, let $v \in V$, $v \neq 0$, be fixed, and let $\epsilon > 0$ be given. Then, since $\{f_k\}$ is Cauchy, there exists a positive integer N such that $\|f_n - f_m\|_{V^*} < \epsilon/\|v\|$ for all $m, n \geq N$. By definition of $\|\cdot\|_{V^*}$, it follows that $|f_n(v) - f_m(v)| < \epsilon$ for all $m, n \geq N$. This proves that $\{f_k(v)\}$ is a Cauchy sequence of real numbers and hence converges. We define $f(v) = \lim_{k\to\infty} f_k(v)$. Since v was an arbitrary element of V, this defines $f : V \to \mathbf{R}$. Moreover, it is easy to show that f is linear:

$$f(\alpha v) = \lim_{k\to\infty} f_k(\alpha v) = \lim_{k\to\infty} \alpha f_k(v) = \alpha \lim_{k\to\infty} f_k(v) = \alpha f(v),$$
$$f(u+v) = \lim_{k\to\infty} f_k(u+v) = \lim_{k\to\infty} (f_k(u) + f_k(v)) = \lim_{k\to\infty} f_k(u) + \lim_{k\to\infty} f_k(v) = f(u) + f(v).$$

We can also show that f is bounded. Since $\{f_k\}$ is Cauchy under $\|\cdot\|_{V^*}$, it is easy to show that $\{\|f_k\|_{V^*}\}$ is a bounded sequence of real numbers, that is, that there exists $M > 0$ such that $\|f_k\|_{V^*} \leq M$ for all k. Therefore, if $v \in V$, $\|v\| \leq 1$, then $|f(v)| = |\lim_{k\to\infty} f_k(v)| = \lim_{k\to\infty} |f_k(v)| \leq \lim_{k\to\infty} \|f_k\|_{V^*}\|v\| \leq M\|v\|$. Thus f is bounded, and hence $f \in V^*$. Finally, we must show that $f_k \to f$ under $\|\cdot\|_{V^*}$, that is, that $\|f_k - f\|_{V^*} \to 0$ as $k \to \infty$. Let $\epsilon > 0$ be given. Since $\{f_k\}$ is Cauchy under $\|\cdot\|_{V^*}$, there exists a positive integer N such that $\|f_m - f_n\|_{V^*} < \epsilon/2$ for all $m, n \geq N$. We will show that $|f_n(v) - f(v)| < \epsilon$ for all $v \in V$, $\|v\| \leq 1$, and all $n \geq N$, which then implies that $\|f_n - f\|_{V^*} < \epsilon$ for all $n \geq N$ and completes the proof. For any $v \in V$, $\|v\| \leq 1$, and all $n, m \geq N$, we have

$$|f_n(v) - f(v)| \leq |f_n(v) - f_m(v)| + |f_m(v) - f(v)| < \frac{\epsilon}{2} + |f_m(v) - f(v)|.$$

Moreover, since $f_m(v) \to f(v)$, there exists $m \geq N$ such that $|f_m(v) - f(v)| \leq \epsilon/2$. Therefore, it follows that $|f_n(v) - f(v)| < \epsilon$ for all $n \geq N$. This holds for all $v \in V$ (notice that N is independent of v), and the proof is complete.

10.4 Weak convergence

5. Let V be a normed linear space over $\mathbf{R}$, let x be any vector in V, and let $S = B_r(x) = \{y \in V : \|y - x\| < r\}$, where $r > 0$. Then, for any $y, z \in S$, $\alpha, \beta \in [0,1]$, $\alpha + \beta = 1$, we have

$$\begin{aligned}\|\alpha y + \beta z - x\| = \|\alpha y + \beta z - \alpha x - \beta x\| &= \|\alpha(y-x) + \beta(z-x)\| \\ &\leq \alpha\|y-x\| + \beta\|z-x\| \\ &< \alpha r + \beta r = r.\end{aligned}$$

This shows that $\alpha y + \beta z \in S$, and hence that S is convex.

7. Let $f : \mathbf{R}^n \to \mathbf{R}$ be convex and continuously differentiable, and let $x, y \in \mathbf{R}^n$ be given. By the previous exercise,

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y),$$
$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x).$$

Adding these two equations yields

$$f(x) + f(y) \geq f(y) + f(x) + \nabla f(y) \cdot (x - y) + \nabla f(x) \cdot (y - x).$$

Canceling the common terms and rearranging yields

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0.$$